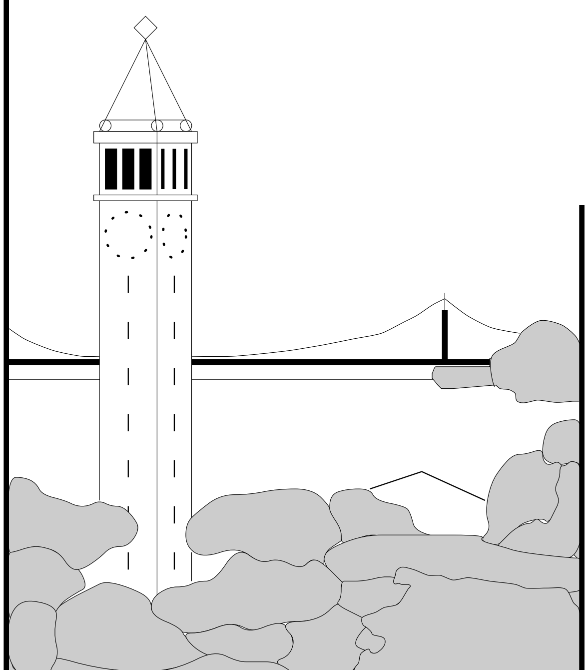
Nash Equilibrium for Upward-Closed Objectives

Krishnendu Chatterjee



Report No. UCB/CSD-5-1407

August 2005

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

Nash Equilibrium for Upward-Closed Objectives *

Krishnendu Chatterjee

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

August 2005

Abstract

We study infinite stochastic games played by n-players on a finite graph with goals specified by sets of infinite traces. The games are *concurrent* (each player simultaneously and independently chooses an action at each round), *stochastic* (the next state is determined by a probability distribution depending on the current state and the chosen actions), *infinite* (the game continues for an infinite number of rounds), *nonzero-sum* (the players' goals are not necessarily conflicting), and undiscounted. We show that if each player has an upward-closed objective, then there exists an ε-Nash equilibrium in memoryless strategies, for every $\varepsilon > 0$; and exact Nash equilibria need not exist. Upward-closure of an objective means that if a set Z of infinitely repeating states is winning, then all supersets of Z of infinitely repeating states are also winning. Memoryless strategies are strategies that are independent of history of plays and depend only on the current state. We also study the complexity of finding values (payoff profile) of an ε-Nash equilibrium. We show that the values of an ε-Nash equilibrium in nonzero-sum concurrent games with upward-closed objectives for all players can be computed by computing ε-Nash equilibrium values of nonzero-sum concurrent games with reachability objectives for all players and a polynomial procedure. As a consequence we establish that values of an ε-Nash equilibrium can be computed in TFNP (total functional NP), and hence in EXPTIME.

*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 grant CCR-0225610.

1 Introduction

Stochastic games. Non-cooperative games provide a natural framework to model interactions between agents [16, 18]. The simplest class of non-cooperative games consists of the "one-step" games — games with single interaction between the agents after which the game ends and the payoffs are decided (e.g., matrix games). However, 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* [20, 9] are a natural model for such 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 other players. 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. In verification and control of finite state reactive systems such games proceed for infinite rounds, generating an 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 [15]. 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 in B_i, and 0 otherwise. In verification, payoff functions are usually index sets of *ω-regular languages.* The ω-regular languages generalize 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. The simplest ω-regular objectives correspond to safety ("closed sets") and reachability ("open sets") objectives.

Zero-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 [25]. 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 [15] shows that all stochastic zero-sum games with Borel payoffs are determined.

Nonzero-sum games. For nonzero-sum games, the fundamental equilibrium concept is a *Nash equilibrium* [11], that is, a strategy profile such that no player can gain by deviating from the profile, assuming the other player continues playing the strategy in the profile. Again, for one-step games, the existence of such equilibria is guaranteed by Nash's theorem [11]. However, the existence of Nash equilibria in infinite games is not immediate: Nash's theorem holds for finite bimatrix games, but in case of stochastic games, the strategy space is not compact. The existence of Nash equilibria is known only in very special cases of stochastic games. In fact, Nash equilibria may not exist, and the best one can hope for is an ε-Nash equilibrium for all $\varepsilon > 0$, where an ε-Nash equilibrium is a strategy profile where unilateral deviation can only increase the payoff of a player by at most ε. Exact Nash equilibria do exist in discounted stochastic games [10]. For concurrent nonzero-sum games with payoffs defined by Borel sets, surprisingly little is known. Secchi and Sudderth [19] showed that exact Nash equilibria do exist when all players have payoffs defined by closed sets ("safety objectives" or Π_1 objectives). In the case of open sets ("reachability objectives" or Σ_1 objectives), the existence of ε-Nash equilibrium for every $\varepsilon > 0$, has been established in [4]. For the special case of two-player games, existence of ε-Nash equilibrium, for every $\varepsilon > 0$, is known for ω-regular objectives [2] and limit-average objectives [23, 24]. The existence of ε-Nash equilibrium in n-player concurrent games with objectives in higher levels of Borel hierarchy than Σ_1 and Π_1 has been an intriguing open problem; existence of ε-Nash equilibrium is not even known even when each player has a Büchi objective.

Result and proof techniques. In this paper we show that ε-Nash equilibrium exists, for every $\varepsilon > 0$, for n-player concurrent games with *upward-closed* objectives. However, exact Nash equilibria need not exist. Informally, an objective Ψ is an upward-closed objective, if a play ω that visits a set Z of states infinitely often is in Ψ, then a play ω' that visits $Z' \supseteq Z$ of states infinitely often is also in Ψ. The class of upward-closed objectives subsumes Büchi and generalized Büchi objectives as special cases. For n-player concurrent games our result extends the existence of ε-Nash equilibrium from the lowest level of Borel hierarchy (open and closed sets) to a class of objectives that lie in the higher levels of Borel hierarchy (upward-closed objectives lie in Π_2) and subsumes several interesting class of objectives. Along with the existence of ε-Nash equilibrium, our result presents a finer characterization of ε-Nash equilibrium showing existence of ε-Nash equilibrium in *memoryless strategies* (strategies that are independent of the history of the play and depend only on the current state). Our result is organized as follows:

1. In Section 3 we develop some results on one player version of concurrent games and n-player concurrent games with reachability objectives.

2. In Section 4 we use induction on the number of players, results of Section 3 and analysis of Markov chains to establish the desired result.

Complexity of ε-Nash equilibrium. Computing the values of a Nash equilibria, when it exists, is another challenging problem [17, 26]. For one-step zero-sum games, equilibrium values and strategies can be computed in polynomial time (by reduction to linear programming) [16]. For one-step nonzero-sum games, no polynomial time algorithm is known to compute an exact Nash equilibrium, even in two-player games [17]. From the computational aspects, a desirable property of an existence proof of Nash equilibrium is its ease of algorithmic analysis. We show that our proof for existence of ε-Nash equilibrium is completely constructive and algorithmic. Our proof shows that the computation of an ε-Nash equilibrium in n-player concurrent games with upward-closed objectives can be achieved by computing ε-Nash equilibrium of games with reachability objectives and a polynomial time procedure. Our result thus shows that computing ε-Nash equilibrium for upward-closed objectives is no harder than solving ε-Nash equilibrium of n-player games with reachability objectives by a polynomial factor. We then prove that an ε-Nash equilibrium can be computed in TFNP (total functional NP) and hence in EXPTIME.

2 Definitions

Notation. For a countable set A, a *probability distribution* on A is a function $\delta\colon A \to [0,1]$ 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 game structures) *An n-player* concurrent game structure $\mathcal{G} = \langle S, A, \Gamma_1, \Gamma_2, \ldots, \Gamma_n, \delta\rangle$ *consists of the following components:*

- *A finite state space S and a finite set A of moves.*
- *Move assignments $\Gamma_1, \Gamma_2, \ldots, \Gamma_n : S \to 2^A \setminus \emptyset$. For $i \in \{ 1, 2, \ldots, n \}$, move assignment Γ_i associates with each state $s \in S$ the non-empty set $\Gamma_i(s) \subseteq A$ of moves available to player i at state s.*
- *A probabilistic transition function $\delta : S \times A \times A \ldots \times A \to \mathcal{D}(S)$, that gives the probability $\delta(s, a_1, a_2, \ldots, a_n)(t)$ of a transition from s to t when player i plays move a_i, for all $s, t \in S$ and $a_i \in \Gamma_i(s)$, for $i \in \{ 1, 2, \ldots, n \}$.* ∎

We define the size of the game structure $\mathcal{G}$ to be equal to the size of the transition function δ; specifically,

$$|\mathcal{G}| = \sum_{s \in S} \sum_{(a_1, a_2, \ldots, a_n) \in \Gamma_1(s) \times \Gamma_2(s) \times \ldots \times \Gamma_n(s)} \sum_{t \in S} |\delta(s, a_1, a_2, \ldots, a_n)(t)|,$$

where $|\delta(s, a_1, a_2, \ldots, a_n)(t)|$ denotes the space to specify the probability distribution. At every state $s \in S$, each player i chooses a move $a_i \in \Gamma_i(s)$, and simultaneously and independently of the other players, and the game then proceeds to the successor state t with probability $\delta(s, a_1, a_2, \ldots, a_n)(t)$, for all $t \in S$. A state s is called an *absorbing state* if for all $a_i \in \Gamma_i(s)$ we have $\delta(s, a_1, a_2, \ldots, a_n)(s) = 1$. In other words, at s for all choices of moves of the players the next state is always s. For all states $s \in S$ and moves $a_i \in \Gamma_i(s)$ we indicate by $\mathrm{Dest}(s, a_1, a_2, \ldots, a_n) = \mathrm{Supp}(\delta(s, a_1, a_2, \ldots, a_n))$ the set of possible successors of s when moves $a_1, a_2, \ldots, a_n$ are selected.

