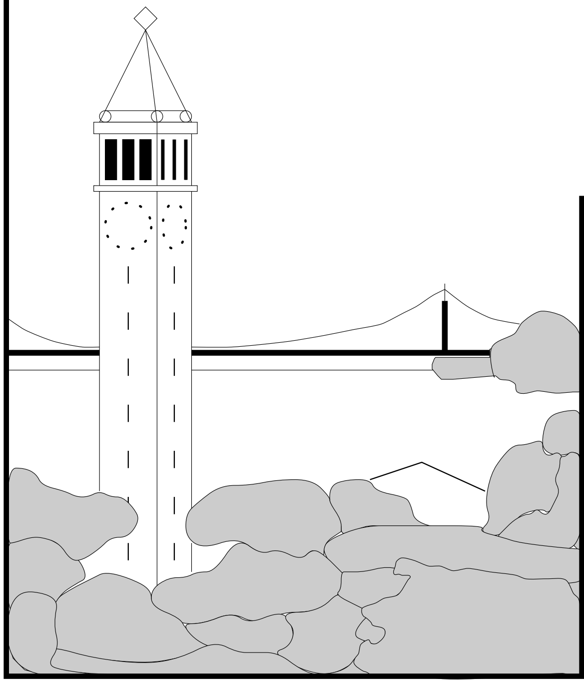
Concurrent Games with Tail Objectives

Krishnendu Chatterjee



Report No. UCB/CSD-5-1385

April 2005

Computer Science Division (EECS)
University of California
Berkeley, California 94720

Concurrent Games with Tail Objectives *

Krishnendu Chatterjee

Department of Electrical Engineering and Computer Sciences
University of California, Berkeley, USA
c_krish@cs.berkeley.edu

April 2005

Abstract

We study infinite stochastic games played by two-players over a finite state space, with objectives specified by sets of infinite traces. The games are *concurrent* (players make moves simultaneously and independently), *stochastic* (the next state is determined by a probability distribution that depends on the current state and chosen moves of the players) and *infinite* (proceeds for infinite number of rounds). The analysis of concurrent stochastic games can be classified into: *quantitative analysis*, analyzing the optimum value of the game and ε-optimal strategies that ensure values within ε of the optimum value; and *qualitative analysis*, analyzing the set of states with optimum value 1 and ε-optimal strategies for the states with optimum value 1. We consider concurrent games with tail objectives, i.e., objectives that are independent of the finite-prefix of traces, and show that the class of tail objectives are strictly richer than the ω-regular objectives. We develop new proof techniques to extend several properties of concurrent games with ω-regular objectives to concurrent games with tail objectives. We prove the *positive limit-one* property for tail objectives, that states for all concurrent games if the optimum value for a player is positive for a tail objective Φ at some state, then there is a state where the optimum value is 1 for objective Φ for the player. We show that the strategies for quantitative winning can be constructed from witnesses of strategies for qualitative winning. The results establish relationship between the quantitative and qualitative analysis of concurrent games

*This research was supported in part by the ONR grant N00014-02-1-0671, the AFOSR MURI grant F49620-00-1-0327, and the NSF ITR grant CCR-0225610.

with tail objectives. We also show that the optimum values of *zero-sum* (strictly conflicting objectives) games with tail objectives can be related to equilibrium values of *nonzero-sum* (not strictly conflicting objectives) games with simpler reachability objectives. A consequence of our analysis presents a polytime reduction of the quantitative analysis of tail objectives to the qualitative analysis for the sub-class of one-player stochastic games (Markov decision processes).

1 Introduction

Stochastic games. Non-cooperative games provide a natural framework to model interactions between agents [15, 16]. A wide class of games progress over time and in stateful manner, and the current game depends on the history of interactions. Infinite *stochastic games* [18, 9] are a natural model for such dynamic games. A stochastic game is played over a finite *state space* and is played in rounds. In concurrent games, in each round, each player chooses an action from a finite set of available actions, simultaneously and independently of the other player. The game proceeds to a new state according to a probabilistic transition relation (stochastic transition matrix) based on the current state and the joint actions of the players. Concurrent games subsume the simpler class of *turn-based games*, where at every state at most one player can choose between multiple actions; and Markov decision processes (MDPs), where only one player can choose between multiple actions at every state. In verification and control of finite state reactive systems such games proceed for infinite rounds, generating a infinite sequence of states, called the *outcome* of the game. The players receive a payoff based on a payoff function that maps every outcome to a real number.

Objectives. Payoffs are generally Borel measurable functions [13]. For example, the payoff set for each player is a Borel set B_i in the Cantor topology on S^ω (where S is the set of states), and player i gets payoff 1 if the outcome of the game is a member of B_i, and 0 otherwise. In verification, payoff functions are usually index sets of *ω-regular languages*. The ω-regular languages generalizes the classical regular languages to infinite strings, they occur in low levels of the Borel hierarchy (they are in $\Sigma_3 \cap \Pi_3$), and they form a robust and expressive language for determining payoffs for commonly used specifications [12]. The simplest ω-regular objectives correspond to safety ("closed sets") and reachability ("open sets") objectives.

Zero-sum games, determinacy and nonzero-sum games. Games may be *zero-sum*, where two players have directly conflicting objectives and the payoff of one player is one minus the payoff of the other, or *nonzero-sum*, where each player has a prescribed payoff function based on the outcome of the game. The fundamental question for games is the existence of equilibrium values. For zero-sum games, this involves showing a *determinacy* theorem that states that the expected optimum value obtained by player 1 is exactly one minus the expected optimum value obtained by player 2. For one-step zero-sum games, this is von Neumann's minmax theorem [21]. For infinite games, the existence of such equilibria is not obvious, in fact, by using the axiom of choice, one can construct games for which determinacy does not hold. However, a remarkable result by Martin [13] shows that all stochastic zero-sum games with Borel payoffs are determined. For nonzero-sum games, the fundamental equilibrium concept is a *Nash equilibrium* [10], that is, a strategy profile such that no player can gain by deviating from the profile, assuming the other player continue playing the strategy in the profile.

Qualitative and quantitative analysis. The analysis of concurrent zero-sum games can be broadly classified into

- *quantitative analysis* that involves analysis of the optimum values of the games, and ε-optimal strategies that ensure values within ε of the optimum value; and
- *qualitative analysis* that involves the simpler analysis of the set of states where the optimum value is 1, and *ε-limit-sure* winning strategies that ensure satisfying the objective with value at least $1 - \varepsilon$.

In general qualitative analysis of concurrent games is simpler as compared to quantitative analysis, as it only considers the case when the value is 1. Optimum values in concurrent games can be irrational even for reachability and safety objectives (with all rational transition probabilities) and hence quantitative analysis requires more involved analysis.

Properties of concurrent games. The result of Martin [13] establishes the determinacy of zero-sum concurrent games for all Borel objectives. The determinacy result sets forth the problem of study and closer understanding of properties and behaviors of concurrent games with different class of objectives. Several interesting questions related to concurrent games are: (1) relationship of qualitative and quantitative analysis; (2) characterizing ε-optimal strategies and ε-limit-sure winning strategies and their relationship; (3) relationship of zero-sum and nonzero-sum games. The results of [6, 7, 1] exhibited several interesting properties for concurrent games with ω-regular objectives specified as parity objectives. The result of [6] showed the positive limit-one property, that states if there is a state with positive optimum value, then there is a state with optimum value 1, for concurrent games with parity objectives. The positive limit-one property and establishing the relation of qualitative and quantitative analysis were key to develop algorithms and improved complexity bound for quantitative analysis concurrent games with parity objectives [1]. The above properties can often be the basic ingredients for the computational complexity analysis of quantitative analysis of concurrent games.

Outline of results. In this work, we consider *tail objectives*, the objectives that do not depend on any finite-prefix of the traces. Tail objectives subsume canonical ω-regular objectives such as parity objectives and Müller objectives, and we show that there exist tail objectives that cannot be expressed as ω-regular objectives. Hence tail objectives are a strictly richer class of objectives than ω-regular objectives. Our result characterizes several properties of concurrent games with tail objectives. The results are as follows:

1. We show the positive limit-one property for concurrent games with tail-objectives. Our result thus extend the result of [6] from parity objectives objectives to a richer class of objective that lie in the higher levels of Borel hierarchy. The result of [6] follows from a complementation argument of quantitative μ-calculus formula. Our proof technique is completely different: it uses a novel strategy construction procedure and a convergence result from martingale theory. It may be noted that the positive limit-one property for concurrent games with Müller objectives follows from the positive limit-one property for parity objectives and the reduction of Müller objectives to parity objectives [20]. Since Müller objectives are tail objectives, our result presents a direct proof for the positive limit-one property for concurrent games with Müller objectives.

2. We establish connection between the complexity of strategies for quantitative winning (ε-optimality) and qualitative winning (ε-limit-sure winning) for tail objectives. We show that witnesses for strategies for quantitative winning can be constructed by composing witnesses of strategies that are qualitative winning in sub-games, and respect certain local conditions.

3. We relate the optimum values of zero-sum games with tail-objectives with Nash-equilibrium values of non-zero sum games with reachabil-

ity objectives. This establishes a relationship between the values of concurrent games with complicated tail objectives and Nash equilibrium of nonzero-sum games with simpler objectives. Our result also presents a polytime reduction of quantitative analysis of tail objectives to qualitative analysis for the special case of MDPs. The above result was previously known for the sub-class of ω-regular objectives [4, 5, 2]. The proof techniques of [4, 5, 2] uses different analysis of the structure of MDPs and is completely different from our proof techniques.

