Qualitative Reachability in Stochastic BPA Games

Tomáš Brázdil[a,1], Václav Brožek[b,1], Antonín Kučera[a,1], Jan Obdržálek[a,1]

[a] *Faculty of Informatics, Masaryk University, Botanická 68a, 60200 Brno, Czech Republic*
[b] *LFCS, School of Informatics, University of Edinburgh, 10 Crichton Street, Edinburgh EH8 9AB Scotland, United Kingdom*

Abstract

We consider a class of infinite-state stochastic games generated by stateless pushdown automata (or, equivalently, 1-exit recursive state machines), where the winning objective is specified by a regular set of target configurations and a qualitative probability constraint '>0' or '=1'. The goal of one player is to maximize the probability of reaching the target set so that the constraint is satisfied, while the other player aims at the opposite. We show that the winner in such games can be determined in **P** for the '>0' constraint, and in **NP** ∩ **co-NP** for the '=1' constraint. Further, we prove that the winning regions for both players are regular, and we design algorithms which compute the associated finite-state automata. Finally, we show that winning strategies can be synthesized effectively.

1. Introduction

Stochastic games are a formal model for discrete systems where the behavior in each state is either controllable, adversarial, or stochastic. Formally, a stochastic game is a directed graph G with a denumerable set of vertices V which is split into three disjoint subsets $V_\Box$, $V_\Diamond$, and $V_\bigcirc$. For every $v \in V_\bigcirc$, there is a fixed probability distribution over the outgoing edges of v. We also require that the set of outgoing edges of every vertex is nonempty. The

[1] All authors are supported by the research center Institute for Theoretical Computer Science, project No. 1M0545. Tomáš Brázdil and Antonín Kučera are also supported by the Czech Science Foundation, project No. P202/10/1469. Václav Brožek is also supported by Newton International Fellowship from the Royal Society.

game is initiated by putting a token on some vertex. The token is then moved from vertex to vertex by two players, $\Box$ and $\Diamond$, who choose the next move in the vertices of $V_\Box$ and $V_\Diamond$, respectively. In the vertices of $V_\bigcirc$, the outgoing edges are chosen according to the associated fixed probability distribution. A *quantitative winning objective* is specified by some Borel set W of infinite paths in G and a probability constraint $\rhd\varrho$, where $\rhd \in \{>, \geq\}$ is a comparison and $\varrho \in [0,1]$. An important subclass of quantitative winning objectives are *qualitative winning objectives* where the constant ϱ must be either 0 or 1. The goal of player $\Box$ is to maximize the probability of all runs that stay in W so that it is $\rhd$-related to ϱ, while player $\Diamond$ aims at the opposite. A *strategy* specifies how a player should play. In general, a strategy may or may not depend on the history of a play (we say that a strategy is *history-dependent (H)* or *memoryless (M)*), and the edges may be chosen deterministically or randomly (*deterministic (D)* and *randomized (R)* strategies). In the case of randomized strategies, a player chooses a probability distribution on the set of outgoing edges. Note that deterministic strategies can be seen as restricted randomized strategies, where one of the outgoing edges has probability 1. Each pair of strategies (σ, π) for players $\Box$ and $\Diamond$ determines a *play*, i.e., a unique Markov chain obtained from G by applying the strategies σ and π in the natural way. The *outcome* of a play initiated in v is the probability of all runs initiated in v that are contained in the set W (this probability is denoted by $\mathcal{P}_v^{\sigma,\pi}(W)$). We say that a play is $(\rhd\varrho)$-won by player $\Box$ if its outcome is $\rhd$-related to ϱ; otherwise, the play is $(\not\rhd\varrho)$-won by player $\Diamond$. A strategy σ of player $\Box$ is *$(\rhd\varrho)$-winning* if for every strategy π of player $\Diamond$, the corresponding play is $(\rhd\varrho)$-won by player $\Box$. Similarly, a strategy π of player $\Diamond$ is *$(\not\rhd\varrho)$-winning* if for every strategy σ of player $\Box$, the corresponding play is $(\not\rhd\varrho)$-won by player $\Diamond$. A natural question is whether the game is *determined*, i.e., for every choice of $\rhd$ and ϱ, either player $\Box$ has a $(\rhd\varrho)$-winning strategy, or player $\Diamond$ has a $(\not\rhd\varrho)$-winning strategy. The answer is somewhat subtle. A celebrated result of Martin [22] (see also [20]) implies that stochastic games with Borel winning conditions are *weakly determined*, i.e., each vertex v has a *value* given by

$$val(v) \quad = \quad \sup_{\sigma} \inf_{\pi} \mathcal{P}_v^{\sigma,\pi}(W) \quad = \quad \inf_{\pi} \sup_{\sigma} \mathcal{P}_v^{\sigma,\pi}(W) \tag{1}$$

Here σ and π range over the sets of all strategies for player $\Box$ and player $\Diamond$, respectively. From this we can immediately deduce the following:

- If both players have *optimal* strategies that guarantee the outcome

$val(v)$ or better against every strategy of the opponent (for example, this holds for *finite-state* stochastic games and the "usual" classes of quantitative/qualitative Borel objectives), then the game is determined for every choice of $\triangleright\varrho$.

- Although optimal strategies are not guaranteed to exists in general, equation 1 implies the existence of ε-optimal strategies (see Definition 2.3) for every $\varepsilon > 0$. Hence, the game is determined for every choice of $\triangleright\varrho$ where $\varrho \neq val(v)$.

The only problematic case is the situation when optimal strategies do not exist and $\varrho = val(v)$. The example given in Figure 1 at page 11 witnesses that such games are generally *not* determined, even for reachability objectives. On the other hand, we show that *finitely-branching* games (such as BPA games considered in this paper) with reachability objectives *are* determined, although an optimal strategy for player $\Box$ in a finitely-branching game does *not* necessarily exist.

Algorithmic issues for stochastic games with quantitative/qualitative winning objectives have been studied mainly for finite-state stochastic games. A lot of attention has been devoted to quantitative *reachability objectives*, including the special case when $\varrho = \frac{1}{2}$. The problem whether player $\Box$ has a $(>\frac{1}{2})$-winning strategy is known to be in **NP** $\cap$ **co-NP**, but its membership to **P** is a long-standing open problems in algorithmic game theory [10, 25]. Later, more complicated qualitative/quantitative ω-regular winning objectives (such as Büchi, co-Büchi, Rabin, Street, Muller etc.) were considered, and the complexity of the corresponding decision problems was analyzed. We refer to [7–9, 11, 24, 26] for more details. As for infinite-state stochastic games, the attention has so far been focused on stochastic games induced by lossy channel systems [1, 4] and by pushdown automata (or, equivalently, recursive state machines) [5, 13–16]. In the next paragraphs, we discuss the latter model in greater detail because these results are closely related to the results presented in this paper.

A *pushdown automaton (PDA)* (see, e.g., [19]) is equipped with a finite control unit and an unbounded stack. The dynamics is specified by a finite set of rules of the form $pX \hookrightarrow q\alpha$, where p, q are control states, X is a stack symbol, and α is a (possibly empty) sequence of stack symbols. A rule of the form $pX \hookrightarrow q\alpha$ is applicable to every configuration of the form $pX\beta$ and produces the configuration $q\alpha\beta$. If there are several rules with the same left-hand side, one of them must be chosen, and the choice is made by player $\Box$, player $\Diamond$, or it is randomized. Technically, the set of all left-hand sides (i.e., pairs of the form pX) is split into three disjoint subsets $H_\Box$, $H_\Diamond$, and $H_\bigcirc$, and for all $pX \in H_\bigcirc$ there is a fixed probability distribution over the set of all rules of the form $pX \hookrightarrow q\alpha$. Thus, each PDA induces the associated infinite-state stochastic game where the vertices are PDA configurations and the edges are determined in the natural way. An important subclass of PDA is obtained by restricting the number of control states to 1. Such PDA are also known as *stateless PDA* or *BPA* (the BPA acronym stands for Basic Process Algebra and it is borrowed from concurrency theory where it refers to a calculus expressively equivalent (up to bisimilarity) to stateless PDA [3]). PDA and BPA correspond to *recursive state machines (RSM)* and *1-exit RSM* respectively, in the sense that their descriptive powers are equivalent, and there are effective linear-time translations between the corresponding models.

The syntax of PDA and BPA is particularly apt for modelling programs with recursive procedure calls. The stack symbols are used to represent procedures and their local data, and the global data are stored in the finite control. A procedure call is modeled by pushing a new symbol onto the stack, and a return to a calling procedure corresponds to popping the associated stack symbol. The configurations where a system interacts with its environment are "adversarial" and belong to player $\Diamond$. Player $\Box$ corresponds to a controller which can influence the program's behaviour in certain configurations. Stochastic choice is used to reflect some known distribution on input data or to model randomization (e.g., "coin flips" in randomized algorithms). A winning strategy for player $\Box$ in a given stochastic PDA or BPA game then corresponds to a controller which guarantees a given property (objective) no matter what the unpredictable environment does.

Stochastic PDA and BPA games are closely related to other well-known stochastic models such as multi-type branching processes [2, 18] or stochastic context-free grammars that are widely used in natural language processing [21] and biology. Another closely related model are backoff processes [17] which correspond to random walks over a finite graph with "back button".

Known results. The existing results about stochastic BPA and PDA games concern *termination objectives* (a restricted form of reachability objectives), and *BPA Markov decision processes* (i.e., stochastic BPA games where $H_\Diamond = \emptyset$ or $H_\Box = \emptyset$) with reachability objectives.

- *Results about stochastic PDA and BPA games with termination ob-*

jectives. A terminating run is a run which hits a configuration with the empty stack. Hence, termination is a special form of reachability. Stochastic PDA and BPA games with quantitative and qualitative objectives are examined in [14]. For BPA, it is shown that the vector of optimal values $(val(X), X \in \Gamma)$, where Γ is the stack alphabet, forms the least solution of an effectively constructible system of min-max equations. Moreover, both players have *optimal* MD strategies which depend only on the topmost stack symbol of a given configuration (such strategies are called SMD, meaning Stackless MD). Hence, stochastic BPA games with quantitative/qualitative termination objectives are determined. Since the least solution of the constructed equational system can be encoded in first order theory of the reals, the existence of a $({\rhd}\varrho)$-winning strategy for player $\Box$ can be decided in polynomial space. In the same paper [14], the $\Sigma_2^P \cap \Pi_2^P$ upper complexity bound for the subclass of qualitative termination objectives is established. As for PDA games, it is shown that for every fixed $\varepsilon > 0$, the problem to distinguish whether the optimal value $val(pX)$ is equal to 1 or less than ε, is undecidable. The $\Sigma_2^P \cap \Pi_2^P$ upper bound for stochastic BPA games with qualitative termination objectives is improved to $\mathbf{NP} \cap \mathbf{co\text{-}NP}$ in [15]. In the same paper, it is also shown that the quantitative reachability problem for finite-state stochastic games (see above) is efficiently reducible to the qualitative termination problem for stochastic BPA games. Hence, the $\mathbf{NP} \cap \mathbf{co\text{-}NP}$ upper bound cannot be further improved without a major breakthrough in algorithmic game theory. In the special case of stochastic BPA games where $H_\Diamond = \emptyset$ or $H_\Box = \emptyset$, the qualitative termination problem is shown to be in $\mathbf{P}$ (observe that if $H_\Diamond = \emptyset$ or $H_\Box = \emptyset$, then a given BPA induces an infinite-state Markov decision process and the goal of the only player is to maximize or minimize the termination probability, respectively).

- *Results about BPA Markov decision processes with qualitative reachability objectives.* BPA Markov decision processes can be seen as stochastic BPA games with only one player (i.e., $H_\Diamond = \emptyset$). In [5], it is shown that the properties of BPA Markov decision processes with reachability objectives are quite different from the ones of termination objectives. In particular, there is no apparent way how to express the vector of optimal values as a solution of some recursive equational system, and the SMD determinacy result (see above) does not hold either. Still, the set of all configurations where player $\Box$ has a strategy such that the probability of all runs satisfying a given reachability objective is $\bowtie$-related to a given $\varrho \in \{0,1\}$, where $\bowtie \in \{<, \leq, >, \geq\}$, is effectively regular and the associated finite-state automaton is computable in polynomial time. Further, one can efficiently compute a finite description of such a strategy.

Our contribution: In this paper, we continue the study initiated in [5, 13–16] and solve the qualitative reachability problem for unrestricted stochastic BPA games. Thus, we obtain a substantial generalization of the previous results.

We start by resolving the determinacy issue in Section 3. We observe that general stochastic games with reachability objectives are *not* determined, and we also show that *finitely branching* stochastic games (such as BPA stochastic games) with quantitative/qualitative reachability objectives *are* determined, i.e., in every vertex, either player $\Box$ has a $(\rhd\varrho)$-winning strategy, or player $\Diamond$ has a $(\not\rhd\varrho)$-winning strategy. This is a consequence of several observations that are specific to reachability objectives and perhaps interesting on their own.

The main results of our paper, presented in Sections 5, 6, and 7 concern stochastic BPA games with qualitative reachability objectives. In the context of BPA, a reachability objective is specified by a *regular* set T of target configurations. We show that the problem of determining the winner in stochastic BPA games with qualitative reachability objectives is in **P** for the '$\rhd 0$' constraint, and in **NP** $\cap$ **co-NP** for the '$\rhd 1$' constraint. Here we rely on the previously discussed results about qualitative termination [15] and use the corresponding algorithms as "black-box procedures" at appropriate places. We also rely on observations presented in [5] which were used to solve the simpler case with only one player. However, the full (two-player) case brings completely new complications that need to be tackled by new methods and ideas. Many "natural" hypotheses turned out to be incorrect (some of the interesting cases are documented in Section 6, see Examples 6.5 and 6.7). We also show that for each $\varrho \in \{0,1\}$, the sets of all configurations where player $\Box$ (or player $\Diamond$) has a $(\rhd\varrho)$-winning (or $(\not\rhd\varrho)$-winning) strategy is effectively regular[2], and the corresponding finite-state automaton is effectively constructible by a deterministic polynomial-time algorithm (for the '$\rhd 1$' constraint, the algorithm needs $\mathbf{NP} \cap \mathbf{co\text{-}NP}$ oracle). Finally, we also give algorithms which *compute* winning strategies if they exist. These strategies are memoryless, and they are also *effectively regular* in the sense that their functionality is effectively expressible by finite-state automata (see Definition 4.3). Hence, winning strategies in stochastic BPA games with qualitative reachability objectives can be effectively implemented.

For the sake of readability, some of the more involved (and long) proofs of Section 6 have been postponed to Section 7. In the main body of the paper, we try to sketch the key ideas and provide some intuition behind the presented technical constructions.

2. Basic Definitions

In this paper, the sets of all positive integers, non-negative integers, rational numbers, real numbers, and non-negative real numbers are denoted by $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{R}^{\geq 0}$, respectively. For every finite or countably infinite set S, the symbol S^* denotes the set of all finite words over S. The length of a given word u is denoted by $|u|$, and the individual letters in u are denoted by $u(0), \cdots, u(|u|-1)$. The empty word is denoted by ε, and we set $|\varepsilon| = 0$. We also use S^+ to denote the set $S^* \smallsetminus \{\varepsilon\}$. For every finite or countably infinite set M, a binary relation $\to \subseteq M \times M$ is *total* if for every $m \in M$ there is some $n \in M$ such that $m \to n$. A *path* in $\mathcal{M} = (M, \to)$ is a finite or infinite sequence $w = m_0, m_1, \ldots$ such that $m_i \to m_{i+1}$ for every i. The *length* of a finite path $w = m_0, \ldots, m_i$, denoted by $|w|$, is $i + 1$. We also use $w(i)$ to denote the element m_i of w, and w_i to denote the path $m_i, m_{i+1}, \ldots$ (by writing $w(i) = m$ or w_i we implicitly impose the condition that $|w| \geq i{+}1$). A given $n \in M$ is *reachable* from a given $m \in M$, written $m \to^* n$, if there is a finite path from m to n. A *run* is an infinite path. The sets of all finite paths and all runs in $\mathcal{M}$ are denoted by $\mathit{FPath}(\mathcal{M})$ and $\mathit{Run}(\mathcal{M})$, respectively. Similarly, the sets of all finite paths and runs that start in a given $m \in M$ are denoted by $\mathit{FPath}(\mathcal{M}, m)$ and $\mathit{Run}(\mathcal{M}, m)$, respectively.

[2]Let us note that if $0 < \varrho < 1$, then the set of all configurations where player $\Box$ (or player $\Diamond$) has a ($\rhd\varrho$)-winning (or ($\not\rhd\varrho$)-winning) strategy is *not* necessarily regular. An example can be found in [6], Proposition 4.7.

Now we recall basic notions of probability theory. Let A be a finite or countably infinite set. A *probability distribution* on A is a function $f : A \to \mathbb{R}^{\geq 0}$ such that $\sum_{a \in A} f(a) = 1$. A distribution f is *rational* if $f(a) \in \mathbb{Q}$ for every $a \in A$, *positive* if $f(a) > 0$ for every $a \in A$, *Dirac* if $f(a) = 1$ for some $a \in A$, and *uniform* if A is finite and $f(a) = \frac{1}{|A|}$ for every $a \in A$. The set of all distributions on A is denoted by $\mathcal{D}(A)$.

A *σ-field* over a set X is a set $\mathcal{F} \subseteq 2^X$ that includes X and is closed under complement and countable union. A *measurable space* is a pair $(X, \mathcal{F})$ where X is a set called *sample space* and $\mathcal{F}$ is a σ-field over X. A *probability measure* over a measurable space $(X, \mathcal{F})$ is a function $\mathcal{P} : \mathcal{F} \to \mathbb{R}^{\geq 0}$ such that, for each countable collection $\{X_i\}_{i \in I}$ of pairwise disjoint elements of $\mathcal{F}$, $\mathcal{P}(\bigcup_{i \in I} X_i) = \sum_{i \in I} \mathcal{P}(X_i)$, and moreover $\mathcal{P}(X) = 1$. A *probability space* is a triple $(X, \mathcal{F}, \mathcal{P})$ where $(X, \mathcal{F})$ is a measurable space and $\mathcal{P}$ is a probability measure over $(X, \mathcal{F})$.