A *path* or a *play* ω of $\mathcal{G}$ 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 $a_i^k \in \Gamma_i(s_k)$ and with $\delta(s_k, a_1^k, a_2^k, \ldots, a_n^k)(s_{k+1}) > 0$. 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_i for player $i \in \{ 1, 2, \ldots, n \}$ is a function $\xi_i : S \to \mathcal{D}(A)$ such that for all $s \in S$ and $a \in A$, if $\xi_i(s)(a) > 0$ then $a \in \Gamma_i(s)$. We denote by Λ_i the set of all selectors for player $i \in \{1, 2, \ldots, n\}$. A *strategy* σ_i for player i is a function $\sigma_i : S^+ \to \Lambda_i$ that associates with every finite non-empty sequence of states, representing the history of the play so far, a selector. 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 σ_i for the selector corresponding to a memoryless strategy σ_i. A memoryless strategy σ_i for player i is *uniform memoryless* if the selector of the memoryless strategy is an uniform distribution over its support, i.e., for all states s we have $\sigma_i(s)(a_i) = 0$ if $a_i \notin \mathrm{Supp}(\sigma_i(s))$ and $\sigma_i(s)(a_i) = \frac{1}{|\mathrm{Supp}(\sigma_i(s))|}$ if $a_i \in \mathrm{Supp}(\sigma_i(s))$. We denote by Σ_i, Σ_i^M and Σ_i^{UM} the set of all strategies, set of all memoryless strategies and the set of all uniform memoryless strategies for player i, respectively. Given strategies σ_i for player i, we denote by $\overline{\sigma}$ the strategy profile $\langle \sigma_1, \sigma_2, \ldots, \sigma_n \rangle$. A strategy profile $\overline{\sigma}$ is memoryless (resp. uniform memoryless) if all the component strategies are memoryless (resp. uniform memoryless).

Given a strategy profile $\overline{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ and a state s, we denote by $\mathrm{Outcome}(s, \overline{\sigma}) = \{ \omega = \langle s_0, s_1, s_2 \ldots \rangle \mid s_0 = s \text{ and } \exists a_i^k.\ \sigma_i(\langle s_0, s_1, \ldots, s_k \rangle)(a_i^k) > 0. \text{ and } \delta(s_k, a_1^k, a_2^k, \ldots, a_n^k)(s_{k+1}) > 0 \}$ the set of all possible plays from s, given $\overline{\sigma}$. Once the starting state s and the strategies σ_i for the 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 $\mathrm{Pr}_s^{\overline{\sigma}}(\mathcal{A})$ the probability that a path belongs to $\mathcal{A}$ when the game starts from s and the players follow the strategies σ_i, and $\overline{\sigma} = \langle \sigma_1, \sigma_2, \ldots, \sigma_n \rangle$.

Objectives. Objectives for the players in nonterminating games are specified by providing the set of *winning plays* $\Psi \subseteq \Omega$ for each player. A general class of objectives are the Borel objectives [13]. A *Borel objective* $\Phi \subseteq S^\omega$ is a Borel set in the Cantor topology on S^ω. The class of ω*-regular objectives* [21], lie in the first $2^1\!/\!_2$ levels of the Borel hierarchy (i.e., in the intersection of Σ_3 and Π_3). The ω-regular objectives, and subclasses thereof, can be specified in the following forms. 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.

1. *Reachability and safety objectives.* Given a game graph $\mathcal{G}$, and a set $T \subseteq S$ of *target* states, the reachability specification $\mathrm{Reach}(T)$ requires that some state in T be visited. The reachability specification $\mathrm{Reach}(T)$ defines the objective $[\![\mathrm{Reach}(T)]\!] = \{ \langle s_0, s_1, s_2, \ldots \rangle \in \Omega \mid \exists k \geq 0.\ s_k \in T \}$ of winning plays. Given a set $F \subseteq S$ of *safe* states, the safety specification $\mathrm{Safe}(F)$ requires that only states in F be visited. The safety specification $\mathrm{Safe}(F)$ defines the objective $[\![\mathrm{Safe}(F)]\!] = \{ \langle s_0, s_1, \ldots \rangle \in \Omega \mid \forall k \geq 0.\ s_k \in F \}$ of winning of plays.

2. *Büchi and generalized Büchi objectives.* Given a game graph $\mathcal{G}$, and a set $B \subseteq S$ of *Büchi* states, the Büchi specification $\mathrm{Büchi}(B)$ requires that states in B be visited infinitely often. The Büchi specification

Büchi(B) defines the objective $[\![\text{Büchi}(B)]\!] = \{\omega \in \Omega \mid \text{Inf}(\omega) \cap B \neq \emptyset\}$ of winning plays. Let $B_1, B_2, \ldots, B_n$ be subset of states, i.e., each $B_i \subseteq S$. The generalized Büchi specification is the requires that every Büchi specification Büchi(B_i) be satisfied. Formally, the generalized Büchi objective is $\bigcap_{i \in \{1,2,\ldots,n\}} [\![\text{Büchi}(B_i)]\!]$.

3. *Müller and upward-closed objectives.* Given a set $M \subseteq 2^S$ of *Müller* set of states, the Müller specification Müller(M) requires that the set of states visited infinitely often in a play is exactly one of the sets in M. The Müller specification Müller(M) defines the objective $[\![\text{Müller}(M)]\!] = \{\omega \in \Omega \mid \text{Inf}(\omega) \in M\}$ of winning plays. The *upward-closed* objectives form a sub-class of Müller objectives, with the restriction that the set M is upward-closed. Formally a set $UC \subseteq 2^S$ is upward-closed if the following condition hold: if $U \in UC$ and $U \subseteq Z$, then $Z \in UC$. Given a upward-closed set $UC \subseteq 2^S$, the upward-closed objective is defined as the set $[\![\text{UpClo}(UC)]\!] = \{\omega \in \Omega \mid \text{Inf}(\omega) \in UC\}$ of winning plays.

Observe that the upward-closed objectives specifies that if a play ω that visits a subset U of states visited infinitely often is winning, then a play ω' that visits a superset of U of states infinitely often is also winning. The upward-closed objectives subsumes Büchi and generalized Büchi (i.e., conjunction of Büchi) objectives. The upward-closed objectives also subsumes disjunction of Büchi objectives. Since the Büchi objectives lie in the second level of the Borel hierarchy (in Π_2), it follows that upward-closed objectives can express objectives that lie in Π_2. Müller objectives are canonical forms to express ω-regular objectives, and the class of upward-closed objectives form a strict subset of Müller objectives and cannot express all ω-regular properties.

We write Ψ for an arbitrary objective. We write the objective of player i as Ψ_i. Given a Müller objective Ψ, the set of paths Ψ is measurable for any choice of strategies for the players [22]. Hence, the probability that a path satisfies a Müller objective Ψ starting from state $s \in S$ under a strategy profile $\overline{\sigma}$ is $\Pr_s^{\overline{\sigma}}(\Psi)$.

Notations. Given a strategy profile $\overline{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_n)$, we denote by $\overline{\sigma}_{-i} = (\sigma_1, \sigma_2, \ldots, \sigma_{i-1}, \sigma_{i+1}, \ldots, \sigma_n)$ the strategy profile with the strategy for player i removed. Given a strategy $\sigma'_i \in \Sigma_i$, and a strategy profile $\overline{\sigma}_{-i}$, we denote by $\overline{\sigma}_{-i} \cup \sigma'_i$ the strategy profile $(\sigma_1, \sigma_2, \ldots, \sigma_{i-1}, \sigma'_i, \sigma_{i+1}, \ldots, \sigma_n)$. We also use the following notations: $\overline{\Sigma} = \Sigma_1 \times \Sigma_2 \times \ldots \times \Sigma_n$; $\overline{\Sigma}^M = \Sigma_1^M \times \Sigma_2^M \times \ldots \times \Sigma_n^M$; $\overline{\Sigma}^{UM} = \Sigma_1^{UM} \times \Sigma_2^{UM} \times \ldots \times \Sigma_n^{UM}$; and $\overline{\Sigma}_{-i} = \Sigma_1 \times \Sigma_2 \times$

$\dots \Sigma_{i-1} \times \Sigma_{i+1} \times \dots \Sigma_n$. The notations for $\overline{\Sigma}_{-i}^{M}$ and $\overline{\Sigma}_{-i}^{UM}$ are similar. For $n \in \mathbb{N}$, we denote by $[n]$ the set $\{ 1, 2, \dots, n \}$.

Concurrent nonzero-sum games. A concurrent nonzero-sum game consists of a concurrent game structure $\mathcal{G}$ with objective Ψ_i for player i. The zero-sum values for the players in concurrent games with objective Ψ_i for player i are defined as follows.

Definition 2 (Zero-sum values) *Let $\mathcal{G}$ be a concurrent game structure with objective Ψ_i for player i. Given a state $s \in S$ we call the maximal probability with which player i can ensure that Ψ_i holds from s against all strategies of the other players is the* zero-sum value *of player i at s. Formally, the zero-sum value for player i is given by the function $val_i^{\mathcal{G}}(\Psi_i) : S \to [0,1]$ defined for all $s \in S$ by*

$$val_i^{\mathcal{G}}(\Psi_i)(s) = \sup_{\sigma_i' \in \Sigma_i} \inf_{\overline{\sigma}_{-i} \in \overline{\Sigma}_{-i}} \Pr_s^{\overline{\sigma}_{-i} \cup \sigma_i'}(\Psi_i). \blacksquare$$

A two-player concurrent game structure $\mathcal{G}$ with objectives Ψ_1 and Ψ_2 for player 1 and player 2, respectively, is *zero-sum* if the objectives of the players are complementary, i.e., $\Psi_1 = \Omega \setminus \Psi_2$. Concurrent zero-sum games satisfy a *quantitative* version of determinacy [15], stating that for all two-player concurrent games with Müller objectives Ψ_1 and Ψ_2, such that $\Psi_1 = \Omega \setminus \Psi_2$, and all $s \in S$, we have

$$val_1^{\mathcal{G}}(\Psi_1)(s) + val_2^{\mathcal{G}}(\Psi_2)(s) = 1.$$

The determinacy also establishes existence of ε-Nash equilibrium, for all $\varepsilon > 0$, in concurrent zero-sum games.