The properties we prove makes it likely that qualitative analysis for concurrent games with tail objectives can be extended to quantitative analysis. The complexity for qualitative analysis of concurrent games and its sub-classes with tail objectives is an open problem.

2 Definitions

Notation. For a countable set A, a *probability distribution* on A is a function $\delta\colon A \mapsto [0,1]$ such that $\sum_{a\in A}\delta(a) = 1$. We denote the set of probability distributions on A by $\mathcal{D}(A)$. Given a distribution $\delta \in \mathcal{D}(A)$, we denote by $\mathrm{Supp}(\delta) = \{x \in A \mid \delta(x) > 0\}$ the *support* of δ.

Definition 1 (Concurrent Games) *A (two-player)* concurrent game structure $G = \langle S, \mathit{Moves}, \mathit{Mv}_1, \mathit{Mv}_2, \delta\rangle$ *consists of the following components:*

- *A finite state space S.*
- *A finite set Moves of moves.*
- *Two move assignments $\mathit{Mv}_1, \mathit{Mv}_2 : S \mapsto 2^{\mathit{Moves}} \setminus \emptyset$. For $i \in \{1,2\}$, assignment Mv_i associates with each state $s \in S$ the non-empty set $\mathit{Mv}_i(s) \subseteq \mathit{Moves}$ of moves available to player i at state s.*
- *A probabilistic transition function $\delta : S \times \mathit{Moves} \times \mathit{Moves} \to \mathcal{D}(S)$, that gives the probability $\delta(s, a_1, a_2)(t)$ of a transition from s to t when player 1 plays a_1 and player 2 plays move a_2, for all $s, t \in S$ and $a_1 \in \mathit{Mv}_1(s)$, $a_2 \in \mathit{Mv}_2(s)$.* ∎

An important special class of concurrent games are Markov decision processes (MDPs). In MDPs at every state s, $|\mathit{Mv}_2(s)| = 1$, i.e., the set of available moves for player 2 is singleton at every state.

At every state $s \in S$, player 1 chooses a move $a_1 \in \mathit{Mv}_1(s)$, and simultaneously and independently player 2 chooses a move $a_2 \in \mathit{Mv}_2(s)$. The game then proceeds to the successor state t with probability $\delta(s, a_1, a_2)(t)$, for all $t \in S$. A state s is called an *absorbing state* if for all $a_1 \in Mv_1(s)$ and $a_2 \in Mv_2(s)$ we have $\delta(s, a_1, a_2)(s) = 1$. In other words, at s for all choice of moves of the players the next state is always s. We assume that the players act *non-cooperatively*, i.e., each player chooses her strategy independently and secretly from the other player, and is only interested in maximizing her own reward. For all states $s \in S$ and moves $a_1 \in Mv_1(s)$ and $a_2 \in Mv_2(s)$, we indicate by $\mathrm{Dest}(s, a_1, a_2) = \mathrm{Supp}(\delta(s, a_1, a_2))$ the set of possible successors of s when moves a_1, a_2 are selected.

A *path* or a *play* ω of G is an infinite sequence $\omega = \langle s_0, s_1, s_2, \ldots \rangle$ of states in S such that for all $k \geq 0$, there are moves $a_1^k \in Mv_1(s_k)$ and $a_2^k \in Mv_2(s_k)$ with $\delta(s_k, a_1^k, a_2^k)(s_{k+1}) > 0$. We denote by Ω the set of all paths and by Ω_s the set of all paths $\omega = \langle s_0, s_1, s_2, \ldots \rangle$ such that $s_0 = s$, i.e., the set of plays starting from state s.

Randomized strategies. A *selector* ξ for player $i \in \{1, 2\}$ is a function $\xi : S \mapsto \mathcal{D}(Moves)$ such that for all $s \in S$ and $a \in Moves$, if $\xi(s)(a) > 0$, then $a \in Mv_i(s)$. We denote by Λ_i the set of all selectors for player $i \in \{1, 2\}$. A selector ξ is *pure* if for every $s \in S$ there is $a \in Moves$ such that $\xi(s)(a) = 1$; we denote by $\Lambda_i^P \subseteq \Lambda_i$ the set of pure selectors for player i. A *strategy* for player 1 is a function $\tau : S^+ \to \Lambda_1$ associates with every finite non-empty sequence of states, representing the history of the play so far, a selector. Similarly we define strategies π for player 2. A strategy τ for player i is *pure* if it yields only pure selectors, that is, is of type $S^+ \to \Lambda_i^P$. A memoryless strategy is independent of the history of the play and depends only on the current state. Memoryless strategies coincide with selectors, and we often write τ for the selector corresponding to a memoryless strategy τ. A strategy is pure memoryless if it is pure and memoryless. We denote by $\Gamma^P, \Gamma^M, \Gamma^{PM}$ the family of pure, memoryless and pure memoryless strategies for player 1 respectively. Analogously we define the families of strategies for player 2. We denote by Γ and Π the set of all strategies for player 1 and player 2, respectively.

Once the starting state s and the strategies τ and π for the two players have been chosen, the game is reduced to an ordinary stochastic process. Hence, the probabilities of events are uniquely defined, where an *event* $\mathcal{A} \subseteq \Omega_s$ is a measurable set of paths. For an event $\mathcal{A} \subseteq \Omega_s$, we denote by $\Pr_s^{\tau,\pi}(\mathcal{A})$ the probability that a path belongs to $\mathcal{A}$ when the game starts from s and the players follows the strategies τ and π. For $i \geq 0$, we also denote by $\Theta_i : \Omega_s \to S$ the random variable denoting the i-th state along a path.

Objectives. We specify objectives for the players by providing the set of *winning plays* $\Phi \subseteq \Omega$ for each player. Given an objective Φ we denote by $\overline{\Phi} = \Omega \setminus \Phi$, the complementary objective of Φ. A concurrent game with objective Φ_1 for player 1 and Φ_2 for player 2 is *zero-sum* if $\Phi_2 = \overline{\Phi_1}$ [17, 9]. A general class of objectives are the Borel objectives [11]. A *Borel objective* $\Phi \subseteq S^\omega$ is a Borel set in the Cantor topology on S^ω. In this paper we consider ω*-regular objectives* [20], which lie in the first $2^1/_2$ levels of the Borel hierarchy (i.e., in the intersection of Σ_3 and Π_3) and *tail objectives* which is a strict superset of ω-regular objectives. The ω-regular objectives, and subclasses thereof, and tail objectives are defined below. For a play $\omega = \langle s_0, s_1, s_2, \ldots \rangle \in \Omega$, we define $\mathrm{Inf}(\omega) = \{ s \in S \mid s_k = s \text{ for infinitely many } k \geq 0 \}$ to be the set of states that occur infinitely often in ω.

- *Reachability and safety objectives.* Given a set $T \subseteq S$ of "target" states, the reachability objective requires that some state of T be visited. The set of winning plays is thus $\mathrm{Reach}(T) = \{ \omega = \langle s_0, s_1, s_2, \ldots \rangle \in \Omega \mid s_k \in T \text{ for some } k \geq 0 \}$. Given a set $F \subseteq S$, the safety objective requires that only states of F be visited. Thus, the set of winning plays is $\mathrm{Safe}(F) = \{ \omega = \langle s_0, s_1, s_2, \ldots \rangle \in \Omega \mid s_k \in F \text{ for all } k \geq 0 \}$.

- *Büchi and coBüchi objectives.* Given a set $B \subseteq S$ of "Büchi" states, the Büchi objective requires that B is visited infinitely often. Formally, the set of winning plays is $\text{Büchi}(B) = \{ \omega \in \Omega \mid \mathrm{Inf}(\omega) \cap B \neq \emptyset \}$. Given $C \subseteq S$, the coBüchi objective requires that all states visited infinitely often are in C. Formally, the set of winning plays is $\text{coBüchi}(C) = \{ \omega \in \Omega \mid \mathrm{Inf}(\omega) \subseteq C \}$.

- *Parity objective.* For $c, d \in \mathbb{N}$, we let $[c..d] = \{ c, c+1, \ldots, d \}$. Let $p : S \mapsto [0..d]$ be a function that assigns a *priority* $p(s)$ to every state $s \in S$, where $d \in \mathbb{N}$. The *Even parity objective* is defined as $\mathrm{Parity}(p) = \{ \omega \in \Omega \mid \min\big(p(\mathrm{Inf}(\omega))\big) \text{ is even } \}$, and the *Odd parity objective* as $\mathrm{coParity}(p) = \{ \omega \in \Omega \mid \min\big(p(\mathrm{Inf}(\omega))\big) \text{ is odd } \}$. Informally we say that a path ω satisfies the parity objective, $\mathrm{Parity}(p)$, if $\omega \in \mathrm{Parity}(p)$.

- *Muller objective.* Given a set $\mathcal{M} \subseteq 2^S$, the *Müller objective* is defined as $\text{Müller}(\mathcal{M}) = \{ \omega \in \Omega \mid \mathrm{Inf}(\omega) \in \mathcal{M} \}$.

- *Tail objective.* Informally the class of tail objectives are the class of objectives that are independent of all finite prefixes. An objective Φ is a tail objective, if the following condition hold: a path $\omega \in \Phi$ if and only if for all $i \geq 0$, $\omega_i \in \Phi$, where ω_i denotes the path ω with the prefix of length i deleted. Formally, let $\mathcal{G}_i = \sigma(\Theta_i, \Theta_{i+1}, \ldots)$ be the σ-field generated by the random-variables $\Theta_i, \Theta_{i+1}, \ldots$. The tail σ-field $\mathcal{T}$ is defined as $\mathcal{T} = \bigcap_{i\geq 0} \mathcal{G}_i$. An objective Φ is a tail objective if and only if Φ belongs to the tail σ-field $\mathcal{T}$, i.e., the tail objectives are indicator functions of events $\mathcal{A} \in \mathcal{T}$.