Definition 2.1. *A* Markov chain *is a triple* $\mathcal{M} = (M, \longrightarrow, \mathit{Prob})$ *where* M *is a finite or countably infinite set of* states*,* $\longrightarrow \subseteq M \times M$ *is a total* transition relation*, and Prob is a function which to each* $s \in M$ *assigns a positive probability distribution over the set of its outgoing transitions.*

In the rest of this paper, we write $s \xrightarrow{x} t$ whenever $s \longrightarrow t$ and $\mathit{Prob}(s)(t) = x$. Each $w \in \mathit{FPath}(\mathcal{M})$ determines a *basic cylinder* $\mathit{Run}(\mathcal{M}, w)$ which consists of all runs that start with w. To every $s \in M$ we associate the probability space $(\mathit{Run}(\mathcal{M}, s), \mathcal{F}, \mathcal{P})$ where $\mathcal{F}$ is the σ-field generated by all basic cylinders $\mathit{Run}(\mathcal{M}, w)$ where w starts with s, and $\mathcal{P} : \mathcal{F} \to \mathbb{R}^{\geq 0}$ is the unique probability measure such that $\mathcal{P}(\mathit{Run}(\mathcal{M}, w)) = \Pi_{i=0}^{m-1} x_i$ where $w = s_0, \cdots, s_m$ and $s_i \xrightarrow{x_i} s_{i+1}$ for every $0 \leq i < m$ (if $m = 0$, we put $\mathcal{P}(\mathit{Run}(\mathcal{M}, w)) = 1$).

Definition 2.2. *A* stochastic game *is a tuple* $G = (V, \mapsto, (V_\Box, V_\Diamond, V_\bigcirc), \mathit{Prob})$ *where* V *is a finite or countably infinite set of* vertices*,* $\mapsto \subseteq V \times V$ *is a total* edge relation*,* $(V_\Box, V_\Diamond, V_\bigcirc)$ *is a partition of* V*, and Prob is a* probability assignment *which to each* $v \in V_\bigcirc$ *assigns a positive probability distribution on the set of its outgoing edges. We say that* G *is* finitely branching *if for each* $v \in V$ *there are only finitely many* $u \in V$ *such that* $v \mapsto u$*.*

A stochastic game G is played by two players, $\Box$ and $\Diamond$, who select the moves in the vertices of $V_\Box$ and $V_\Diamond$, respectively. Let $\odot \in \{\Box, \Diamond\}$. A *strategy* for player $\odot$ in G is a function which to each $wv \in V^* V_\odot$ assigns a probability distribution on the set of outgoing edges of v. The sets of all strategies for player $\Box$ and player $\Diamond$ in G are denoted by Σ_G and Π_G (or just by Σ and Π if G is understood), respectively. We say that a strategy τ is *memoryless (M)* if $\tau(wv)$ depends just on the last vertex v, and *deterministic (D)* if $\tau(wv)$ is a Dirac distribution for all wv. Strategies that are not necessarily memoryless are called *history-dependent (H)*, and strategies that are not necessarily deterministic are called *randomized (R)*. Thus, we define the following four classes of strategies: MD, MR, HD, and HR, where MD $\subseteq$ HD $\subseteq$ HR and MD $\subseteq$ MR $\subseteq$ HR, but MR and HD are incomparable.

Each pair of strategies $(\sigma, \pi) \in \Sigma \times \Pi$ determines a unique *play* of the game G, which is a Markov chain $G(\sigma, \pi)$ where V^+ is the set of states, and $wu \xrightarrow{x} wuu'$ iff $u \mapsto u'$ and one of the following conditions holds:

- $u \in V_\Box$ and $\sigma(wu)$ assigns x to $u \mapsto u'$, where $x > 0$;
- $u \in V_\Diamond$ and $\pi(wu)$ assigns x to $u \mapsto u'$, where $x > 0$;
- $u \in V_\bigcirc$ and $u \overset{x}{\mapsto} u'$.

Let $T \subseteq V$ be a set of *target* vertices. For each pair of strategies $(\sigma, \pi) \in \Sigma \times \Pi$ and every $v \in V$, let $\mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T, G))$ be the probability of all $w \in \mathit{Run}(G(\sigma, \pi), v)$ such that w visits some $u \in T$ (technically, this means that $w(i) \in V^*T$ for some $i \in \mathbb{N}_0$). We write $\mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T))$ instead of $\mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T, G))$ if G is understood.

We say that a given $v \in V$ *has a value* in G if $\sup_{\sigma\in\Sigma} \inf_{\pi\in\Pi} \mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T)) = \inf_{\pi\in\Pi} \sup_{\sigma\in\Sigma} \mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T))$. If v has a value, then $\mathit{val}(v, G)$ denotes the *value of* v defined by this equality (we write just $\mathit{val}(v)$ instead of $\mathit{val}(v, G)$ if G is understood). Since the set of all runs that visit a vertex of T is obviously Borel, we can apply the powerful result of Martin [22] (see also Theorem 3.1) and conclude that *every* $v \in V$ has a value.

Definition 2.3. *Let $\varepsilon \geq 0$ and $v \in V$. We say that*

- $\sigma \in \Sigma$ *is* ε-optimal *(or* ε-optimal maximizing*) in v if $\mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T)) \geq \mathit{val}(v) - \varepsilon$ for all $\pi \in \Pi$;*
- $\pi \in \Pi$ *is* ε-optimal *(or* ε-optimal minimizing*) in v if $\mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T)) \leq \mathit{val}(v) + \varepsilon$ for all $\sigma \in \Sigma$.*

A 0-*optimal strategy is called* optimal. *A* (quantitative) reachability objective *is a pair* $(T, \rhd\varrho)$ *where* $T \subseteq V$ *and* $\rhd\varrho$ *is a probability constraint, i.e.,* $\rhd \in \{>, \geq\}$ *and* $\varrho \in [0,1]$. *If* $\varrho \in \{0,1\}$, *then the objective is* qualitative. *We say that*

- $\sigma \in \Sigma$ *is* $(T, \rhd\varrho)$-winning *in* v *if* $\mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T)) \rhd \varrho$ *for all* $\pi \in \Pi$*;*
- $\pi \in \Pi$ *is* $(T, \not\rhd\varrho)$-winning *in* v *if* $\mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T)) \not\rhd \varrho$ *for all* $\sigma \in \Sigma$.

The $(T, \rhd\varrho)$-winning region *of player* $\Box$, *denoted by* $[T]_{\Box}^{\rhd\varrho}$, *is the set of all* $v \in V$ *such that player* $\Box$ *has a* $(T, \rhd\varrho)$*-winning strategy in* v. *Similarly, the* $(T, \not\rhd\varrho)$-winning region *of player* $\Diamond$, *denoted by* $[T]_{\Diamond}^{\not\rhd\varrho}$, *consists of all* $v \in V$ *such that player* $\Diamond$ *has a* $(T, \not\rhd\varrho)$*-winning strategy in* v.

When writing down probability constraints, we usually use $<$1, $=$1, *and* $=$0 *instead of* $\not\geq$1, $\geq$1, *and* $\not>$0, *respectively.*

3. Determinacy of Stochastic Games with Reachability Objectives

In this section we show that finitely-branching stochastic games with quantitative/qualitative reachability objectives are *determined* in the sense that for every quantitative reachability objective $(T, \rhd\varrho)$, each vertex of the game belongs either to $[T]_{\Box}^{\rhd\varrho}$ or to $[T]_{\Diamond}^{\not\rhd\varrho}$ (see Definition 2.3). Let us note that this result cannot be extended to general (infinitely-branching) stochastic games. A counterexample is given in Figure 1, where $T = \{t\}$ is the set of target vertices. Observe that $\mathit{val}(s) = 0$, $\mathit{val}(u) = 1$, $\mathit{val}(v) = \frac{1}{2}$, and none of the two players has an optimal strategy in v. Now suppose that player $\Box$ has a $(T, \rhd\frac{1}{2})$-winning strategy $\hat{\sigma}$ in v. Obviously, there is some fixed $\varepsilon > 0$ such that for every $\pi \in \Pi$ we have that $\mathcal{P}_u^{\hat{\sigma},\pi}(\mathit{Reach}(T)) = 1 - \varepsilon$. Further, player $\Diamond$ has a strategy $\hat{\pi}$ which is $\frac{\varepsilon}{2}$-optimal in every vertex. Hence, $\mathcal{P}_v^{\hat{\sigma},\hat{\pi}}(\mathit{Reach}(T)) < \frac{1}{2}$, which is a contradiction. Similarly, one can show that there is no $(T, \not\rhd\frac{1}{2})$-winning strategy for player $\Diamond$ in v.

For the rest of this section, let us fix a game $G = (V, \mapsto, (V_{\Box}, V_{\Diamond}, V_{\bigcirc}), \mathit{Prob})$ and a set of target vertices T. Also, for every $n \in \mathbb{N}_0$ and every pair of strategies $(\sigma, \pi) \in \Sigma \times \Pi$, let $\mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}_n(T))$ be the probability of all runs $w \in \mathit{Run}(G(\sigma,\pi), v)$ such that w visits some $u \in T$ in at most n transitions (clearly, $\mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T)) = \lim_{n\to\infty} \mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}_n(T))$).

To keep this paper self-contained, we include an elementary proof of Martin's weak determinacy result (see (1)) for the special case of games with reachability objectives. Observe that the game G fixed above is not required to be finite or finitely-branching.

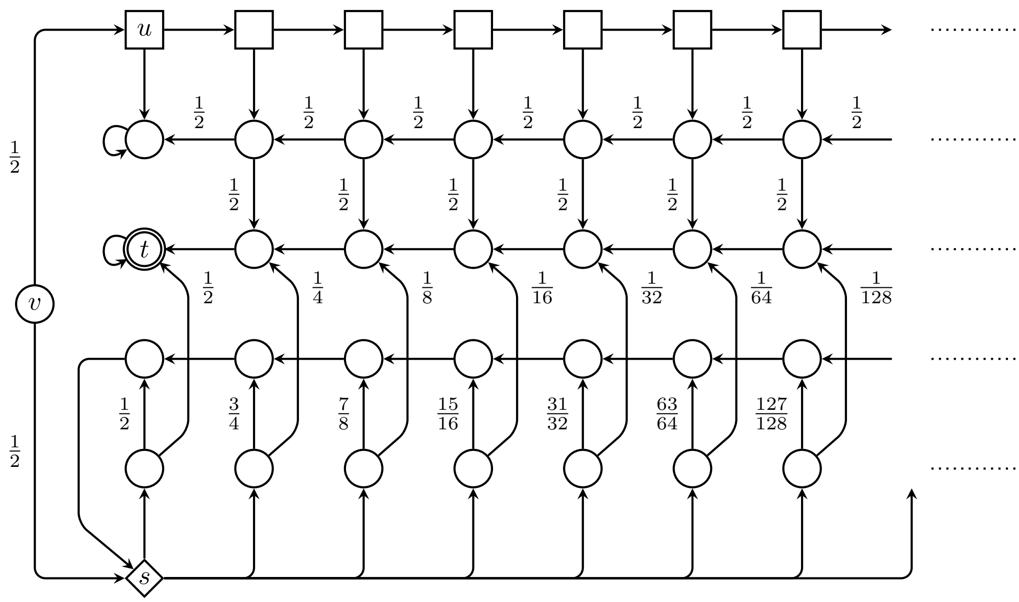


Figure 1: A game which is not determined.

Theorem 3.1. *Every $v \in V$ has a value. Moreover, if G is finitely-branching, then there is a MD strategy $\pi \in \Pi$ which is optimal minimizing in every vertex.*

Theorem 3.1 is proven by showing that the tuple of all values is the least fixed-point of the following (Bellman) functional $\mathcal{V} : (V \to [0,1]) \to (V \to [0,1])$ defined by

$$\mathcal{V}(f)(v) \quad = \quad \begin{cases} 1 & \text{if } v \in T \\ \sup\{f(u) \mid v \mapsto u\} & \text{if } v \in V_\Box \setminus T \\ \inf\{f(u) \mid v \mapsto u\} & \text{if } v \in V_\Diamond \setminus T \\ \sum_{v \overset{x}{\mapsto} u} x \cdot f(u) & \text{if } v \in V_\bigcirc \setminus T \end{cases}$$

Technical details of this proof are given is Section 7.1. Our next lemma says that the ability of player $\Diamond$ to prolong a play in a finitely-branching game is somewhat limited.

Lemma 3.2. *If G is finitely-branching, then for every $v \in V$ we have that*

$$\forall \varepsilon{>}0 \;\; \exists \sigma \in \Sigma \;\; \exists n \in \mathbb{N} \;\; \forall \pi \in \Pi \;:\; \mathcal{P}_v^{\sigma,\pi}(Reach_n(T)) > val(v) - \varepsilon$$

Proof. For all $v \in V$ and $i \in \mathbb{N}_0$, we use $\mathcal{V}_i(v)$ to denote the value of v in G with "reachability in at most i-steps" objective. More precisely, we put $\mathcal{V}_i(v) = 1$ for all $v \in T$ and $i \in \mathbb{N}_0$. If $v \notin T$, we define $\mathcal{V}_i(v)$ inductively as follows: $\mathcal{V}_0(v) = 0$, and $\mathcal{V}_{i+1}(v)$ is equal either to $\max\{\mathcal{V}_i(u) \mid v \mapsto u\}$, $\min\{\mathcal{V}_i(u) \mid v \mapsto u\}$, or $\sum_{v \stackrel{x}{\mapsto} u} x \cdot \mathcal{V}_i(u)$, depending on whether $v \in V_\Box$, $v \in V_\Diamond$, or $v \in V_\bigcirc$, respectively.

A straightforward induction on i reveals that

$$\mathcal{V}_i(v) \quad = \quad \max_{\sigma \in \Sigma} \min_{\pi \in \Pi} \mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}_i(T))$$

Also observe that, for every $i \in \mathbb{N}_0$, there is a fixed HD strategy $\sigma_i \in \Sigma$ such that for every $\pi \in \Pi$ and every $v \in V$ we have that $\mathcal{V}_i(v) \leq \mathcal{P}_v^{\sigma_i,\pi}(\mathit{Reach}_i(T))$. Further, put $\mathcal{V}_\infty(v) = \lim_{i \to \infty} \mathcal{V}_i(v)$ (note that the limit exists because the sequence $\mathcal{V}_0(v), \mathcal{V}_1(v), \ldots$ is non-decreasing and bounded). We show that $\mathcal{V}_\infty$ is a fixed-point of the functional $\mathcal{V}$ defined in the proof of Theorem 3.1. Hence, $\mu\mathcal{V}(v) \leq \mathcal{V}_\infty(v)$ for every $v \in V$, which implies that for every $\varepsilon > 0$ there is $n \in \mathbb{N}$ such that for every $\pi \in \Pi$ we have that

$$\mathcal{P}_v^{\sigma_n,\pi}(\mathit{Reach}_n(T)) \geq V_n(v) > \mu\mathcal{V}(v) - \varepsilon = \mathit{val}(v) - \varepsilon$$

So, it remains to prove that $\mathcal{V}(\mathcal{V}_\infty) = \mathcal{V}_\infty$. We distinguish three cases.

(a) $v \in V_\Box$. Then

$$\mathcal{V}(\mathcal{V}_\infty)(v) = \max_{v \mapsto u} \lim_{i \to \infty} \mathcal{V}_i(u) = \lim_{i \to \infty} \max_{v \mapsto u} \mathcal{V}_i(u) = \lim_{i \to \infty} \mathcal{V}_{i+1}(v) = \mathcal{V}_\infty(v)$$

In the second equality, the '$\leq$' direction is easy, and the '$\geq$' direction can be justified as follows: For every $u \in V$, the sequence $\mathcal{V}_1(u), \mathcal{V}_2(u), \ldots$ is non-decreasing. Hence, for all $i \in \mathbb{N}$ and $u \in V$ we have that $\lim_{j \to \infty} \mathcal{V}_j(u) \geq \mathcal{V}_i(u)$ and thus $\max_{v \mapsto u} \lim_{j \to \infty} \mathcal{V}_j(u) \geq \max_{v \mapsto u} \mathcal{V}_i(u)$ which implies the '$\geq$' direction.

(b) $v \in V_\Diamond$. Then

$$\mathcal{V}(\mathcal{V}_\infty)(v) = \min_{v \mapsto u} \lim_{i \to \infty} \mathcal{V}_i(u) = \lim_{i \to \infty} \min_{v \mapsto u} \mathcal{V}_i(u) = \lim_{i \to \infty} \mathcal{V}_{i+1}(v) = \mathcal{V}_\infty(v)$$

In the second equality, the '$\geq$' direction is easy, and the '$\leq$' direction can be justified as follows: For every $\delta > 0$ there is $i \in \mathbb{N}$ such that for every $v \mapsto u$ we have that $\lim_{j \to \infty} V_j(u) - \delta \leq V_i(u)$ (remember that G is finitely-branching). It follows that $\min_{v \mapsto u} \lim_{j \to \infty} V_j(u) - \delta \leq \min_{v \mapsto u} V_i(u)$ and thus $\min_{v \mapsto u} \lim_{j \to \infty} V_j(u) - \delta \leq \lim_{i \to \infty} \min_{v \mapsto u} V_i(u)$ which implies the '$\leq$' direction because δ was chosen arbitrarily.

(c) $v \in V_\bigcirc$. Then

$$\mathcal{V}(\mathcal{V}_\infty)(v) = \sum_{v \overset{x}{\mapsto} u} x \cdot \lim_{i\to\infty} \mathcal{V}_i(u) = \lim_{i\to\infty} \sum_{v \overset{x}{\mapsto} u} x \cdot \mathcal{V}_i(u) = \lim_{i\to\infty} \mathcal{V}_{i+1}(v) = \mathcal{V}_\infty(v)$$

by linearity of the limit. □

Now we can state and prove the promised determinacy theorem.

Theorem 3.3 (Determinacy)**.** *Assume that G is finitely branching. Let $(T, \rhd\varrho)$ be a quantitative reachability objective. Then V is a disjoint union of $[T]_\Box^{\rhd\varrho}$ and $[T]_\Diamond^{\not\rhd\varrho}$.*

Proof. First, note that we may safely assume that for each $t \in T$ there is only one outgoing edge $t \mapsto t$ (this assumption simplifies some of the claims presented below). Let $v \in V$. If $\varrho > \mathit{val}(v)$, then $v \in [T]_\Box^{\rhd\varrho}$ because player □ has an ε-optimal strategy for an arbitrarily small $\varepsilon > 0$ (see Theorem 3.1). Similarly, if $\varrho < \mathit{val}(v)$, then $v \in [T]_\Diamond^{\not\rhd\varrho}$. Now assume that $\varrho = \mathit{val}(v)$. Obviously, it suffices to show that if player ◇ *does not* have a $(T, \not\rhd\varrho)$-winning strategy in v, then player □ has a $(T, \rhd\varrho)$-winning strategy in v. This means to show that

$$\forall \pi \in \Pi \;\; \exists \sigma \in \Sigma \;:\; \mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T)) \rhd \varrho \tag{2}$$

implies

$$\exists \sigma \in \Sigma \; \forall \pi \in \Pi \;:\; \mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T)) \rhd \varrho$$

If $\rhd$ is $>$ or $\mathit{val}(v) = 0$, then the above implication follows easily: Observe that

- if $\rhd$ is $>$, then (2) does not hold, because player ◇ has an optimal minimizing strategy by Theorem 3.1;
- for the constraint ≥ 0 the statement is trivial.