Definition 3 (ε-Nash equilibrium) *Let $\mathcal{G}$ be a concurrent game structure with objective Ψ_i for player i. For $\varepsilon \geq 0$, a strategy profile $\overline{\sigma}^* = (\sigma_1^*, \dots, \sigma_n^*) \in \overline{\Sigma}$ is an ε-Nash equilibrium for a state $s \in S$ iff the following condition hold for all $i \in [n]$:*

$$\sup_{\sigma_i \in \Sigma_i} \Pr_s^{\overline{\sigma}^*_{-i} \cup \sigma_i}(\Psi_i) \leq \Pr_s^{\overline{\sigma}^*}(\Psi_i) + \varepsilon.$$

A Nash equilibrium is an ε-Nash equilibrium with $\varepsilon = 0$. $\blacksquare$

Example 1 (ε-Nash equilibrium) *Consider the two-player game structure shown in Fig. 1.(a). The state s_1 and s_2 are absorbing states and the* *set of available moves for player 1 and player 2 at* s_0 *is* $\{ a, b \}$ *and* $\{ c, d \}$, *respectively. The transition function is defined as follows:*

$$\delta(s_0, a, c)(s_0) = 1; \qquad \delta(s_0, b, d)(s_2) = 1; \qquad \delta(s_0, a, d)(s_1) = \delta(s_0, b, c)(s_1) = 1.$$

The objective of player 1 is an upward-closed objective $[\![UpClo(UC_1)]\!]$ *with* $UC_1 = \{ \{ s_1 \}, \{ s_1, s_2 \}, \{ s_0, s_1 \}, \{ s_0, s_1, s_2 \} \}$, *i.e., the objective of player 1 is to visit* s_1 *infinitely often (i.e.,*$[\![Büchi(\{ s_1 \})]\!]$*). Since* s_1 *is an absorbing state in the game shown, the objective of player 1 is equivalent to* $[\![Reach(\{ s_1 \})]\!]$. *The objective of player 2 is an upward-closed objective* $[\![UpClo(UC_2)]\!]$ *with* $UC_2 = \{ \{ s_2 \}, \{ s_1, s_2 \}, \{ s_0, s_2 \}, \{ s_0, s_1, s_2 \}, \{ s_0 \}, \{ s_0, s_1 \} \}$. *Observe that any play* ω *such that* $\mathrm{Inf}(\omega) \neq \{ s_1 \}$ *is winning for player 2. Hence the objective of player 1 and player 2 are complementary. For* $\varepsilon > 0$, *consider the memoryless strategy* $\sigma_1^\varepsilon \in \Sigma^M$ *that plays move* a *with probability* $1 - \varepsilon$, *and move* b *with probability* ε. *The game starts at* s_0, *and in each round if player 2 plays move* c, *then the play reaches* s_1 *with probability* ε *and stays in* s_0 *with probability* $1 - \varepsilon$; *whereas if player 2 plays move* d, *then the game reaches state* s_1 *with probability* $1 - \varepsilon$ *and state* s_2 *with probability* ε. *Hence it is easy to argue against all strategies* σ_2 *for player 2, given the strategy* σ_1^ε *of player 1, the game reaches* s_1 *with probability at least* $1 - \varepsilon$. *Hence for all* $\varepsilon > 0$, *there exists a strategy* σ_1^ε *for player 1, such that against all strategies* σ_2, *we have* $\mathrm{Pr}_{s_0}^{\sigma_1^\varepsilon, \sigma_2}([\![Reach(\{ s_1 \})]\!]) \geq 1 - \varepsilon$; *hence* $(1 - \varepsilon, \varepsilon)$ *is an* ε*-Nash equilibrium value profile at* s_0. *However, we argue that* $(1, 0)$ *is not an Nash equilibrium at* s_0. *To prove the claim, given a strategy* σ_1 *for player 1 consider the counter strategy* σ_2 *for player 2 as follows: for* $k \geq 0$, *at round* k, *if player 1 plays move* a *with probability 1, then player 2 chooses the move* c *and ensures that the state* s_1 *is reached with probability 0; otherwise if player 1 plays move* b *with positive probability at round* k, *then player 2 chooses move* d, *and the play reaches* s_2 *with positive probability. That is either* s_2 *is reached with positive probability or* s_0 *is visited infinitely often. Hence player 1 cannot satisfy* $[\![UpClo(UC_1)]\!]$ *with probability 1. This shows that in game structures with upward-closed objectives, Nash equilibrium need not exist and* ε*-Nash equilibrium, for all* $\varepsilon > 0$, *is the best one can achieve.*

Consider the game shown in Fig. 1.(b). The transition function at state s_0 *is same as in Fig 1.(a). The state* s_2 *is an absorbing state and from state* s_1 *the next state is always* s_0. *The objective for player 1 is same as in the previous example, i.e.,* $[\![Büchi(\{ s_1 \})]\!]$. *Consider any upward-closed objective* $[\![UpClo(UC_2)]\!]$ *for player 2. We claim that the following strategy profile* (σ_1, σ_2) *is a Nash equilibrium at* s_0: σ_1 *is memoryless strategy that plays* a

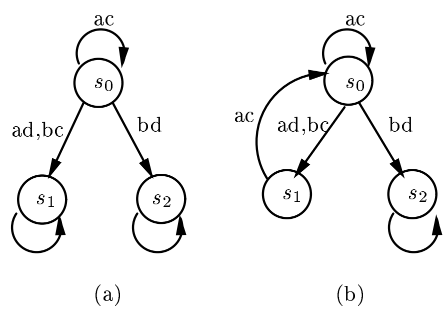


Figure 1: Examples of Nash equilibrium in two-player concurrent game structures

with probability 1, and σ_2 *is a memoryless strategy that plays c and d with probability* $1/2$. *Given* σ_1 *and* σ_2 *we have that the states* s_0 *and* s_1 *are visited infinitely often with probability 1. If* $\{ s_0, s_1 \} \in UC_2$, *then the objectives of both players are satisfied with probability 1. If* $\{ s_0, s_1 \} \notin UC_2$, *then no subset of* $\{ s_0, s_2 \}$ *is in* UC_2. *Given strategy* σ_1, *for all strategies* σ_2' *of player 2, the plays under* σ_1 *and* σ_2' *visits subsets of* $\{ s_0, s_1 \}$ *infinitely often. Hence, if* $\{ s_0, s_1 \} \notin UC_2$, *then given* σ_1, *for all strategies of player 2, the objective* $[\![UpClo(UC_2)]\!]$ *is satisfied with probability 0, hence player 2 has no incentive to deviate from* σ_2. *The claim follows. Note that the present example can be contrasted to the zero-sum game on the same game structure with objective* $[\![\textit{Büchi}(\{ s_1 \})]\!]$ *for player 1 and the complementary objective for player 2 (which is not upward-closed). In the zero-sum case,* ε*-optimal strategies require infinite-memory (see [7]) for player 1. In the case of nonzero-sum game with upward-closed objectives (which do not generalize the zero-sum case) we exhibited existence of memoryless Nash equilibrium.* ∎

3 Markov Decision Processes and Nash Equilibrium for Reachability Objectives

The section is divided in two parts: subsection 3.1 develops facts about one player concurrent game structures and subsection 3.2 develops facts about n-player concurrent game structures with reachability objectives. The facts developed in this section will play a key role in the analysis of the later sections.

3.1 Markov decision processes

In this section we develop some facts about one player versions of concurrent game structures, known as *Markov decision processes* (MDPs) [1]. For $i \in [n]$, a player i-MDP is a concurrent game structure where for all $s \in S$, for all $j \in [n] \setminus \{ i \}$ we have $|\Gamma_j(s)| = 1$, i.e., at every state only player i can choose between multiple moves and the choice for the other players are singleton. If for all states $s \in S$, for all $i \in [n]$, $|\Gamma_i(s)| = 1$, then we have a *Markov chain*. Given a concurrent game structure $\mathcal{G}$, if we fix a memoryless strategy profile $\overline{\sigma}_{-i} = (\sigma_1, \ldots, \sigma_{i-1}, \sigma_{i+1}, \ldots, \sigma_n)$ for players in $[n] \setminus \{ i \}$, the game structure is equivalent to a player i-MDP $\mathcal{G}_{\overline{\sigma}_{-i}}$ with transition function

$$\delta_{\overline{\sigma}_{-i}}(s, a_i)(t) = \sum_{(a_1, a_2, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n)} \delta(s, a_1, a_2, \ldots, a_n)(t) \cdot \prod_{j \in ([n] \setminus \{ i \})} \sigma_j(s)(a_j),$$

for all $s \in S$ and $a_i \in \Gamma_i(s)$. Similarly, if we fix a memoryless strategy profile $\overline{\sigma} \in \overline{\Sigma}^M$ for a concurrent game structure $\mathcal{G}$, we obtain a Markov chain, which we denote by $\mathcal{G}_{\overline{\sigma}}$. In an MDP, the sets of states that play an equivalent role to the closed recurrent set of states in Markov chains [14] are called *end components* [5, 6]. Without loss of generality, we consider player 1-MDPs and since the set $\overline{\Sigma}_{-1}$ is singleton for player 1-MDPs we only consider strategies for player 1.

Definition 4 (End components and maximal end components) *Given a player 1-MDP $\mathcal{G}$, an* end component *(EC) in $\mathcal{G}$ is a subset $C \subseteq S$ such that there is a memoryless strategy $\sigma_1 \in \Sigma_1^M$ for player 1 under which C forms a closed recurrent set in the resulting Markov chain, i.e., in the Markov chain $\mathcal{G}_{\sigma_1}$. Given a player 1-MDP $\mathcal{G}$, an end component C is a maximal end component, if the following condition hold: if $C \subseteq Z$ and Z is an end component, then $C = Z$, i.e., there is no end component that encloses C.*

Graph of a MDP. Given a player 1-MDP $\mathcal{G}$, the graph of $\mathcal{G}$ is a directed graph (S, E) with the set E of edges defined as follows: $E = \{ (s,t) \mid s,t \in S.\ \exists a_1 \in \Gamma_1(s).\ t \in \mathrm{Dest}(s, a_1) \}$, i.e., $E(s) = \{ t \mid (s,t) \in E \}$ denotes the set of possible successors of the state s in the MDP $\mathcal{G}$.

Equivalent characterization. An equivalent characterization of an end component C is as follows: for each $s \in C$, there is a subset of moves $M_1(s) \subseteq \Gamma_1(s)$ such that:

1. when a move in $M_1(s)$ is chosen at s, all the states that can be reached with non-zero probability are in C, i.e., for all $s \in C$, for all $a \in M_1(s)$, $\mathrm{Dest}(s,a) \subseteq C$;

2. the graph (C, E) is strongly connected, where E consists of the transitions that occur with non-zero probability when moves in $M_1(\cdot)$ are taken, i.e., $E = \{\, (s,t) \mid s,t \in S.\ \exists a \in M_1(s).\ t \in \mathrm{Dest}(s,a) \,\}$.