The Müller and parity objectives are canonical forms to represent ω-regular objectives [14, 19]. Observe that Müller and parity objectives are tail objectives. Note that for a priority function $p : V \to \{ 0, 1 \}$, an even parity objective $\mathrm{Parity}(p)$ is equivalent to the Büchi objective $\mathrm{Büchi}(p^{-1}(0))$, i.e., the Büchi set consists of the states with priority 0. Büchi and coBüchi objectives are special cases of parity objectives and hence tail objectives. Reachability objectives are not necessarily tail objectives, but for a set $T \subseteq S$ of states, if every state $s \in T$ is an absorbing state, then the objective $\mathrm{Reach}(T)$ equivalent to $\mathrm{Büchi}(T)$ and hence is a tail objective. It may be noted that since σ-fields are closed under complementation the class of tail objectives are closed under complementation. We give an example to show that the class of tail objectives are richer than ω-regular objectives.

Example 1 *Let r be a reward function that maps every state s to a real-valued reward $r(s)$, i.e., $r : S \to \mathbb{R}$. For a constant $c \in \mathbb{N}$ consider the objective $\Phi_c = \mathbf{1}_{\limsup_c}$ defined as follows:*

$$\Phi_c = \{ \omega \in \Omega \mid \omega = \langle s_1, s_2, s_3, \ldots \rangle, \lim \sup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} r(s_i) \geq c \}.$$

Intuitively, Φ_c accepts the set of paths such that the "long-run" average of the rewards in the path is at least the constant c. The $\limsup$ *condition lie in the third-level of the Borel-hierarchy (i.e., in Π_3 and Π_3-complete) and cannot be expressed as an ω-regular objective. Hence the class $\cup_{c\in\mathbb{N}} \mathbf{1}_{\limsup_c}$ of objectives cannot be expressed as ω-regular objectives. It may be noted that the "long-run" average of a path is independent of all finite-prefixes of the path. Formally, the class $\cup_{c\in\mathbb{N}} \mathbf{1}_{\limsup_c}$ of objectives are tail objectives. Since* $\limsup_c$ *are Π_3-complete objectives, it follows that tail objectives lie in higher levels of Borel hierarchy than ω-regular objectives.* ∎

Values. The probability that a path satisfies an objective Φ starting from state $s \in S$, given strategies τ, π for the players is $\mathrm{Pr}_s^{\tau,\pi}(\Phi)$. Given a state $s \in S$ and an objective, Φ, we are interested in the maximal probability with which player 1 can ensure that Φ and player 2 can ensure that $\overline{\Phi}$ holds from s. We call such probability the *value of the game* G at s for player $i \in \{1, 2\}$. The value for player 1 and player 2 are given by the function $\langle\langle 1 \rangle\rangle_{val}(\Phi) : S \mapsto [0, 1]$ and $\langle\langle 2 \rangle\rangle_{val}(\overline{\Phi}) : S \mapsto [0, 1]$, defined for all $s \in S$ by

$$\langle\langle 1 \rangle\rangle_{val}(\Phi)(s) = \sup_{\tau \in \Gamma} \inf_{\pi \in \Pi} \Pr_s^{\tau,\pi}(\Phi)$$

$$\langle\langle 2 \rangle\rangle_{val}(\overline{\Phi})(s) = \sup_{\pi \in \Pi} \inf_{\tau \in \Gamma} \Pr_s^{\tau,\pi}(\overline{\Phi}).$$

Note that the objectives of the player are complementary and hence we have a zero-sum game. Concurrent games satisfy a *quantitative* version of determinacy [13], stating that for all Borel-objectives Φ, and all $s \in S$, we have

$$\langle\langle 1 \rangle\rangle_{val}(\Phi)(s) + \langle\langle 2 \rangle\rangle_{val}(\overline{\Phi})(s) = 1.$$

A strategy τ for player 1 is *optimal* for objective Φ if for all $s \in S$ we have

$$\inf_{\pi \in \Pi} \Pr_s^{\tau,\pi}(\Phi) = \langle\langle 1 \rangle\rangle_{val}(\Phi)(s).$$

For $\varepsilon > 0$, a strategy τ for player 1 is *ε-optimal* for objective Φ if for all $s \in S$ we have

$$\inf_{\pi \in \Pi} \Pr_s^{\tau,\pi}(\Phi) \geq \langle\langle 1 \rangle\rangle_{val}(\Phi)(s) - \varepsilon.$$

We define optimal and ε-optimal strategies for player 2 symmetrically. For $\varepsilon > 0$, an objective Φ for player 1 and $\overline{\Phi}$ for player 2, we denote by $\Gamma_\varepsilon(\Phi)$ and $\Pi_\varepsilon(\overline{\Phi})$ the set of ε-optimal strategies for player 1 and player 2, respectively. Note that the quantitative determinacy of concurrent games is equivalent to the existence of ε-optimal strategies for objective Φ for player 1 and $\overline{\Phi}$ for player 2, for all $\varepsilon > 0$, at all states $s \in S$, i.e., for all $\varepsilon > 0$, $\Gamma_\varepsilon(\Phi) \neq \emptyset$ and $\Pi_\varepsilon(\overline{\Phi}) \neq \emptyset$.

We refer to the analysis of *limit-sure winning* states (the set of states s such that $\langle\langle 1 \rangle\rangle_{val}(\Phi)(s) = 1$) and ε-limit-sure winning strategies (ε-optimal strategies for the limit-sure winning states) as the *qualitative* analysis of objective Φ. We refer to the analysis of the values and the ε-optimal strategies as the *quantitative* analysis of objective Φ.

3 Positive Limit-one Property

The *positive limit-one* property for concurrent games for a class $\mathcal{C}$ of objectives states that for all objectives $\Phi \in \mathcal{C}$, for all concurrent games G, if there is a state s such that the value for player 1 is positive at s for objective Φ, then there is a state s' where the value for player 1 is 1 for objective Φ. The property means if a player can win with positive value from some state, then from some state she can win with value 1. The positive limit-one property was proved for parity objectives in [6] and has been one of the key properties used in the algorithmic analysis of concurrent games with parity objectives [1]. In this section we prove the *positive limit-one* property for concurrent games with tail-objectives, and thereby extend the positive limit-one property from parity objectives to a richer class of objectives that subsume several canonical ω-regular objectives. Our proof uses a result from martingale theory and a novel strategy construction, whereas the proof for the sub-class of parity objectives [6] followed from complementation arguments of quantitative μ-calculus formula.

Notation. In the setting of concurrent games the natural filtration sequence $(\mathcal{F}_n)$ for the stochastic process under any pair of strategies is defined as

$$\mathcal{F}_n = \sigma(\Theta_1, \Theta_2, \ldots, \Theta_n)$$

i.e., the σ-field generated by the random-variables $\Theta_1, \Theta_2, \ldots, \Theta_n$.

Lemma 1 (Lévy's 0-1 law) *Suppose $\mathcal{H}_n \uparrow \mathcal{H}_\infty$, i.e.,$\mathcal{H}_n$ is a sequence of increasing σ-fields and $\mathcal{H}_\infty = \sigma(\cup_n \mathcal{H}_n)$. For all events $\mathcal{A} \in \mathcal{H}_\infty$ we have*

$$\mathrm{E}(\mathbf{1}_{\mathcal{A}} \mid \mathcal{H}_n) = \Pr(\mathbf{1}_{\mathcal{A}} \mid \mathcal{H}_n) \to \mathbf{1}_{\mathcal{A}} \textit{ almost-surely, (i.e., with probability 1),}$$

where $\mathbf{1}_{\mathcal{A}}$ is the indicator function of event $\mathcal{A}$.

The proof of the lemma is available in Durrett (page 262—263) [8]. An immediate consequence of Lemma 1 in the setting of concurrent games is the following Lemma.

Lemma 2 (0-1 law in concurrent games) *For all events $\mathcal{A} \in \mathcal{F}_\infty = \sigma(\cup_n \mathcal{F}_n)$, for all strategies $(\tau, \pi) \in \Gamma \times \Pi$, for all states s, we have*

$$\Pr_s^{\tau,\pi}(\mathbf{1}_{\mathcal{A}} \mid \mathcal{F}_n) \to \mathbf{1}_{\mathcal{A}} \textit{ almost-surely,}$$

where $\mathbf{1}_{\mathcal{A}}$ is the indicator function of event $\mathcal{A}$.

Intuitively, the lemma means that the probability $\Pr_s^{\tau,\pi}(\mathbf{1}_{\mathcal{A}} \mid \mathcal{F}_n)$ converges almost-surely (i.e., with probability 1) to 0 or 1 (since indicator functions take values in the range $\{ 0, 1 \}$). Note that the tail σ-field $\mathcal{T}$ is a subset of $\mathcal{F}_\infty$, i.e., $\mathcal{T} \subseteq \mathcal{F}_\infty$, and hence the result of Lemma 2 holds for all $\mathcal{A} \in \mathcal{T}$.