Hence, it suffices to consider the case when $\rhd$ is $\geq$ and $\varrho = \mathit{val}(v) > 0$. Assume that (2) holds. We say that a vertex $u \in V$ is *good* if

$$\forall \pi \in \Pi \;\; \exists \sigma \in \Sigma \;:\; \mathcal{P}_u^{\sigma,\pi}(\mathit{Reach}(T)) \geq \mathit{val}(u) \tag{3}$$

Note that the vertex v fixed above is good by (2). Further, we say that an edge $u \mapsto u'$ of G is *optimal* if either $u \in V_\bigcirc$, or $u \in V_\Box \cup V_\Diamond$ and $\mathit{val}(u) = \mathit{val}(u')$.

Observe that for every $u \in V_\Box \cup V_\Diamond$ there is at least one optimal edge $u \mapsto u'$, because G is finitely branching (recall that the tuple of all values is the least fixed-point of the functional $\mathcal{V}$ defined in the proof of Theorem 3.1). Further, note that if $u \in V_\Box$ is a good vertex, then there is at least one optimal edge $u \mapsto u'$ where u' is good (otherwise we immediately obtain a contradiction with (3); also observe that if $u \in T$, then $u \mapsto u$ by the technical assumption above). Similarly, if $u \in V_\Diamond$ is good then for every optimal edge $u \mapsto u'$ we have that u' is good, and if $u \in V_\bigcirc$ is good and $u \mapsto u'$ then u' is good. Hence, we can define a game $\bar{G}$, where the set of vertices $\bar{V}$ consists of all good vertices of G, and for all $u, u' \in \bar{V}$ we have that (u, u') is an edge of $\bar{G}$ iff $u \mapsto u'$ is an optimal edge of G. The edge probabilities in $\bar{G}$ are the same as in G. The rest of the proof proceeds by proving the following three claims:

(a) For every $u \in \bar{V}$ we have that $\mathit{val}(u, \bar{G}) = \mathit{val}(u, G)$.
(b) There is $\bar{\sigma} \in \Sigma_{\bar{G}}$ such that for every $\bar{\pi} \in \Pi_{\bar{G}}$ we have that $\mathcal{P}_v^{\bar{\sigma},\bar{\pi}}(\mathit{Reach}(T, \bar{G})) \geq \mathit{val}(v, \bar{G}) = \varrho$.
(c) The strategy $\bar{\sigma}$ can be modified into a strategy $\sigma \in \Sigma_G$ such that for every $\pi \in \Pi_G$ we have that $\mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T, G)) \geq \varrho$.

We start by proving Claim (a). Let $u \in \bar{V}$. Due to Theorem 3.1, there is a MD strategy $\pi \in \Pi_G$ which is optimal minimizing in every vertex of G (particularly in u) and selects only the optimal edges. Hence, the strategy π can also be used in the restricted game $\bar{G}$ and thus we obtain $\mathit{val}(u, \bar{G}) \leq \mathit{val}(u, G)$. Now suppose that $\mathit{val}(u, \bar{G}) < \mathit{val}(u, G)$. By applying Theorem 3.1 to $\bar{G}$, there is an optimal minimizing MD strategy $\bar{\pi} \in \Pi_{\bar{G}}$. Further, for every vertex t of G which is not good there is a strategy $\pi_t \in \Pi_G$ such that for every $\sigma \in \Sigma_G$ we have that $\mathcal{P}_t^{\sigma,\pi_t}(\mathit{Reach}(T, G)) < \mathit{val}(u, G)$ (this follows immediately from (3)). Now consider a strategy $\pi' \in \Pi_G$ which for every play of G initiated in u behaves in the following way:

- As long as player $\Box$ uses only the edges of G that are preserved in $\bar{G}$, the strategy π' behaves exactly like the strategy $\bar{\pi}$.
- When player $\Box$ uses an edge $r \mapsto r'$ which is not an edge in $\bar{G}$ for the first time, then the strategy π' starts to behave either like the optimal minimizing strategy π or the strategy $\pi_{r'}$, depending on whether r' is good or not (observe that if r' is good, then $\mathit{val}(r', G) < \mathit{val}(r, G)$).

Now it is easy to check that for every $\sigma \in \Sigma_G$ we have that $\mathcal{P}_u^{\sigma,\pi'}(\mathit{Reach}(T, G)) < \mathit{val}(u, G)$, which contradicts the assumption that u is good.

Now we prove Claim (b). Due to Lemma 3.2, for every $u \in \bar{V}$ we can fix a strategy $\bar{\sigma}_u \in \Sigma_{\bar{G}}$ and $n_u \in \mathbb{N}$ such that for every $\bar{\pi} \in \Pi_{\bar{G}}$ we have that $\mathcal{P}_u^{\bar{\sigma}_u,\bar{\pi}}(\mathit{Reach}_{n_u}(T,\bar{G})) > \mathit{val}(u,\bar{G})/2$. For every $k \in \mathbb{N}_0$, let $B(k)$ be the set of all vertices u reachable from v in $\bar{G}$ via a path of length exactly k which does not visit T. Observe that $B(k)$ is finite because $\bar{G}$ is finitely-branching. Further, for every $i \in \mathbb{N}_0$ we define a bound $m_i \in \mathbb{N}$ inductively as follows: $m_0 = 1$, and $m_{i+1} = m_i + \max\{n_u \mid u \in B(m_i)\}$. Now we define a strategy $\bar{\sigma} \in \Sigma_{\bar{G}}$ which turns out to be $(T,{\geq}\varrho)$-winning in the vertex v of $\bar{G}$. For every $w \in \bar{V}^*\bar{V}_\Box$ such that $m_i \leq |w| < m_{i+1}$ we put $\bar{\sigma}(w) = \bar{\sigma}_u(uw_2)$, where $w = w_1uw_2$, $|w_1| = m_i - 1$ and $u \in \bar{V}$. Now it is easy to check that for every $i \in \mathbb{N}$ and every strategy $\bar{\pi} \in \Pi_{\bar{G}}$ we have that $\mathcal{P}_v^{\bar{\sigma},\bar{\pi}}(\mathit{Reach}_{m_i}(T,\bar{G})) > (1 - \frac{1}{2^i})\varrho$. This means that the strategy $\bar{\sigma}$ is $(T,{\geq}\varrho)$-winning in v.

It remains to prove Claim (c). Consider a strategy $\sigma \in \Sigma_G$ which for every play of G initiated in v behaves as follows:

- As long as player $\Diamond$ uses only the optimal edges, the strategy σ behaves exactly like the strategy $\bar{\sigma}$.
- When player $\Diamond$ uses a non-optimal edge $r \mapsto r'$ for the first time, the strategy σ starts to behave like an ε-optimal maximizing strategy in r', where $\varepsilon = (\mathit{val}(r',G) - \mathit{val}(r,G))/2$. Note that since $r \mapsto r'$ is not optimal, we have that $\mathit{val}(r',G) > \mathit{val}(r,G)$.

It is easy to check that σ is $(T,{\geq}\varrho)$-winning in v. □

4. Stochastic BPA Games

Stochastic BPA games correspond to stochastic games induced by stateless pushdown automata or 1-exit recursive state machines (see Section 1). A formal definition follows.

Definition 4.1. *A* stochastic BPA *game is a tuple* $\Delta = (\Gamma, \hookrightarrow, (\Gamma_\Box, \Gamma_\Diamond, \Gamma_\bigcirc), \mathit{Prob})$ *where* Γ *is a finite* stack alphabet, $\hookrightarrow\ \subseteq \Gamma \times \Gamma^{\leq 2}$ *is a finite set of* rules *(where* $\Gamma^{\leq 2} = \{w \in \Gamma^* : |w| \leq 2\}$*) such that for each* $X \in \Gamma$ *there is some rule* $X \hookrightarrow \alpha$*,* $(\Gamma_\Box, \Gamma_\Diamond, \Gamma_\bigcirc)$ *is a partition of* Γ*, and* Prob *is a* probability assignment *which to each* $X \in \Gamma_\bigcirc$ *assigns a rational positive probability distribution on the set of all rules of the form* $X \hookrightarrow \alpha$*.*

A *configuration* of Δ is a word $\alpha \in \Gamma^*$, which can intuitively be interpreted as the current stack content where the leftmost symbol of α is on top of the stack. Each stochastic BPA game $\Delta = (\Gamma, \hookrightarrow, (\Gamma_\Box, \Gamma_\Diamond, \Gamma_\bigcirc), Prob)$ determines a unique stochastic game $G_\Delta = (\Gamma^*, \mapsto, (\Gamma_\Box\Gamma^*, \Gamma_\Diamond\Gamma^*, \Gamma_\bigcirc\Gamma^* \cup \{\varepsilon\}), Prob_\Delta)$, where the edges of $\mapsto$ are determined as follows: $\varepsilon \mapsto \varepsilon$, and $X\beta \mapsto \alpha\beta$ iff $X \hookrightarrow \alpha$. The probability assignment $Prob_\Delta$ is the natural extension of $Prob$, i.e., $\varepsilon \overset{1}{\mapsto} \varepsilon$ and for all $X \in \Gamma_\bigcirc$ we have that $X\beta \overset{x}{\mapsto} \alpha\beta$ iff $X \overset{x}{\hookrightarrow} \alpha$. The size of Δ, denoted by $|\Delta|$, is the length of the corresponding binary encoding.

In this section we consider stochastic BPA games with qualitative reachability objectives $(T, \rhd\varrho)$ where $T \subseteq \Gamma^*$ is a *regular* set of configurations. For technical convenience, we define the size of T as the size of the minimal deterministic finite-state automaton $\mathscr{A}_T = (Q, q_0, \delta, F)$ which recognizes the *reverse* of T (if we view configurations as stacks, this corresponds to the bottom-up direction). Note that the automaton $\mathscr{A}_T$ can be simulated on-the-fly in Δ by employing standard techniques (see, e.g., [12]). That is, the stack alphabet is extended to $\Gamma \times Q$ and the rules are adjusted accordingly (for example, if $X \hookrightarrow YZ$, then for every $q \in Q$ the extended BPA game has a rule $(X, q) \hookrightarrow (Y, r)(Z, q)$ where $\delta(q, Z) = r$). Note that the on-the-fly simulation of $\mathscr{A}_T$ in Δ does not affect the way how the game is played, and the size of the extended game is polynomial in $|\Delta|$ and $|\mathscr{A}_T|$. The main advantage of this simulation is that the information whether a current configuration belongs to T or not can now be deduced just by looking at the symbol on top of the stack. This leads to an important technical simplification in the definition of T.

Definition 4.2. *We say that $T \subseteq \Gamma^*$ is* simple *if $\varepsilon \notin T$ and there is $\Gamma_T \subseteq \Gamma$ such that for every $X\alpha \in \Gamma^+$ we have that $X\alpha \in T$ iff $X \in \Gamma_T$.*

Note that the requirement $\varepsilon \notin T$ in the previous definition is not truly restrictive, because each BPA can be equipped with a fresh bottom-of-the-stack symbol which cannot be removed. Hence, we can safely restrict ourselves just to simple sets of target configurations. All of the obtained results (including the complexity bounds) are valid also for regular sets of target configurations.

Since stochastic BPA games have infinitely many vertices, even memoryless strategies are not necessarily finitely representable. It turns out that the winning strategies for both players in stochastic BPA games with qualitative reachability objectives are (effectively) *regular* in the following sense:

Definition 4.3. *Let $\Delta = (\Gamma, \hookrightarrow, (\Gamma_\Box, \Gamma_\Diamond, \Gamma_\bigcirc), Prob)$ be a stochastic BPA game, and let $\odot \in \{\Box, \Diamond\}$. We say that a strategy τ for player $\odot$ is* regular *if there is a deterministic finite-state automaton $\mathscr{A}$ over the alphabet Γ such that, for every $X\alpha \in \Gamma_\odot\Gamma^*$, the value of $\tau(X\alpha)$ depends just on the control state entered by $\mathscr{A}$ after reading the reverse of $X\alpha$ (i.e., the automaton $\mathscr{A}$ reads the stack bottom-up). Note that regular strategies are not necessarily deterministic.*

A special type of regular strategies are stackless MD (SMD) *strategies, where $\tau(X\alpha)$ depends just on the symbol X on top of the stack. Note that SMD strategies are deterministic.*

We use T_ε to denote the set $T \cup \{\varepsilon\}$, and we also slightly abuse the notation by writing ε instead of $\{\varepsilon\}$ (particularly in expressions such as $Reach(\varepsilon)$ or $[\varepsilon]_\Diamond^{<1}$).

In the next sections, we consider the two meaningful qualitative probability constraints >0 and $=1$. We show that the winning regions $[T]_\Box^{>0}$, $[T]_\Diamond^{=0}$, $[T]_\Box^{=1}$, and $[T]_\Diamond^{<1}$ are effectively regular. Further, we show that the membership to $[T]_\Box^{>0}$ and $[T]_\Diamond^{=0}$ is in **P**, and the membership to $[T]_\Box^{=1}$ and $[T]_\Diamond^{<1}$ is in **NP** $\cap$ **co-NP**. Finally, we show that the associated winning strategies are regular and effectively constructible (for both players).

5. Computing the Regions $[T]_\Box^{>0}$ and $[T]_\Diamond^{=0}$

For the rest of this section, we fix a stochastic BPA game $\Delta = (\Gamma, \hookrightarrow, (\Gamma_\Box, \Gamma_\Diamond, \Gamma_\bigcirc), Prob)$ and a simple set T of target configurations. Since we are interested only in reachability objectives, we can safely assume that for every $R \in \Gamma_T$, the only rule where R appears on the left-hand side is $R \hookrightarrow R$ (this assumption simplifies the formulation of some claims).

We start by observing that the sets $[T]_\Box^{>0}$ and $[T]_\Diamond^{=0}$ are regular, and the associated finite-state automata have a fixed number of control states.

Proposition 5.1. *Let $\mathscr{A} = [T]_\Box^{>0} \cap \Gamma$ and $\mathscr{B} = [T_\varepsilon]_\Box^{>0} \cap \Gamma$. Then $[T]_\Box^{>0} = \mathscr{B}^*\mathscr{A}\Gamma^*$ and $[T_\varepsilon]_\Box^{>0} = \mathscr{B}^*\mathscr{A}\Gamma^* \cup \mathscr{B}^*$. Consequently, $[T]_\Diamond^{=0} = \Gamma^* \smallsetminus [T]_\Box^{>0} = (\mathscr{B}\smallsetminus\mathscr{A})^* \cup (\mathscr{B}\smallsetminus\mathscr{A})^*(\Gamma\smallsetminus\mathscr{B})\Gamma^*$ and $[T_\varepsilon]_\Diamond^{=0} = \Gamma^* \smallsetminus [T_\varepsilon]_\Box^{>0} = (\mathscr{B}\smallsetminus\mathscr{A})^*(\Gamma\smallsetminus\mathscr{B})\Gamma^*$.*

Proof. Note that $\mathscr{A} \subseteq \mathscr{B}$. We start by introducing some notation. For every strategy $\sigma \in \Sigma$ and every $\alpha \in \Gamma^*$, let

- $\sigma[-\alpha]$ be a strategy such that for every finite sequence of configurations $\gamma_1, \ldots, \gamma_n, \gamma$, where $n \geq 0$ and $\gamma \in \Gamma_\Box\Gamma^*$, and every edge $\gamma \mapsto \delta$ we have that $\sigma[-\alpha](\gamma_1, \ldots, \gamma_n, \gamma)(\gamma \mapsto \delta) = \sigma(\gamma_1\alpha, \ldots, \gamma_n\alpha, \gamma\alpha)(\gamma\alpha \mapsto \delta\alpha)$

- $\sigma[+\alpha]$ be a strategy such that for every finite sequence of configurations $\gamma_1\alpha, \ldots, \gamma_n\alpha, \gamma\alpha$, where $n \geq 0$ and $\gamma\alpha \in \Gamma_\Box\Gamma^*$, and every edge $\gamma\alpha \mapsto \delta\alpha$ we have that $\sigma[+\alpha](\gamma_1\alpha, \ldots, \gamma_n\alpha, \gamma\alpha)(\gamma\alpha \mapsto \delta\alpha) = \sigma(\gamma_1, \ldots, \gamma_n, \gamma)(\gamma \mapsto \delta)$

By induction on the length of $\beta \in \Gamma^*$, we prove that $\beta \in [T]_\Box^{>0}$ iff $\beta \in \mathscr{B}^*\mathscr{A}\Gamma^*$. For $\beta = \varepsilon$, both sides of the equivalence are false. Now assume that the equivalence holds for all configurations of length k and consider an arbitrary $X\alpha \in \Gamma^+$ where $|\alpha| = k$. If $X\alpha \in [T]_\Box^{>0}$, then there are two possibilities:

- There is a strategy $\sigma \in \Sigma$ such that for all $\pi \in \Pi$, the probability of reaching T without prior reaching α is positive in the play $G_\Delta(\sigma, \pi)$ initiated in $X\alpha$. Then $\sigma[-\alpha]$ is $(T, {>}0)$-winning in X, which means that $X \in [T]_\Box^{>0}$, i.e., $X \in \mathscr{A}$.

- There is a strategy $\sigma \in \Sigma$ such that for all $\pi \in \Pi$, the probability of reaching T is positive in the play $G_\Delta(\sigma, \pi)$ initiated in $X\alpha$, but for some $\hat{\pi} \in \Pi$, the configuration α is always reached before reaching T. In this case, consider again the strategy $\sigma[-\alpha]$. Then $\sigma[-\alpha]$ is $(T_\varepsilon, {>}0)$-winning in X, which means $X \in [T_\varepsilon]_\Box^{>0}$, i.e., $X \in \mathscr{B}$. Moreover, observe that the strategy σ is $(T, {>}0)$-winning in α. Thus, $\alpha \in [T]_\Box^{>0}$ and by induction hypothesis we obtain $\alpha \in \mathscr{B}^*\mathscr{A}\Gamma^*$.