Given a set $\mathcal{F} \subseteq 2^S$ of subset of states we denote by $\mathrm{InfSt}(\mathcal{F})$ the event $\{\, \omega \mid \mathrm{InfSt}(\omega) \in \mathcal{F} \,\}$. The following lemma states that in a player 1-MDP, for all strategies of player 1, the set of states visited infinitely often is an end component with probability 1. Lemma 2 follows easily from Lemma 1.

Lemma 1 ([6, 5]) *Let $\mathcal{C}$ be the set of end components of a player 1-MDP $\mathcal{G}$. For all strategies $\sigma_1 \in \Sigma_1$ and all states $s \in S$, we have $\mathrm{Pr}_s^{\sigma_1}(\mathrm{InfSt}(\mathcal{C})) = 1$.*

Lemma 2 *Let $\mathcal{C}$ be the set of end components and $\mathcal{Z}$ be the set of maximal end components of a player 1-MDP $\mathcal{G}$. Then the following assertions hold:*

- $L = \bigcup_{C \in \mathcal{C}} C = \bigcup_{Z \in \mathcal{Z}} Z$; *and*
- *for all strategies $\sigma_1 \in \Sigma_1$ and all states $s \in S$, we have $\mathrm{Pr}_s^{\sigma_1}([\![Reach(L)]\!]) = 1$.*

Lemma 3 *Given a player 1-MDP $\mathcal{G}$ and an end component C, there is a uniform memoryless strategy $\sigma_1 \in \Sigma_1^{UM}$, such that for all states $s \in C$, we have $\mathrm{Pr}_s^{\sigma_1}(\{\, \omega \mid \mathrm{Inf}(\omega) = C \,\}) = 1$.*

Proof. For a state $s \in C$, let $M_1(s) \subseteq \Gamma_1(s)$, be the subset of moves such that the conditions of the equivalent characterization of end components hold. Consider the uniform memoryless strategy σ_1 defined as follows: for all states $s \in C$,

$$\sigma_1(s)(a) = \begin{cases} \frac{1}{|M_1(s)|}, & \text{if } a \in M_1(s) \\ 0 & \text{otherwise.} \end{cases}$$

Given the strategy σ_1, in the Markov chain $\mathcal{G}_{\sigma_1}$, the set C is a closed recurrent set of states. Hence the result follows. ∎

3.2 Nash equilibrium for reachability objectives

Memoryless Nash equilibrium in discounted games. We first prove the existence of Nash equilibrium in memoryless strategies in n-player discounted games with reachability objective $[\![\mathrm{Reach}(R_i)]\!]$ for player i, for $R_i \subseteq S$. We then characterize ε-Nash equilibrium in memoryless strategies with some special property in n-player discounted games with reachability objectives.

Definition 5 (β-discounted games) *Given an n-player game structure $\mathcal{G}$ we write $\mathcal{G}^\beta$ to denote a β-discounted version of the game structure $\mathcal{G}$. The game $\mathcal{G}^\beta$ at each step halts with probability β (goes to a special absorbing state* halt *such that* halt *is not in R_i for all i), and continues as the game $\mathcal{G}$ with probability $1-\beta$. We refer to β as the discount-factor.* ■

In this paper we write $\mathcal{G}^\beta$ to denote a β-discounted game.

Definition 6 (Stopping time of history in β-discounted games) *Consider the stopping time τ defined on histories $h = \langle s_0, s_1, \ldots \rangle$ by*

$$\tau(h) = \inf\{k \geq 0 \mid s_k = halt\}$$

where the infimum of the empty set is $+\infty$. ■

Lemma 4 *Let $\mathcal{G}^\beta$ be an n-player β-discounted game structure, with $\beta > 0$. Then, for all states $s \in S$ and all strategy profiles $\overline{\sigma}$ we have*

$$\Pr_s^{\overline{\sigma}}[\tau > m] \leq (1-\beta)^m.$$

Proof. At each step of the game $\mathcal{G}^\beta$ the game reaches the halt state with probability β. Hence the probability of not reaching the halt state in m steps is at most $(1-\beta)^m$. ■

The proof of the following lemma is similar to the proof of Lemma 2.2 of [19].

Lemma 5 *For every n-player β-discounted game structure $\mathcal{G}^\beta$, with $\beta > 0$, with reachability objective $[\![Reach(R_i)]\!]$ for player i, there exist memoryless strategies σ_i for $i \in [n]$, such that the memoryless strategy profile $\overline{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ is a Nash equilibrium in $\mathcal{G}^\beta$ for every $s \in S$.*

Proof. Regard each n-tuple $\overline{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ of memoryless strategies as a vector in a compact, convex subset K of the appropriate Eucledian space. Then define a correspondence λ that maps each element $\overline{\sigma}$ of K to the set $\lambda(\overline{\sigma})$ of all elements $\overline{g} = (g_1, g_2, \ldots, g_n)$ of K such that, for all $i \in [n]$ and all $s \in S$, g_i is an optimal response for player i in $\mathcal{G}^\beta$ against $\overline{\sigma}_{-i}$.

Clearly, it suffices to show that there is a $\overline{\sigma} \in K$ such that $\lambda(\overline{\sigma}) = \overline{\sigma}$. To show this, we will verify the Kakutani's Fixed Point Theorem [12]:

1. For every $\overline{\sigma} \in K$, $\lambda(\overline{\sigma})$ is closed, convex and nonempty;
2. If, for $\overline{g}^1, \overline{g}^2, \ldots, \overline{g}^{(k)} \in \lambda(\overline{\sigma}^{(k)}), \lim_{k\to\infty} \overline{g}^{(k)} = \overline{g}$ and $\lim_{k\to\infty} \overline{\sigma}^{(k)} = \overline{\sigma}$, then $\overline{g} \in \lambda(\overline{\sigma})$.

To verify condition 1, fix $\overline{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_n) \in K$ and $i \in \{1, 2, \ldots, n\}$. For each $s \in S$, let $v_i^\beta(s)$ be the maximal payoff that player i can achieve in $\mathcal{G}^\beta$ against $\overline{\sigma}_{-i}$, i.e., $v_i^\beta(s) = val_i^{\mathcal{G}^\beta_{\overline{\sigma}_{-i}}}([\![\mathrm{Reach}(R_i)]\!])(s)$. Since fixing the strategies for all the other players the game structure becomes a MDP, we know that g_i is an optimal response to $\overline{\sigma}_{-i}$ if and only if, for each $s \in S$, $g_i(s)$ puts positive probability only on actions $a_i \in A_i$ that maximize the expectation of $v_i^\beta(s)$, namely,

$$\sum_{s'} v_i^\beta(s')\delta_{\overline{\sigma}_{-i}}(s, a_i)(s').$$

The fact that any convex combination of optimal responses is again an optimal response in MDPs with reachability objectives follows from the fact that MDPs with reachability objectives can be solved by a linear program and the convex combination of optimal responses satisfy the constraints of the linear program with optimal values. Hence condition 1 follows.

Condition 2 is an easy consequence of the continuity mapping

$$\overline{\sigma} \to \mathrm{Pr}_s^{\overline{\sigma}}([\![\mathrm{Reach}(R_i)]\!])$$

from K to the real line. It follows from Lemma 4 that the mapping is continuous. The desired result follows. ∎

Definition 7 (Difference of two MDP's) *Let $\mathcal{G}_1$ and $\mathcal{G}_2$ be two player i-MDPs defined on the same state space S with the same set A of moves. The difference of the two MDPs, denoted $\mathit{diff}(\mathcal{G}_1, \mathcal{G}_2)$, is defined as:*

$$\mathit{diff}(\mathcal{G}_1, \mathcal{G}_2) = \sum_{s' \in S} \sum_{s \in S} \sum_{a \in A} |\delta_1(s,a)(s') - \delta_2(s,a)(s')|.$$

That is, $\mathit{diff}(\mathcal{G}_1, \mathcal{G}_2)$ is the sum of the difference of the probabilities of all the edges of the MDPs. ∎

Observe that in context of player i-MDPs $\mathcal{G}$, for all objectives Ψ_i, we have $val_i^{\mathcal{G}}(\Psi_i)(s) = \sup_{\sigma_i \in \Sigma_i} \Pr_s^{\sigma_i}(\Psi_i)$. The following lemma follows from Theorem 4.3.7 (page 185) of Filar-Vrieze [9].

Lemma 6 (Lipschitz continuity for reachability objectives) *Let $\mathcal{G}_1^\beta$ and $\mathcal{G}_2^\beta$ be two β-discounted player i-MDPs, for $\beta > 0$, on the same state space and with the same set of moves. For $j \in \{ 1, 2 \}$, let $v_j^\beta(s) = val_i^{\mathcal{G}_j^\beta}([\![Reach(R_i)]\!])(s)$, for $R_i \subseteq S$, i.e., $v_j^\beta(s)$ denotes the value for player i for the reachability objective $[\![Reach(R_i)]\!]$ in the β-discounted MDP $\mathcal{G}_j^\beta$. Then the following assertion hold:*

$$|v_1^\beta(s) - v_2^\beta(s)| \leq \mathit{diff}(\mathcal{G}_1^\beta, \mathcal{G}_2^\beta).$$

Lemma 7 (Nash equilibrium with full support) *Let $\mathcal{G}^\beta$ be an n-player β-discounted game structure, with $\beta > 0$, and reachability objective $[\![Reach(R_i)]\!]$ for player i. For every $\varepsilon > 0$, there exist memoryless strategies σ_i for $i \in [n]$, such that for all $i \in [n]$, for all $s \in S$, $\mathrm{Supp}(\sigma_i(s)) = \Gamma_i(s)$ (i.e., all the moves of player i is played with positive probability), and the memoryless strategy profile $\overline{\sigma} = (\sigma_1, \sigma_2, \dots, \sigma_n)$ is an ε-Nash equilibrium in $\mathcal{G}^\beta$ for every $s \in S$.*