Theorem 1 (Positive limit-one property) *For all concurrent games, for all tail objectives Φ, if there exists a state s such that $\langle\langle 1\rangle\rangle_{val}(\Phi)(s) > 0$, then there exists a state s' such that $\langle\langle 1\rangle\rangle_{val}(\Phi)(s') = 1$.*

Proof. Assume towards contradiction that there exists a state s such that $\langle\langle 1\rangle\rangle_{val}(\Phi)(s) > 0$, but for all states s' we have $\langle\langle 1\rangle\rangle_{val}(\Phi)(s') < 1$. Since $\langle\langle 2\rangle\rangle_{val}(\overline{\Phi})(s') = 1 - \langle\langle 1\rangle\rangle_{val}(\Phi)(s')$, we have $\langle\langle 2\rangle\rangle_{val}(\overline{\Phi})(s') > 0$, for all states s'. Fix η such that $0 < \eta = \min_{s\in S}\langle\langle 2\rangle\rangle_{val}(\overline{\Phi})(s)$. Let $0 < 2\varepsilon < \min\{\, \eta, \langle\langle 1\rangle\rangle_{val}(\Phi)(s) \,\}$. Fix an ε-optimal strategy τ_ε for player 1. We define a sequence of strategies π_i for player 2: let π_0 be an ε-optimal strategy for player 2. The strategy π_{i+1} is defined as follows. For a history $\langle s_1, s_2, \ldots, s_k\rangle$ we have

$$\pi_{i+1}(\langle s_1, s_2, \ldots, s_k\rangle) = \begin{cases} \pi_i(\langle s_1, s_2, \ldots, s_k\rangle) & \text{if } \Pr_s^{\tau_\varepsilon,\pi_i}(\overline{\Phi} \mid \langle s_1, s_2, \ldots, s_k\rangle) \geq \varepsilon \\ \widetilde{\pi}_i(\langle s_1, s_2, s_3, \ldots, s_k\rangle) & \text{if } \Pr_s^{\tau_\varepsilon,\pi_i}(\overline{\Phi} \mid \langle s_1, s_2, \ldots, s_k\rangle) < \varepsilon \\ & \text{where } \widetilde{\pi}_i \text{ is an } \varepsilon\text{-optimal strategy} \\ & \text{from state } s_k \end{cases}$$

Intuitively, the strategy π_{i+i} is as follows: for a history $\langle s_1, s_2, \ldots, s_k\rangle$ if the strategy π_i ensures value greater than ε, then π_{i+1} follows strategy π_i; else it switches to an ε-optimal strategy $\widetilde{\pi}_i$ from state s_k. Since $\widetilde{\pi}_i$ is an ε-optimal strategy from state s_k, $\langle\langle 2\rangle\rangle_{val}(\overline{\Phi})(s) > 2\varepsilon$ for all states s, and $\overline{\Phi}$ is a tail objective that is independent of all finite-prefixes we have $\Pr_s^{\tau_\varepsilon,\pi_{i+1}}(\overline{\Phi}) \geq \Pr_s^{\tau_\varepsilon,\pi_i}(\overline{\Phi})$. Let

$$\Pr_s^{\tau_\varepsilon,\pi_\infty}(\overline{\Phi}) = \lim_{i\to\infty} \Pr_s^{\tau_\varepsilon,\pi_i}(\overline{\Phi});$$

(the limit exists since it is a non-decreasing sequence of values bounded by 1) where $\pi_\infty = \lim_{i\to\infty} \pi_i$. Since $\overline{\Phi}$ is a tail objective (i.e., $\overline{\Phi} \in \lim_{n\to\infty} \sigma(\Theta_n, \Theta_{n+1}, \ldots,))$, it follows that $\Pr_s^{\tau_\varepsilon,\pi_\infty}(\overline{\Phi}) \geq \Pr_s^{\tau_\varepsilon,\pi_i}(\overline{\Phi})$, for all $i \geq 0$. Again by the construction of the sequence of strategies we have

$$\Pr_s^{\tau_\varepsilon,\pi_\infty}(\overline{\Phi} \mid \langle s_1, s_2, s_3, \ldots, s_n\rangle) \geq \varepsilon,$$

for all histories $\langle s_1, s_2, s_3, \ldots, s_n\rangle$. It follows from Lemma 2 that $\Pr_s^{\tau_\varepsilon,\pi_\infty}(\overline{\Phi} \mid \mathcal{F}_n) \to \{\, 0, 1 \,\}$ almost-surely. Hence we conclude that $\Pr_s^{\tau_\varepsilon,\pi_\infty}(\overline{\Phi} \mid \mathcal{F}_n) \to 1$ almost-surely, i.e., $\Pr_s^{\tau_\varepsilon,\pi_\infty}(\Phi \mid \mathcal{F}_n) \to 0$ almost-surely. Since τ_ε is an ε-optimal strategy, we get that $\langle\langle 1\rangle\rangle_{val}(\Phi)(s) < \varepsilon$. But by assumption, $2\varepsilon < \langle\langle 1\rangle\rangle_{val}(\Phi)(s)$ and hence we have a contradiction. Thus the desired result is established. ∎

4 Strategy Characterization for Tail Objectives

In this section we show that in concurrent games with tail objectives, witnesses for ε-optimal strategies can be constructed using witnesses for ε-limit-sure winning strategies of sub-games, that respect certain *local optimality* conditions. The result characterizes the strategy complexity for quantitative optimality for tail objectives in terms of qualitative optimality and local optimality.

We relate the values of zero-sum games with tail-objectives with the Nash equilibrium values of nonzero-sum games with reachability objectives. The result shows that the values of a zero-sum game with complicated objectives can be related to equilibrium values of a nonzero-sum game with simpler objectives. We also show that for MDPs the value function for a tail objective Φ can be computed by computing the maximal probability of reaching the set of states with value 1. As an immediate consequence of the above analysis, we obtain a polytime reduction of the quantitative analysis of MDPs with tail objectives, to the qualitative analysis.

Local optimality. A key notion that will play an important role in the construction of ε-optimal strategies is the notion of *local optimality*. Informally, a selector function ξ is *locally optimal* if it is optimal in the one-step matrix game where each state is assigned a reward value $\langle\langle 1\rangle\rangle_{val}(\Phi)(s)$. A *locally optimal strategy* is a strategy that consists of locally optimal selectors. A *locally ε-optimal strategy* is a strategy that has a total deviation from locally-optimal selectors of at most ε. We note that *local ε-optimality* and *ε-optimality* are very different notions. *Local ε-optimality* consists in the approximation of a local selector; a locally ε-optimal strategy provides no guarantee of yielding a probability of winning the game close to the optimal one.

Definition 2 (Locally ε-optimal selectors and strategies) *A selector ξ is* locally optimal *for objective Φ if for all $s \in S$ and $a_2 \in Mv_s(s)$ we have*

$$\mathrm{E}[\langle\langle 1\rangle\rangle_{val}(\Phi)(\Theta_1) \mid s, \xi(s), a_2] \geq \langle\langle 1\rangle\rangle_{val}(\Phi)(s).$$

We denote by $\Lambda^{\ell}(\Phi)$ the set of locally-optimal selectors for objective Φ. A strategy τ is locally optimal *for objective Φ if for every history $\langle s_0, s_1, \ldots, s_k\rangle$ we have $\tau(\langle s_0, s_1, \ldots, s_k\rangle) \in \Lambda^{\ell}(\Phi)$, i.e., player 1 plays a locally optimal selector at every stage of the play. We denote by $\Gamma^{\ell}(\Phi)$ the set of locally optimal strategies for objective Φ. A strategy τ_ε is* locally ε-optimal *for objective Φ if for every strategy $\pi \in \Pi$, for all $k \geq 1$, for all states s we have*

$$\langle\langle 1\rangle\rangle_{val}(\Phi)(s) - \mathrm{E}_s^{\tau,\pi}[\langle\langle 1\rangle\rangle_{val}(\Phi)(\Theta_k)] \leq \varepsilon.$$

Observe that a strategy that at each round i chooses a locally optimal selector with probability at least $(1-\varepsilon_i)$, with $\sum_{i=0}^{\infty}\varepsilon_i \leq \varepsilon$, is a locally ε-optimal strategy. We denote by $\Gamma_{\varepsilon}^{\ell}(\Phi)$ the set of locally ε-optimal strategies for objective Φ. ∎

We first show that for all tail objectives, for all $\varepsilon > 0$, there exist strategies that are ε-optimal and ε-locally optimal as well.

Lemma 3 *For all tail objectives Φ, for all $\varepsilon > 0$,*

1. $\Gamma_{\frac{\varepsilon}{2}}(\Phi) \subseteq \Gamma_{\varepsilon}^{\ell}(\Phi)$.
2. $\Gamma_{\varepsilon}(\Phi) \cap \Gamma_{\varepsilon}^{\ell}(\Phi) \neq \emptyset$.