In both cases, we obtained $X\alpha \in \mathscr{B}^*\mathscr{A}\Gamma^*$. If $X\alpha \in \mathscr{B}^*\mathscr{A}\Gamma^*$, we can again distinguish two possibilities:

- $X \in \mathscr{A}$ and there is a $(T, {>}0)$-winning strategy $\sigma \in \Sigma$ for the initial configuration X. Then the strategy $\sigma[+\alpha]$ is $(T, {>}0)$-winning in $X\alpha$. Thus, $X\alpha \in [T]_\Box^{>0}$.

- $X \in \mathscr{B}$ and $\alpha \in \mathscr{B}^*\mathscr{A}\Gamma^*$. Then there exists a $(T_\varepsilon, {>}0)$-winning strategy $\sigma_1 \in \Sigma$ in X. By induction hypothesis, there is a $(T, {>}0)$-winning strategy $\sigma_2 \in \Sigma$ in α. We construct a strategy σ' which behaves like $\sigma_1[+\alpha]$ until α is reached, and from that point on it behaves like σ_2. Obviously, σ' is $(T, {>}0)$-winning, which means that $X\alpha \in [T]_\Box^{>0}$.

The proof of $[T_\varepsilon]_\Box^{>0} = \mathscr{B}^*\mathscr{A}\Gamma^* \cup \mathscr{B}^*$ is similar. Observe that the sets $[T]_\Box^{>0}$ and $[T]_\Diamond^{=0}$ are recognizable by deterministic finite-state automata with three control states and total transition functions. □

Our next proposition says how to compute the sets $\mathscr{A}$ and $\mathscr{B}$.

Proposition 5.2. *The pair* $(\mathscr{A}, \mathscr{B})$ *is the least fixed-point of the function* $F : (2^\Gamma \times 2^\Gamma) \to (2^\Gamma \times 2^\Gamma)$ *defined as follows:* $F(A, B) = (\hat{A}, \hat{B})$*, where*

$$\begin{aligned}
\hat{A} &= \Gamma_T \cup A \cup \{X \in \Gamma_\Box \cup \Gamma_\bigcirc \mid \textit{ there is } X \hookrightarrow \beta \textit{ such that } \beta \in B^* A \Gamma^*\} \\
&\cup \ \{X \in \Gamma_\Diamond \mid \textit{ for all } X \hookrightarrow \beta \textit{ we have that } \beta \in B^* A \Gamma^*\} \\
\hat{B} &= \Gamma_T \cup B \cup \{X \in \Gamma_\Box \cup \Gamma_\bigcirc \mid \textit{ there is } X \hookrightarrow \beta \textit{ such that } \beta \in B^* A \Gamma^* \cup B^*\} \\
&\cup \ \{X \in \Gamma_\Diamond \mid \textit{ for all } X \hookrightarrow \beta \textit{ we have that } \beta \in B^* A \Gamma^* \cup B^*\}
\end{aligned}$$

Proof. For every $i \in \mathbb{N}_0$, let $(A_i, B_i) = F^i(\emptyset, \emptyset)$. The set $2^\Gamma \times 2^\Gamma$ with the component-wise inclusion forms a finite lattice. The longest chain in this lattice has length $2|\Gamma| + 1$. Since F is clearly monotone, by Knaster-Tarski theorem $(\mathscr{A}_F, \mathscr{B}_F) = (\bigcup_{i=0}^{2|\Gamma|} A_i, \bigcup_{i=0}^{2|\Gamma|} B_i)$ is the least fixed-point of F. We show that $(\mathscr{A}_F, \mathscr{B}_F) = (\mathscr{A}, \mathscr{B})$.

We start with the "$\subseteq$" direction. We use the following notation:

- for every $X \in \mathscr{A}_F$, let $I_A(X)$ be the least $i \in \mathbb{N}$ such that $X \in A_i$;
- for every $X \in \mathscr{B}_F$, let $I_B(X)$ be the least $i \in \mathbb{N}$ such that $X \in B_i$;
- for every $\alpha Y \in \mathscr{B}_F^* \mathscr{A}_F$, let $I(\alpha Y) = \max(\{I_A(Y)\} \cup \{I_B(Z) \mid Z \text{ appears in } \alpha\})$;
- for every $\beta \in \Gamma^*$, let $\mathit{price}(\beta) = \min\{I(\gamma) \mid \gamma \text{ is a prefix of } \beta, \gamma \in \mathscr{B}_F^* \mathscr{A}_F\}$, where $\min(\emptyset) = \infty$.

First observe that Γ_T is a subset of both $\mathscr{A}$ and $\mathscr{B}$. For every $X \in (\mathscr{A}_F \cap \Gamma_\Box) \smallsetminus \Gamma_T$, we fix some $X \hookrightarrow \alpha$ (the "A-rule") such that $\mathit{price}(\alpha) < I_A(X)$. It follows directly from the definition of F that there must by such a rule. Similarly, for every $X \in (\mathscr{B}_F \cap \Gamma_\Box) \smallsetminus \Gamma_T$, we fix some $X \hookrightarrow \alpha$ (the "B-rule") such that either $\mathit{price}(\alpha) < I_B(X)$, or $\alpha \in \mathscr{B}_F^*$ and $I_B(Y) < I_B(X)$ for every Y of α.

Now consider a MD strategy $\sigma \in \Sigma$ which for a given $X\alpha \in \mathscr{B}_F^* \mathscr{A}_F \Gamma^* \cap \Gamma_\Box \Gamma^*$ selects

- an arbitrary outgoing rule if $X \in \Gamma_T$;
- the A-rule of X if $X \in \mathscr{A}_F$ and $I_A(X) = \mathit{price}(X\alpha)$;
- the B-rule of X otherwise.

We claim that σ is $(T, >0)$-winning in every configuration of $\mathscr{B}_F^* \mathscr{A}_F \Gamma^*$. In particular, this means that $\mathscr{A}_F \subseteq \mathscr{A}$. To see this, realize that for every $\pi \in \Pi$, the play $G_\Delta(\sigma, \pi)$ contains a path along which every transition either decreases the price, or maintains the price but decreases either the length or replaces the first symbol with a sequence of symbols whose I_B-value is strictly smaller. Hence, this path must inevitably visit T after performing a finite number of transitions.

Similar arguments show that σ is $(T, >0)$-winning in every configuration of $\mathscr{B}_F^* \mathscr{A}_F \Gamma^* \cup \mathscr{B}_F^*$. In particular, this means that $\mathscr{B}_F \subseteq \mathscr{B}$.

Now we prove the "$\supseteq$" direction, i.e., $\mathscr{A}_F \supseteq \mathscr{A}$ and $\mathscr{B}_F \supseteq \mathscr{B}$. Let us define the *$\mathscr{A}$-norm* of a given $X \in \Gamma$, $N_A(X)$, to be the least n such that for some $\sigma \in \Sigma$ and for all $\pi \in \Pi$ there is a path in $G_\Delta(\sigma, \pi)$ of length at most n from X to T. Similarly, define the *$\mathscr{B}$-norm* of a given $X \in \Gamma$, $N_B(X)$, to be the least n such that for some $\sigma \in \Sigma$ and for all $\pi \in \Pi$ there is a path in $G_\Delta(\sigma, \pi)$ of length at most n from X to T_ε (if there are no such paths, then we put $N_A(X) = \infty$ and $N_B(X) = \infty$, respectively).

It follows from König's lemma and the fact that the game is finitely branching that $N_A(X)$ is finite for every $X \in \mathscr{A}$, and $N_B(X)$ is finite for every $X \in \mathscr{B}$. Also note that for all $X \in \Gamma$ we have that $N_A(X) \geq N_B(X)$.

We show, by induction on n, that every $X \in \mathscr{A}$ s.t. $N_A(X) = n$ belongs to A_n, and that every $X \in \mathscr{B}$ s.t. $N_B(X) = n$ belongs to B_n. The base case is easy since $N_A(X) = 1$ iff $N_B(X) = 1$ iff $X \in \Gamma_T$, and $(A_1, B_1) = (\Gamma_T, \Gamma_T)$. The inductive step follows:

- $X \in \mathscr{A}$. If $X \in \Gamma_\Box$ (or $X \in \Gamma_\Diamond$), then some (or every) rule of the form $X \hookrightarrow \beta Y \gamma$ satisfies $\beta \in \mathscr{B}^*$, $Y \in \mathscr{A}$, $N_A(Y) < n$, and $N_B(Z) < n$ for all Z which appear in β. By induction hypothesis, $\beta \in B_{n-1}^*$ and $Y \in A_{n-1}$. Hence, $X \in A_n$.

- $X \in \mathscr{B}$. If $X \in \Gamma_\Box$ (or $X \in \Gamma_\Diamond$), then some (or every) rule of the form $X \hookrightarrow \bar{\beta}$ satisfies one of the following conditions:
  - $\bar{\beta} = \beta Y \gamma$ where $\beta \in \mathscr{B}^*$, $Y \in \mathscr{A}$, $N_A(Y) < n$, and $N_B(Z) < n$ for all Z which appear in β. By induction hypothesis, $\beta \in B_{n-1}^*$ and $Y \in A_{n-1}$. Hence, $X \in A_n \subseteq B_n$.
  - $\bar{\beta} \in \mathscr{B}^*$ where $N_B(Z) < n$ for all Z which appear in $\bar{\beta}$. By induction hypothesis, $\bar{\beta} \in B_{n-1}^*$, and hence $X \in B_n$.

$\Box$

Since the least fixed-point of the function F defined in Proposition 5.2 is computable in polynomial time, the finite-state automata recognizing the sets $[T]_{\Box}^{>0}$ and $[T]_{\Diamond}^{=0}$ are computable in polynomial time. Thus, we obtain the following theorem:

Theorem 5.3. *The membership to $[T]_{\Box}^{>0}$ and $[T]_{\Diamond}^{=0}$ is decidable in polynomial time. Both sets are effectively regular, and the associated finite-state automata are constructible in polynomial time. Further, there is a regular strategy $\sigma \in \Sigma$ and a SMD strategy $\pi \in \Pi$ constructible in polynomial time such that σ and π is $(T, >0)$-winning and $(T, =0)$-winning in every configuration of $[T]_{\Box}^{>0}$ and $[T]_{\Diamond}^{=0}$, respectively.*

Proof. Due to Proposition 5.2, it only remains to show that σ is regular, π is SMD, and both σ and π are effectively constructible in polynomial time. Observe that the MD strategy σ defined in the proof of Proposition 5.2 is $(T, >0)$-winning for player $\Box$. Moreover, σ is regular, because the *price* of a given configuration can be determined by an effectively constructible finite-state automaton which reads configurations from right to left. Since the *price* of a given configuration is bounded by $2|\Gamma|$, the automaton needs only $\mathcal{O}(|\Gamma|)$ control states and can be easily computed in polynomial time.

A SMD $(T, =0)$-winning strategy π for player $\Diamond$ is easy to construct. Consider a strategy π such that for every $X\alpha \in \Gamma_{\Diamond}\Gamma^*$ we have that

- if $X \in (\mathscr{B} \setminus \mathscr{A})$, then $\pi(X\alpha)$ selects an edge $X\alpha \mapsto \beta\alpha$ where $X \hookrightarrow \beta$ and $\beta \in (\mathscr{B} \setminus \mathscr{A})^* \cup (\mathscr{B} \setminus \mathscr{A})^*(\Gamma \setminus \mathscr{B})\Gamma^*$;
- if $X \in (\Gamma \setminus \mathscr{B})$, then $\pi(X\alpha)$ selects an edge $X\alpha \mapsto \beta\alpha$ where $X \hookrightarrow \beta$ and $\beta \in (\mathscr{B} \setminus \mathscr{A})^*(\Gamma \setminus \mathscr{B})\Gamma^*$;
- otherwise, π is defined arbitrarily.

It is easy to check that π is $(T, =0)$-winning in every configuration of $[T]_{\Diamond}^{=0} = (\mathscr{B} \setminus \mathscr{A})^* \cup (\mathscr{B} \setminus \mathscr{A})^*(\Gamma \setminus \mathscr{B})\Gamma^*$. □

Remark 5.4. *Note that Theorem 5.3 holds also for the winning regions $[T_\varepsilon]_{\Box}^{>0}$ and $[T_\varepsilon]_{\Diamond}^{=0}$. The argument is particularly simple in the case of $[T_\varepsilon]_{\Diamond}^{=0}$, where we only need to modify the strategy π constructed in the proof of Theorem 5.3 so that if $X \in (\mathscr{B} \setminus \mathscr{A})$, then $\pi(X\alpha)$ selects an edge $X\alpha \mapsto \beta\alpha$ where $X \hookrightarrow \beta$ and $\beta \in (\mathscr{B} \setminus \mathscr{A})^*(\Gamma \setminus \mathscr{B})\Gamma^*$.*

6. Computing the Regions $[T]^{=1}_{\Box}$ and $[T]^{<1}_{\Diamond}$

The results presented in this subsection constitute the very core of this paper. The problems are more complicated than in the case of $[T]^{>0}_{\Box}$ and $[T]^{=0}_{\Diamond}$, and several deep observations are needed to tackle them. As in Section 5, we fix a stochastic BPA game $\Delta = (\Gamma, \hookrightarrow, (\Gamma_{\Box}, \Gamma_{\Diamond}, \Gamma_{\bigcirc}), \mathit{Prob})$ and a simple set T of target configurations such that, for every $R \in \Gamma_T$, the only rule where R appears on the left-hand side is $R \hookrightarrow R$.

The regularity of the sets $[T]^{=1}_{\Box}$ and $[T]^{<1}_{\Diamond}$ is revealed in the next proposition.

Proposition 6.1. *Let* $\mathscr{A} = [T_\varepsilon]^{<1}_{\Diamond} \cap \Gamma$, $\mathscr{B} = [T_\varepsilon]^{=1}_{\Box} \cap \Gamma$, $\mathscr{C} = [T]^{<1}_{\Diamond} \cap \Gamma$, *and* $\mathscr{D} = [T]^{=1}_{\Box} \cap \Gamma$. *Then* $[T]^{=1}_{\Box} = \mathscr{B}^* \mathscr{D} \Gamma^*$ *and* $[T]^{<1}_{\Diamond} = \mathscr{C}^* \mathscr{A} \Gamma^* \cup \mathscr{C}^*$.

Proof. Due to Theorem 3.3, we have that $[T]^{<1}_{\Diamond} = \Gamma^* \smallsetminus [T]^{=1}_{\Box}$, and hence it suffices to prove the equality $[T]^{=1}_{\Box} = \mathscr{B}^* \mathscr{D} \Gamma^*$. By induction on the length of $\beta \in \Gamma^*$, we show that $\beta \in [T]^{=1}_{\Box}$ iff $\beta \in \mathscr{B}^* \mathscr{D} \Gamma^*$, using the notation $\sigma[-\alpha]$ and $\sigma[+\alpha]$ that was introduced in the proof of Proposition 5.1. For $\beta = \varepsilon$, both sides of the equivalence are false. Now assume that the equivalence holds for all configurations of length k, and consider an arbitrary $X\alpha \in \Gamma^+$ where $|\alpha| = k$. If $X\alpha \in [T]^{=1}_{\Box}$, we distinguish two possibilities:

- There is a strategy $\sigma \in \Sigma$ such that for all $\pi \in \Pi$, the probability of reaching T from $X\alpha$ without prior reaching α is 1 in $G_\Delta(\sigma, \pi)$. Then $\sigma[-\alpha]$ is $(T,=1)$-winning in X, which means that $X \in [T]^{=1}_{\Box}$, i.e., $X \in \mathscr{D}$.

- There is a strategy $\sigma \in \Sigma$ such that for all $\pi \in \Pi$, the probability of reaching T from $X\alpha$ in the play $G_\Delta(\sigma, \pi)$ is 1, but for some $\hat{\pi} \in \Pi$, the configuration α is reached with a positive probability before reaching T. In this case, consider again the strategy $\sigma[-\alpha]$, which is $(T_\varepsilon,=1)$-winning in X and hence $X \in \mathscr{B}$. Moreover, observe that the strategy σ is $(T,=1)$-winning in α. Hence, $\alpha \in [T]^{=1}_{\Box}$ and by applying induction hypothesis we obtain $\alpha \in \mathscr{B}^* \mathscr{D} \Gamma^*$.

For the opposite direction, we assume $X\alpha \in \mathscr{B}^* \mathscr{D} \Gamma^*$, and distinguish the following possibilities:

- $X \in \mathscr{D}$ and there is a $(T,=1)$-winning strategy $\sigma \in \Sigma$ in X. Then $\sigma[+\alpha]$ is $(T,=1)$-winning in $X\alpha$. Thus, $X\alpha \in [T]^{=1}_{\Box}$.

- $X \in \mathscr{B}$ and $\alpha \in \mathscr{B}^*\mathscr{D}\Gamma^*$. Then there is a $(T_\varepsilon,{=}1)$-winning strategy $\sigma_1 \in \Sigma$ in X. By applying induction hypothesis, there is a $(T,{=}1)$-winning strategy $\sigma_2 \in \Sigma$ in α. Now we can set up a $(T,{=}1)$-winning strategy in $X\alpha$, which behaves like $\sigma_1[+\alpha]$ until α is reached, and from that point on it behaves like σ_2. Hence, $X\alpha \in [T]^{=1}_{\Box}$. □

By Theorem 3.3, $\mathscr{B} = \Gamma \setminus \mathscr{A}$ and $\mathscr{D} = \Gamma \setminus \mathscr{C}$. Hence, it suffices to compute the sets $\mathscr{A}$ and $\mathscr{C}$. In the next definition we introduce the crucial notion of a *terminal* set of stack symbols, which plays a key role in our considerations.

Definition 6.2. *A set $M \subseteq \Gamma$ is* terminal *if the following conditions are satisfied:*

- $\Gamma_T \cap M = \emptyset$*;*
- *for every $Z \in M \cap (\Gamma_{\Box} \cup \Gamma_{\bigcirc})$ and every rule of the form $Z \hookrightarrow \alpha$ we have that $\alpha \in M^*$;*
- *for every $Z \in M \cap \Gamma_{\Diamond}$ there is a rule $Z \hookrightarrow \alpha$ such that $\alpha \in M^*$.*

Since the empty set is terminal and the union of two terminal sets is terminal, there is the greatest terminal set that will be denoted by C in the rest of this section. Also note that C determines a stochastic BPA game Δ_C obtained from Δ by restricting the set of stack symbols to C and including all rules $X \hookrightarrow \alpha$ where $X, \alpha \in C^*$. The set of rules of Δ_C is denoted by $\hookrightarrow_C$. The probability of stochastic rules in Δ_C is the same as in Δ.