Proof. Fix an Nash equilibrium $\overline{\sigma}' = (\sigma_1', \sigma_2', \dots, \sigma_n')$ as obtained from Lemma 5. For all $i \in [n]$, define

$$\sigma_i(s)(a) = \frac{\varepsilon}{|A|^n \cdot |\Gamma_i(s)| \cdot n \cdot |S|^2} + \big(1 - \frac{\varepsilon}{|A|^n \cdot |\Gamma_1(s)| \cdot n \cdot |S|^2}\big) \cdot \sigma_i'(s)(a),$$

for all $s \in S$ and for all $a \in \Gamma_i(s)$. Let $\overline{\sigma} = (\sigma_1, \sigma_2, \dots, \sigma_n)$. Note that for all $s \in S$, we have $\mathrm{Supp}(\sigma_i(s)) = \Gamma_i(s)$. For all $i \in [n]$, consider the two player i-MDPs $\mathcal{G}^\beta_{\overline{\sigma}_{-i}}$ and $\mathcal{G}^\beta_{\overline{\sigma}'_{-i}}$: for all $s, s' \in S$, and $a_i \in \Gamma_i(s)$ we have from the construction that: $|\delta_{\overline{\sigma}_{-i}}(s, a_i)(s') - \delta_{\overline{\sigma}'_{-i}}(s, a_i)(s')| \leq \frac{\varepsilon}{|S|^2 \cdot |\Gamma_i(s)|}$. It follows that $\mathit{diff}(\mathcal{G}^\beta_{\overline{\sigma}_{-i}}, \mathcal{G}^\beta_{\overline{\sigma}'_{-i}}) \leq \varepsilon$. Since $\overline{\sigma}'$ is a Nash equilibrium, and for all $i \in [n]$, $\mathcal{G}^\beta_{\overline{\sigma}_{-i}}$ and $\mathcal{G}^\beta_{\overline{\sigma}'_{-i}}$ are β-discounted player i-MDPs with diff bounded by ε, the result follows from Lemma 6. ∎

Lemma 8 (ε-Nash equilibrium for reachability games [4]) *For every n-player game structure $\mathcal{G}$, with reachability objective $[\![Reach(R_i)]\!]$ for player i, for every $\varepsilon > 0$, there exists $\beta > 0$, such that a memoryless ε-Nash equilibrium in $\mathcal{G}^\beta$, is an 2ε-Nash equilibrium in $\mathcal{G}$.*

Lemma 8 follows from the results in [4]. Lemma 7 and Lemma 8 yield Theorem 1.

Theorem 1 (ε-Nash equilibrium of full support) *For every n-player game structure $\mathcal{G}$, with reachability objective $[\![Reach(R_i)]\!]$ for player i, for every $\varepsilon > 0$, there exists a memoryless ε-Nash equilibrium $\overline{\sigma}^* = (\sigma_1^*, \sigma_2^*, \ldots, \sigma_n^*)$ such that for all $s \in S$, for all $i \in [n]$, we have $\mathrm{Supp}(\sigma_i^*(s)) = \Gamma_i(s)$.*

4 Nash Equilibrium for Upward-closed Objectives

In this section we prove existence of memoryless ε-Nash equilibrium, for all $\varepsilon > 0$, for all n-player concurrent game structures, with upward-closed objectives for all players. The key arguments use induction on the number of players, the results of Section 3 and analysis of Markov chains and MDPs. We present some definitions required for the analysis of the rest of the section.

MDP and graph of a game structure. Given an n-player concurrent game structure $\mathcal{G}$, we define an associated MDP $\overline{\mathcal{G}}$ of $\mathcal{G}$ and an associated graph of $\mathcal{G}$. The MDP $\overline{\mathcal{G}} = (\overline{S}, \overline{A}, \overline{\Gamma}, \overline{\delta})$ is defined as follows:

- $\overline{S} = S$; $\overline{A} = A \times A \times \ldots \times A = A^n$; and $\overline{\Gamma}(s) = \{ (a_1, a_2, \ldots, a_n) \mid a_i \in \Gamma_i(s) \}$.
- $\overline{\delta}(s, (a_1, a_2, \ldots, a_n)) = \delta(s, a_1, a_2, \ldots, a_n)$.

The graph of the game structure $\mathcal{G}$ is defined as the graph of the MDP $\overline{\mathcal{G}}$.

Games with absorbing states. Given a game structure $\mathcal{G}$ we partition the state space of $\mathcal{G}$ as follows:

1. The set of absorbing states in S are denoted as T, i.e., $T = \{ s \in C \mid s \text{ is an absorbing state} \}$.
2. A set U of states that consists of states s such that $|\Gamma_i(s)| = 1$ for $i \in [n]$ and $(U \times S) \cap E \subseteq U \times T$. In other words, at states in U there is no non-trivial choice of moves for the players and thus for any state s in U the game proceeds to the set T according to the probability distribution of the transition function δ at s.
3. $C = S \setminus (U \cup T)$.

Reachable sets. Given a game structure $\mathcal{G}$ and a state $s \in S$, we define $\mathit{Reachable}(s, \mathcal{G}) = \{ t \in S \mid \text{ there is a path from } s \text{ to } t \text{ in the graph of } \mathcal{G} \}$ as the set of states that are reachable from s in the graph of the game structure. For a set $Z \subseteq S$, we denote by $\mathit{Reachable}(Z, \mathcal{G})$ the set of states reachable from a state in Z, i.e., $\mathit{Reachable}(Z, \mathcal{G}) = \bigcup_{s \in Z} \mathit{Reachable}(s, \mathcal{G})$. Given a set Z, let $Z_R = \mathit{Reachable}(Z, \mathcal{G})$. We denote by $\mathcal{G} \restriction Z_R$, the sub-game induced by the set Z_R of states. Similarly, given a set $\mathcal{F} \subseteq 2^S$, we denote by $\mathcal{F} \restriction Z_R$ the set $\{ U \mid \exists F \in \mathcal{F}.\ U = F \cap Z_R \}$.

Terminal non-absorbing maximal end components (TNEC). Given a game structure $\mathcal{G}$, let $\mathcal{Z}$ be the set of maximal end components of $\mathcal{G}$. Let $\mathcal{L} = \mathcal{Z} \setminus T$ be the set of maximal non-absorbing end components and let $H = \bigcup_{L \in \mathcal{L}} L$. A maximal end component $Z \subseteq C$, is a terminal non-absorbing maximal end component (TNEC), if $\mathit{Reachable}(Z, \mathcal{G}) \cap (H \setminus Z) = \emptyset$, i.e., no other non-absorbing maximal end component is reachable from Z.

We consider in this section game structures $\mathcal{G}$ with upward-closed objective $[\![\mathrm{UpClo}(UC_i)]\!]$ for player i. We also denote by $R_i = \{\{s\} \in T \mid s \in UC_i\}$ the set of the absorbing states in T that are in UC_i. We now prove the following key result.

Theorem 2 *For all n-player concurrent game structures $\mathcal{G}$, with upward-closed objective $[\![UpClo(UC_i)]\!]$ for player i, one of the following conditions (condition C1 or C2) hold:*

1. *(Condition C1) There exists a memoryless strategy profile $\overline{\sigma} \in \overline{\Sigma}^M$ such that in the Markov chain $\mathcal{G}_{\overline{\sigma}}$ there is closed recurrent set $Z \subseteq C$, such that $\overline{\sigma}$ is a Nash equilibrium for all states $s \in Z$.*

2. *(Condition C2) There exists a state $s \in C$, such that for all $\varepsilon > 0$, there exists a memoryless ε-Nash equilibrium $\overline{\sigma} \in \overline{\Sigma}^M$ for state s, such that $\Pr_s^{\overline{\sigma}}([\![Reach(T)]\!]) = 1$, and for all $s \in S$, and for all $i \in [n]$, we have $\mathrm{Supp}(\sigma_i(s)) = \Gamma_i(s)$.*

The proof of Theorem 2 is by induction on the number of players. We first analyze the base case.

Base Case. (One player game structures or MDPs) We consider player 1-MDPs and analyze the following cases:

- (Case 1.) If there in no TNEC in C, then it follows from Lemma 2 that for all states $s \in C$, for all strategies $\sigma_1 \in \Sigma_1$, we have $\Pr_s^{\sigma_1}([\![\mathrm{Reach}(T)]\!]) = 1$, and $\Pr_s^{\sigma_1}([\![\mathrm{Reach}(R_1)]\!]) = \Pr_s^{\sigma_1}([\![\mathrm{UpClo}(UC_1)]\!])$ (recall $R_1 = \{ \{ s \} \in T \mid s \in UC_1 \}$). The result of Theorem 1 yields an ε-Nash equilibrium σ_1 that satisfies condition C2 of Theorem 2, for all states $s \in C$.

- (Case 2.) Else let $Z \subseteq C$ be a TNEC.

  1. If $Z \in UC_1$, fix a uniform memoryless strategy $\sigma_1 \in \Sigma_1^{UM}$ such that for all $s \in Z$, $\Pr_s^{\sigma_1}(\{ \omega \mid \mathrm{Inf}(\omega) = Z \}) = 1$ and $\Pr_s^{\sigma_1}([\![\mathrm{UpClo}(UC_1)]\!]) = 1$ (such a strategy exists by Lemma 3, since C is an end component). In other words, Z is a closed recurrent set in the Markov chain $\mathcal{G}_{\sigma_1}$ and the objective of player 1 is satisfied with probability 1. Hence condition C1 of Theorem 2 is satisfied.
  2. If $Z \notin UC_1$, then since UC_1 is upward-closed, for all set $Z_1 \subseteq Z$, $Z_1 \notin UC_1$. Hence for any play ω, such that $\omega \in [\![\mathrm{Safe}(Z)]\!]$, we have $\mathrm{Inf}(\omega) \subseteq Z$, and hence $\omega \notin [\![\mathrm{UpClo}(UC_1)]\!]$. Hence we have for all states $s \in Z$,
  $$\sup_{\sigma_1 \in \Sigma_1} \Pr_s^{\sigma_1}([\![\mathrm{UpClo}(UC_1)]\!]) = \sup_{\sigma_1 \in \Sigma_1} \Pr_s^{\sigma_1}([\![\mathrm{Reach}(R_1)]\!]).$$
  If the set of edges from Z to $U \cup T$ is empty, then for all strategies σ_1 we have $\Pr_s^{\sigma_1}([\![\mathrm{UpClo}(UC_1)]\!]) = 0$, and hence any uniform memoryless strategy can be fixed and condition C1 of Theorem 2 can be satisfied. Otherwise, the set of edges from Z to $U \cup T$ is non-empty, and then for $\varepsilon > 0$, consider an ε-Nash equilibrium for reachability objective $[\![\mathrm{Reach}(R_1)]\!]$ satisfying the conditions of Theorem 1. Since Z is an end component, for all states $s \in Z$, $\mathrm{Supp}(\sigma_1(s)) = \Gamma_1(s)$, and the set of edges to Z to $U \cup T$ is non-empty it follows that for all states $s \in Z$, we have $\Pr_s^{\sigma_1}([\![\mathrm{Reach}(T)]\!]) = 1$. Thus condition C2 of Theorem 2 is satisfied.

