# On Nash Equilibria in Stochastic Games*

Krishnendu Chatterjee[1], Rupak Majumdar[2], and Marcin Jurdziński[3]

[1]EECS, University of California, Berkeley,USA
[2]CS, University of California, Los Angeles,USA
[3]CS, University of Warwick, UK
{c_krish,mju}@eecs.berkeley.edu, rupak@cs.ucla.edu

**Abstract.** We study infinite stochastic games played by $n$-players on a finite graph with goals given by sets of infinite traces. The games are *stochastic* (each player simultaneously and independently chooses an action at each round, and 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 a reachability objective, that is, if the goal for each player $i$ is to visit some subset $R_i$ of the states, then there exists an $\epsilon$-Nash equilibrium in memoryless strategies, for every $\epsilon > 0$. However, exact Nash equilibria need not exist. We study the complexity of finding such Nash equilibria, and show that the payoff of some $\epsilon$-Nash equilibrium in memoryless strategies can be $\epsilon$-approximated in NP.
We study the important subclass of $n$-player *turn-based probabilistic* games, where at each state at most one player has a nontrivial choice of moves. For turn-based probabilistic games, we show the existence of $\epsilon$-Nash equilibria in pure strategies for games where the objective of player $i$ is a *Borel* set $B_i$ of infinite traces. However, exact Nash equilibria may not exist. For the special case of $\omega$-regular objectives, we show exact Nash equilibria exist, and can be computed in NP when the $\omega$-regular objectives are expressed as parity objectives.

## 1 Introduction

The interaction of several agents is naturally modeled as non-cooperative games [19, 21]. The simplest, and most common interpretation of a non-cooperative game is that there is a single interaction among the players ("one-shot"), after which the payoffs are decided and the game ends. However, many, if not all, strategic endeavors occur over time, and in stateful manner. That is, the games progress over time, and the current game is decided based on the history of the interactions. Infinite *stochastic games* [24, 9, 6] form a natural model for such interactions. A stochastic game is played over a finite *state space*, and is played in rounds. In each round, each player chooses an available action out of a finite

* This research was supported in part by the AFOSR MURI grant F49620-00-1-0327, ONR grant N00014-02-1-0671, NSF grants CCR-9988172 and CCR-0225610.

set of actions simultaneously with and independently from all other players, and the game moves to a new state under a possibly probabilistic transition relation based on the current state and the joint actions. For the verification and control of reactive systems, such games are infinite: play continues for an infinite number of rounds, giving rise to an infinite sequence of states, called the *outcome* of the game. The players receive a payoff based on a payoff function mapping infinite outcomes to a real in $[0, 1]$.

Payoffs are generally Borel measurable functions [18]. For example, the payoff set for each player is a Borel set $B_i$ in the Cantor topology on $S^\omega$ (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 $\omega$-*regular languages*. $\omega$-regular sets occur in low levels of the Borel hierarchy (they are in $\Sigma_3 \cap \Pi_3$), but they form a robust and expressive language for determining payoffs for commonly used specifications [16]. The simplest $\omega$-regular games correspond to safety (closed sets) or reachability (open sets) objectives.

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 actions of all the other players. 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-shot zero sum games, this is von Neumann's minmax theorem [30]. 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 [18] shows that all stochastic zero sum games with Borel payoffs are determined.

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 all other players continue playing their strategies in the profile. Again, for one-shot 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 possible number of strategies is infinite. 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 $\epsilon$-Nash equilibrium for all $\epsilon > 0$, where an $\epsilon$-Nash equilibrium is a strategy profile where unilateral deviation can only increase the payoff of a player by at most $\epsilon$. Exact Nash equilibria do exist in discounted stochastic games [10], and other special cases [26, 27]. For limit average payoffs, exact Nash equilibria need not exist even for two-player games [1]. Recently the existence of $\epsilon$-Nash equilibria for all $\epsilon > 0$ was proved in [28, 29] for the two-player case, and the general case remains an important open question. For games with payoffs defined by Borel sets, surprisingly little is known. Secchi and Sudderth [23] showed that exact Nash equilibria do exist when all players have payoffs defined by closed sets ("safety games"), where the objective of each player is to stay within a certain set of good states. Formally, each player $i$ has a subset of states $S_i$ as their safe states, and gets a payoff 1 if the play never leaves the set $S_i$ and gets payoff 0 otherwise. This result was generalized to general state and action spaces [23, 15], where only $\epsilon$-equilibria exist.

However, not much more is known: for example, the existence of $\epsilon$-Nash equilibria when the payoff for each player is given by an open set ("reachability games") remained open. In the open (or reachability) game, each player $i$ has a subset of states $R_i$ as reachability targets. Player $i$ gets payoff 1 if the outcome visits some state from $R_i$ at some point, and 0 otherwise. In this case, an $\epsilon$-equilibrium is the best one can hope for: there exist two-player reachability games for which no exact Nash equilibria exist [13]. In this paper, we answer this question in the affirmative: we show that every $n$-player reachability game has an $\epsilon$-Nash equilibrium, for every $\epsilon > 0$. Moreover, there is an $\epsilon$-Nash equilibrium profile of *memoryless strategies* (a strategy is memoryless if it only depends on the current state of the game, and not on the history). However, strategies in general may require randomization. We achieve our result by going through discounted games. Our proof has three main technical ingredients: the existence of equilibria in certain discounted games, the approximability of such equilibria with simple strategies, and the approximability of certain undiscounted payoffs in a Markov decision process (MDP) by discounted payoffs. First, we use the existence of exact Nash equilibria in memoryless strategies in discounted reachability games, that is, reachability games where play stops at each round with a probability $\beta < 1$, and continues to the next round with probability $1 - \beta$. Second, we show that the exact Nash equilibrium in discounted games can be approximated by *simple* strategy profiles, where each player plays uniformly over a multi-set of actions, and the size of the multi-set can be bounded in terms of the number of players and the size of the game. Third, using a result of Condon [3], we show that for an MDP, for any $\epsilon > 0$, we can choose the discount factor $\beta$ such that the difference between the payoffs in the original MDP and the discounted MDP is bounded by $\epsilon$. It follows that on the MDP obtained by fixing these simple strategies for all but player $i$, player $i$ cannot significantly increase his payoff by switching to a different strategy. This construction yields an $\epsilon$-Nash equilibrium in memoryless strategies. In contrast, the proof for safety games [23] proceeds by induction on the number of players, and does not give a Nash equilibrium in memoryless strategies. It is not known if exact Nash equilibria in memoryless strategies exist for safety games.