Definition 6.3. *A stack symbol $Y \in \Gamma$ is a* witness *if one of the following conditions is satisfied:*

(1) $Y \in [T_\varepsilon]^{=0}_{\Diamond}$*;*

(2) $Y \in C$ and $Y \in [\varepsilon]^{<1}_{\Diamond}$, where the set $[\varepsilon]^{<1}_{\Diamond}$ is computed in Δ_C.

The set of all witnesses is denoted by W.

In the next lemma we show that every witness belongs to the set $\mathscr{A}$.

Lemma 6.4. *The problem whether $Y \in W$ for a given $Y \in \Gamma$ is in* ***NP*** $\cap$ ***co-NP****. Further, there is a SMD strategy $\pi \in \Pi$ constructible by a deterministic polynomial-time algorithm with* ***NP*** $\cap$ ***co-NP*** *oracle such that for all $Y \in W$ and $\sigma \in \Sigma$ we have that $\mathcal{P}^{\sigma,\pi}_Y(\mathit{Reach}(T_\varepsilon)) < 1$.*

Proof. Let W_2 be the set of all type (2) witnesses of Δ, and let W_1 be the set of all type (1) witnesses that are not type (2) witnesses (see Definition 6.3).

Let us first consider the BPA game Δ_C (note that Δ_C is constructible in polynomial time). By the results of [15], there are SMD strategies σ' and π' in $G(\Delta_C)$ such that σ' is $(\varepsilon,{=}1)$-winning in every configuration of $[\varepsilon]_\Box^{=1}$ and π' is $(\varepsilon,{<}1)$-winning in every configuration of $[\varepsilon]_\Diamond^{<1}$ (here the sets $[\varepsilon]_\Box^{=1}$ and $[\varepsilon]_\Diamond^{<1}$ are considered in Δ_C). In [15], it is also shown that the problem whether a given SMD strategy is $(\varepsilon,{=}1)$-winning (or $(\varepsilon,{<}1)$-winning) in every configuration of $[\varepsilon]_\Box^{=1}$ (or $[\varepsilon]_\Diamond^{<1}$) is decidable in polynomial time. Hence, the problem whether a given $Y \in \Gamma$ belongs to W_2 is in **NP**$\cap$**co-NP**, and the strategy π' is constructible by an algorithm which successively fixes one of the available rules for every $Y \in \Gamma_\Diamond \cap C$ so that the set $[\varepsilon]_\Diamond^{<1}$ remains unchanged when all of the other rules with Y on the left-hand side are removed from Δ_C. Obviously, this algorithm needs only $\mathcal{O}(|\Delta_C|)$ time to fix such a rule for every $Y \in \Gamma_\Diamond \cap C$ (i.e., to construct the strategy π') if it is equipped with a **NP** $\cap$ **co-NP** oracle which can be used to verify that the currently considered rule is a correct one.

The strategy π' can also be applied in the game $G(\Delta)$ (for every $Z \in \Gamma_\Diamond \smallsetminus C$ we just define $\pi'(Z)$ arbitrarily). Since $\Gamma_T \cap C = \emptyset$, for all $Y \in W_2$ and $\sigma \in \Sigma$ we have that $\mathcal{P}_Y^{\sigma,\pi'}(\mathit{Reach}(T_\varepsilon)) < 1$.

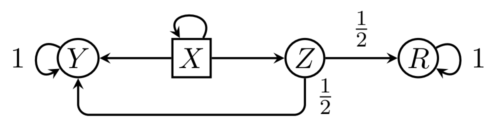
The remaining witnesses of W_1 can be discovered in polynomial time, and there is a SMD strategy $\pi'' \in \Pi$ constructible in polynomial time such that for all $Y \in W_1$ and $\sigma \in \Sigma$ we have that $\mathcal{P}_Y^{\sigma,\pi''}(\mathit{Reach}(T_\varepsilon)) = 0$ or $\mathcal{P}_Y^{\sigma,\pi''}(\mathit{Reach}(W_2\Gamma^*)) > 0$. This follows directly from Theorem 5.3 and Remark 5.4.

The strategy π is constructed simply by "combining" the strategies π' and π''. That is, π behaves like π' (or π'') in all configurations $Y\alpha$ where $Y \in W_2$ (or $Y \in W_1$). $\square$

Due to Lemma 6.4, we have that $W \subseteq \mathscr{A}$. One may be tempted to think that the set $\mathscr{A}$ is just the *attractor* of W, denoted $\mathit{Att}(W)$, which consists of all stack symbols from which player $\Diamond$ can enforce visiting a witness with a positive probability. However, this (natural) hypothesis is false, as demonstrated by the following example:

Example 6.5. *Consider a stochastic BPA game* $\hat{\Delta} = (\{X,Y,Z,R\}, \hookrightarrow, (\{X\},\emptyset,\{Y,Z,R\}), \mathit{Prob})$, *where* $X \hookrightarrow X$, $X \hookrightarrow Y$, $X \hookrightarrow Z$, $Y \overset{1}{\hookrightarrow} Y$, $Z \overset{1/2}{\hookrightarrow} Y$, $Z \overset{1/2}{\hookrightarrow} R$, $R \overset{1}{\hookrightarrow} R$, *and the set* T_Γ *contains just* R.

The game is initiated in X, and the relevant part of $G_{\hat{\Delta}}$ (reachable from X) is shown in the following figure:



Observe that $\mathscr{A} = \{X, Y, Z\}$, $C = W = \{Y\}$, but $Att(\{Y\}) = \{Z, Y\}$.

The problem is that, in general, player $\Box$ cannot be "forced" to enter $Att(W)$ (in Example 6.5, player $\Box$ can always select the rule $X \hookrightarrow X$ and thus avoid entering $Att(\{Y\})$). Nevertheless, observe that player $\Box$ has essentially only two options: she either enters a symbol of $Att(W)$, or avoids visiting the symbols of $Att(W)$ completely. The second possibility is analyzed by "cutting off" the set $Att(W)$ from the considered BPA game, and recomputing the set of all witnesses together with its attractor in the resulting BPA game which is smaller than the original one. In Example 6.5, we "cut off" the attractor $Att(\{Y\})$ and thus obtain a smaller BPA game with just one symbol X and the rule $X \hookrightarrow X$. Since that X is a witness in this game, it can be safely added to the set $\mathscr{A}$. In general, the algorithm for computing the set $\mathscr{A}$ proceeds by putting $\mathscr{A} := \emptyset$ and then repeatedly computing the set $Att(W)$, setting $\mathscr{A} := \mathscr{A} \cup Att(W)$, and "cutting off" the set $Att(W)$ from the game. This goes on until the set $Att(W)$ becomes empty.

We start by demonstrating that if $\mathscr{A} \neq \emptyset$ then there is at least one witness. This is an important (and highly non-trivial) result, whose proof is postponed to Section 7.2.

Proposition 6.6. *If $\mathscr{A} \neq \emptyset$, then $W \neq \emptyset$.*

In other words, the non-emptiness of $\mathscr{A}$ is always certified by at least one witness, and hence each stochastic BPA game with a non-empty $\mathscr{A}$ can be made smaller by "cutting off" $Att(W)$. The procedure which "cuts off" the symbols $Att(W)$ is not completely trivial. A naive idea of removing the symbols of $Att(W)$ together with the rules where they appear (this was used for the stochastic BPA game of Example 6.5) does not always work. This is illustrated in the following example:

Example 6.7. *Consider a stochastic BPA game $\hat{\Delta} = (\{X, Y, Z, R\}, \hookrightarrow, (\{X\}, \emptyset, \{Y, Z, R\}), Prob)$, where*

$$X \hookrightarrow X,\ X \hookrightarrow Y,\ X \hookrightarrow ZY,\ Y \overset{1}{\hookrightarrow} Y,\ Z \overset{1/2}{\hookrightarrow} X,\ Z \overset{1/2}{\hookrightarrow} R,\ R \overset{1}{\hookrightarrow} R$$

and $\hat{\Gamma}_T = \{R\}$. *The game is initiated in* X *(see Fig. 2). We have that* $\mathscr{A} = \{Y\}$ *(observe that* $X, Z, R \in [T_\varepsilon]^{=1}_{\Box}$*, because the strategy* σ *of player* $\Box$ *which always selects the rule* $X \hookrightarrow ZY$ *is* $(T,=1)$*-winning). Further, we have that* $C = W = Att(W) = \{Y\}$*. If we remove* Y *together with all rules where* Y *appears, we obtain the game* $\Delta' = (\{X, Z, R\}, \hookrightarrow, (\{X\}, \emptyset, \{Z, R\}), Prob)$*, where* $X \hookrightarrow X$*,* $Z \stackrel{1/2}{\hookrightarrow} X$*,* $Z \stackrel{1/2}{\hookrightarrow} R$*,* $R \stackrel{1}{\hookrightarrow} R$*. In the game* Δ'*,* X *becomes a witness and hence the algorithm would incorrectly put* X *into* $\mathscr{A}$*.*

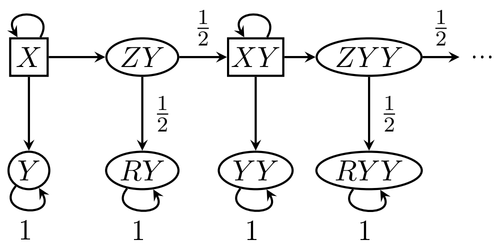


Figure 2: The game of Example 6.7

Hence, the "cutting" procedure must be designed more carefully. Intuitively, we do not remove rules of the form $X \hookrightarrow ZY$, where $Y \in Att(W)$, but change them into $X \hookrightarrow \widetilde{Z}$, where the plays initiated in $\widetilde{Z}$ behave like the ones initiated in Z with the exception that ε cannot be reached whatever the players do.

Now we show how to compute the set $\mathscr{A}$, formalizing the intuition given above. To simplify the proofs of our claims, we adopt some additional (safe) assumptions about the considered BPA game Δ.

Definition 6.8. *We say that* Δ *is in* special normal form (SNF) *if all of the following conditions are satisfied:*

- *For every* $R \in \Gamma_T$ *we have that* $R \in \Gamma_{\bigcirc}$ *and* $R \stackrel{1}{\hookrightarrow} R$*.*
- *For every rule* $X \hookrightarrow \alpha$ *where* $X \in \Gamma_{\Diamond} \cup \Gamma_{\bigcirc}$ *we have that* $\alpha \in \Gamma$*.*
- *The set* $\Gamma_{\Box}$ *can be partitioned into three disjoint subsets* $\Gamma[1]$*,* $\Gamma[2]$*, and* $\Gamma[3]$ *so that*
  - *if* $X \in \Gamma[1]$ *and* $X \hookrightarrow \alpha$*, then* $\alpha \in \Gamma$*;*
  - *if* $X \in \Gamma[2]$*, then* $X \hookrightarrow \varepsilon$ *and there is no other rule of the form* $X \hookrightarrow \alpha$*;*

– *if* $X \in \Gamma[3]$*, then* $X \hookrightarrow YZ$ *for some* $Y, Z \in \Gamma$*, and there is no other rule of the form* $X \hookrightarrow \alpha$*.*

Note that every BPA game can be efficiently transformed into an "equivalent" BPA game in SNF by introducing fresh stack symbols (which belong to player $\Box$) and adding the corresponding dummy rules. For example, if the original BPA game contains the rules $X \hookrightarrow \varepsilon$ and $X \hookrightarrow YZ$, then the newly constructed BPA game in SNF contains the rules $X \hookrightarrow E$, $X \hookrightarrow P$, $E \hookrightarrow \varepsilon$, $P \hookrightarrow YZ$, where E, P are fresh stack symbols that belong to player $\Box$. Obviously, the set $\mathscr{A}$ of the original BPA game is the set $\mathscr{A}$ of the newly constructed BPA game restricted to the stack symbols of the original BPA game.

So, from now on we assume that the considered BPA game Δ is in SNF. In particular, note that only player $\Box$ can change the height of the stack; and if she can do it, then she cannot do anything else for the given stack symbol.

Our algorithm for computing the set $\mathscr{A}$ consists of two parts, the procedure `Init` and the procedure `Main`. The procedure `Init` transforms the BPA game Δ into another BPA game $\bar{\Delta}$, which is then used as an input for the procedure `Main` which computes the set $\mathscr{A}$ of $\bar{\Delta}$.

For every $X \in \Gamma$, let $\widetilde{X}$ be a fresh "twin" of X, and let $\widetilde{\Gamma} = \{\widetilde{X} \mid X \in \Gamma\}$. Similarly, for every $\odot \in \{\bigcirc, \Diamond, \Box\}$ we put $\widetilde{\Gamma}_\odot = \{\widetilde{X} \mid X \in \Gamma_\odot\}$. The procedure `Init` inputs the BPA game Δ and outputs another BPA game $\bar{\Delta} = (\bar{\Gamma}, \hookrightarrow, (\bar{\Gamma}_\Box, \bar{\Gamma}_\Diamond, \bar{\Gamma}_\bigcirc), \mathit{Prob})$ where $\bar{\Gamma} = \Gamma \cup \widetilde{\Gamma}$, $\bar{\Gamma}_\odot = \Gamma_\odot \cup \widetilde{\Gamma}_\odot$ for every $\odot \in \{\bigcirc, \Diamond, \Box\}$, and the rules are constructed as follows:

- if $X \hookrightarrow \varepsilon$ is a rule of Δ, then $X \hookrightarrow \varepsilon$ and $\widetilde{X} \hookrightarrow \widetilde{X}$ are rules of $\bar{\Delta}$;
- if $X \hookrightarrow Y$ is a rule of Δ, then $X \hookrightarrow Y$ and $\widetilde{X} \hookrightarrow \widetilde{Y}$ are rules of $\bar{\Delta}$;
- if $X \hookrightarrow YZ$ is a rule of Δ, then $X \hookrightarrow YZ$ and $\widetilde{X} \hookrightarrow Y\widetilde{Z}$ are rules of $\bar{\Delta}$;
- $\bar{\Delta}$ has no other rules.

Further, if $X \stackrel{x}{\hookrightarrow} Y$ in Δ, then $X \stackrel{x}{\hookrightarrow} Y$ and $\widetilde{X} \stackrel{x}{\hookrightarrow} \widetilde{Y}$ in $\bar{\Delta}$. We put $\bar{\Gamma}_T = \{R, \widetilde{R} \mid R \in \Gamma_T\}$.

Intuitively, the only difference between X and $\widetilde{X}$ is that $\widetilde{X}$ can never be fully removed from the stack. Also observe that the newly added stack symbols of $\widetilde{\Gamma}$ are unreachable from the original stack symbols of Γ. Hence, the set $\mathscr{A}$ of Δ is obtained simply by restricting the set $\mathscr{A}$ of $\bar{\Delta}$ to the symbols of Γ. In the rest of this section, we adopt the following convention:

the elements of Γ are denoted by $X, Y, Z, \ldots$, the corresponding elements of $\widetilde{\Gamma}$ are denoted by $\widetilde{X}, \widetilde{Y}, \widetilde{Z}, \ldots$, and for every $X \in \Gamma$, the symbol $\bar{X}$ denotes either X or $\widetilde{X}$ (sometimes we use even $\hat{X}$, which also denotes either X or $\widetilde{X}$).

The set $\mathscr{A}$ of $\bar{\Delta}$ is computed by the procedure `Main` (see page 28). At line 3, we assign to $\mathcal{M}$ the least fixed-point of the function $\mathtt{Att}_{\Theta,W} : 2^{\bar{\Gamma}} \to 2^{\bar{\Gamma}}$, where Θ is an auxiliary BPA game maintained by the procedure `Main` and W is the set of all witnesses of Θ. Intuitively, the least fixed-point of $\mathtt{Att}_{\Theta,W}$ consists of all stack symbols where player $\Diamond$ "clearly" wins. Formally, $\mathtt{Att}_{\Theta,W}$ is defined as follows (the set of rules of Θ is denoted by $\leadsto$):

$$
\begin{aligned}
\mathtt{Att}_{\Theta,W}(S) \;=\;& W \\
\cup\;& \{\bar{A} \in \bar{\Gamma}_{\bigcirc} \cup \bar{\Gamma}_{\Diamond} \mid \text{there is a rule } \bar{A} \leadsto \bar{B} \text{ where } \bar{B} \in S\} \\
\cup\;& \{\bar{A} \in \bar{\Gamma}[1] \mid \bar{B} \in S \text{ for all } \bar{A} \leadsto \bar{B}\} \\
\cup\;& \{\bar{A} \in \bar{\Gamma}[3] \mid \bar{A} \leadsto Y\bar{C} \text{ where } Y \in S \text{ or } \widetilde{Y}, \bar{C} \in S\}
\end{aligned}
$$

Note that the procedure `Main` actually computes the sets $\mathscr{A}$ and $\mathscr{C}$ of $\bar{\Delta}$ simultaneously, as stated in the following proposition. A proof is postponed to Section 7.3.

Procedure `Main`

Data: A BPA game $\bar{\Delta} = (\bar{\Gamma}, \hookrightarrow, (\bar{\Gamma}_{\Box}, \bar{\Gamma}_{\Diamond}, \bar{\Gamma}_{\bigcirc}), \mathit{Prob})$.
Result: The sets $\mathcal{W}$ and $\mathcal{U}$.