We prove the following lemma, that will be useful for the analysis of the inductive case.

Lemma 9 *Consider a player i-MDP $\mathcal{G}$ with an upward-closed objective $[\![UpClo(UC_i)]\!]$ for player i. Let $\sigma_i \in \Sigma_i^M$ be a memoryless strategy such that for all $s \in S$, we have* $\mathrm{Supp}(\sigma_i(s)) = \Gamma_i(s)$. *Let $Z \subseteq S$ be a closed recurrent set in the Markov chain $\mathcal{G}_{\sigma_i}$. Then σ_i is a Nash equilibrium for all states $s \in Z$.*

Proof. The proof follows from the analysis of two cases.

1. If $Z \in UC_i$, then since Z is a closed recurrent set in $\mathcal{G}_{\sigma_i}$, for all states $s \in S$ we have $\Pr_s^{\sigma_i}(\{ \omega \mid \mathrm{Inf}(\omega) = Z \}) = 1$. Hence we have $\Pr_s^{\sigma_i}([\![\mathrm{UpClo}(UC_i)]\!]) = 1 \geq \sup_{\sigma_i' \in \Sigma_i} \Pr_s^{\sigma_i'}([\![\mathrm{UpClo}(UC_i)]\!])$. The result follows.

2. We now consider the case such that $Z \notin UC_i$. Since for all $s \in Z$, we have $\mathrm{Supp}(\sigma_i(s)) = \Gamma_i(s)$, it follows that for all strategies $\sigma_i' \in \Sigma_i$ and for all $s \in Z$, we have $\mathrm{Outcome}(s, \sigma_i') \subseteq \mathrm{Outcome}(s, \sigma_i) \subseteq [\![\mathrm{Safe}(Z)]\!]$ (since Z is a closed recurrent set in $\mathcal{G}_{\sigma_i}$). It follows that for all strategies σ_i' we have $\Pr_s^{\sigma_i'}([\![\mathrm{Safe}(Z)]\!]) = 1$. Hence for all strategies σ_i', for all states $s \in Z$ we have $\Pr_s^{\sigma_i'}(\{\omega \mid \mathrm{Inf}(\omega) \subseteq Z\}) = 1$. Since $Z \notin UC_i$, and UC_i is upward-closed, it follows that for all strategies σ_i', for all states $s \in Z$ we have $\Pr_s^{\sigma_i'}([\![\mathrm{UpClo}(UC_i)]\!]) = 0$. Hence for all states $s \in Z$, we have $\sup_{\sigma_i' \in \Sigma_i} \Pr_s^{\sigma_i'}([\![\mathrm{UpClo}(UC_i)]\!]) = 0 = \Pr_s^{\sigma_i}([\![\mathrm{UpClo}(UC_i)]\!])$. Thus the result follows. ∎

Inductive case. Given a game structure $\mathcal{G}$, consider the MDP $\overline{\mathcal{G}}$: if there are no TNEC in C, then the result follows from analysis similar to Case 1 of the base case. Otherwise consider a TNEC $Z \subseteq C$ in $\overline{\mathcal{G}}$. If for every player i we have $Z \in UC_i$, then fix a uniform memoryless strategy $\overline{\sigma} \in \overline{\Sigma}^{UM}$ such that for all $s \in Z$, $\Pr_s^{\overline{\sigma}}(\{ \omega \mid \mathrm{Inf}(\omega) = Z \}) = 1$ (such a strategy exists by Lemma 3, since C is an end component in $\overline{\mathcal{G}}$). Hence, for all $i \in [n]$ we have $\Pr_s^{\overline{\sigma}}([\![\mathrm{UpClo}(UC_i)]\!]) = 1$. That is Z is a closed recurrent set in the Markov chain $\mathcal{G}_{\overline{\sigma}}$ and the objective of each player is satisfied with probability 1 from all states $s \in Z$. Hence condition C1 of Theorem 2 is satisfied. Otherwise, there exists $i \in [n]$, such that $Z \notin UC_i$, and without loss of generality we assume that this holds for player 1, i.e., $Z \notin UC_1$. If $Z \notin UC_1$, then we prove Lemma 10 to prove Theorem 2.

Lemma 10 *Consider an n-player concurrent game structure $\mathcal{G}$, with upward-closed objective $[\![UpClo(UC_i)]\!]$ for player i. Let Z be a* TNEC *in $\overline{\mathcal{G}}$ such that $Z \notin UC_1$ and let $Z_R = Reachable(Z, \mathcal{G})$. The following assertions hold:*

1. *If there exists $\sigma_1 \in \Sigma_1^M$, such that for all $s \in Z$, $\mathrm{Supp}(\sigma_1(s)) = \Gamma_1(s)$, and condition C1 of Theorem 2 holds in $\mathcal{G}_{\sigma_1} \upharpoonright Z_R$, then condition C1 Theorem 2 holds in $\mathcal{G}$.*

2. *Otherwise, condition C2 of Theorem 2 holds in $\mathcal{G}$.*

Proof. Given a memoryless strategy σ_1, fixing the strategy σ_1 for player 1, we get an $n-1$-player game structure and by inductive hypothesis either condition C1 or C2 of Theorem 2 holds.

- Case 1. If there is a memoryless strategy $\sigma_1 \in \Sigma_1^M$, such that for all $s \in Z$, $\mathrm{Supp}(\sigma_1(s)) = \Gamma_1(s)$, and condition C1 of Theorem 2 holds in $\mathcal{G}_{\sigma_1}$, then let $\overline{\sigma}_{-1} = (\sigma_2, \sigma_3, \ldots, \sigma_n)$ be the memoryless Nash equilibrium and $Z_1 \subseteq Z$ be the closed recurrent set in $\mathcal{G}_{\overline{\sigma}_{-1} \cup \sigma_1}$ satisfying the condition C1 of Theorem 2 in $\mathcal{G}_{\sigma_1}$. Observe that (σ_1, Z_1) satisfy the conditions of Lemma 9 in the MDP $\mathcal{G}_{\overline{\sigma}_{-i}}$. Hence, an application of Lemma 9 yields that σ_1 is a Nash equilibrium for all states $s \in Z_1$, in the MDP $\mathcal{G}_{\overline{\sigma}_{-1}}$. Since $\overline{\sigma}_{-1}$ is a Nash equilibrium for all states in Z_1 in $\mathcal{G}_{\sigma_1}$, it follows that $\overline{\sigma} = \overline{\sigma}_{-1} \cup \sigma_1$ and Z_1 satisfy condition C1 of Theorem 2.

- For $\varepsilon > 0$, consider a memoryless ε-Nash equilibrium $\overline{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ in $\mathcal{G}$ with objective $[\![\mathrm{Reach}(R_i)]\!]$ for player i, such that for all $s \in S$, for all $i \in [n]$, we have $\mathrm{Supp}(\sigma_i(s)) = \Gamma_i(s)$ (such an ε-Nash equilibrium exists from Theorem 1). We now prove the desired result analyzing two sub-cases:

  1. Suppose there exits $j \in [n]$, and $Z_j \subseteq Z$, such that $Z_j \in UC_j$, and Z_j is an end component in $\mathcal{G}_{\overline{\sigma}_{-j}}$, then let σ'_j be a memoryless strategy for player j, such that Z_j is a closed recurrent set of states in the Markov chain $\mathcal{G}_{\overline{\sigma}_{-j} \cup \sigma'_j}$. Let $\overline{\sigma}' = \overline{\sigma}_{-j} \cup \sigma'_j$. Since $Z_j \in UC_j$, it follows that for all states $s \in Z_j$, we have $\Pr_s^{\overline{\sigma}'}([\![\mathrm{UpClo}(UC_j)]\!]) = 1$, and hence player j has no incentive to deviate from $\overline{\sigma}'$. Since for all σ_i, for $i \neq j$, and for all states $s \in S$, we have $\mathrm{Supp}(\sigma_i)(s) = \Gamma_i(s)$, and Z_j is a closed recurrent set in $\mathcal{G}_{\overline{\sigma}'}$, it follows from Lemma 9 that for all $j \neq i$, σ_i is a Nash equilibrium in $\mathcal{G}_{\overline{\sigma}'_{-i}}$. Hence we have $\overline{\sigma}'$ is a Nash equilibrium for all states $s \in Z_j$ in $\mathcal{G}$ and condition C1 of Theorem 2 is satisfied.
  2. Hence it follows that if Case 1 fails, for all $i \in [n]$, all end components $Z_i \subseteq Z$, in $\mathcal{G}_{\overline{\sigma}_{-i}}$, we have $Z_i \notin UC_i$. Hence for all $i \in [n]$, for all $s \in Z$, for all $\sigma'_i \in \Sigma_i$, $\Pr_s^{\overline{\sigma}_{-i} \cup \sigma'_i}([\![\mathrm{UpClo}(UC_i)]\!]) = \Pr_s^{\overline{\sigma}_{-i} \cup \sigma'_i}([\![\mathrm{Reach}(R_i)]\!])$. Since $\overline{\sigma}$ is an ε-Nash equilibrium with objectives $[\![\mathrm{Reach}(R_i)]\!]$ for player i in $\mathcal{G}$, it follows that $\overline{\sigma}$ is an ε-Nash equilibrium in $\mathcal{G}$ with objectives $[\![\mathrm{UpClo}(UC_i)]\!]$ for player i. Moreover, if there is an closed recurrent set $Z' \subseteq Z$ in the Markov chain $\mathcal{G}_{\overline{\sigma}}$, then case 1 would have been true (follows

from Lemma 9). Hence if case 1 fails, then it follows that there is no closed recurrent set $Z' \subseteq Z$ in $\mathcal{G}_{\overline{\sigma}}$, and hence for all states $s \in Z$, we have $\Pr_s^{\overline{\sigma}}([\![\text{Reach}(T)]\!]) = 1$. Hence condition C2 of Theorem 2 holds, and the desired result is established. ∎

Inductive application of Theorem 2. Given a game structure $\mathcal{G}$, with upward-closed objective $[\![\text{UpClo}(UC_i)]\!]$ for player i, to prove existence of ε-Nash equilibrium for all states $s \in S$, for $\varepsilon > 0$, we apply Theorem 2 recursively. We convert the game structure $\mathcal{G}$ to a game structure $\mathcal{G}'$ as follows:

1. **Transformation 1.** If condition C1 of Theorem 2 holds, then let Z be the closed recurrent set that satisfy the condition C1 of Theorem 2.