Proof. For $\varepsilon > 0$, fix an $\frac{\varepsilon}{2}$-optimal strategy τ for player 1. By definition τ is an ε-optimal strategy as well. We argue that $\tau \in \Gamma_{\varepsilon}^{\ell}(\Phi)$. Assume towards contradiction that $\tau \notin \Gamma_{\varepsilon}^{\ell}(\Phi)$, i.e., there exists a player 2 strategy π, a state s, and k such that

$$\langle\langle 1\rangle\rangle_{val}(\Phi)(s) - \mathrm{E}_s^{\tau,\pi}[\langle\langle 1\rangle\rangle_{val}(\Phi)(\Theta_k)] \geq \varepsilon.$$

Fix a strategy $\pi^* = (\pi + \widetilde{\pi})$ for player 2 as follows: play π for k steps, then switch to an $\frac{\varepsilon}{4}$-optimal strategy $\widetilde{\pi}$. Formally for a history $\langle s_1, s_2, \ldots, s_n\rangle$ we have

$$\pi^*(\langle s_1, s_2, \ldots, s_n\rangle) = \begin{cases} \pi(\langle s_1, s_2, \ldots, s_n\rangle) & \text{if } n \leq k \\ \widetilde{\pi}(\langle s_{k+1}, s_{k+2}, \ldots, s_n\rangle) & \text{if } n > k \\ & \text{where } \widetilde{\pi} \text{ is an } \frac{\varepsilon}{4}\text{-optimal strategy.} \end{cases}$$

Since Φ is a tail objective we have

$$\mathrm{Pr}_s^{\tau,\pi^*}(\Phi) \leq \mathrm{E}_s^{\tau,\pi}[\langle\langle 1\rangle\rangle_{val}(\Phi)(\Theta_k)] + \frac{\varepsilon}{4} \qquad (\text{since } \widetilde{\pi} \text{ is an } \frac{\varepsilon}{4}\text{-optimal strategy}).$$

Hence we have

$$\mathrm{Pr}_s^{\tau,\pi^*}(\Phi) \leq (\langle\langle 1\rangle\rangle_{val}(\Phi)(s) - \varepsilon) + \frac{\varepsilon}{4} = \langle\langle 1\rangle\rangle_{val}(\Phi)(s) - \frac{3\varepsilon}{4} < \langle\langle 1\rangle\rangle_{val}(\Phi)(s) - \frac{\varepsilon}{2}.$$

Since by assumption τ is an $\frac{\varepsilon}{2}$-optimal strategy we have a contradiction. This establishes the desired result. ∎

A *value class* of the game is the set of all states where the game has a given value.

Definition 3 (Value class) *A value class $\mathrm{VC}(r)$ is the set of states s such that the value for player 1 is r. Formally, $\mathrm{VC}(r) = \{s \mid \langle\langle 1\rangle\rangle_{val}(\Phi)(s) = r\}$. Note that for any game there are at most $|S|$ many value classes. By $\mathrm{VC}^{<r}$ we denote the set $\{s \mid \langle\langle 1\rangle\rangle_{val}(\Phi)(s) < r\}$ and similarly we use $\mathrm{VC}^{>r}$ to denote the set $\{s \mid \langle\langle 1\rangle\rangle_{val}(\Phi)(s) > r\}$.* ∎

Reduction. We present a reduction from every value class $\mathrm{VC}(r)$ to a concurrent game $\widetilde{G}_r$ and then establish a few key properties of the game $\widetilde{G}_r$. Let $G = (S, Moves, Mv_1, Mv_2, \delta)$ be a concurrent game with a tail objective Φ for player 1. For a state s we define the set of allowable actions as follows

$$\begin{aligned}\mathrm{OptSupp}(s) \quad = \quad & \{\gamma \subseteq Mv_1(s) \mid \text{ such that there is an optimal selector} \\ & \xi_1^\ell \in \Lambda^\ell(\Phi) \text{ and } \mathrm{Supp}(\xi_1^\ell) = \gamma\}.\end{aligned}$$

Consider a value class $\mathrm{VC}(r)$ with $0 < r < 1$. We construct a concurrent game $\widetilde{G}_r = (\widetilde{S}_r, \widetilde{Moves}, \widetilde{Mv}_1, \widetilde{Mv}_2, \widetilde{\delta})$ as follows:

1. **State space.** Given a state s let $\mathrm{OptSupp}(s) = \{\gamma_1, \gamma_2, \ldots, \gamma_k\}$. Then we have
$$\widetilde{S}_r = \{\, \widetilde{s} \mid s \in \mathrm{VC}(r)\,\} \cup \{\, w_1, w_2\,\}.$$
2. **Moves assignment.** $\widetilde{Mv}_1(\widetilde{s}) = \{1, 2, \ldots, k\}$ such that $\mathrm{OptSupp}(s) = \{\gamma_1, \gamma_2, \ldots, \gamma_k\}$ and $\widetilde{Mv}_2(\widetilde{s}) = Mv_2(s)$.
3. **Transition function.**
   (a) The states w_1 and w_2 are absorbing states such that player 1 have value 1 and 0 at state w_1 and w_2, respectively.
   (b) *Transition function at state $\widetilde{s}$.*
      i. For any move $a_2 \in Mv_2(s)$, if there is a move $a_1 \in \gamma_i$ such that $\sum_{s' \notin \mathrm{VC}(r)} \delta(s, a_1, a_2)(s') > 0$, then $\widetilde{\delta}(\widetilde{s}, i, a_2)(w_1) = 1$. The above transition specifies that if for a move a_2 for player 2 and a move $a_1 \in \gamma_i$ for player 1, if the game G proceeds to a different value class with positive probability then in $\widetilde{G}_r$ the game proceeds to the state w_1, which has value 1 for player 1, with probability 1. Note, that since $a_1 \in \gamma_i$ and $\gamma_i \in \mathrm{OptSupp}(s)$, if in G the game proceeds to a different value class with positive probability it also proceeds to $\mathrm{VC}^{>r}$ with positive probability.
      ii. For any move $a_2 \in Mv_2(s)$, if for every move $a_1 \in \gamma_i$ we have $\sum_{s' \in \mathrm{VC}(r)} \delta(s, a_1, a_2)(s') = 1$, then
$$\widetilde{\delta}(\widetilde{s}, i, a_2)(\widetilde{s'}) = \sum_{a_1 \in \gamma_i} \xi_1^\ell(a_1) \cdot \delta(s, a_1, a_2)(s')$$
where ξ_1^ℓ is an locally optimal selector with $\mathrm{Supp}(\xi_1^\ell) = \gamma_i$.

iii. For any move $a_1 \in (Mv_1(s) \setminus \gamma_i)$, for any move $a_2 \in Mv_2(s)$ we have:

$$\widetilde{\delta}(\widetilde{s}, a_1, a_2)(\widetilde{s'}) = \delta(s, a_1, a_2)(s') \text{ for } s' \in \mathrm{VC}(r);$$

$$\widetilde{\delta}(\widetilde{s}, a_1, a_2)(w_2) = \sum_{s' \notin \mathrm{VC}(r)} \delta(s, a_1, a_2)(s').$$

Notation. Let $U_{>0} = \{ \widetilde{s} \in \widetilde{S}_r \setminus \{ w_2 \} \mid \langle\langle 2 \rangle\rangle_{val}(\overline{\Phi})(\widetilde{s}) > 0 \}$ and $U_1 = \{ \widetilde{s} \in \widetilde{S}_r \setminus \{ w_2 \} \mid \langle\langle 2 \rangle\rangle_{val}(\overline{\Phi})(\widetilde{s}) = 1 \}$.

Strategy maps. We define two strategy maps; $\widetilde{t} : \Gamma \to \widetilde{\Gamma}$ that maps a strategy in the game G to a strategy $\widetilde{\tau} = \widetilde{t}(\tau)$ in the game $\widetilde{G}_r$; and $t : \widetilde{\Pi} \to \Pi$ that maps a strategy in the game $\widetilde{G}_r$ to a strategy $\pi = t(\widetilde{\pi})$ in the game G. The strategy maps are defined as follows:

1. Given a strategy τ in the game G the strategy $\widetilde{\tau} = \widetilde{t}(\tau)$ in the game $\widetilde{G}_r$ is as follows:
   - $\widetilde{\tau}(\widetilde{s_0}, \widetilde{s_1}, \ldots, \widetilde{s_k})(j) = \sum_{a \in \gamma_j} \tau(s_0, s_1, \ldots, s_k)(a)$ where $\gamma_j = \arg\max_{\gamma \in \mathrm{OptSupp}(s_k)} \sum_{a \in \gamma} \tau(s_0, s_1, \ldots, s_k)(a)$; and for all $a' \notin \gamma_j$ we have $\widetilde{\tau}(\widetilde{s_0}, (\widetilde{s_0}, i_0), \widetilde{s_1}, (\widetilde{s_1}, i_1), \ldots, \widetilde{s_k}, (\widetilde{s_k}, j))(a') = \tau(s_0, s_1, \ldots, s_k)(a')$.
2. Given a strategy $\widetilde{\pi}$ in the game $\widetilde{G}_r$ the strategy $\pi = t(\widetilde{\pi})$ in the game G is as follows:

$$\pi(s_0, s_1, \ldots, s_k) = \begin{cases} \widetilde{\pi}(\widetilde{s_0}, \widetilde{s_1}, \ldots, \widetilde{s_k}) & \text{if } \forall i.\ 0 \leq i \leq k. s_i \in \mathrm{VC}(r) \\ \pi' & \text{otherwise; where } \pi' \text{ is an arbitrary strategy} \end{cases}$$

Fact 1. It follows from the construction of the game $\widetilde{G}_r$ that for all strategies $\tau \in \Gamma_\varepsilon^\ell(\Phi)$, for all states $\widetilde{s} \in \widetilde{S}_r \setminus \{ w_2 \}$, for all strategies $\widetilde{\pi}$ for player 2 we have $\mathrm{Pr}_{\widetilde{s}}^{\widetilde{\tau}, \widetilde{\pi}}(\mathrm{Reach}(\{ w_2 \})) \leq \varepsilon$; where $\widetilde{\tau} = \widetilde{t}(\tau)$.