Computing the values of a Nash equilibrium, when it exists, is another challenging problem [20, 31]. For one-shot zero sum games, equilibrium values and strategies can be computed in polynomial time (by reduction to linear programming) [19]. For zero-sum stochastic games with $\omega$-regular payoffs, doubly exponential time algorithms can be obtained to approximate the value to within $\epsilon$ by expressing the value as nested fixpoints, and then reducing to the validity of a sentence in the theory of reals [6]. For one-shot nonzero sum games, no polynomial time algorithm to compute an exact Nash equilibrium in a two-player game is known [20].

We study the complexity of approximating Nash equilibria in reachability games. We give an NP algorithm to approximate the value of some $\epsilon$-Nash equilibrium in memoryless strategies, for any constant $\epsilon > 0$. We show that the problem of finding a Nash equilibrium in a reachability game in *pure* (i.e., not requiring randomization) memoryless strategies, if it exists, where each player gets at least some (specified) payoff is NP-complete. In contrast, for matrix games, finding a *pure* Nash equilibrium, if it exists, can always be achieved in polynomial time in size of the input. Related NP-hardness results appear in [4], but our results do not follow from theirs as their hardness proof does not hold for pure strategy Nash equilibria. For two-player zero-sum games with reachability objective, values can be algebraic and there are simple examples when they are irrational [6]. Hence one can only hope to compute it to an $\epsilon$-precision.

Interestingly, the techniques we develop for the nonzero sum games can be used to get better complexities for the zero sum case. In particular, we show an improved NP $\cap$ co-NP upper bound to approximate the valued for two-player zero sum reachability games within $\epsilon$-tolerance for any constant $\epsilon$. This improves the previously best known EXPTIME upper bound for reachability games [6]. This also generalizes a result of Condon [3]. Open and closed sets form the lowest level of the Borel hierarchy, and together with [23], this paper answers positively the existence of ($\epsilon$) Nash equilibria in such games. We leave the generalization of these results to higher levels of the Borel hierarchy as an interesting open problem.

An important class of stochastic games are *turn-based probabilistic games* [27, 22, 9], where at each stage, at most one player has a nontrivial choice of actions. In verification, such games model asynchronous interactions among components. For turn-based probabilistic games, we prove the existence of $\epsilon$-Nash equilibria for payoffs specified by arbitrary Borel sets. Moreover, we show that there is an $\epsilon$-equilibrium strategy profile in *pure* strategies, that is, strategies that do not require randomization. Our proof proceeds in two steps. First, we prove a pure strategy determinacy theorem for turn-based probabilistic zero sum games and Borel payoff functions. The proof is a specialization of Martin's determinacy proof for stochastic games with Borel payoffs [18]. Second, using this and a general construction of *threat strategies* [19], we show that $\epsilon$-Nash equilibria exist for all turn based probabilistic nonzero sum games with arbitrary Borel set payoffs. We show this result is optimal: there exist turn-based probabilistic games with Borel payoffs for which exact Nash equilibria do not exist. As a special case, we study turn-based probabilistic games with *parity* winning objectives [7]. Parity objectives form a canonical form for all $\omega$-regular winning objectives [7]. Using existence of pure memoryless optimal strategy for zero sum turn-based probabilistic parity games, we show the existence of pure strategy exact Nash equilibria for parity payoffs. This proves that (exact) Nash equilibria exist for turn based probabilistic games with $\omega$-regular payoffs. We give corresponding complexity results to compute Nash equilibrium solutions. Using an NP $\cap$ co-

| Model | Type | Payoffs | Existence | Strategy | Complexity |
|---|---|---|---|---|---|
| General | Z | Borel | [18] | | 2EXPTIME to $\epsilon$-approximate for parity payoffs [6] |
| TB | Z | Borel | [18] | (*) pure | NP $\cap$ co-NP for parity payoffs [2] |
| General | NZ | Safety | NE [23] | ? | ? |
| General | NZ | Reachability | (*) $\epsilon$-NE | (*) memoryless | (*) NP to $\epsilon$-approximate |
| General | NZ | Borel | ? | ? | ? |
| General | Z | Reachability | [8] | memoryless $\epsilon$-optimal [8] | (*) NP $\cap$ co-NP to $\epsilon$-approximate |
| TB | NZ | Borel | (*) $\epsilon$-NE | (*) pure | ? |
| TB | NZ | $\Sigma_3 \cap \Pi_3$ | (*) NE | (*) pure | (*) NP for parity payoffs |

**Fig. 1.** Summary of known and new/our (asterisk (*)) results. "General" is for stochastic games, "TB" for turn-based probabilistic games. "Z" denotes zero sum, "NZ" denotes nonzero sum. A '?' denotes open problem. The "Strategy" column indicates existence of an (exact or $\epsilon$) Nash equilibrium with that class of strategies, for example, there exists an $\epsilon$-Nash equilibrium profile in memoryless strategies for $n$-player nonzero sum reachability games.

NP strategy construction algorithm for probabilistic turn-based zero sum parity games [2], we get an NP algorithm to find an exact Nash equilibrium in these games. Figure 1 gives a summary of known and new (our) results in this area.

## 2 Definitions

An *n-player stochastic game* $G$ consists of a finite, nonempty set of states $S$, $n$ players $1, 2, \ldots, n$, finite sets of action sets $A_1, A_2, \ldots, A_n$ for the players, a conditional probability distribution $p$ on $S \times (A_1 \times A_2 \times \cdots \times A_n)$ called the law of motion, and bounded, real valued payoff functions $\phi_1, \phi_2, \ldots, \phi_n$ defined on the history space $H = S \times A \times S \times A \cdots$, where $A = A_1 \times A_2 \times \cdots A_n$. The game is called a $n$-player deterministic game if for all states $s \in S$ and action choices $a = (a^1, a^2, \ldots, a^n)$ there is a unique state $s'$ such that $p(s'|s, a) = 1$.

Play begins at an initial state $s_0 = s \in S$. Each player independently and concurrently selects a mixed action $a_i^1$ with a probability distribution $\sigma_i(s)$ belonging to $\mathcal{P}(A_i)$, the set of probability measures on $A_i$. Given $s_0$ and the chosen mixed actions $a^1 = (a_1^1, a_2^1, \ldots a_n^1) \in A$, the next state $s_1$ has the probability distribution $p(\cdot|s_0, a^1)$. Then again each player $i$ independently selects $a_i^2$ with a distribution $\sigma_i((s_0, a^1, s_1))$ and given $a^2 = (a_1^2, a_2^2, \ldots, a_n^2)$, the next state $s_2$ has the probability distribution $p(\cdot|s_1, a^2)$. Play continues in this way for an infinite number of steps and generates a random history $h = (s_0, a^1, s_1, a^2, \ldots) \in H$. The payoff is decided based on the infinite history.