   - In $\mathcal{G}'$ convert every state $s \in Z$ to an absorbing state;
   - if $Z \notin UC_i$, for player i, then the objective for player i in $\mathcal{G}'$ is UC_i;
   - if $Z \in UC_i$ for player i, the objective for player i in $\mathcal{G}$ is modified to include every state $s \in Z$, i.e., for all $Q \subseteq S$, if $s \in Q$, for some $s \in Z$, then Q is included in UC_i.

   Observe that the states in Z are converted to absorbing states and will be interpreted as states in T in $\mathcal{G}'$.

2. **Transformation 2.** If condition C2 of Theorem 2 holds, then let $\overline{\sigma}^*$ be an $\frac{\varepsilon}{|S|}$-Nash equilibrium from state s, such that $\Pr_s^{\overline{\sigma}^*}([\![\text{Reach}(T)]\!]) = 1$. The state is converted as follows: for all $i \in [n]$, the available moves for player i at s is reduced to 1, i.e., for all $i \in [n]$, $\Gamma_i(s) = \{ a_i \}$, and the transition function δ' in $\mathcal{G}'$ at s is defined as:

$$\delta(s, a_1, a_2, \ldots, a_n)(t) = \begin{cases} \Pr_s^{\overline{\sigma}^*}([\![\text{Reach}(t)]\!]) & \text{if } t \in T \\ 0 & \text{otherwise.} \end{cases}$$

   Note that the state s can be interpreted as a state in U in $\mathcal{G}'$.

To obtain a ε-Nash equilibrium for all states $s \in S$ in $\mathcal{G}$, it suffices to obtain an ε-Nash equilibrium for all states in $\mathcal{G}'$. Also observe that for all states in $U \cup T$, Nash equilibrium exists by definition. Applying the transformations recursively on $\mathcal{G}'$, we proceed to convert every state to a state in $U \cup T$, and the desired result follows. This yields Theorem 3.

Theorem 3 *For all n-player concurrent game structures $\mathcal{G}$, with upward-closed objective $[\![UpClo(UC_i)]\!]$ for player i, for all $\varepsilon > 0$, for all states $s \in S$, there exists a memoryless strategy profile $\overline{\sigma}^*$, such that $\overline{\sigma}^*$ is an ε-Nash equilibrium for state s.*

Remark 1 *It may be noted that upward-closed objectives are not closed under complementation, and hence Theorem 3 is not a generalization of determinacy result for concurrent zero-sum games with upward-closed objective for one player. For example in concurrent zero-sum games with Büchi objective for a player, ε-optimal strategies require infinite-memory in general, but the complementary objective of a Büchi objective is not upward-closed (recall Example 1). In contrast, we show the existence of memoryless ε-Nash equilibrium for n-player concurrent games where each player has an upward-closed objective. For the special case of zero-sum turn-based games, with upward-closed objective for a player, existence of memoryless optimal strategies was proved in [3]; however note that the memoryless strategies require randomization as pure or deterministic strategies require memory even for turn-based games with generalized Büchi objectives.*

5 Computational Complexity

In this section we present an algorithm to compute an ε-Nash equilibrium for n-player game structures with upward-closed objectives, for $\varepsilon > 0$. A key result for the algorithmic analysis is Lemma 11.

Lemma 11 *Consider an n-player concurrent game structure $\mathcal{G}$, with upward-closed objective $[\![UpClo(UC_i)]\!]$ for player i. Let Z be a* TNEC *in $\overline{\mathcal{G}}$ such that $Z \notin UC_n$ and let $Z_R = Reachable(Z, \mathcal{G})$. The following assertion hold:*

1. *Suppose there exists $\sigma_n \in \Sigma_n^M$, such that for all $s \in Z$, $\mathrm{Supp}(\sigma_n(s)) = \Gamma_n(s)$, and condition C1 of Theorem 2 holds in $\mathcal{G}_{\sigma_n} \upharpoonright Z_R$. Let $\sigma_n^* \in \Sigma_n^{UM}$ such that for all $s \in Z$ we have $\mathrm{Supp}(\sigma_n^*(s)) = \Gamma_n(s)$ (i.e., σ_n^* is a uniform memoryless strategy that plays all available moves at all states in Z). Then condition C1 holds in $\mathcal{G}_{\sigma_n^*} \upharpoonright Z_R$.*

Proof. The result follows from the observation that for any strategy profile $\overline{\sigma}_{-n} \in \overline{\Sigma}_{-n}^M$, the closed recurrent set of states in $\mathcal{G}_{\overline{\sigma}_{-n} \cup \sigma_n}$ and $\mathcal{G}_{\overline{\sigma}_{-n} \cup \sigma_n^*}$ are the same. ∎

Lemma 11 presents the basic principle to identify if condition C1 of Theorem 2 holds in a game structure $\mathcal{G}$ with upward-closed objective

Algorithm 1 UpCloCondC1

Input : An n-player game structure $\mathcal{G}$, with upward-closed objective $[\![\mathrm{UpClo}(UC_i)]\!]$ for player i, for all $i \in [n]$.
Output: Either $(Z, \overline{\sigma})$ satisfying condition C1 of Theorem 2 or else $(\emptyset, \emptyset)$.

1. **if** $n = 0$,
   1.1 **if** there is a non-absorbing closed recurrent set Z in the Markov chain $\mathcal{G}$, **return** $(Z, \emptyset)$.
   1.2 **else return** $(\emptyset, \emptyset)$.
2. $\mathcal{Z} =$**ComputeMaximalEC**$(\overline{\mathcal{G}})$
   (i.e., $\mathcal{Z}$ is the set of maximal end components in the MDP of $\mathcal{G}$).
3. **if** there is no TNEC in $\overline{\mathcal{G}}$, **return** $(\emptyset, \emptyset)$.
4. **if** there exists $Z \in \mathcal{Z}$ such that for all $i \in [n]$, $Z \in UC_i$,
   4.1. **return** $(Z, \overline{\sigma})$ such that $\overline{\sigma} \in \overline{\Sigma}^{UM}$ and Z is closed recurrent set in $\mathcal{G}_{\overline{\sigma}}$.
5. Let Z be a TNEC in $\overline{\mathcal{G}}$, and let $Z_R = \mathit{Reachable}(Z, \mathcal{G})$.
6. **else** without loss of generality let $Z \notin UC_n$.
   6.1. Let $\sigma_n \in \Sigma_n^{UM}$ such that for all states $s \in Z_R$, $\sigma_n(s) = \Gamma_n(s)$.
   6.2. $(Z_1, \overline{\sigma}) =$ **UpCloCondC1** $(\mathcal{G}_{\sigma_n} \upharpoonright Z_R, n-1, [\![\mathrm{UpClo}(UC_i \upharpoonright Z_R)]\!]$ for $i \in [n-1])$
   6.3. **if** $(Z_1 = \emptyset)$ **return** $(\emptyset, \emptyset)$;
   6.4. **else return** $(Z_1, \overline{\sigma}_{-n} \cup \sigma_n)$.

$[\![\mathrm{UpClo}(UC_i)]\!]$ for player i. An informal description of the algorithm (Algorithm 1) is as follows: the algorithm takes as input a game structure $\mathcal{G}$ of n-players, objectives $[\![\mathrm{UpClo}(UC_i)]\!]$ for player i, and it either returns $(Z, \overline{\sigma}) \in S \times \overline{\Sigma}^{M}$ satisfying the condition C1 of Theorem 2 or returns $(\emptyset, \emptyset)$. Let $\overline{\mathcal{G}}$ be the MDP of $\mathcal{G}$, and let $\mathcal{Z}$ be the set of maximal end components in $\overline{\mathcal{G}}$ (computed in Step 2 of Algorithm 1). If there is no TNEC in $\overline{\mathcal{G}}$, then condition C1 of Theorem 2 fails and $(\emptyset, \emptyset)$ is returned (Step 3 of Algorithm 1). If there is a maximal end component $Z \in \mathcal{Z}$ such that for all $i \in [n]$, $Z \in UC_i$, then fix a uniform memoryless strategy $\overline{\sigma} \in \overline{\Sigma}^{UM}$ such that Z is a closed recurrent set in $\mathcal{G}_{\overline{\sigma}}$ and return $(Z, \overline{\sigma})$ (Step 4 of Algorithm 1). Else let Z be a TNEC and without of loss of generality let $Z \notin UC_n$. Let $Z_R = \mathit{Reachable}(Z, \mathcal{G})$, and fix a strategy $\sigma_n \in \Sigma_n^{UM}$, such that for all $s \in Z_R$, $\mathrm{Supp}(\sigma_n(s)) = \Gamma_n(s)$. The $n-1$-player game structure $\mathcal{G}_{\sigma_n} \upharpoonright Z_R$ is solved by an recursive call (Step 6.3) and the result of the recursive call is returned. It follows from Lemma 11 that if Algorithm 1

Algorithm 2 NashEqmCompute

Input : An n-player game structure $\mathcal{G}$, with upward-closed objective $[\![\mathrm{UpClo}(UC_i)]\!]$ for player i, for all $i \in [n]$.
Output: Either $(Z, \overline{\sigma})$ satisfying condition C1 of Theorem 2 or else $(s, \overline{\sigma})$ satisfying condition C2 of Theorem 2.

1. $\mathcal{Z} =$**ComputeMaximalEC**$(\overline{\mathcal{G}})$
   (i.e., $\mathcal{Z}$ is the set of maximal end components in the MDP of $\mathcal{G}$).
2. **if** there is no TNEC in $\overline{\mathcal{G}}$,
   return $(s,$**ReachEqmFull**$(\mathcal{G}, n, \varepsilon))$ for some $s \in C$.
3. Let Z be a TNEC in $\overline{\mathcal{G}}$, and let $Z_R = \mathit{Reachable}(Z, \mathcal{G})$.
4. Let $(Z_1, \overline{\sigma}) =$ **UpCloCondC1** $(\mathcal{G}_{\sigma_n} \upharpoonright Z_R, n-1, [\![\mathrm{UpClo}(UC_i \upharpoonright Z_R)]\!]$ for $i \in [n-1])$
5. **if** $(Z_1 \neq \emptyset)$ **return** $(Z_1, \overline{\sigma})$;
6. Let $\overline{\sigma} =$ **ReachEqmFull**$(\mathcal{G}, n, \varepsilon)$.
7. For $s \in C$, if $\overline{\sigma}$ is an ε-Nash equilibrium for s, with objective $[\![\mathrm{UpClo}(UC_i)]\!]$ for player i,
   return $(s, \overline{\sigma})$.

returns $(\emptyset, \emptyset)$, then condition C2 of Theorem 2 holds for some state $s \in C$. Let $T(|\mathcal{G}|, n)$ denote the running time of Algorithm 1 on a game structure $\mathcal{G}$ with n-players. Step 2 of the algorithm can be computed in $O(|\mathcal{G}|^2)$ time (see [8] for a $O(|\mathcal{G}|^2)$ time algorithm to compute maximal end components of a MDP). Step 4 can be achieved in time linear in the size of the game structure. Thus we obtain the following recurrence

$$T(|\mathcal{G}|, n) = O(|\mathcal{G}|^2) + T(|\mathcal{G}|, n-1).$$

Hence we have $T(|\mathcal{G}|, n) = O(n \cdot |\mathcal{G}|^2)$.