Lemma 4 *Let $\tau_\varepsilon \in \Gamma_\varepsilon^\ell(\Phi) \cap \Gamma_\varepsilon^Q(\Phi)$, for $0 < 2\varepsilon < r$. For all strategies $\widetilde{\pi} \in \widetilde{\Pi}$, for all states $\widetilde{s} \in \widetilde{S}_r$ we have*

$$\mathrm{Pr}_{\widetilde{s}}^{\widetilde{\tau}_\varepsilon, \widetilde{\pi}}(\Phi) \geq r - 2\varepsilon,$$

where $\widetilde{\tau}_\varepsilon = \widetilde{t}(\tau_\varepsilon)$.

Proof. Consider a strategy $\widetilde{\pi}$ in the game $\widetilde{G}_r$ and let $\pi = t(\widetilde{\pi})$. Since τ_ε is an ε-optimal strategy and $\langle\langle 1\rangle\rangle_{val}(\Phi)(s) = r$, we have $\Pr_s^{\tau_\varepsilon,\pi}(\Phi) \geq r - \varepsilon$. Thus we have

$$\begin{array}{rcl} r - \varepsilon \leq \Pr_s^{\tau_\varepsilon,\pi}(\Phi) & = & \Pr_s^{\tau_\varepsilon,\pi}(\Phi \cap \mathrm{Safe}(\mathrm{VC}(r)) \\ & + & \Pr_s^{\tau_\varepsilon,\pi}(\Phi \mid \mathrm{Reach}(\mathrm{VC}^{>r} \cup \mathrm{VC}^{<r})) \cdot \Pr_s^{\tau_\varepsilon,\pi}(\mathrm{Reach}(\mathrm{VC}^{>r} \cup \mathrm{VC}^{<r})) \\ & \leq & \Pr_s^{\tau_\varepsilon,\pi}(\Phi \cap \mathrm{Safe}(\mathrm{VC}(r)) + \Pr_s^{\tau_\varepsilon,\pi}(\mathrm{Reach}(\mathrm{VC}^{>r} \cup \mathrm{VC}^{<r})) \end{array}$$

It follows from the construction of the game $\widetilde{G}_r$ that we have

$$\Pr_s^{\tau_\varepsilon,\pi}(\Phi \cap \mathrm{Safe}(\mathrm{VC}(r)) = \Pr_{\widetilde{s}}^{\widetilde{\tau}_\varepsilon,\widetilde{\pi}}(\Phi \cap \mathrm{Safe}(\widetilde{S}_r)).$$

Since τ_ε is a locally ε-optimal strategy, from Fact 1 and the construction of the game $\widetilde{G}_r$ we have

$$\Pr_{\widetilde{s}}^{\widetilde{\tau}_\varepsilon,\widetilde{\pi}}(\mathrm{Reach}(\{w_1\}) \geq \Pr_{\widetilde{s}}^{\widetilde{\tau}_\varepsilon,\widetilde{\pi}}(\mathrm{Reach}(\{w_1, w_2\})) - \varepsilon = \Pr_s^{\tau_\varepsilon,\pi}(\mathrm{Reach}(\mathrm{VC}^{>r} \cup \mathrm{VC}^{<r})) - \varepsilon.$$

Since w_1 is a winning absorbing state for player 1 we have

$$\begin{array}{rcl} \Pr_{\widetilde{s}}^{\widetilde{\tau}_\varepsilon,\widetilde{\pi}}(\Phi) & = & \Pr_{\widetilde{s}}^{\widetilde{\tau}_\varepsilon,\widetilde{\pi}}(\Phi \cap \mathrm{Safe}(\widetilde{S}_r)) \\ & + & \Pr_{\widetilde{s}}^{\widetilde{\tau}_\varepsilon,\widetilde{\pi}}(\Phi \mid \mathrm{Reach}(\{\, w_1 \,\})) \cdot \Pr_{\widetilde{s}}^{\widetilde{\tau}_\varepsilon,\widetilde{\pi}}(\mathrm{Reach}(\{\, w_1 \,\})) \\ & = & \Pr_{\widetilde{s}}^{\widetilde{\tau}_\varepsilon,\widetilde{\pi}}(\Phi \cap \mathrm{Safe}(\widetilde{S}_r)) + \Pr_{\widetilde{s}}^{\widetilde{\tau}_\varepsilon,\widetilde{\pi}}(\mathrm{Reach}(\{\, w_1 \,\})) \\ & \geq & \Pr_s^{\tau_\varepsilon,\pi}(\Phi \cap \mathrm{Safe}(\mathrm{VC}(r)) + \Pr_s^{\tau_\varepsilon,\pi}(\mathrm{Reach}(\mathrm{VC}^{>r} \cup \mathrm{VC}^{<r})) - \varepsilon \geq r - 2\varepsilon. \end{array}$$

The Lemma follows. ∎

Lemma 5 *If $U_{>0} \neq \emptyset$, then $U_1 \neq \emptyset$.*

Proof. The argument is similar to the proof of Theorem 1. Assume towards contradiction that $U_1 = \emptyset$, $U_{>0} \neq \emptyset$, and let $\widetilde{s} \in U_{>0}$. Fix $0 < 3\varepsilon < \min\{\, r, \langle\langle 2\rangle\rangle_{val}(\overline{\Phi})(\widetilde{s}) \,\}$. Let $\widetilde{\pi}_\varepsilon$ be an ε-optimal strategy for player 2. We construct a sequence of strategies $\widehat{\tau}_\varepsilon^i$ for player 1 as follows:

1. Let $\tau_\varepsilon^0 \in \Gamma_\varepsilon^\ell(\Phi) \cap \Gamma_\varepsilon(\Phi)$, i.e., τ_ε^0 is an ε-optimal and locally ε-optimal strategy in the game G (the fact that $\Gamma_\varepsilon^\ell(\Phi) \cap \Gamma_\varepsilon(\Phi) \neq \emptyset$ follows from Lemma 3). Then $\widehat{\tau}_\varepsilon^0 = \widetilde{\tau}_\varepsilon^0 = \widetilde{t}(\tau_\varepsilon^0)$.

2. The strategy $\widehat{\tau}_\varepsilon^{i+1}$ is inductively defined as follows:

$$\widehat{\tau}_\varepsilon^{i+1}(\langle \widetilde{s_0}, \widetilde{s_1}, \ldots, \widetilde{s_k} \rangle) = \begin{cases} \widehat{\tau}_\varepsilon^i(\langle \widetilde{s_0}, \widetilde{s_1}, \ldots, \widetilde{s_k} \rangle) & \Pr_{\widetilde{s}}^{\widetilde{\tau}_\varepsilon^i,\widetilde{\pi}_\varepsilon}(\Phi \mid \langle \widetilde{s_0}, \widetilde{s_1}, \ldots, \widetilde{s_k} \rangle) \geq \varepsilon \\ \widetilde{\tau'}_\varepsilon(\langle \widetilde{s_0}, \widetilde{s_1}, \ldots, \widetilde{s_k} \rangle) & \Pr_{\widetilde{s}}^{\widetilde{\tau}_\varepsilon^i,\widetilde{\pi}_\varepsilon}(\Phi \mid \langle \widetilde{s_0}, \widetilde{s_1}, \ldots, \widetilde{s_k} \rangle) < \varepsilon, \\ & \text{where } \tau'_\varepsilon \text{ is an } \varepsilon\text{-optimal and} \\ & \text{locally } \varepsilon\text{-optimal strategy} \\ & \text{from } s_k \text{ in } G \text{ and } \widetilde{\tau'}_\varepsilon = \widetilde{t}(\tau'_\varepsilon). \end{cases}$$

Let $\widehat{\tau}_\varepsilon^\infty = \lim_{i\to\infty} \widehat{\tau}_\varepsilon^i$. Since $3\varepsilon < r$ (i.e., $r - 2\varepsilon > \varepsilon$), it follows from Lemma 4 and arguments similar to Theorem 1, that for all histories $\langle s_0, s_1, s_2, \ldots, s_n\rangle$, we have $\Pr_{\widetilde{s}}^{\widehat{\tau}_\varepsilon^\infty, \widetilde{\pi}_\varepsilon}(\Phi \mid \langle s_0, s_1, s_2, \ldots, s_n\rangle) \geq \varepsilon$. It follows from arguments similar to Theorem 1 that we have $\Pr_{\widetilde{s}}^{\widehat{\tau}_\varepsilon^\infty, \widetilde{\pi}_\varepsilon}(\Phi \mid \mathcal{F}_n) \to 1$ almost-surely, i.e., $\Pr_{\widetilde{s}}^{\widehat{\tau}_\varepsilon^\infty, \widetilde{\pi}_\varepsilon}(\overline{\Phi} \mid \mathcal{F}_n) \to 0$ almost-surely. Hence we have a contradiction that $\widetilde{s} \in U_{>0}$. ∎

Lemma 6 *For every $r > 0$, for every state $s \in \mathrm{VC}(r)$, the state $\widetilde{s}$ is a limit-sure winning state in the game $\widetilde{G}_r$ for player 1, i.e., from state $\widetilde{s}$ player 1 can win with probability arbitrarily close to 1.*

Proof. To prove the desired result we show that $U_{>0} = \emptyset$. It follows from Lemma 5 that it suffices to show that $U_1 = \emptyset$. Fix $0 < 2\varepsilon < r$, and let τ_ε be a locally ε-optimal and ε-optimal strategy in G, i.e., $\tau_\varepsilon \in \Gamma_\varepsilon^\ell(\Phi) \cap \Gamma_\varepsilon(\Phi)$ (the fact that $\Gamma_\varepsilon^\ell(\Phi) \cap \Gamma_\varepsilon(\Phi) \neq \emptyset$ follows from Lemma 3). Assume for the sake of contradiction that U_1 is non-empty. Let $\widetilde{s} \in U_1$ and $\widetilde{\pi}_\varepsilon$ be an ε-optimal strategy for player 2 from $\widetilde{s}$. We construct a strategy $\widetilde{\tau}_\varepsilon$ for player 1 in $\widetilde{G}_r$ and a strategy π_ε for player 2 in G as follows:

1. Strategy $\widetilde{\tau}_\varepsilon$ in the game $\widetilde{G}_r$ is defined as $\widetilde{\tau}_\varepsilon = \widetilde{t}(\tau_\varepsilon)$.
2. Strategy π_ε in the game G is defined as $\pi_\varepsilon = t(\widetilde{\pi}_\varepsilon)$.