A function $\pi_i$ that specifies for each partial history $h' = (s_0, a^1, s_1, a^2, \ldots, s_k)$ the conditional distribution $\pi_i(h') \in \mathcal{P}(A_i)$ for player $i$'s next action $a_i^{k+1}$ is called a *strategy* for player $i$. A strategy profile $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$ consists of a strategy $\pi_i$ for each player $i$. A *selector* for a player $i$ is a mapping $\sigma_i : S \to \mathcal{P}(A_i)$. A selector profile $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ consists of a selector

$\sigma_i$ for each player $i$. A selector $\sigma_i$ defines a *memoryless* strategy $\sigma_i^\infty$ for player $i$: $\sigma_i^\infty$ chooses the mixed action $\sigma_i(s')$ each time the play visits $s'$. A strategy profile $\sigma^\infty = (\sigma_1^\infty, \sigma_2^\infty, \ldots, \sigma_n^\infty)$ is a memoryless strategy profile if all the strategies $\sigma_1^\infty, \sigma_2^\infty, \ldots, \sigma_n^\infty$ are memoryless. Given a memoryless strategy profile $\sigma^\infty = (\sigma_1^\infty, \sigma_2^\infty, \ldots, \sigma_n^\infty)$ we write $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ to denote the corresponding selector profile for the players. An initial state $s$ and a strategy profile $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$ together with the law of motion $p$ determine a probability distribution $P_{s,\pi}$ on the history space. We write $E_{s,\pi}$ for the expectation operator associated with $P_{s,\pi}$. A strategy $\pi_i$ for player $i$ is *pure* if for every history $h = (s_0, a^1, s_1, \ldots, a^k, s_k)$ there is a action $a_k \in A_{is_k}$ such that $\pi_i(a) = 1$. In other words, a strategy is pure if for every history the strategy chooses one action rather than a probability distribution over the action set. A strategy profile is pure if all the strategies of the profile are pure.

Given a strategy profile $\tau = (\tau_1, \tau_2, \ldots, \tau_n)$ the strategy profile $\tau_{-i} = (\tau_1, \ldots, \tau_{i-1}, \tau_{i+1}, \ldots, \tau_n)$ is the strategy profile obtained by deleting the strategy $\tau_i$ from $\tau$ whereas for any strategy $\mu_i$ of player $i$, $\rho(\tau_{-i}, \mu_i) = (\tau_1, \ldots, \tau_{i-1}, \mu_i, \tau_{i+1}, \ldots, \tau_n)$ denotes the strategy profile where player $i$ follows $\mu_i$ and the other players follows the strategy of $\tau_{-i}$. Similar definitions hold for selector profiles as well.

Assume now that the payoff functions $\phi_i : H \to \mathbb{R}$ are bounded and measurable, where $\mathbb{R}$ is the set of reals. If the initial state of the game is $s$ and each player $i$ choses a strategy $\pi_i$, then the payoff to each player $i$ is the expectation $E_{s,\pi}\phi_i$, where $\pi$ is the strategy profile $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$. For $\epsilon \geq 0$, an $\epsilon$-equilibrium at the initial state $s$ is a profile $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$ such that, for all $i = 1, 2, \ldots, n$ we have $E_{s,\pi}\phi_i \geq \sup_{\mu_i} E_{s,\rho(\pi_{-i},\mu_i)}\phi_i - \epsilon$, where $\mu_i$ ranges over the set of all strategies for player $i$. In other words, each $\pi_i$ guarantees an expected payoff for player $i$ which is within $\epsilon$ of the best possible expected payoff for player $i$ when every other player $j \neq i$ plays $\pi_j$. A 0-equilibrium is called a *Nash equilibrium* and for every $\epsilon > 0$ an $\epsilon$-equilibrium is called an $\epsilon$-*Nash equilibrium* [11]. A strategy profile $\pi$ for an $\epsilon$-Nash equilibrium is called an $\epsilon$-equilibrium profile. A strategy profile $\pi$ for a Nash equilibrium is called a Nash equilibrium profile.

Let $R_1, R_2, \ldots, R_n$ be subsets of the state space $S$. The subset of states $R_i$ is referred as the *target set* for player $i$. Then let $R_1^\infty, R_2^\infty, \ldots, R_n^\infty$ be the subsets of $H$ defined by $R_i^\infty = \{h = (s_0, a^1, s_1, a^2, \ldots) \mid \exists k \in \mathbb{N}. s_k \in R_i\}$ and take the payoff function $\phi_i^{R_i}$ to be the indicator function of $R_i^\infty$ for $i = 1, 2, \ldots, n$. Thus each player receives a payoff of 1 if the process of states $s_0, s_1, \ldots$ reaches a state in $R_i$ and receives payoff 0 otherwise. We call stochastic games with the payoff functions of this form *reach-a-set games.*

A two-player zero sum reachability game (also called *concurrent reachability game* [5]) $G$ is a two-player stochastic game with $R_1 \subseteq S$ as a target set of states for player 1. Given a random history $h = (s_0, a^1, s_1, a^2, \ldots)$ player 1 gets a payoff 1 if the history contains a state in $R_1$, else the player 2 gets an payoff 1. In other words, player 1 plays a reachability game with target set $R_1$ and player 2 has a strictly competing objective of keeping the game out of $R_1$ forever. Notice that this is not a two-player reach-a-set game.

## 3 Existence of $\epsilon$-Nash Equilibria

The main result of this section is the existence of $\epsilon$-Nash equilibria in $n$-player reach-a-set games for every $\epsilon > 0$.

**Theorem 1** **($\epsilon$-equilibrium).** *An $n$-player reach-a-set game $G$ with finite state and action spaces has an $\epsilon$-Nash equilibrium at every initial state $s \in S$ for every $\epsilon > 0$. Moreover, there is a memoryless $\epsilon$-equilibrium strategy profile for every $\epsilon > 0$.*

Already for two-player zero sum reachability games optimal strategies need not exist [13, 6]. The example can be easily adapted to show that Nash equilibria need not exist even for 2-player reach-a-set games. Hence an $\epsilon$-Nash equilibrium for all $\epsilon > 0$, is the best one can achieve for $n$-player reach-a-set games.