1 $\mathcal{W} := \emptyset$; $\mathcal{U} := \emptyset$; $\Theta := \bar{\Delta}$;
2 **while** *the greatest set* W *of witnesses in* Θ *is not empty* **do**
3     $\mathcal{M} :=$ the least fixed-point of $\mathtt{Att}_{\Theta,W}$;
4   **for** *every* $\bar{A} \in \mathcal{M}$ **do**
5   remove the symbol $\bar{A}$ and all rules with $\bar{A}$ on the left-hand side;
6   **for** *every rule* $\bar{A} \leadsto \bar{B}$ *where* $\bar{A} \in \bar{\Gamma}_{\Box} \setminus \mathcal{M}$ *and* $\bar{B} \in \mathcal{M}$ **do**
7   remove the rule $\bar{A} \leadsto \bar{B}$;
8   **for** *every rule* $\bar{A} \leadsto Y\bar{C}$ *where* $\bar{A} \in \bar{\Gamma}_{\Box} \setminus \mathcal{M}$ *and* $\bar{C} \in \mathcal{M}$ **do**
9   replace the rule $\bar{A} \leadsto Y\bar{C}$ with the rule $\bar{A} \leadsto \widetilde{Y}$;
10     $\mathcal{W} := \mathcal{W} \cup \mathcal{M}$;
11     $\mathcal{U} := \mathcal{U} \cup \{\bar{Y} \mid \widetilde{Y} \in \mathcal{W}\}$;
12 **return** $\mathcal{W}, \mathcal{U}$;

Proposition 6.9. *The sets $\mathcal{W}$ and $\mathcal{U}$ computed by the procedure* `Main` *are exactly the sets $\mathscr{A}$ and $\mathscr{C}$ of the BPA game $\bar{\Delta}$, respectively.*

Now, let us analyze the complexity of the procedure `Main`. Obviously, the main loop initiated at line 2 terminates after $\mathcal{O}(|\bar{\Delta}|)$ iterations. In each iteration, we need to compute the greatest set of witnesses W of the current game, which is the only step that needs exponential time. Hence, the running time of the procedure `Main` is *exponential* in the size of $\bar{\Delta}$. Nevertheless, the procedure `Main` can be easily modified into its *non-deterministic* variant `Main-NonDet` where every computation terminates after a polynomial number of steps, and all "successful" computations of `Main-NonDet` output the same sets $\mathcal{W}, \mathcal{U}$ as the procedure `Main`. This means that the membership problem as well as the non-membership problem for the set $\mathscr{A}$ is in **NP**, which implies that both problems are in fact in **NP** $\cap$ **co-NP**. The same applies to the set $\mathscr{C}$. The only difference between the procedures `Main` and `Main-NonDet` is the way of computing the greatest set of witnesses W. Due to Lemma 6.4, the problem whether $Y \in W$ for a given $Y \in \Gamma$ is in **NP** $\cap$ **co-NP**. Hence, the membership as well as the non-membership to W is certified by certificates of polynomial size that are verifiable in polynomial time (in the proof of Lemma 6.4, we indicated how to construct these certificates, but this is not important now). The procedure `Main-NonDet` guesses the set W together with a tuple of certificates that are supposed to prove that the guess was fully correct (i.e., the guessed set is exactly the set of all witnesses). Then, all of these certificates are verified. If some of them turns out to be invalid, the procedure `Main-NonDet` terminates immediately (this type of termination is considered "unsuccessful"). Otherwise, the procedure `Main-NonDet` proceeds by performing the same instructions as the procedure `Main`.

Since the membership problem for the sets $\mathscr{A}, \mathscr{C}$ is in **NP** $\cap$ **co-NP**, the membership problem for the sets $\mathscr{B}, \mathscr{D}$ is also in **NP** $\cap$ **co-NP** (see the discussion at page 23). Hence, an immediate consequence of the previous observations and Proposition 6.1 is the following:

Theorem 6.10. *The membership to $[T]^{=1}_{\Box}$ and $[T]^{<1}_{\Diamond}$ is in* ***NP*** $\cap$ ***co-NP***. *Both sets are effectively regular, and the associated finite-state automata are constructible by a deterministic polynomial-time algorithm with* ***NP*** $\cap$ ***co-NP*** *oracle.*

Since the arguments used in the proof of Proposition 6.9 are mostly constructive, the winning strategies for both players are effectively regular. This is stated in our final theorem (a proof can be found in Section 7.3).

Theorem 6.11. *There are regular strategies* $\sigma \in \Sigma$ *and* $\pi \in \Pi$ *such that* σ *is* $(T,{=}1)$*-winning in every configuration of* $[T]_{\Box}^{=1}$ *and* π *is* $(T,{<}1)$*-winning in every configuration of* $[T]_{\Diamond}^{<1}$*. Moreover, the strategies* σ *and* π *are constructible by a deterministic polynomial-time algorithm with* $\boldsymbol{NP} \cap \boldsymbol{co\text{-}NP}$ *oracle.*

7. Proofs

In this section we present the proofs that were omitted in Sections 3 and 6.

7.1. A Proof of Theorem 3.1

Let us fix a game $G = (V, \mapsto, (V_\Box, V_\Diamond, V_\bigcirc), \mathit{Prob})$ and a set of target vertices T. Let $(V \to [0,1], \sqsubseteq)$ be the complete lattice of all functions $f : V \to [0,1]$ with pointwise ordering. We show that the tuple of all values is the least fixed-point of the following (Bellman) functional $\mathcal{V} : (V \to [0,1]) \to (V \to [0,1])$ defined by

$$\mathcal{V}(f)(v) \quad = \quad \begin{cases} 1 & \text{if } v \in T \\ \sup\{f(u) \mid v \mapsto u\} & \text{if } v \in V_\Box \setminus T \\ \inf\{f(u) \mid v \mapsto u\} & \text{if } v \in V_\Diamond \setminus T \\ \sum_{v \overset{x}{\mapsto} u} x \cdot f(u) & \text{if } v \in V_\bigcirc \setminus T \end{cases}$$

Since $\mathcal{V}$ is monotone, by Knaster-Tarski theorem [23] there is the least fixed-point $\mu\mathcal{V}$ of $\mathcal{V}$. Let $\mathcal{A} : V \to [0,1]$ be a function defined by $\mathcal{A}(v) = \sup_{\sigma\in\Sigma} \inf_{\pi\in\Pi} \mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T))$. We prove the following:

(i) $\mathcal{A}$ is a fixed-point of $\mathcal{V}$.

(ii) For every $\varepsilon > 0$ there is $\pi \in \Pi$ such that for every $v \in V$ we have that

$$\sup_{\sigma\in\Sigma} \mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T)) \le \mu\mathcal{V}(v) + \varepsilon \qquad (4)$$

Observe that (i) implies $\mu\mathcal{V}(v) \le \sup_{\sigma\in\Sigma} \inf_{\pi\in\Pi} \mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T))$. Obviously,

$$\sup_{\sigma\in\Sigma} \inf_{\pi\in\Pi} \mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T)) \quad \le \quad \inf_{\pi\in\Pi} \sup_{\sigma\in\Sigma} \mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T))$$

and due to (ii) we further have that $\inf_{\pi\in\Pi}\sup_{\sigma\in\Sigma}\mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T)) \leq \mu\mathcal{V}(v)$. Hence, (i) and (ii) together imply that $\mu\mathcal{V}(v)$ is the value of v for every $v \in V$. It remains to prove (i) and (ii).

Ad (i). Let $v \in V$. If $v \in T$, then clearly $\mathcal{A}(v) = 1 = \mathcal{V}(\mathcal{A})(v)$. If $v \notin T$, we can further distinguish three cases.

(a) $v \in V_\Box$. Then

$$\begin{array}{rcl} \mathcal{V}(\mathcal{A})(v) & = & \sup\{\mathcal{A}(u) \mid v \mapsto u\} \\ & = & \sup\{\sup_{\sigma\in\Sigma}\inf_{\pi\in\Pi}\mathcal{P}_u^{\sigma,\pi}(\mathit{Reach}(T)) \mid v \mapsto u\} \\ & = & \sup_{\sigma\in\Sigma}\inf_{\pi\in\Pi}\mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T)) \\ & = & \mathcal{A}(v) \end{array}$$

(b) $v \in V_\Diamond$. Let us denote by $\mathcal{D}(v)$ the set of all positive probability distributions on the set of outgoing edges of v. Then

$$\begin{array}{rcl} \mathcal{V}(\mathcal{A})(v) & = & \inf\{\mathcal{A}(u) \mid v \mapsto u\} \\ & = & \inf\{\sup_{\sigma\in\Sigma}\inf_{\pi\in\Pi}\mathcal{P}_u^{\sigma,\pi}(\mathit{Reach}(T)) \mid v \mapsto u\} \\ & = & \inf_{\eta\in\mathcal{D}(v)}\sum_{v\mapsto u}\eta(v\mapsto u)\cdot\sup_{\sigma\in\Sigma}\inf_{\pi\in\Pi}\mathcal{P}_u^{\sigma,\pi}(\mathit{Reach}(T)) \\ & =^{(*)} & \sup_{\sigma\in\Sigma}\inf_{\eta\in\mathcal{D}(v)}\sum_{v\mapsto u}\eta(v\mapsto u)\cdot\inf_{\pi\in\Pi}\mathcal{P}_u^{\sigma,\pi}(\mathit{Reach}(T)) \\ & = & \sup_{\sigma\in\Sigma}\inf_{\pi\in\Pi}\mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T)) \\ & = & \mathcal{A}(v) \end{array}$$

In the equality $(*)$, the '$\geq$' direction is easy, and the '$\leq$' direction can be justified as follows: For every $\delta > 0$, there is a strategy $\bar{\sigma} \in \Sigma$ such that for every $u \in V$ we have that

$$\sup_{\sigma\in\Sigma}\inf_{\pi\in\Pi}\mathcal{P}_u^{\sigma,\pi}(\mathit{Reach}(T)) \quad \leq \quad \inf_{\pi\in\Pi}\mathcal{P}_u^{\bar{\sigma},\pi}(\mathit{Reach}(T)) + \delta$$

This means that, for every $\eta \in \mathcal{D}(v)$

$$\sum_{v\mapsto u}\eta(v\mapsto u)\cdot\sup_{\sigma\in\Sigma}\inf_{\pi\in\Pi}\mathcal{P}_u^{\sigma,\pi}(\mathit{Reach}(T)) \leq \sum_{v\mapsto u}\eta(v\mapsto u)\cdot\inf_{\pi\in\Pi}\mathcal{P}_u^{\bar{\sigma},\pi}(\mathit{Reach}(T))+\delta$$

and thus

$$\inf_{\eta\in\mathcal{D}(v)}\sum_{v\mapsto u}\eta(v\mapsto u)\cdot\sup_{\sigma\in\Sigma}\inf_{\pi\in\Pi}\mathcal{P}_u^{\sigma,\pi}(\mathit{Reach}(T)) \leq \inf_{\eta\in\mathcal{D}(v)}\sum_{v\mapsto u}\eta(v\mapsto u)\cdot\inf_{\pi\in\Pi}\mathcal{P}_u^{\bar{\sigma},\pi}(\mathit{Reach}(T))+\delta$$

which implies $(*)$ because δ was chosen arbitrarily.

(c) $v \in V_{\bigcirc}$. Then

$$\begin{array}{rcl} \mathcal{V}(\mathcal{A})(w) & = & \sum_{v \stackrel{x}{\mapsto} u} x \cdot \mathcal{A}(u) \\ & = & \sum_{v \stackrel{x}{\mapsto} u} x \cdot \sup_{\sigma \in \Sigma} \inf_{\pi \in \Pi} \mathcal{P}_u^{\sigma,\pi}(\mathit{Reach}(T)) \\ & =^{(**)} & \sup_{\sigma \in \Sigma} \inf_{\pi \in \Pi} \sum_{v \stackrel{x}{\mapsto} u} x \cdot \mathcal{P}_u^{\sigma,\pi}(\mathit{Reach}(T)) \\ & = & \sup_{\sigma \in \Sigma} \inf_{\pi \in \Pi} \mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T)) \\ & = & \mathcal{A}(v) \end{array}$$

Note that the equality $(**)$ can be justified similarly as $(*)$ above.

Ad (ii). Let us fix some $\varepsilon > 0$. For every $j \in \mathbb{N}_0$, we define a strategy π_j as follows: For a given $wv \in V^*V_\diamond$, we choose (some) edge $u \in V$ such that $\mu\mathcal{V}(u) \leq \mu\mathcal{V}(v) + \frac{\varepsilon}{2^{|w|+j+1}}$ and put $\pi_j(wv)(v \mapsto u) = 1$. Note that such an edge must exist, and if G is finitely-branching, then there is even an edge $v \mapsto u$ such that $\mu\mathcal{V}(u) = \mu\mathcal{V}(v)$ (i.e., when G is finitely-branching, we can also consider the case when $\varepsilon = 0$). In the sequel we also write π instead of π_0. We prove that for all $\sigma \in \Sigma$, $v \in V$, and $i \geq 0$ we have that

$$\mathcal{P}_v^{\sigma,\pi_j}(\mathit{Reach}_i(T)) \quad \leq \quad \mu\mathcal{V}(v) + \sum_{k=j+1}^{j+i} \frac{\varepsilon}{2^k}$$

In particular, for $j = 0$ we get

$$\mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}_i(T)) \quad \leq \quad \mu\mathcal{V}(v) + \sum_{k=1}^{i} \frac{\varepsilon}{2^k}$$

and hence

$$\sup_{\sigma \in \Sigma} \mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}(T)) \quad = \quad \sup_{\sigma \in \Sigma} \lim_{i \to \infty} \mathcal{P}_v^{\sigma,\pi}(\mathit{Reach}_i(T)) \quad \leq \quad \mu\mathcal{V}(v) + \varepsilon$$

If $v \in T$, then $\mathcal{P}_v^{\sigma,\pi_j}(\mathit{Reach}_i(T)) = 1 = \mu\mathcal{V}(v)$ for all $j \in \mathbb{N}_0$. If $v \notin T$, we proceed by induction on i. If $i = 0$, then $\mathcal{P}_v^{\sigma,\pi_j}(\mathit{Reach}_0(T)) = 0 \leq \mu\mathcal{V}(v)$ for all $j \in \mathbb{N}_0$. Now assume that $i \geq 1$. For every $\sigma \in \Sigma$, we use σ_v to denote the strategy such that $\sigma_v(wu) = \sigma(vwu)$ for all $wu \in \mathcal{V}^*\mathcal{V}_\Box$. We distinguish three cases.

(a) $v \in V_\Box$. Then

$$\begin{array}{rcl} \mathcal{P}_v^{\sigma,\pi_j}(\mathit{Reach}_i(T)) & = & \sum_{v \mapsto u} \sigma(v)(v \mapsto u) \cdot \mathcal{P}_u^{\sigma_v,\pi_{j+1}}(\mathit{Reach}_{i-1}(T)) \\ & \leq & \sum_{v \mapsto u} \sigma(v)(v \mapsto u) \cdot \left(\mu\mathcal{V}(u) + \sum_{k=j+2}^{j+i} \frac{\varepsilon}{2^k}\right) \\ & = & \left(\sum_{v \mapsto u} \sigma(v)(v \mapsto u) \cdot \mu\mathcal{V}(u)\right) + \sum_{k=j+2}^{j+i} \frac{\varepsilon}{2^k} \\ & \leq & \mu\mathcal{V}(v) + \sum_{k=j+2}^{j+i} \frac{\varepsilon}{2^k} \end{array}$$

(b) $v \in V_\Diamond$. Then

$$\begin{array}{rcl} \mathcal{P}_v^{\sigma,\pi_j}(Reach_i(T)) & = & \sum_{v\mapsto u} \pi_j(v)(v\mapsto u)\cdot \mathcal{P}_u^{\sigma_v,\pi_{j+1}}(Reach_{i-1}(T)) \\ & \leq & \sum_{v\mapsto u} \pi_j(v)(v\mapsto u)\cdot \left(\mu\mathcal{V}(u) + \sum_{k=j+2}^{j+i} \frac{\varepsilon}{2^k}\right) \\ & = & \left(\sum_{v\mapsto u} \pi_j(v)(v\mapsto u)\cdot \mu\mathcal{V}(u)\right) + \sum_{k=j+2}^{j+i} \frac{\varepsilon}{2^k} \\ & \leq & \mu\mathcal{V}(v) + \frac{\varepsilon}{2^{j+1}} + \sum_{k=j+2}^{j+i} \frac{\varepsilon}{2^k} \\ & \leq & \mu\mathcal{V}(v) + \sum_{k=j+1}^{j+i} \frac{\varepsilon}{2^k} \end{array}$$

(c) $v \in V_\bigcirc$. Then

$$\begin{array}{rcl} \mathcal{P}_v^{\sigma,\pi_j}(Reach_i(T)) & = & \sum_{v\stackrel{x}{\mapsto}u} x\cdot \mathcal{P}_u^{\sigma_v,\pi_j}(Reach_{i-1}(T)) \\ & \leq & \sum_{v\stackrel{x}{\mapsto}u} x\cdot \left(\mu\mathcal{V}(u) + \sum_{k=j+2}^{j+i} \frac{\varepsilon}{2^k}\right) \\ & = & \left(\sum_{v\stackrel{x}{\mapsto}u} x\cdot \mu\mathcal{V}(u)\right) + \sum_{k=j+2}^{j+i} \frac{\varepsilon}{2^k} \\ & = & \mu\mathcal{V}(v) + \sum_{k=j+2}^{j+i} \frac{\varepsilon}{2^k} \end{array}$$

If G is finitely branching, then an optimal minimizing strategy π is obtained by considering $\varepsilon = 0$ in the above proof of (ii).

7.2. A Proof of Proposition 6.6

We start by formulating a simple corollary to Proposition 5.2, which turns out to be useful at several places.

Proposition 7.1. *Let $\sigma \in \Sigma$ be a strategy of player $\Box$ which always returns a uniform probability distribution over the available outgoing edges. Then for every $X \in [T]_\Box^{\geq 0} \cap \Gamma$ (or $X \in [T_\varepsilon]_\Box^{\geq 0} \cap \Gamma$) and every $\pi \in \Pi$ there is a path w from X to T (to T_ε, resp.) in $G_\Delta(\sigma,\pi)$ such that*

1. *the length of w is at most $2^{2|\Gamma|}$;*
2. *the length of all configurations visited by w is at most $2|\Gamma|$.*

Proof. Let us consider the sets A_i and B_i from the proof of Proposition 5.2. Recall that $[T]_\Box^{\geq 0} \cap \Gamma = \bigcup_{i=0}^{2|\Gamma|} A_i$ and $[T_\varepsilon]_\Box^{\geq 0} \cap \Gamma = \bigcup_{i=0}^{2|\Gamma|} B_i$. By induction on i, we prove that for every $X \in A_i$ (or $X \in B_i$) and every $\pi \in \Pi$ there is a path w from X to T (or to T_ε, resp.) in $G_\Delta(\sigma,\pi)$ such that

(1) the length of w is at most 2^i;
(2) the length of all configurations visited by w is at most i.

The case $i = 1$ is trivial, as $\mathscr{A}_1 = \mathscr{B}_1 = \Gamma_T$. Now assume that $i > 1$. If $X \in A_i \cap (\Gamma_\square \cup \Gamma_\bigcirc)$, then by the definition of A_i, there is a transition $X \hookrightarrow \gamma$ such that $\gamma \in \Gamma_T \cup A_{i-1}\Gamma \cup B_{i-1}A_{i-1} \cup A_{i-1}$. By induction hypothesis, there is a path w' from γ to T in $G_\Delta(\sigma, \pi)$ of length at most $2^i + 2^i = 2^{i+1}$ such that the length of all configurations entered by w' is at most $\max\{i+1, i\} = i + 1$. The rest follows from the fact that σ always returns a uniform probability distribution, and if $X \in A_i \cap \Gamma_\diamond$, then all outgoing transitions of X have the form $X \hookrightarrow \gamma$ where $\gamma \in \Gamma_T \cup A_{i-1}\Gamma \cup B_{i-1}A_{i-1} \cup A_{i-1}$ (we use induction hypothesis to obtain the desired result). The case when $X \in B_i$ follows similarly. □

Proposition 6.6 is obtained as a corollary to the following (stronger) claim that will also be used later when synthesizing a regular $(T, {=}1)$-winning strategy for player □.