Since τ_ε is locally ε-optimal, we have $\Pr_{\widetilde{s}}^{\widetilde{\tau}_\varepsilon, \widetilde{\pi}_\varepsilon}(\mathrm{Reach}(\{ w_2 \})) \leq \varepsilon$. Since $\langle\langle 2\rangle\rangle_{val}(\overline{\Phi})(\widetilde{s}) = 1$, and $\widetilde{\pi}_\varepsilon$ is an ε-optimal strategy we have that $\Pr_{\widetilde{s}}^{\widetilde{\tau}_\varepsilon, \widetilde{\pi}_\varepsilon}(\overline{\Phi} \cap \mathrm{Safe}(\widetilde{S}_r \setminus \{ w_1, w_2 \})) \geq 1 - \varepsilon$. Hence it follows that

$$\Pr_s^{\tau_\varepsilon, \pi_\varepsilon}(\overline{\Phi}) \geq \Pr_s^{\tau_\varepsilon, \pi_\varepsilon}(\overline{\Phi} \cap \mathrm{Safe}(\mathrm{VC}(r))) \geq 1 - \varepsilon.$$

Hence we have $\Pr_s^{\tau_\varepsilon, \pi_\varepsilon}(\Phi) \leq \varepsilon$. Since τ_ε is an ε-optimal strategy and $r > 2\varepsilon$, and $\langle\langle 1\rangle\rangle_{val}(\Phi)(s) = r$, we get a contradiction. Thus the desired result follows. ∎

Definition 4 (Value-class qualitative ε-optimal strategy) *For an objective Φ, for $\varepsilon > 0$, a strategy τ_ε is a value-class qualitative ε-optimal strategy for a value-class $\mathrm{VC}(r)$, with $0 < r < 1$, if*

1. *τ_ε is locally ε-optimal.*
2. *for all strategies $\pi \in \Pi$, for all states s, for all histories $\langle s_0, s_1, s_2, \ldots, s_k\rangle$ such that $s_k \in \mathrm{VC}(r)$, $\Pr_s^{\tau_\varepsilon, \pi}(\Phi \mid \langle s_0, s_1, s_2, \ldots, s_k\rangle, \mathit{Safe}(\mathrm{VC}(r))) \geq 1 - \varepsilon$.*

A strategy τ_ε is value-class qualitative ε-optimal for objective Φ if it is value-class qualitative ε-optimal for all value classes $\mathrm{VC}(r)$, for $0 < r < 1$. Value-class qualitative ε-optimal strategies for player 2 are defined similarly. We denote by $\Gamma_\varepsilon^Q(\Phi)$ and $\Pi_\varepsilon^Q(\overline{\Phi})$ the set of value-class qualitative ε-optimal strategies for objectives Φ and $\overline{\Phi}$ for player 1 and player 2, respectively. ∎

Observe that the ε-limit-sure winning strategies in the game $\widetilde{G}_r$ satisfies the requirement for value-class qualitative ε-optimal strategies for value-class $\mathrm{VC}(r)$. The existence of value-class qualitative ε-optimal strategies for tail objectives follows from Lemma 6.

Lemma 7 *For all tail objectives Φ, for all $\varepsilon > 0$, $\Gamma_\varepsilon^Q(\Phi) \neq \emptyset$ and $\Pi_\varepsilon^Q(\overline{\Phi}) \neq \emptyset$.*

We denote by $W_1 = \{ s \mid \langle\langle 1 \rangle\rangle_{val}(\Phi)(s) = 1 \}$ and $W_2 = \{ s \mid \langle\langle 2 \rangle\rangle_{val}(\overline{\Phi})(s) = 1 \}$, the set of states where player 1 and player 2 have values 1, respectively.

Lemma 8 *Let τ_ε be a locally ε-optimal strategy. For all strategies π for player 2, if $\Pr_s^{\tau_\varepsilon,\pi}(Reach(W_1 \cup W_2)) = 1$, then $\Pr_s^{\tau_\varepsilon,\pi}(Reach(W_1)) \geq \langle\langle 1 \rangle\rangle_{val}(\Phi)(s) - \varepsilon$.*

Proof. The results follows from the fact that the sequence $(\langle\langle 1 \rangle\rangle_{val}(\Phi)(\Theta_i))_i$ is a sub-martingale under τ_ε and π. ∎

The following Lemma shows that the value-class qualitative ε-optimal strategies for different value classes can be "stitched" or composed together to produce an ε-optimal strategy. The key argument is as follows: if a play stays in $S \setminus (W_1 \cup W_2)$ then by properties of value-class qualitative ε-optimal strategies player 1 wins with probability 1; else the play reaches $W_1 \cup W_2$ and then ε-optimality is guaranteed by local ε-optimality and Lemma 8. The details of the argument is similar to Lemma 14 in [1].

Lemma 9 (Stitching Lemma) *Let τ_ε be a value-class qualitative ε-optimal strategy and τ_ε is an ε-optimal strategy for all states in W_1. Then τ_ε is an ε-optimal strategy.*

Lemma 7, Lemma 9 and the characterization of value-class qualitative ε-optimal strategies as ε-limit-sure winning strategies in sub-games and locally ε-optimal strategies establishes the following Theorem. The Theorem states that witnesses for ε-optimal strategies can be constructed from witnesses of ε-limit-sure winning strategies in sub-games and locally ε-optimal strategies.

Theorem 2 (Limit-sure to ε-optimality) *For all tail objectives Φ, for $\varepsilon > 0$, let τ_ε be a strategy such that*

1. *τ_ε is locally ε-optimal, i.e., $\tau_\varepsilon \in \Gamma^\ell_\varepsilon(\Phi)$;*
2. *for all value-classes $\mathrm{VC}(r)$, with $r > 0$, for all strategies $\widetilde{\pi}$ in $\widetilde{G}_r$, for all states $\widetilde{s} \in \widetilde{S}_r$, we have $\Pr_{\widetilde{s}}^{\widetilde{\tau}_\varepsilon,\widetilde{\pi}}(\Phi) \geq 1 - \varepsilon$, where $\widetilde{\tau}_\varepsilon = \widetilde{t}(\tau_\varepsilon)$.*

Then τ_ε is ε-optimal, i.e., $\tau_\varepsilon \in \Gamma_\varepsilon(\Phi)$.

Zero-sum tail games and nonzero-sum reachability games. Given a gamegraph G with a tail objective Φ consider the gamegraph G_A such that every state $s \in W_1 \cup W_2$ is transformed to an absorbing state and the states in W_1 are winning for player 1 and the states in W_2 are winning for player 2, i.e., $\langle\langle 1\rangle\rangle_{val}(\Phi)(s) = 1$ for $s \in W_1$ and $\langle\langle 2\rangle\rangle_{val}(\overline{\Phi})(s) = 1$ for $s \in W_2$. Note that for every state s the value for state s in G and G_A are the same. In the following Lemma we show that there exist ε-optimal strategies which if the players follow, then the play reaches $W_1 \cup W_2$ with probability 1. We then extend the result to relate the values of game with tail objectives to equilibrium values of nonzero-sum games with simple reachability objectives.

Lemma 10 *In the gamegraph G_A, let $(\tau,\pi) \in \Gamma^Q_\varepsilon(\Phi) \times \Pi^Q_\varepsilon(\overline{\Phi})$, for sufficiently small ε. Then for all states s we have $\Pr_s^{\tau,\pi}(Reach(W_1 \cup W_2)) = 1$.*

Proof. We first prove that there exists constant $\eta > 0$, such that for all histories $\langle s_0, s_1, s_2, \ldots, s_k\rangle$,

$$\Pr_s^{\tau,\pi}(\mathrm{Reach}(W_1 \cup W_2) \mid \langle s_0, s_1, s_2, \ldots, s_k\rangle) \geq \eta > 0.$$

For all histories $\langle s_0, s_1, s_2, \ldots, s_k\rangle$, such that $s_k \in \mathrm{VC}(r)$ we must have $\Pr_s^{\tau,\pi}(\mathrm{Safe}(\mathrm{VC}(r)) \mid \langle s_0, s_1, s_2 \ldots, s_k\rangle) = 0$. If $\Pr_s^{\tau,\pi}(\mathrm{Safe}(\mathrm{VC}(r)) \mid \langle s_0, s_1, s_2, \ldots, s_k\rangle) > 0$, then by properties of value-class qualitative ε-optimal strategies we have

$$\Pr_s^{\tau,\pi}(\Phi \mid \langle s_0, s_1, s_2, \ldots, s_k\rangle, \mathrm{Safe}(\mathrm{VC}(r))) \geq 1 - \varepsilon,$$

$$\Pr_s^{\tau,\pi}(\overline{\Phi} \mid \langle s_0, s_1, s_2, \ldots, s_k\rangle, \mathrm{Safe}(\mathrm{VC}(r))) \geq 1 - \varepsilon;$$

which is a contradiction for $\varepsilon < 1/2$. Hence for all histories $\langle s_0, s_1, s_2, \ldots, s_k\rangle$ such that $s_k \in \mathrm{VC}(r)$ we have $\Pr_s^{\tau,\pi}(\mathrm{Safe}(\mathrm{VC}(r)) \mid \langle s_0, s_1, s_2, \ldots, s_k\rangle) = 0$. Since value-class qualitative ε-optimal strategies are ε-locally optimal, it follows that there exists constant $\eta' > 0$, such that for all histories $\langle s_0, s_1, s_2, \ldots, s_k\rangle$, if $s_k \in \mathrm{VC}(r)$, then we have

$$\mathrm{Pr}_s^{\tau,\pi}(\mathrm{Reach}(\mathrm{VC}^{>r}) \mid \langle s_0, s_1, s_2, \ldots, s_k\rangle) \geq \eta' > 0,$$

i.e., the play goes to a greater value class with positive probability η'. Since the number of value classes are finite it follows that there exists constant $\eta > 0$ such that

$$\mathrm{Pr}_s^{\tau,\pi}(\mathrm{Reach}(W_1 \cup W_2) \mid \langle s_0, s_1, s_2, \ldots, s_k\rangle) \geq \eta > 0. \tag{1}$$