**Memoryless Nash equilibrium in discounted games.** We first prove the existence of a Nash equilibrium in memoryless strategies in a *discounted* $n$-player reach-a-set game. Given a $n$-player game $G$ we use $G^\beta$ to denote a $\beta$-discounted version of the game $G$. The game $G^\beta$ at each step halts with probability $\beta$ (goes to a special sink state *halt* which has a reward 0 for every player) and continues as the game $G$ with probability $1 - \beta$. We refer to $\beta$ as the discount-factor. The proof of the next lemma uses Kakutani's Fixed point theorem to show the existence of Nash equilibria in discounted games [25, 23].

**Lemma 1.** *For every $\beta$-discounted $n$-player reach-a-set game $G^\beta$ there exist selectors $\sigma_i : S \to \mathcal{P}(A_i), i = 1, 2, \ldots, n$, such that the memoryless profile $\sigma^\infty = (\sigma_1^\infty, \sigma_2^\infty, \ldots, \sigma_n^\infty)$ is a Nash equilibrium profile in $G^\beta$ for every $s \in S$.*

**$k$-uniform strategies.** Next, we show that a Nash equilibrium in the discounted game can be approximated using "simple" strategy profiles. We start with some definitions. An MDP is a 1-player stochastic game. A MDP reach-a-set game is a 1-player stochastic reach-a-set game. Fix an $n$-player stochastic game $G$. Let $\sigma^\infty = (\sigma_1^\infty, \sigma_2^\infty, \ldots, \sigma_n^\infty)$ be a memoryless strategy profile and $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ be the corresponding selector profile. Then the game $G_\sigma$ is a Markov chain whose law of motion $p_\sigma$ is defined by the functions in the selector profile $\sigma$ and the law of motion $p$ of the game $G$. Similarly, $G_{\sigma_{-i}}$ is an MDP where the mixed action of each player $j \neq i$ at a state $s$ is fixed according to the selector function $\sigma_j(s)$. The law of the motion $p_{\sigma_{-i}}$ of the MDP is determined by the selectors in $\sigma_{-i}$ and law of motion $p$ of $G$. Given a MDP reach-a-set game $G$ with the target $R$, the *value* of the game at state $s$ is denoted by $v(s) = \sup_\pi E_{s,\pi}\phi_1^R$, where $\pi$ ranges over all strategies and $\phi_1^R$ is the reach-a-set game payoff for the player in the game $G$. Similarly, we use $v^\beta(s) = \sup_\pi E_{s,\pi}^\beta \phi_1^R$, to denote the value at state $s$ in the game $G^\beta$, where $G^\beta$ is the $\beta$-discounted version of the game $G$.

Given a $n$-player reach-a-set game $G$ let $|S|$ denote the size of the state space and $l = \max_i |A_i|$ denote the maximum number of actions available to any player at any state of $G$. A selector function $\sigma_i$ for player $i$ is *pure* if for all states $s \in S$ there is an action $a_i \in A_i$ such that $\sigma_i(s)(a_i) = 1$. A selector function $\sigma_i^k$ for player $i$ is $k$-*uniform* if for all states $s \in S$ there exists a multiset $M$ of pure selectors with $|M| \leq k$ such that $\sigma_i^k$ is the uniform distribution over $M$. A selector profile $\sigma^k = (\sigma_1^k, \sigma_2^k, \ldots, \sigma_n^k)$ is $k$-uniform if all the selectors $\sigma_i^k$ are $k$-uniform for all $i \in \{1, 2, \ldots, n\}$. A memoryless strategy profile $\sigma^{k,\infty} = (\sigma_1^{k,\infty}, \sigma_2^{k,\infty}, \ldots, \sigma_n^{k,\infty})$ is $k$-uniform if the selector profile $\sigma^k = (\sigma_1^k, \sigma_2^k, \ldots, \sigma_n^k)$ corresponding to the strategy profile $\sigma^{k,\infty}$ is $k$-uniform. We use a technical lemma by Lipton et. al. [14] for matrix games (Lemma 2), and a Lipschitz continuity property for MDPs.

**Lemma 2 ([14]).** *Let $J$ be a matrix game with $n$-players and each player has at most $l$ actions. Let $\pi$ be a Nash-equilibrium strategy profile. Then for every $\epsilon > 0$, for every $k \geq \frac{3n^2 \ln n^2 l}{\epsilon^2}$, there exists a $k$-uniform strategy profile $\pi'$ such that: (a) for every action $a$, if $\pi(a) = 0$ then $\pi'(a) = 0$; (b) for every action $a$, if $\pi(a) > 0$ then $|\pi(a) - \pi'(a)| \leq \epsilon$.*

Let $G_1$ and $G_2$ be two MDP's defined on the same state space $S$ and action space $A$ with laws of motion $p_1$ and $p_2$ respectively. The *difference* of the two MDP's, denoted $err(G_1, G_2)$, is defined as: $err(G_1, G_2) = \sum_{s,s' \in S^2, a \in A} |p_1(s|s', a) - p_2(s|s', a)|$.

**Lemma 3.** *Let $G^\beta$ be a $\beta$-discounted $n$-player reach-a-set game. Let $\sigma^\infty$ be a memoryless Nash equilibrium profile with selector profile $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$. For every $\epsilon > 0$, for every $k \geq \frac{3n^2 \ln n^2 l}{(\frac{\epsilon}{n \cdot |S|^2 \cdot l})^2}$, there exists a $k$-uniform memoryless strategy profile $\sigma^{k,\infty}$ (with selector profile $\sigma^k$) such that for all players $i$, the MDPs $G_{\sigma_{-i}}$ and $G_{\sigma^k_{-i}}$ satisfy that $err(G_{\sigma_{-i}}, G_{\sigma^k_{-i}}) \leq \epsilon$.*

*Proof.* Consider any player $i$ and a player $j \neq i$. Consider the selector profile $\sigma_j$ for player $j$. It follows from Lemma 2 that there is a $k$-uniform selector profile $\sigma^k$, where $k \geq \frac{3n^2 \ln n^2 l}{(\frac{\epsilon}{n \cdot |S|^2 \cdot l})^2}$, such that for player $j$ the following conditions hold: (a) every action that is played with probability 0 in $\sigma_j$ is played with probability 0 in $\sigma_j^k$, (b) for every action $a$ that is played with positive probability in $\sigma_j$, the difference in probability $|\sigma_j(s)(a) - \sigma_j^k(s)(a)|$ for any $s \in S$ is at most $\frac{\epsilon}{n \cdot |S|^2 \cdot l}$. Consider the MDPs $G_{\sigma_{-i}}$ and $G_{\sigma^k_{-i}}$, with laws of motion $p_1$ and $p_2$ respectively. Since there are $n$ players, for any pair of states $s$, $s'$, and action $a$, the difference in probabilities $|p_1(s' \mid s, a) - p_2(s' \mid s, a)|$ is at most $\frac{\epsilon}{|S|^2 \cdot l}$. Since size of the state space in $S$ and the number of actions of player $i$ is at most $l$, there can be at most $|S|^2 \cdot l$ edges. Hence the result follows. ∎

Lemma 3 and Lipschitz continuity of values of MDP reach-a-set game with respect to *err* [9] gives the following lemma.