Proposition 7.2. *Let W be the set of all witnesses (see Definition 6.3). If $W = \emptyset$, then there is a regular strategy σ of player □, computable in polynomial time, which is $(T_\varepsilon, {=}1)$-winning in every configuration of Δ.*

In particular, if $W = \emptyset$ then $\mathscr{A} = \emptyset$, and thus we obtain Proposition 6.6. Now we prove Proposition 7.2, relying on further technical observations that are formulated and proved at appropriate places.

As $W = \emptyset$, the two conditions of Definition 6.3 are not satisfied by any $Y \in \Gamma$. This means that for all $Y \in C$ we have that $Y \in [\varepsilon]_\square^{=1}$, where the set $[\varepsilon]_\square^{=1}$ is computed in Δ_C (we again use Theorem 3.3). Due to [14], there exists a SMD strategy σ_T for player □ in G_{Δ_C} such that for every $Y \in C$ and every strategy π of player ◇ in G_{Δ_C} we have that $\mathcal{P}_Y^{\sigma_T,\pi}(\mathit{Reach}(\varepsilon)) = 1$.

Let σ_U be the SMD strategy of player □ which always returns the uniform probability distribution over the available edges. In the proof we use the following simple property of σ_U, which follows easily from Proposition 7.1.

Lemma 7.3. *There is $\xi > 0$ such that for every $X \in \Gamma$ and every $\pi \in \Pi$ there is a path w from X to a configuration of T_ε in $G_\Delta(\sigma_U, \pi)$ satisfying the following: The length of all configurations visited by w is bounded by $2|\Gamma|$, and the probability of $\mathit{Run}(w)$ in $G_\Delta(\sigma_U, \pi)$ is at least ξ.*

Proof. Since $W = \emptyset$, there are no type (1) witnesses (see Definition 6.3), i.e., $\Gamma \cap [T_\varepsilon]_\diamond^{=0} = \emptyset$, which means that $\Gamma \subseteq [T_\varepsilon]_\square^{>0}$ by Theorem 3.3. Let $\pi \in \Pi$ be an arbitrary (possibly randomized) strategy. We define the associated deterministic strategy $\hat{\pi}$, which for every finite sequence of configurations

$\alpha_1, \ldots, \alpha_n$ selects an edge $\alpha_n \mapsto \beta$ such that $\alpha_n \mapsto \beta$ is assigned a maximal probability in the distribution assigned to $\alpha_1, \ldots, \alpha_n$ by the strategy π. In other words, $\alpha_n \mapsto \beta$ is an edge selected with a maximal probability by π. If there are several candidates for $\alpha_n \mapsto \beta$, any of them can be chosen. Obviously, every path in $G_\Delta(\sigma_U, \hat{\pi})$ initiated in X is also a path in $G_\Delta(\sigma_U, \pi)$ initiated in X. Due to Proposition 7.1, there is a path $\hat{w}$ from X to T_ε in $G_\Delta(\sigma_U, \hat{\pi})$ such that the length of $\hat{w}$ is bounded by $2^{2|\Gamma|}$ and the stack height of all configurations visited by $\hat{w}$ is bounded by $2|\Gamma|$. Now consider the corresponding path w in $G_\Delta(\sigma_U, \pi)$. The only difference between w and $\hat{w}$ is that the probability of the transitions selected by player $\Diamond$ is not necessarily one in w. However, due to the definition of $\hat{w}$ we immediately obtain that the probability of each such transition is at least $\frac{1}{|\hookrightarrow|}$ (this bound is not tight but sufficient for our purposes). Since σ_U is uniform, the same bound is valid also for the probability of transitions selected by player $\Box$. Let μ be the least probability weight of a probabilistic rule assigned by *Prob*. We put

$$\xi = \left(\min\{\mu, \frac{1}{|\hookrightarrow|}\}\right)^{2^{2|\Gamma|}}$$

Obviously, $\mathcal{P}(Run(w)) \geq \xi$ and we are done. $\Box$

Now we are ready to define the regular strategy $\sigma \in \Sigma$ whose existence was promised in Proposition 7.2. Recall that regular strategies are memoryless, and hence they can be formally understood as functions which assign to a given configuration β a probability distribution on the outgoing edges of β. For a given $X\alpha \in \Gamma_\Box\Gamma^*$, we put $\sigma(X\alpha) = \sigma_T(X\alpha)$ if $X\alpha$ starts with some $\beta \in C^*$ where $|\beta| > 2|\Gamma|$. Otherwise, we put $\sigma(X\alpha) = \sigma_U(X\alpha)$.

Remark 7.4. *Note that for every configuration $\alpha Z\beta$ such that $Z \notin [\varepsilon]_\Box^{=1}$ we have that $\sigma(\alpha Z\beta) = \sigma(\alpha Z)$, because $Z \notin C$. This trivial observation is used later in Section 7.3.*

Observe that the strategy σ can easily be represented by a finite state automaton with $\mathcal{O}(|\Gamma|)$ states in the sense of Definition 4.3. Moreover, such an automaton is easily constructible in polynomial time because the set C is computable in polynomial time. So, it remains to prove that σ is $(T_\varepsilon, {=}1)$-winning in every configuration of Δ.

Let us fix some strategy $\pi \in \Pi$. Our goal is to show that for every $\alpha \in \Gamma^+$ we have that $\mathcal{P}_\alpha^{\sigma,\pi}(Reach(T_\varepsilon)) = 1$. Assume the converse, i.e., there is some

$\alpha \in \Gamma^+$ such that $\mathcal{P}^{\sigma,\pi}_{\alpha}(\mathit{Reach}(T_\varepsilon)) < 1$. We fix such an α and show the following:

(1) The set of all runs initiated in α that do not visit T_ε contains a subset V of positive probability such that all runs of V eventually use only the rules of Δ_C.

(2) The set V has zero probability. This is achieved by demonstrating that player $\Box$, who plays according to the strategy σ, selects the rules of Δ_C in such a way that almost all runs that use only the rules of Δ_C eventually terminate (i.e., visit the configuration ε).

Since (2) contradicts (1), we are virtually done.

Now we formulate (1) and (2) precisely and develop their proofs. Let w be a run of $G_\Delta(\sigma,\pi)$. We say that given rule of Δ is *used infinitely often in* w if the rule was used to derive infinitely many transitions of w. Further, we say that w *eventually uses only a given subset* $\rightsquigarrow$ *of* $\hookrightarrow$ if there is some $i \in \mathbb{N}$ such that all transitions $w(j) \longrightarrow w(j{+}1)$, where $j \geq i$, were derived using a rule of $\rightsquigarrow$.

Lemma 7.5. *There is a set of runs* $V \subseteq \mathit{Run}(G_\Delta(\sigma,\pi),\alpha)$ *such that* $\mathcal{P}^{\sigma,\pi}_{\alpha}(V) > 0$, *and for every* $w \in V$ *we have that* w *does not visit* T_ε *and all rules that are used infinitely often in* w *belong to* $\hookrightarrow_C$.

Proof. Let A be the set of all $w \in \mathit{Run}(G_\Delta(\sigma,\pi),\alpha)$ such that w does not visit T_ε. By our assumption, $\mathcal{P}^{\sigma,\pi}_{\alpha}(A) > 0$. The runs of A can be split into finitely many disjoint subsets according to the set of rules which are used infinitely often. Since $\mathcal{P}^{\sigma,\pi}_{\alpha}(A) > 0$, at least one of these subsets V must have positive probability. Let $\hookrightarrow_V$ be the associated set of rules that are used infinitely often in the runs of V.

We prove that $\hookrightarrow_V \subseteq \hookrightarrow_C$. Let $L \subseteq \Gamma$ be the set of all symbols that appear on the left-hand side of some rule in $\hookrightarrow_V$. To show that $\hookrightarrow_V \subseteq \hookrightarrow_C$, it suffices to prove that

(a) for every $Y \in (L \smallsetminus C) \cap (\Gamma_\bigcirc \cup \Gamma_\Box)$ we have that if $Y \hookrightarrow \beta$, then also $Y \hookrightarrow_V \beta$;

(b) for all rules $Y \hookrightarrow_V \beta$ we have that $\beta \in (L \cup C)^*$.

Observe that (a) and (b) together imply that $L \cup C$ is a terminal set. Hence, $L \cup C = C$ by the maximality of C, and thus $\hookrightarrow_V \subseteq \hookrightarrow_C$ as needed.

Claim (a) follows from the fact that player $\Box$, who plays according to the strategy σ, selects edges uniformly at random in all configurations of $((L \smallsetminus C) \cap \Gamma_\Box) \cdot \Gamma^*$. Then every rule $Y \hookrightarrow \beta$, where $Y \in (L \smallsetminus C) \cap (\Gamma_\bigcirc \cup \Gamma_\Box)$, has the probability of being selected greater than some fixed non-zero constant, which means that $Y \hookrightarrow_V \beta$ (otherwise, the probability of V would be zero).

Now we prove Claim (b). Assume that $Y \hookrightarrow_V \gamma$. If $\gamma = \varepsilon$, then $\gamma \in (L \cup C)^*$. If $\gamma = P$, then surely $P \in L$ because configurations with P on the top of the stack occur infinitely often in all runs of V. If $\gamma = PQ$, then $P \in L$ by applying the previous argument. If $Q \in C$, we are done. Now assume that $Q \notin C$. Note that then player $\Box$ selects edges uniformly at random in all configurations of the form $\beta Q\delta$ where $|\beta| \leq 2|\Gamma|$. By Lemma 7.3, there is $0 < \xi < 1$ such that for every configuration of the form $PQ\delta$ there is a path w from $PQ\delta$ to $T \cup \{Q\delta\}$ in $G_\Delta(\sigma, \pi)$ satisfying the following:

- all configurations in w are of the form $\hat{\beta} Q\delta$ where $|\hat{\beta}| \leq 2|\Gamma|$;
- the probability of following w in $G_\Delta(\sigma, \pi)$ is at least ξ.

It follows that almost every run of V enters configurations of $\{Q\} \cdot \Gamma^*$ infinitely many times because every run of V contains infinitely many occurrences of configurations of the form $PQ\delta$ and no run of V enters T. Hence, $Q \in L$. $\Box$

Lemma 7.6. *Let $V \subseteq Run(G_\Delta(\sigma, \pi), \alpha)$ such that for every $w \in V$ we have that w does not visit T_ε and all rules that are used infinitely often in w belong to $\hookrightarrow_C$. Then $\mathcal{P}_\alpha^{\sigma,\pi}(V) = 0$.*

Proof. Since all runs of V eventually use only the rules of $\hookrightarrow_C$, each run $w \in V$ uniquely determines its shortest prefix v_w after which no rules of $\hookrightarrow \smallsetminus \hookrightarrow_C$ are used and the length of each configuration visited after the prefix v_w is at least as large as the length of the last configuration visited by v. For a given finite path v initiated in α, let $U_v = \{w \in V \mid v_w = v\}$. Obviously, V is the (disjoint) union of all U_v. Since there are only countably many v's, it suffices to prove that $\mathcal{P}_\alpha^{\sigma,\pi}(U_v) = 0$ for every v. So, let us fix a finite path v initiated in α, and let $Y\beta$ be the last configuration visited by v. Intuitively, we show that after performing the prefix v, the strategies σ and π can be "simulated" by suitable strategies σ' and π' in the game G_{Δ_C} so that the set of runs U_v is "projected" (by ignoring the prefix v and cutting off β from the bottom of the stack) onto the set of runs U in the play $G_{\Delta_C}(\sigma', \pi')$ so that

$$\mathcal{P}_\alpha^{\sigma,\pi}(U_v) \;=\; \mathcal{P}_\alpha^{\sigma,\pi}(Run(v)) \cdot \mathcal{P}_Y^{\sigma',\pi'}(U)$$

Then, we show that $\mathcal{P}_Y^{\sigma',\pi'}(U) = 0$. This is because the strategy σ' is "sufficiently similar" to the strategy σ_T, and hence the probability of visiting ε in $G_{\Delta_C}(\sigma', \pi')$ is 1.

Now we formalize the above intuition. First, let us realize that every probability distribution f on the outgoing edges of a BPA configuration α determines a unique *rule distribution* f_r on the rules of the considered BPA game such that for every $\alpha \mapsto \alpha'$ we have that $f(\alpha \mapsto \alpha') = f_r(Z \hookrightarrow \gamma)$, where $Z \hookrightarrow \gamma$ is the rule used to derive the edge $\alpha \mapsto \alpha'$.

Observe that $Y \in C$ by the definition of U_v. Let σ' be a MR strategy for player $\Box$ in G_{Δ_C} such that for every $\gamma \in C^+$ we have that $\sigma'(\gamma) = \sigma(\gamma\beta)$. Further, let π' be a strategy for player $\Diamond$ in G_{Δ_C} such that for all $n \in \mathbb{N}$ and all $\alpha_1, \ldots, \alpha_n \in C^*$ we have that the rule distribution of $\pi'(Y, \alpha_1, \ldots, \alpha_n)$ is the same as the rule distribution of $\pi(v, \alpha_1\beta, \ldots, \alpha_n\beta)$. Observe that every run $w \in U_v$ determines a unique run $w_C \in Run(Y)$ in $G_{\Delta_C}(\sigma', \pi')$ obtained from w by first deleting the prefix $v(0), \ldots, v(|v|-2)$ and then "cutting off" β from all configurations in the resulting run. Let $U = \{w_C \mid w \in U_v\}$. Now it is easy to see that $\mathcal{P}_\alpha^{\sigma,\pi}(U_v) = \mathcal{P}_\alpha^{\sigma,\pi}(Run(v)) \cdot \mathcal{P}_Y^{\sigma',\pi'}(U)$. Note that all runs of U avoid visiting ε. However, we show that almost all runs of $G_{\Delta_C}(\sigma', \pi')$ reach ε, which implies $\mathcal{P}_Y^{\sigma',\pi'}(U) = 0$ and hence also $\mathcal{P}_\alpha^{\sigma,\pi}(U_v) = 0$.

Observe that the strategy σ' works as follows. There is a constant $k \leq 2|\Gamma|$ such that in every $\gamma \in C^+$, where $|\gamma| \leq k$, player $\Box$ selects edges uniformly at random. Otherwise, player $\Box$ selects the same edges as if she was playing according to σ_T. We show that there is $0 < \xi < 1$ such that for every γ, where $|\gamma| \leq k$, the probability of reaching ε from γ in $G_{\Delta_C}(\sigma', \pi')$ is at least ξ. Note that if player $\Box$ was playing uniformly in all configurations, the existence of such a ξ would be guaranteed by Lemma 7.3. However, playing according to σ_T in configurations whose length exceeds k can only *increase* the probability of reaching ε. Now note that almost all runs of $Run(Y)$ in $G_{\Delta_C}(\sigma', \pi')$ visit configurations of the form $\gamma \in C^+$, where $|\gamma| \leq k$, infinitely often. From this we obtain that almost all runs of $Run(Y)$ in $G_{\Delta_C}(\sigma', \pi')$ reach ε. $\Box$

7.3. Proofs of Proposition 6.9 and Theorem 6.11

The procedure `Main` (see page 28) starts by initializing $\mathcal{W}$ and $\mathcal{U}$ to $\emptyset$, and the auxiliary BPA game Θ to $\bar{\Delta}$ (the set of rules of Θ is denoted by $\rightsquigarrow$). In the main loop initiated at line 2 we first compute the greatest set W of witnesses in the current game Θ. At line 3, we assign to $\mathcal{M}$ the least fixed-point of the function $\mathtt{Att}_{\Theta,W}$. The BPA game Θ is then modified by "cutting off" the set $\mathcal{M}$ at lines 4–9. Note that the resulting BPA game is again in SNF and it is strictly smaller than the original Θ. Then, the current sets $\mathcal{W}$ and $\mathcal{U}$ are enlarged at lines 10,11, and the new (strictly smaller) game Θ is processed in the same way. This goes on until $\mathcal{W}$ and $\mathcal{U}$ stabilize, which obviously requires only $\mathcal{O}(|\bar{\Delta}|)$ iterations of the main loop.

Let K be the number of iterations of the main loop. For every $0 \leq i \leq K$, let Θ_i, $\mathcal{W}_i$, and $\mathcal{U}_i$ be the values of Θ, $\mathcal{W}$, and $\mathcal{U}$ after executing exactly i iterations. Further, W_i denotes the set of all witnesses in Θ_i, and $\mathcal{M}_i$ denotes the least fixed-point of $\mathtt{Att}_{\Theta_i,W_i}$. The symbols Σ_i and Π_i denote the set of all strategies for player $\Box$ and player $\Diamond$ in G_{Θ_i}, respectively. Finally, $\bar{\Gamma}_i$ and $\rightsquigarrow_i$ denote the stack alphabet and the set of all rules of Θ_i, and $\mapsto_i$ denotes the edge relation of G_{Θ_i}. Observe that $\Theta_0 = \bar{\Delta}$, $\mathcal{W}_0 = \mathcal{U}_0 = \emptyset$, $W_K = \emptyset$, and $\mathcal{W}_K, \mathcal{U}_K$ is the result of the procedure `Main`. Let us note that in this section, the sets $[T_\varepsilon]_\Diamond^{<1}$ and $[T]_\Diamond^{<1}$ are always considered in the game $\bar{\Delta} = \Theta_0$.

We start by a simple observation about the relationship between the symbols X and $\widetilde{X}$ in $\bar{\Delta} = \Theta_i$, which is easy to prove.

Lemma 7.7. *For every $0 \leq i \leq K$ and $X \in \Gamma$ we have that if $X \in \mathcal{W}_i$, then also $\widetilde{X} \in \mathcal{W}_i$. In particular, $\mathcal{W}_i \subseteq \mathcal{U}_i$.*

Now we show that $\mathcal{W}_K = \mathscr{A}_0$ and $\mathcal{U}_K = \mathscr{C}_0$. For every $0 \leq i \leq K$, let $[T_\varepsilon, i]_\Diamond^{<1} = \mathcal{U}_i^* \mathcal{W}_i \bar{\Gamma}^*$ and $[T, i]_\Diamond^{<1} = \mathcal{U}_i^* \mathcal{W}_i \bar{\Gamma}^* \cup \mathcal{U}_i^*$.