Basic principle of Algorithm 2. Consider a game structure $\mathcal{G}$ with objective $[\![\mathrm{UpClo}(UC_i)]\!]$ for player i. Let $\overline{\sigma}$ be a memoryless strategy profile such for all states $s \in S$, for all $i \in [n]$, we have $\mathrm{Supp}(\sigma_i(s)) = \Gamma_i(s)$, and $(s, \overline{\sigma})$ satisfy condition C2 of Theorem 2 for some state $s \in C$. Let $Z_s = \mathit{Reachable}(s, \mathcal{G})$. It follows from the base case analysis of Theorem 2 and Lemma 10, that for all $i \in [n]$, in the MDP $\mathcal{G}_{\overline{\sigma}_{-i}} \upharpoonright Z_s$, for all end components $Z \subseteq Z_s$, $Z \notin UC_i$, and hence in $\mathcal{G}_{\overline{\sigma}_{-i}} \upharpoonright Z_s$, the objective $[\![\mathrm{UpClo}(UC_i)]\!]$ is equivalent to $[\![\mathrm{Reach}(R_i)]\!]$. It follows that if condition C2 of Theorem 2 holds at a state s, then for every $\varepsilon > 0$, any memoryless ε-Nash equilibrium $\overline{\sigma}$ in $\mathcal{G}$ with objective $[\![\mathrm{Reach}(R_i)]\!]$ for player i, such that for all

$s \in S$, for all $i \in [n]$, $\mathrm{Supp}(\sigma_i(s)) = \Gamma_i(s)$, is also an ε-Nash equilibrium in $\mathcal{G}$ with objective $[\![\mathrm{UpClo}(UC_i)]\!]$ for player i. This observation is formalized in Lemma 12. Lemma 12 and Algorithm 1 is the basic principle to obtain a memoryless ε-Nash equilibrium at a non-empty set of states in C.

Lemma 12 *Consider a game structure $\mathcal{G}$ with objective $[\![UpClo(UC_i)]\!]$ for player i. Let $\overline{\sigma}$ be a memoryless strategy profile such for all states $s \in S$, for all $i \in [n]$, we have $\mathrm{Supp}(\sigma_i(s)) = \Gamma_i(s)$, and $(s, \overline{\sigma})$ satisfy condition C2 of Theorem 2 for some state $s \in C$. For $\varepsilon > 0$, any memoryless ε-Nash equilibrium $\overline{\sigma}'$ in $\mathcal{G}$ for state s with objective $[\![Reach(R_i)]\!]$ for player i, such that for all $s \in S$, for all $i \in [n]$, $\mathrm{Supp}(\sigma_i'(s)) = \Gamma_i(s)$, is also an ε-Nash equilibrium in $\mathcal{G}$ for state s with objective $[\![UpClo(UC_i)]\!]$ for player i.*

Description of Algorithm 2. We now describe Algorithm 2 that compute an ε-Nash equilibrium at some state s of a game structure $\mathcal{G}$, with upward-closed objective $[\![\mathrm{UpClo}(UC_i)]\!]$ for player i, for $\varepsilon > 0$. In the algorithm the procedure **ReachEqmFull** returns a strategy $\overline{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ such that for all s, $\mathrm{Supp}(\sigma_i(s)) = \Gamma_i(s)$, and $\overline{\sigma}$ is an ε-Nash equilibrium in $\mathcal{G}$ with reachability objective $[\![\mathrm{Reach}(R_i)]\!]$ for player i, from all states in S. The algorithm first computes the set of maximal end components in $\mathcal{G}$. If there is no TNEC in $\mathcal{G}$, then it invokes **ReachEqmFull**. Otherwise, for some TNEC Z and $Z_R = \mathit{Reachable}(Z, \mathcal{G})$, it invokes Algorithm 1 on the sub-game $\mathcal{G} \restriction Z_R$. If Algorithm 1 returns a non-empty set (i.e., condition C1 of Theorem 2 holds), then the returned value of Algorithm 1 is returned. Otherwise, the algorithm invokes **ReachEqmFull** and returns $(s, \overline{\sigma})$ satisfying condition C2 of Theorem 2. Observe that the procedure **ReachEqmFull** is invoked when: either there is no TNEC in $\mathcal{G}$, or condition C2 holds in $\mathcal{G} \restriction Z_R$. It suffices to compute a memoryless $\frac{\varepsilon}{2}$-Nash equilibrium $\overline{\sigma}' = (\sigma_1', \sigma_2', \ldots, \sigma_n')$ in $\mathcal{G} \restriction Z_R$ with reachability objective $[\![\mathrm{Reach}(R_i)]\!]$ for player i, and then apply the construction of Lemma 7 replacing ε by $\frac{\varepsilon}{2}$ to obtain $(s, \overline{\sigma})$ as desired. Hence it follows that the complexity of **ReachEqmFull** can be bounded by the complexity of a procedure to compute memoryless ε-Nash equilibrium in game structures with reachability objectives. Thus we obtain that the running time of Algorithm 2 is bounded by $O(n \cdot |\mathcal{G}|^2) + \mathbf{ReachEqm}(|\mathcal{G}|, n, \varepsilon)$, where **ReachEqm** is the complexity of a procedure to compute memoryless ε-Nash equilibrium in games with reachability objectives.

The inductive application of Theorem 2 to obtain Theorem 3 using transformation 1 and transformation 2 shows that Algorithm 2 can be applied $|S|$-times to compute a memoryless ε-Nash equilibrium for all states $s \in S$. For all constants $\varepsilon > 0$, existence of polynomial witness and polynomial time verification procedure for $\mathbf{ReachEqm}(\mathcal{G}, n, \varepsilon)$ has been proved in [4]. It follows that for all constants $\varepsilon > 0$, $\mathbf{ReachEqm}(\mathcal{G}, n, \varepsilon)$ is in the complexity class TFNP. The above analysis yields Theorem 4.

Theorem 4 *Given an n-player game structure $\mathcal{G}$ with upward-closed objective $[\![UpClo(UC_i)]\!]$ for player i, a memoryless ε-Nash equilibrium for all $s \in S$ can be computed*

- *in TFNP for all constants $\varepsilon > 0$; and*
- *in time $O(|S| \cdot n \cdot |\mathcal{G}|^2) + |S| \cdot \mathbf{ReachEqm}(\mathcal{G}, n, \varepsilon)$.*

6 Conclusion

In this paper we establish existence of memoryless ε-Nash equilibrium, for all $\varepsilon > 0$, for all n-player concurrent game structures, with upward-closed objectives for all players. We also showed that computation of a memoryless ε-Nash equilibrium can be achieved by a polynomial procedure and solving memoryless ε-Nash equilibrium of n-player concurrent game structures with reachability objectives. The existence of ε-Nash equilibrium, for all $\varepsilon > 0$, in n-player concurrent game structures with ω-regular objectives, and other class of objectives in the higher levels of Borel hierarchy are interesting open problems.

References

[1] D.P. Bertsekas. *Dynamic Programming and Optimal Control.* Athena Scientific, 1995. Volumes I and II.

[2] K. Chatterjee. Two-player nonzero-sum ω-regular games. In *CONCUR'05*, pages 413–427. LNCS 3653, Springer, 2005. Technical Report: UCB/CSD-04-1364.

[3] K. Chatterjee, L. de Alfaro, and T.A. Henzinger. Trading memory for randomness. In *QEST'04*, pages 206–217. IEEE, 2004.

[4] K. Chatterjee, R. Majumdar, and M. Jurdziński. On Nash equilibria in stochastic games. In *CSL'04*, pages 26–40. LNCS 3210, Springer, 2004.

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

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

[7] L. de Alfaro and T.A. Henzinger. Concurrent omega-regular games. In *LICS'00*, pages 141–154. IEEE, 2000.

[8] Luca de Alfaro. Computing minimum and maximum reachability times in probabilistic systems. In *CONCUR'99*, pages 66–81, 1999.

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

[10] A.M. Fink. Equilibrium in a stochastic n-person game. *Journal of Science of Hiroshima University*, 28:89–93, 1964.

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

[12] S. Kakutani. A generalization of Brouwer's fixed point theorem. *Duke Journal of Mathematics*, 8:457–459, 1941.

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

[14] J.G. Kemeny, J.L. Snell, and A.W. Knapp. *Denumerable Markov Chains.* D. Van Nostrand Company, 1966.

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

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

[17] C.H. Papadimitriou. On the complexity of the parity argument and other inefficient proofs of existence. *Journal of Computer and Systems Sciences*, 48(3):498–532, 1994.

[18] C.H. Papadimitriou. Algorithms, games, and the internet. In *STOC'01*, pages 749–753. ACM Press, 2001.

[19] P. Secchi and W.D. Sudderth. Stay-in-a-set games. *International Journal of Game Theory*, 30:479–490, 2001.

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

[21] 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.

[22] M.Y. Vardi. Automatic verification of probabilistic concurrent finite-state systems. In *FOCS'85*, pages 327–338. IEEE Computer Society Press, 1985.

[23] N. Vieille. Two player stochastic games I: a reduction. *Israel Journal of Mathematics*, 119:55–91, 2000.

[24] N. Vieille. Two player stochastic games II: the case of recursive games. *Israel Journal of Mathematics*, 119:93–126, 2000.

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

[26] B. von Stengel. Computing equilibria for two-person games. *Chapter 45, Handbook of Game Theory*, 3:1723–1759, 2002. (editors R.J. Aumann and S. Hart).