**Lemma 4.** *Let $G^\beta$ be a $\beta$-discounted $n$-player reach-a-set game. For every $\epsilon > 0$, for every $k \geq \frac{3n^4 |S|^4 l^2 \ln n^2 l}{\epsilon^2}$, there exists a $k$-uniform memoryless strategy* *profile $\sigma^{k,\infty} = (\sigma_1^{k,\infty}, \sigma_2^{k,\infty}, \ldots, \sigma_n^{k,\infty})$ such that $\sigma^{k,\infty}$ is an $\epsilon$-Nash equilibrium profile in the game $G^\beta$.*

Given any game $G$ we denote by $p_{\min}$ the minimum non-zero transition probability in the law of motion $p$ of $G$; formally $p_{\min} = \min\{p(t \mid s, a_1 \ldots a_n) \mid p(t \mid s, a_1 \ldots a_n) > 0\}$. The following results follows from Lemma 1 and Lemma 4.

**Lemma 5.** *Given a $n$-player $\beta$-discounted stochastic reach-a-set game $G^\beta$, for every $\epsilon > 0$ there exists a memoryless strategy profile $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$ such that $\pi$ is an $\epsilon$-Nash equilibrium profile in the game $G^\beta$ and for every player $i$ the minimum transition probability in the MDP $G_{\sigma^k_{-i}}$ is at least $p_{\min} \cdot \left(\frac{\epsilon^2}{3n^4|S|^4 l^2 \ln n^2 l}\right)^n$.*

Our final ingredient is a result by Condon [3] that relates $p_{\min}$ with the discount factor.

**Lemma 6.** *Let $G$ be a MDP reach-a-set game such that the minimum nonzero transition probability $p_{\min}$ of the MDP is at least $\alpha$. For every $\epsilon > 0$ there is a discount factor $\beta = \epsilon\alpha^{O(|S|)}$, such that for all states $s \in S$ of the $\beta$-discounted game $G^\beta$ we have $v(s) - v^\beta(s) \leq \epsilon$.*

**Nash equilibrium.** We now prove Theorem 1 using the results on discounted reach-a-set games above. Given the game $G$ and $\epsilon > 0$ let $\alpha$ be $p_{\min}\left(\frac{\frac{\epsilon}{2}}{3n^4|S|^4 l^2 \ln n^2 l}\right)^n$. We construct a game $G^\beta$ which is a discounted version of $G$ with discount-factor $\beta$, such that $\beta = (\frac{\epsilon}{2})(\alpha)^{O(|S|)}$. Let $\sigma^{k,\infty} = (\sigma_1^{k,\infty}, \sigma_2^{k,\infty}, \ldots, \sigma_n^{k,\infty})$ be a memoryless strategy profile such that $\sigma^{k,\infty}$ is an $\frac{\epsilon}{2}$-Nash equilibrium profile in the game $G^\beta$ and for every player $i$ the minimum transition probability in the MDP $G_{\sigma^k_{-i}}$ is at least $\alpha$ (existence follows from Lemma 5). Consider any player $i$ and the strategy profile $\sigma_{-i}^{k,\infty}$. The game $G_{\sigma^k_{-i}}$ is a MDP where the mixed actions of all the other players are fixed according to the $\sigma^k_{-i}$. Also, $G^\beta_{\sigma^k_{-i}}$ is the MDP which is the $\beta$-discounted version of the game $G_{\sigma^k_{-i}}$. Hence for every player $i$ we have

$$\begin{aligned}
\sup_{\mu_i} E_{s,\rho(\sigma^k_{-i},\mu_i)} \phi_i^{R_i} &\leq v^\beta(s) + \tfrac{\epsilon}{2} \text{ (from Lemma 6)} \\
&\leq \left(E^\beta_{s,\sigma^{k,\infty}} \phi_i^{R_i} + \tfrac{\epsilon}{2}\right) + \tfrac{\epsilon}{2} \\
&\leq E_{s,\sigma^{k,\infty}} \phi_i^{R_i} + \epsilon
\end{aligned}$$

Hence $\sigma^{k,\infty} = (\sigma_1^{k,\infty}, \sigma_2^{k,\infty}, \ldots, \sigma_n^{k,\infty})$ is an $\epsilon$-Nash equilibrium profile in $G$.

## 4 Computational Complexity

Let $\pi$ be an $\epsilon$-equilibrium profile. Let $v_i^\pi(s) = E_{s,\pi}\phi_i^{R_i}$ denote the value at state $s$ for player $i$ for the strategy profile $\pi$. The value of an $\epsilon$-equilibrium profile $\pi$ at a state $s$ is the value vector $\boldsymbol{v}^\pi(s) = (v_1^\pi(s), v_2^\pi(s), \ldots, v_n^\pi(s))$. Our main results about the computational complexity of computing the value of any $\epsilon$-equilibrium profile within a tolerance of $\epsilon$ are summarized below.

**Theorem 2 (Computing values of a memoryless equilibrium profile).** *Let $G$ be an $n$-player reach-a-set game, let $s$ be a state in $G$, and let $\boldsymbol{v} = (v_1, v_2, \ldots, v_n) \in [0,1]^n$ be a value vector.*
**(1)** *For any constant $\epsilon > 0$, the complexity of computing if there is an $\epsilon$-equilibrium profile $\pi$ in memoryless strategies such that for all player $i$ we have $v_i^\pi(s) \geq v_i - \epsilon$ is NP.*
**(2)** *The complexity of determining whether there is a pure Nash equilibrium strategy profile $\pi$ such that the value $v_i^\pi(s)$ for every player $i$ from the state $s$ for the profile $\pi$ is greater than equal to $v_i$ is NP-complete. The problem is NP-hard already for deterministic reach-a-set games.*

**Approximating in NP.** We will prove that for every fixed $\epsilon > 0$ the value of some $\epsilon$-Nash equilibrium can be approximated in NP.