Since all states in $W_1 \cup W_2$ are absorbing states it follows that $\mathrm{Reach}(W_1 \cup W_2)$ is a tail objective. Hence by Lemma 2 we have $\mathrm{Pr}_s^{\tau,\pi}(\mathrm{Reach}(W_1 \cup W_2) \mid \mathcal{F}_n) \to \{0, 1\}$ almost-surely. It follows from (1) that $\mathrm{Pr}_s^{\tau,\pi}(\mathrm{Reach}(W_1 \cup W_2) \mid \mathcal{F}_n) \to 1$ almost-surely. The desired result follows. ∎

The above Lemma states that if the players play value-class qualitative ε-optimal strategies, for sufficiently small ε, then the play reaches $W_1 \cup W_2$ with probability 1. Since value-class qualitative ε-optimal strategies are ε-optimal strategies (Lemma 9) the following lemma is immediate.

Lemma 11 *Given a gamegraph G with objectives Φ for player 1 and $\overline{\Phi}$ for player 2 we have*

$$\lim_{\varepsilon \to 0} \sup_{\tau \in \Gamma} \inf_{\pi \in \Pi_\varepsilon^Q(\overline{\Phi})} \mathrm{Pr}_s^{\tau,\pi}(\mathit{Reach}(W_1)) = \langle\!\langle 1 \rangle\!\rangle_{val}(\Phi)(s);$$

$$\lim_{\varepsilon \to 0} \sup_{\pi \in \Pi} \inf_{\tau \in \Gamma_\varepsilon^Q(\Phi)} \mathrm{Pr}_s^{\tau,\pi}(\mathit{Reach}(W_2)) = \langle\!\langle 2 \rangle\!\rangle_{val}(\overline{\Phi})(s).$$

Consider a non-zero sum reachability game G_R such that the objectives of both players are reachability objectives: the objective for player 1 is $\mathrm{Reach}(W_1)$ and the objective for player 2 is $\mathrm{Reach}(W_2)$. Note that the game G_R is not zero-sum in the following sense: there are infinite paths ω such that $\omega \not\in \mathrm{Reach}(W_1)$ and $\omega \not\in \mathrm{Reach}(W_2)$ and each player gets a payoff 0 for the path ω. We define ε-Nash equilibrium of the game G_R and relate some special ε-Nash equilibrium of G_R with the values of G.

Definition 5 (ε-Nash equilibrium in G_R) *A strategy profile $(\tau^*, \pi^*) \in \Gamma \times \Pi$ is an ε-Nash equilibrium at state s if the following two conditions hold:*

$$\mathrm{Pr}_s^{\tau^*,\pi^*}(\mathit{Reach}(W_1)) \geq \sup_{\tau \in \Gamma} \mathrm{Pr}_s^{\tau,\pi^*}(\mathit{Reach}(W_1)) - \varepsilon$$

$$\mathrm{Pr}_s^{\tau^*,\pi^*}(\mathit{Reach}(W_2)) \geq \sup_{\pi \in \Pi} \mathrm{Pr}_s^{\tau^*,\pi}(\mathit{Reach}(W_2)) - \varepsilon \ ∎$$

Theorem 3 (Nash equilibrium of reachability game G_R) *The following assertion holds for the game G_R.*

1. *For all $\varepsilon > 0$, there is an ε-Nash equilibrium $(\tau_\varepsilon^*, \pi_\varepsilon^*) \in \Gamma_\varepsilon^Q(\Phi) \times \Pi_\varepsilon^Q(\overline{\Phi})$ such that for all states s we have*

$$\lim_{\varepsilon \to 0} \mathrm{Pr}_s^{\tau_\varepsilon^*, \pi_\varepsilon^*}(\mathit{Reach}(W_1)) = \langle\langle 1 \rangle\rangle_{val}(\Phi)(s)$$

$$\lim_{\varepsilon \to 0} \mathrm{Pr}_s^{\tau_\varepsilon^*, \pi_\varepsilon^*}(\mathit{Reach}(W_2)) = \langle\langle 2 \rangle\rangle_{val}(\overline{\Phi})(s).$$

Proof. It follows from Lemma 11. ∎

Note that in case of MDPs the strategy for player 2 is trivial, i.e., player 2 has only one strategy. Hence in context of MDPs we drop the strategy π of player 2. A specialization of Theorem 3 in case of MDPs yields the following Theorem.

Theorem 4 *For all MDPs G_M, for all tail-objectives Φ, we have*

$$\langle\langle 1 \rangle\rangle_{val}(\Phi)(s) = \sup_{\tau \in \Gamma} \mathrm{Pr}_s^{\tau}(\mathit{Reach}(W_1)) = \langle\langle 1 \rangle\rangle_{val}(\mathit{Reach}(W_1))(s)$$

Since the values in MDPs with reachability objectives can be computed in polynomial time (by linear-programming) [3, 9], our result presents a polytime reduction of quantitative analysis of tail objectives in MDPs to qualitative analysis.

Acknowledgments. I thank Tom Henzinger for many insightful discussions. I am grateful to Luca de Alfaro for several key insights on concurrent games that I received from him. I thank David Aldous for an useful discussion.

References

[1] K. Chatterjee, L. de Alfaro, and T.A. Henzinger. The complexity of quantitative concurrent parity games. Technical Report UCB/CSD-3-1354, UC Berkeley, 2004. (Available also at http://www-cad.eecs.berkeley.edu/∼c_krish/publications/concurrent-complexity.ps).

[2] K. Chatterjee, L. de Alfaro, and T.A. Henzinger. Trading memory for randomness. In *Proceedings of the First Annual Conference on Quantitative Evaluation of Systems*. IEEE Computer Society Press, 2004.

[3] A. Condon. The complexity of stochastic games. *Information and Computation*, 96(2):203–224, 1992.

[4] C. Courcoubetis and M. Yannakakis. The complexity of probabilistic verification. *Journal of the ACM*, 42(4):857–907, 1995.

[5] L. de Alfaro. *Formal Verification of Probabilistic Systems*. PhD thesis, Stanford University, 1997.

[6] L. de Alfaro and T.A. Henzinger. Concurrent omega-regular games. In *Proceedings of the 15th Annual Symposium on Logic in Computer Science*, pages 141–154. IEEE Computer Society Press, 2000.

[7] L. de Alfaro and R. Majumdar. Quantitative solution of omega-regular games. In *STOC 01: Symposium on Theory of Computing*, pages 675–683. ACM Press, 2001. Full version to appear in JCSS.

[8] Richard Durrett. *Probability: Theory and Examples*. Duxbury Press, 1995.

[9] J. Filar and K. Vrieze. *Competitive Markov Decision Processes*. Springer-Verlag, 1997.

[10] J.F. Nash Jr. Equilibrium points in n-person games. *Proceedings of the National Academny of Sciences USA*, 36:48–49, 1950.

[11] A. Kechris. *Classical Descriptive Set Theory*. Springer, 1995.

[12] Z. Manna and A. Pnueli. *The Temporal Logic of Reactive and Concurrent Systems: Specification*. Springer-Verlag, 1992.

[13] D.A. Martin. The determinacy of Blackwell games. *The Journal of Symbolic Logic*, 63(4):1565–1581, 1998.

[14] A.W. Mostowski. Regular expressions for infinite trees and a standard form of automata. In *5th Symposium on Computation Theory*, LNCS 208, pages 157–168. Springer-Verlag, 1984.

[15] G. Owen. *Game Theory*. Academic Press, 1995.

[16] C.H. Papadimitriou. Algorithms, games, and the internet. In *STOC 01: Symposium on Theory of Computing*, pages 749–753. ACM Press, 2001.

[17] T.E.S. Raghavan and J.A. Filar. Algorithms for stochastic games — a survey. *ZOR — Methods and Models of Operations Research*, 35:437–472, 1991.

[18] L.S. Shapley. Stochastic games. *Proc. Nat. Acad. Sci. USA*, 39:1095–1100, 1953.

[19] W. Thomas. On the synthesis of strategies in infinite games. In *STACS 95: Theoretical Aspects of Computer Science*, volume 900 of *Lecture Notes in Computer Science*, pages 1–13. Springer-Verlag, 1995.

[20] W. Thomas. Languages, automata, and logic. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, volume 3, Beyond Words, chapter 7, pages 389–455. Springer, 1997.

[21] J. von Neumann and O. Morgenstern. *Theory of games and economic behavior*. Princeton University Press, 1947.