The following Lemma follows by approximating a memoryless $\epsilon$-Nash equilibrium (existence follows from Theorem 1) by $k$-uniform memoryless equilibrium using arguments similar to Lemmas 3 and 4.

**Lemma 7.** *Given a $n$-player stochastic reach-a-set game $G$ for every $\epsilon > 0$, there exists, for every $k \geq \frac{12n^4|S|^4 l^2 \ln n^2 l}{\epsilon^2}$ a $k$-uniform memoryless strategy profile $\sigma^{k,\infty}$ such that $\sigma^{k,\infty}$ is a $\epsilon$-equilibrium profile for the game $G$.*

**Lemma 8.** *Given a constant $\epsilon$ the value of an $\epsilon$-equilibrium with a memoryless strategy profile of a $n$-player stochastic reach-a-set game can be approximated within $\epsilon$ tolerance by an NP algorithm.*

*Proof.* The NP algorithm guesses a $k$-uniform selector $\sigma^k$ for a $k$-uniform memoryless $\epsilon$-equilibrium strategy profile $\sigma^{k,\infty}$. It then verifies that the value for the MDP's $G_{\sigma^k_{-i}}$ for every state $s \in S$ and each player $i$ is within $\epsilon$-tolerance as compared to the value of the Markov chain define by $G_{\sigma^k}$. Since the computation of values of a MDP can be achieved in polynomial time (using linear programming) it follows that the approximation within $\epsilon$ tolerance can be achieved by a NP-algorithm. ∎

**NP-completeness.** We first prove it is NP-hard to compute a pure, memoryless Nash equilibrium profile for $n$-player deterministic reach-a-set games by reduction from 3-SAT. Given a 3-SAT formula $\psi$ with $n$-clauses and $m$-variables we will construct a $n$-player deterministic reach-a-set game $G_\psi$. Let the variables in the formula $\psi$ be $x_1, x_2, \ldots, x_m$ and the clauses be $C_1, C_2, \ldots, C_n$. In the game $G_\psi$ each clause is a player. The state space $S$, the law of motion and the target states are defined as follows. There is a state $i$ for each variable $x_i$, two sink states $sink$ and $m+1$, and states $(i,j)$ for $i = 1, \ldots, m$ and $j = 0, 1$ for an assignment $j$ to the variable $x_i$: $S = \{ 1, 2, \ldots, m, m+1 \} \cup \{ (i,j) \mid i = 1, \ldots, m, \text{ and } j = 0, 1 \} \cup \{sink\}$. For any state $(i,j)$ the game always moves to the state $i+1$. Let $C_i = \{c_{i_1}, c_{i_2}, \ldots, c_{i_k}\}$ be the set of clauses in which variable $x_i$ occurs. Then, in state $i$ players $i_1, i_2, \ldots, i_k$ have a choice of moves between $\{0, 1\}$. If all the players chose move 0 the game proceeds to state $(i, 0)$, if all the players chose move 1 the game proceeds to state $(i, 1)$, else the game goes to the sink state. Once the game reaches the sink state or the state $m+1$ it remains there for ever. Let $C_i^0 = \{c^0_{k_1}, c^0_{k_2}, \ldots, c^0_{k_l}\}$ be the set of clauses that are satisfied assigning $x_i = 0$, then the state $(i,0)$ is a target state for players $k_1, k_2, \ldots, k_l$. Similarly, let $C_i^1 = \{c^1_{k'_1}, c^1_{k'_2}, \ldots, c^1_{k'_j}\}$ be the set of clauses that are satisfied by assigning the variable $x_i = 1$ then the state $(i,1)$ is a target state for players $k'_1, k'_2, \ldots, k'_j$. States $1, 2, \ldots, m+1$ and the *sink* state is not a target state for any player.

We reduce the 3-SAT problem $\psi$ to the problem of determining whether there is an equilibrium in $G_\psi$ such that each player has a value $\geq v_i$ at state $s$. Each player gets a value 1 at state 1 iff the formula $\psi$ is satisfiable. If the formula is satisfiable then consider a satisfying assignment to the variables. Then at each state $i$ all the players chose the move as specified by the satisfying assignment and hence every player get a payoff 1. If there is a pure memoryless strategy profile such that all the players get a payoff 1 in the game $G_\psi$ then it follows from the construction of $G_\psi$ that there is an assignment such that every clause is satisfied and hence the 3-SAT formula $\psi$ is satisfiable.

Inclusion in NP is proved by guessing a pure memoryless strategy profile (i.e., at each state, guess an action for each player), and verifying that in the resulting Markov chain, each player $i$ gets at least the payoff $v_i$, and also that in the Markov decision process obtained by fixing the strategies of all but player $i$, player $i$ cannot improve his expected payoff. This proves Theorem 2(2).

**Concurrent reachability games.** The techniques developed for $n$-player nonzero sum games can be used to find better bounds in two-player zero sum reachability games. Our proof goes through a special case of two-player (nonzero sum) reach-a-set games, namely two-player *constant-sum* games. For a two-player reach-a-set game $G$ let $\Pi$ denote the set of all $\epsilon$-equilibrium strategy profiles for $\epsilon \geq 0$. We use the following notation: $v_1(s) = \sup_{\pi\in\Pi} E_{s,\pi}\phi_1^{R_1}$ and $v_2(s) = \sup_{\pi\in\Pi} E_{s,\pi}\phi_2^{R_2}$. A two-player reach-a-set game is constant-sum if for all states $s \in S$ we have (a) $v_1(s)+v_2(s) = 1$ and (b) for all $\pi \in \Pi$, $E_{s,\pi}\phi_1^{R_1} + E_{s,\pi}\phi_1^{R_2} = 1$. For this special case, we prove a NP $\cap$ co-NP bound to approximate the value of a $\epsilon$-equilibrium profile, for any fixed $\epsilon$.

**Lemma 9.** *Let $G$ be a two-player constant-sum reach-a-set game, $s$ an initial state and $v^1$ and $v^2$ be two values. For a fixed $\epsilon$ it can be determined in NP $\cap$ co-NP whether $v_1(s) \geq v^1 - \epsilon$ and $v_2(s) \geq v^2 - \epsilon$.*

*Proof.* Inclusion in NP follows as a special case of Lemma 8. To prove inclusion in co-NP, consider the case when $v_1(s) < v^1 - \epsilon$. The following NP algorithm determines if $v_1(s) < v^1 - \epsilon$. It guesses the $k$-uniform selector $\sigma_2^k$ for player 2 and verifies that the value of player 1 in the state $s$ in the MDP $G_{\sigma^k_{-1}}$ is less than $v^1 - \epsilon$. From Lemma 7 we know that there is a $k$-uniform memoryless $\epsilon$-equilibrium profile $\sigma^{k,\infty} = (\sigma_1^{k,\infty}, \sigma_2^{k,\infty})$ with selector profile $\sigma^k = (\sigma_1^k, \sigma_2^k)$. Since the value of a MDP at any state can be computed in polynomial time (using linear programming) the required result follows. ∎

We now reduce two-player zero sum reach-a-set games to two player constant-sum reach-a-set games in the following way. First, we compute the set of states where the players have value 1 (i.e., for each player $i \in \{1, 2\}$, the set of states $W_i$ where player $i$ can ensure that she gets payoff arbitrarily close to 1). This can be done in polynomial time [5]. Second, consider the constant-sum reach-a-set game $G'$ constructed from the concurrent reachability game $G$ where the set of states $W_1$ and $W_2$ are converted to *sink* states and the objective for player $i$ is to reach the set $W_i$. Then, we can show the value obtained by player 1 in the game $G$ is equal to the value $v_1(s)$ obtained by player 1 in $G'$. This gives an NP $\cap$ co-NP algorithm for two player zero sum reachability games, improving the previously known EXPTIME bound [6].

**Theorem 3.** *The value of a two-player concurrent reachability game can be approximated within $\epsilon$-tolerance in NP $\cap$ co-NP, for a fixed $\epsilon > 0$.*

The natural question at this point is whether there is a polynomial time algorithm for concurrent zero sum reachability games. Since simple stochastic games [3] can be easily reduced to concurrent reachability games, a polynomial time algorithm for this problem will imply a polynomial time algorithm for simple stochastic games. This is a long standing open problem.

## 5 Games with Turns

An $n$-player stochastic game is *turn-based* (or perfect information) if at each state, there is exactly one player who determines the next state. Formally, we extend the action sets $A_i$ for $i = 1, \ldots, n$ to be state dependent, that is, for each state $s \in S$, there are action sets $A_{is}$ for $i = 1, \ldots, n$, and we restrict the action sets so that for any $s \in S$, there is at most one $i \in \{1, \ldots, n\}$ such that $|A_{is}| > 1$.

We consider payoff functions that are index sets of *Borel sets* (see e.g., [12] for definitions), that is, given a Borel set $B$, we consider a payoff function $\chi_B$ that assigns a payoff 1 to a play that is in the set $B$, and 0 to a play that is not in the set $B$. With abuse of notation, we identify the set $B$ with the payoff function $\chi_B$. We consider turn based games in which each player is given a Borel payoff $B_i$. A two-player Borel game is *zero sum* if the payoff set $B$ of one player is the complement $S^\omega \setminus B$ of the other player, that is, the players have strictly opposing objectives. A deep result by Martin shows that two player zero sum infinite stochastic games with Borel payoffs have a value [18]. The proof constructs, for each real $v \in (0, 1]$ a zero sum turn-based deterministic infinite-state game with Borel payoff such that a (pure) winning strategy for player 1 in this game can be used to construct a (mixed) winning strategy in the original game that assures player 1 a payoff of at least $v$. From the determinacy of turn-based deterministic games with Borel payoffs [17], the existence of value in zero sum stochastic games with Borel payoffs follows. Moreover, the proof constructs $\epsilon$-optimal mixed winning strategies. A careful inspection of Martin's proof in the special case of turn-based probabilistic games shows that the $\epsilon$-optimal strategies of player 1 are pure. This is because the mixed strategies are derived from solving certain one-shot concurrent games at each round. In our

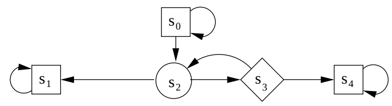


**Fig. 2.** A turn-based probabilistic game

special case these one-shot games have pure winning strategies since only one player has a choice of actions.

**Lemma 10.** *For each $\epsilon > 0$ there is a pure strategy $\pi_1$ of player 1 such that for all strategies $\pi_2$ of player 2 we have $E_s^{\pi_1,\pi_2}\{ f \} \geq v - \epsilon$.*

**Theorem 4.** *For each $\epsilon > 0$ there exists an $\epsilon$-Nash equilibrium in every $n$-player turn based probabilistic game with Borel payoffs.*

*Proof.* Our construction is based on a general construction from repeated games. The basic idea is that player $i$ plays optimal strategies in the zero sum game against all other players, and any deviation by player $i$ is punished indefinitely by the other players by playing $\epsilon$-optimal spoiling strategies in the zero sum game against player $i$ (see, e.g., [19, 28]). Let player $i$ have the payoff set $B_i$, for $i = 1, \ldots, n$. Consider the $n$ zero sum games played between $i$ and the team $[n] \setminus \{ i \}$, with the winning objective $B_i$ for $i$. By Lemma 10 here is a pure $\epsilon$-optimal strategy $\pi_i^i$ for player $i$ in this game, and a pure $\epsilon$-optimal spoiling strategy for players $j \neq i$. This spoiling strategy induces a strategy $\pi_i^j$ for each player $j \neq i$. Now consider the strategy $\tau^i$ for player $i$ as follows. Player $i$ plays the strategy $\pi_i^i$ as long as all the other players $j$ play $\pi_j^j$ and switch to $\pi_j^i$ as soon as some player $j$ deviates. Since the strategies are pure, any deviation is immediately noted. The strategies $\tau^i$ for $i = 1, \ldots, n$ form an $\epsilon$-Nash equilibrium. ∎

Note that the construction above for probabilistic Borel games guarantees only $\epsilon$-optimality. Example 1 shows that there are two player turn-based probabilistic zero sum games where only $\epsilon$-optimal strategies exist, and optimal strategies do not exist. Hence Theorem 4 cannot be strengthened to Nash equilibrium.

*Example 1.* Consider the turn-based probabilistic game shown in Figure 2. At state $s_0$ player 1 chooses between two actions: $a_1 = s_0 \to s_1$ and $a_2 = s_0 \to s_2$. At state $s_1$ the play actions to state $s_2$ and $s_3$ with probability $\frac{1}{2}$. At state $s_3$ player 2 chooses between two actions: $b_1 = s_3 \to s_2$ and $b_2 = s_3 \to s_4$. The game is zero sum, and the objective $B$ for player 1 is:

$$\begin{aligned} B = & \{\pi_1\pi_2 \ldots \mid \pi_1 = s_0 \wedge \exists k.\ \ \pi_k = s_1\} \\ & \cup \{\pi_1\pi_2 \ldots \mid \pi_1 = s_0 \wedge \exists k.\ \ \pi_1 = \pi_2 = \ldots = \pi_k = s_0\ \wedge\ \pi_{k+1} \neq s_0 \\ & \qquad \wedge\ \exists j.\ \ k+1 \leq j \leq 3k.\ \pi_j = s_4\} \end{aligned}$$

Informally, the winning condition for player 1 is as follows. Player 1 wins if either the game reaches the state $s_1$, or the game reaches $s_4$ and the number of visits to $s_0$ is greater than the number of visits to $s_3$ before $s_4$ is hit. Player 1 can get payoff arbitrarily close to 1 in the following way. For any $\epsilon > 0$, consider the strategy $\sigma_1$ for player 1 that chooses the action $a_1 = s_0 \to s_0$ for $k$ times and then chooses the action $a_2 = s_0 \to s_2$, where $\epsilon > \frac{1}{2^k}$. The strategy $\sigma_1$ ensures that player 1 wins with probability at least $1 - \epsilon$. On the other hand, player 1 has no optimal strategy. Consider any strategy $\sigma_1$ for player 1 that chooses the action $a_2 = s_0 \to s_2$ after $k$ steps. The strategy for player 2 is as follows: choose action $b_1 = s_3 \to s_2$ for $k+1$ times and then choose the action $b_2 = s_3 \to s_4$. The probability that player 1 wins is at most $1 - \frac{1}{2^{k+1}}$. Hence, player 1 has $\epsilon$-optimal strategies for every $\epsilon > 0$, but no optimal strategy. ∎

Exact Nash equilibria do exist in special cases. First, the determinacy result for zero sum turn based deterministic games with payoffs corresponding to Borel sets [17] shows the existence of optimal winning strategies. Hence, the construction of Theorem 4 gives exact Nash equilibria.

**Corollary 1.** *Every $n$-player turn-based deterministic game with Borel payoffs has a Nash equilibrium with a pure strategy profile.*

Second, in the special case of turn-based probabilistic games with parity winning conditions, pure and memoryless optimal winning strategies exist for two player zero-sum case [2]. Moreover, the pure memoryless optimal strategies can be computed in NP $\cap$ co-NP. (Notice that the winning condition in Example 1 is not $\omega$-regular.) Therefore we have the following.

**Theorem 5.** *There exists a Nash equilibrium with pure strategy profile in every turn-based probabilistic game with parity payoff conditions. The value profile of some Nash equilibrium can be computed in FNP.*

**Acknowledgments.** We are grateful to Prof. William Sudderth for pointing out an error in the earlier version of the paper. We thank Antar Bandyopadhyay and Tom Henzinger for many insightful discussions.

## References

1. D. Blackwell and T.S. Ferguson. The big match. *Annals of Mathematical Statistics*, 39:159–163, 1968.
2. K. Chatterjee, M. Jurdziński, and T.A. Henzinger. Quantitative stochastic parity games. In *SODA 04*, pages 114–123, 2004.
3. A. Condon. The complexity of stochastic games. *Information and Computation*, 96(2):203–224, 1992.
4. V. Conitzer and T. Sandholm. Complexity results about Nash equilibria. In *IJCAI 03*, pages 765–771, 2003.
5. L. de Alfaro, T.A. Henzinger, and O. Kupferman. Concurrent reachability games. In *FOCS 98*, pages 564–575. IEEE Computer Society Press, 1998.
6. L. de Alfaro and R. Majumdar. Quantitative solution of omega-regular games. In *STOC 01*, pages 675–683. ACM Press, 2001.

7. E.A. Emerson and C. Jutla. Tree automata, mu-calculus and determinacy. In *FOCS 91*, pages 368–377. IEEE Computer Society Press, 1991.
8. H. Everett. Recursive games. In *Contributions to the Theory of Games III*, volume 39 of *Annals of Mathematical Studies*, pages 47–78, 1957.
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. A. Kechris. *Classical Descriptive Set Theory*. Springer, 1995.
13. P.R. Kumar and T.H. Shiau. Existence of value and randomized strategies in zero-sum discrete-time stochastic dynamic games. *SIAM J. Control and Optimization*, 19(5):617–634, 1981.
14. R.J. Lipton, E. Markakis, and A. Mehta. Playing large games using simple strategies. In *EC 03*, pages 36–41. ACM Press, 2003.
15. A. Maitra and W.D. Sudderth. Borel stay-in-a-set games. *International Journal of Game Theory*, 32:97–108, 2003.
16. Z. Manna and A. Pnueli. *The Temporal Logic of Reactive and Concurrent Systems: Specification*. Springer-Verlag, 1992.
17. D.A. Martin. Borel determinacy. *Annals of Mathematics*, 102(2):363–371, 1975.
18. D.A. Martin. The determinacy of Blackwell games. *The Journal of Symbolic Logic*, 63(4):1565–1581, 1998.
19. G. Owen. *Game Theory*. Academic Press, 1995.
20. C.H. Papadimitriou. On the complexity of the parity argument and other inefficient proofs of existence. *JCSS*, 48(3):498–532, 1994.
21. C.H. Papadimitriou. Algorithms, games, and the internet. In *STOC 01*, pages 749–753. ACM Press, 2001.
22. 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.
23. P. Secchi and W.D. Sudderth. Stay-in-a-set games. *International Journal of Game Theory*, 30:479–490, 2001.
24. L.S. Shapley. Stochastic games. *Proc. Nat. Acad. Sci. USA*, 39:1095–1100, 1953.
25. M. Sobel. Noncooperative stochastic games. *Ann. Math. Stat.*, 42:1930–1935, 1971.
26. F. Thuijsman. *Optimality and Equilibria in Stochastic Games*. CWI-Tract 82, CWI, Amsterdam, 1992.
27. F. Thuijsman and T.E.S. Raghavan. Perfect information stochastic games and related classes. *International Journal of Game Theory*, 26:403–408, 1997.
28. N. Vieille. Two player stochastic games I: a reduction. *Israel Journal of Mathematics*, 119:55–91, 2000.
29. N. Vieille. Two player stochastic games II: the case of recursive games. *Israel Journal of Mathematics*, 119:93–126, 2000.
30. J. von Neumann and O. Morgenstern. *Theory of games and economic behavior*. Princeton University Press, 1947.
31. 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).