Lemma 7.8. *There are regular MD strategies $\pi[T], \pi[T_\varepsilon] \in \Pi_0$ constructible by a deterministic polynomial-time algorithm with **NP** $\cap$ **co-NP** oracle such that $\pi[T]$ is $(T,<1)$-winning in every configuration of $[T,K]_\Diamond^{<1}$, and $\pi[T_\varepsilon]$ is $(T_\varepsilon,<1)$-winning in every configuration of $[T_\varepsilon,K]_\Diamond^{<1}$.*

Proof. Let $\mathcal{T}$ be either T or T_ε. We construct a regular MD strategy $\pi \in \Pi_0$ such that π is $(\mathcal{T},<1)$-winning in every configuration of $[\mathcal{T},K]_\Diamond^{<1}$.

Due to Lemma 6.4, for every $0 \leq i < K$ there is a SMD strategy $\pi_i \in \Pi_i$ constructible by a deterministic polynomial-time algorithm with **NP** $\cap$ **co-NP** oracle such that for every $\bar{Z} \in \mathcal{M}_i$ and every $\sigma_i \in \Sigma_i$ we have that $\mathcal{P}_{\bar{Z}}^{\sigma_i,\pi_i}(\mathit{Reach}(T_\varepsilon, G_{\Theta_i})) < 1$. (Strictly speaking, Lemma 6.4 guarantees the existence of a SMD strategy $\pi_i \in \Pi_i$ such that the above condition is satisfied just for all $\bar{Z} \in W_i$. However, the strategy π_i of Lemma 6.4 can be easily modified so that it works for all $\bar{Z} \in \mathcal{M}_i = \bigcup_{j=0}^{|\bar{\Gamma}|} \mathtt{Att}_{\Theta_i,W_i}^j(\emptyset)$; whenever a new symbol $\bar{A} \in \bar{\Gamma}_\Diamond$ appears in $\mathtt{Att}_{\Theta_i,W_i}^{j+1}(\emptyset)$, we fix one of the rules $\bar{A} \rightsquigarrow_i \bar{B}$ which witness the membership of $\bar{A}$ to $\mathtt{Att}_{\Theta_i,W_i}^{j+1}(\emptyset)$.)

For every $\alpha \in [\mathcal{T},K]^{\leq 1}_{\diamond}$, let $I(\alpha)$ be the least ℓ such that $\alpha \in [\mathcal{T},\ell]^{\leq 1}_{\diamond}$. The strategy π is constructed so that for every $\bar{Y}\beta \in [\mathcal{T},K]^{\leq 1}_{\diamond}$, where $\bar{Y} \in \bar{\Gamma}_{\diamond}$ and $I(\bar{Y}\beta) = i{+}1$, the following conditions are satisfied:

- If $\bar{Y} \in \mathcal{M}_i$ and $\pi_i(\bar{Y})$ selects the edge $\bar{Y} \rightsquigarrow_i \bar{Z}$, then $\pi(\bar{Y}\beta)$ selects the edge $\bar{Y}\beta \rightsquigarrow_0 \bar{Z}\beta$.
- If $\bar{Y} \notin \mathcal{M}_i$ and $\pi_j(\widetilde{Y})$ selects the edge $\widetilde{Y} \rightsquigarrow_j \widetilde{Z}$, then $\pi(\bar{Y}\beta)$ selects the edge $\bar{Y}\beta \rightsquigarrow_0 \bar{Z}\beta$. Here $j \leq i$ is the unique index such that $\widetilde{Y} \in \mathcal{M}_j$. (Note that since $\bar{Y} \notin \mathcal{M}_i \subseteq \mathcal{W}_{i+1}$ and $\bar{Y}\beta \in [\mathcal{T},i{+}1]^{\leq 1}_{\diamond}$, we have that $\bar{Y} \in \mathcal{U}_{i+1}$, hence $\widetilde{Y} \in \mathcal{W}_{i+1}$, which means that $\widetilde{Y} \in \mathcal{M}_j$ for some $j \leq i$.)

If $\bar{Y}\beta \notin [\mathcal{T},K]^{\leq 1}_{\diamond}$, then $\pi(\bar{Y}\beta)$ is defined arbitrarily.

We show that π is $(\mathcal{T},{<}1)$-winning in every configuration of $[\mathcal{T},K]^{\leq 1}_{\diamond}$. We proceed by induction with respect to a well-founded strict ordering $\sqsubset$ over $[\mathcal{T},K]^{\leq 1}_{\diamond}$ defined as follows

- For every $\alpha \in \bar{\Gamma}^*$, let $I_{\mathcal{U}}(\gamma)$ be the least $0 \leq \ell \leq K$ such that $\alpha \in \mathcal{U}^*_{\ell}$ (if there is no such ℓ, we put $I_{\mathcal{U}}(\gamma) = \infty$). Also note that $I_{\mathcal{U}}(\varepsilon) = 0$.
- For all $\alpha,\beta \in [\mathcal{T},K]^{\leq 1}_{\diamond}$, we put $\alpha \sqsubset \beta$ if either $I(\alpha) < I(\beta)$, or $I(\alpha) = I(\beta)$ and $I_{\mathcal{U}}(\gamma_1) < I_{\mathcal{U}}(\gamma_2)$, where $\alpha = \gamma_1\eta$, $\beta = \gamma_2\eta$, and η is the longest common suffix of α and β.

One can easily verify that $\sqsubset$ is well-founded.

Let $\alpha \in [\mathcal{T},K]^{\leq 1}_{\diamond}$ such that $I(\alpha) = i+1$. We show that for every $\sigma \in \Sigma_0$ we have that either $\mathcal{P}^{\sigma,\pi}_{\alpha}(\mathit{Reach}(\mathcal{T}), G_{\Theta_0}) < 1$, or $\mathcal{P}^{\sigma,\pi}_{\alpha}(\mathit{Reach}(\sqsubset\alpha), G_{\Theta_0}) > 0$, where $\sqsubset\alpha$ is the set of all $\gamma \in [\mathcal{T},K]^{\leq 1}_{\diamond}$ such that $\gamma \sqsubset \alpha$. Since $\sqsubset$ is well-founded, it follows immediately that π is $(\mathcal{T},{<}1)$-winning in every configuration of $[\mathcal{T},K]^{\leq 1}_{\diamond}$.

If $\mathcal{P}^{\sigma,\pi}_{\alpha}(\mathit{Reach}(\sqsubset\alpha), G_{\Theta_0}) > 0$, we are done. Now assume that

$$\mathcal{P}^{\sigma,\pi}_{\alpha}(\mathit{Reach}(\sqsubset\alpha), G_{\Theta_0}) = 0 \tag{5}$$

where $\alpha = \bar{X}\beta$. We distinguish two cases.

- $\bar{X} \in \mathcal{M}_i$. Then we put $k = i$ and $\mathcal{S} = (\bar{\Gamma} \setminus \bar{\Gamma}_k)\bar{\Gamma}^*$.
- $\bar{X} \in \mathcal{U}_{i+1} \setminus \mathcal{W}_{i+1}$ and $\beta \in [\mathcal{T},i{+}1]^{\leq 1}_{\diamond}$. Then $k = j$, where $j \leq i$ is the unique index such that $\widetilde{X} \in \mathcal{M}_j$, and $\mathcal{S} = (\bar{\Gamma} \setminus \bar{\Gamma}_k)\bar{\Gamma}^* \cup \{\varepsilon\}$.

Intuitively, $\mathcal{S}$ is a set of "safe" prefixes that cannot be visited without violating assumption (5)—observe that for every $\gamma \in \mathcal{S}$ we have that $\gamma\beta \sqsubset \bar{X}\beta$.

Let $f : \bar{\Gamma}^* \to \bar{\Gamma}_k^*$ be a function defined inductively as follows (where $\alpha = \bar{X}\beta$ still denotes the initial configuration fixed above):

- $f(\varepsilon) = \varepsilon$;
- $f(\bar{A}\delta) = \begin{cases} f(\delta) & \text{if } \bar{A} \notin \bar{\Gamma}_k; \\ \bar{A}f(\delta) & \text{if } \bar{A} \in \bar{\Gamma}_k \text{ and } \delta \notin \mathcal{S}; \\ \widetilde{A}f(\delta) & \text{if } \widetilde{A} \in \bar{\Gamma}_k \text{ and } \delta \in \mathcal{S}. \end{cases}$

Note that $\widetilde{A} \in \bar{\Gamma}_k$ implies $\bar{A} \in \bar{\Gamma}_k$ by Lemma 7.7, and hence the function f is well-defined.

Since π_k can be defined arbitrarily for all $\bar{X} \notin \mathcal{M}_k \cap \bar{\Gamma}_\diamond$ (see Lemma 6.4), we may safely redefine the strategy π_k so that

- If $\bar{A}\delta \notin \mathcal{S}$, $\delta \notin \mathcal{S}$, $\bar{A} \in \bar{\Gamma}_\diamond$, $\bar{A} \in \Gamma_k \smallsetminus \mathcal{M}_k$, and $\pi(\bar{A}\delta\beta)$ selects an edge $\bar{A}\delta\beta \leadsto_0 \bar{B}\delta\beta$, then $\pi_k(\bar{A}f(\delta))$ selects the edge $\bar{A}f(\delta) \leadsto_k \bar{B}f(\delta)$.

- If $\bar{A}\delta \notin \mathcal{S}$, $\delta \in \mathcal{S}$, $\bar{A} \in \bar{\Gamma}_\diamond$, $\widetilde{A} \in \Gamma_k \smallsetminus \mathcal{M}_k$, and $\pi(\bar{A}\delta\beta)$ selects an edge $\bar{A}\delta\beta \leadsto_0 \bar{B}\delta\beta$, then $\pi_k(\widetilde{A}f(\delta))$ selects the edge $\widetilde{A}f(\delta) \leadsto_k \widetilde{B}f(\delta)$.

From now on, π_k denotes this modified strategy.

For every reachable state $\gamma_0\beta, \ldots, \gamma_j\beta$ of $G_{\Theta_0}(\sigma, \pi)$ we put $\mathcal{F}(\gamma_0\beta, \ldots, \gamma_j\beta) = f(\gamma_0), \ldots, f(\gamma_j)$, where f is the function defined above. Our aim is to construct a strategy $\sigma_k \in \Sigma_k$ so that $\mathcal{F}$ becomes an isomorphism between the reachable parts of $G_{\Theta_0}(\sigma, \pi)$ and $G_{\Theta_k}(\sigma_k, \pi_k)$ initiated in $\bar{X}\beta$ and $f(\bar{X})$, respectively, where the configuration β of Θ_0 is considered as being isomorphic to ε. From this we obtain

$$\mathcal{P}_{\alpha}^{\sigma,\pi}(\mathit{Reach}(\mathcal{T}), G_{\Theta_0}) \;\leq\; \mathcal{P}_{f(\bar{X})}^{\sigma_k,\pi_k}(\mathit{Reach}(T_\varepsilon), G_{\Theta_k}) \tag{6}$$

Since $f(\bar{X}) \in \mathcal{M}_k$, we have that $\mathcal{P}_{f(\bar{X})}^{\sigma_k,\pi_k}(\mathit{Reach}(T_\varepsilon), G_{\Theta_i}) < 1$, and we are done.

It remains to construct the strategy σ_k and verify that $\mathcal{F}$ is an isomorphism. Let $\gamma_0\beta, \ldots, \gamma_j\beta$ be a reachable state of $G_{\Theta_0}(\sigma, \pi)$ such that $f(\gamma_0), \ldots, f(\gamma_j)$ is a reachable state of $G_{\Theta_k}(\sigma_k, \pi_k)$, and $\gamma_n \neq \varepsilon$ for all $0 \leq n \leq j$. We show that if $\gamma_0\beta, \ldots, \gamma_j\beta \xrightarrow{x} \gamma_0\beta, \ldots, \gamma_{j+1}\beta$ where $x > 0$, then $f(\gamma_0), \ldots, f(\gamma_j) \xrightarrow{x} f(\gamma_0), \ldots, f(\gamma_{j+1})$. Note that then $f(\gamma_0), \ldots, f(\gamma_j)$

cannot have any "additional" outgoing transitions because $G_{\Theta_k}(\sigma_k, \pi_k)$ is a Markov chain.

If $\gamma_0\beta, \ldots, \gamma_j\beta \xrightarrow{x} \gamma_0\beta, \ldots, \gamma_{j+1}\beta$, where $\gamma_j \neq \varepsilon$ and $x > 0$, then $\gamma_j \mapsto_0 \gamma_{j+1}$ is an edge in G_{Θ_0}, which is assigned the probability x either by *Prob*, π, or σ, depending on whether the first symbol of γ_j belongs to $\bar{\Gamma}_\bigcirc$, $\bar{\Gamma}_\Diamond$, or $\bar{\Gamma}_\Box$, respectively. It suffices to show that $f(\gamma_j) \mapsto_k f(\gamma_{j+1})$ is an edge in G_{Θ_k}, which is assigned the same probability x by *Prob*, π_k, or the newly constructed σ_k, respectively. Let $\gamma_j = \bar{A}\delta$. Note that since $\bar{A}\delta\beta \not\sqsubseteq \alpha$, we have that $\bar{A}\delta \notin \mathcal{S}$ and hence $f(\gamma_j) = f(\bar{A}\delta) = \hat{A}f(\delta)$, where $\hat{A} = \widetilde{A}$ or $\hat{A} = \bar{A}$, depending on whether $\delta \in \mathcal{S}$ or not, respectively. We distinguish three possibilities.

- $\bar{A} \in \bar{\Gamma}_\bigcirc$. Then $\gamma_{j+1} = \bar{B}\delta$ for some $\bar{B}$ such that $\bar{A} \overset{x}{\rightsquigarrow}_0 \bar{B}$. But then also $\hat{A} \overset{x}{\rightsquigarrow}_k \hat{B}$, where $\hat{B}$ is either B or $\widetilde{B}$ depending on whether $\hat{A} = A$ or $\hat{A} = \widetilde{A}$, respectively. Hence, $f(\gamma_j) = \hat{A}f(\delta) \overset{x}{\mapsto}_k \hat{B}f(\delta) = f(\bar{B}\delta) = f(\gamma_{j+1})$ as needed.

- $\bar{A} \in \bar{\Gamma}_\Diamond$. Then we consider two cases.

  - $\delta \notin \mathcal{S}$. Then $\hat{A} = \bar{A}$. If $\bar{A} \in \mathcal{M}_k$, then $k = I(\alpha) - 1$, because otherwise $\bar{A}\delta\beta \sqsubset \alpha$. This means that $\pi(\bar{A}\delta\beta)$ selects the edge $\bar{A}\delta\beta \rightsquigarrow_0 \bar{B}\delta\beta$ where $\bar{A} \rightsquigarrow_k \bar{B}$ is the edge selected by π_k in $\bar{A}$. Hence, π_k selects the edge $\bar{A}f(\delta) \rightsquigarrow_k \bar{B}f(\delta)$ in $\bar{A}f(\delta)$, and thus $f(\gamma_j) = \bar{A}f(\delta) \overset{x}{\mapsto}_k \bar{B}f(\delta) = f(\bar{B}\delta) = f(\gamma_{j+1})$.

    If $\bar{A} \notin \mathcal{M}_k$, then $\pi_k(\bar{A}f(\delta))$ "mimics" $\pi(\bar{A}\delta\beta)$ and we again obtain $f(\gamma_j) \overset{x}{\mapsto}_k f(\gamma_{j+1})$.
  - $\delta \in \mathcal{S}$. Then $\hat{A} = \widetilde{A}$. If $\widetilde{A} \in \mathcal{M}_k$, then $k = I(\alpha) - 1$, because otherwise $\bar{A}\delta\beta \sqsubset \alpha$. Hence, $\pi(\bar{A}\delta\beta)$ "mimics" $\pi_k(\widetilde{A})$. If $\widetilde{A} \notin \mathcal{M}_k$, then $\pi_k(\widetilde{A}f(\delta))$ "mimics" $\pi(\bar{A}\delta\beta)$. The details are almost the same as in the previous case.

- $\bar{A} \in \bar{\Gamma}_\Box$. This is the most complicated case. It suffices to show that $f(\gamma_j) = \hat{A}f(\delta) \mapsto_k f(\gamma_{j+1})$. The distribution $\sigma_k(f(\gamma_0), \ldots, f(\gamma_j))$ can then safely select the edge $f(\gamma_j) \mapsto_k f(\gamma_{j+1})$ with probability x. According to Definition 6.8, we can distinguish the following three possibilities:

  - $\bar{A} \in \bar{\Gamma}[1]$. Then $\gamma_j = \bar{B}\delta$ for some $\bar{B}$ such that $\bar{A} \rightsquigarrow_0 \bar{B}$. If $\delta \notin \mathcal{S}$, then $\hat{A} = \bar{A}$. Further, $\bar{B} \in \Gamma_k$ because otherwise $\bar{B}\delta \in \mathcal{S}$

and hence $\bar{B}\delta\beta \sqsubset \alpha$, which contradicts assumption (5). Thus, $f(\gamma_j) = \bar{A}f(\delta) \mapsto_k \bar{B}f(\delta) = f(\gamma_{j+1})$ as needed.

If $\delta \in \mathcal{S}$, then $\hat{A} = \widetilde{A}$ and $\widetilde{B} \in \Gamma_k$, because otherwise $\bar{B} \in \mathcal{U}_k$ and $\bar{B}\delta\beta \sqsubset \alpha$, which contradicts assumption (5). Hence, we obtain $f(\gamma_j) = \widetilde{A}f(\delta) \mapsto_k \widetilde{B}f(\delta) = f(\gamma_{j+1})$.

- $\bar{A} \in \bar{\Gamma}[2]$. Then $\bar{A} = A$ and $\gamma_{j+1} = \delta$. If $\delta \notin \mathcal{S}$, then $f(\gamma_j) = Af(\delta) \mapsto_k f(\delta) = f(\gamma_{j+1})$. If $\delta \in \mathcal{S}$, we obtain a contradiction with assumption (5).
- $\bar{A} \in \bar{\Gamma}[3]$. Then $\gamma_{j+1} = B\bar{C}\delta$ where $\bar{A} \rightsquigarrow_0 B\bar{C}$ is the only available rule with $\bar{A}$ on the left-hand side. If $\delta \notin \mathcal{S}$, then $\hat{A} = \bar{A}$ and we have that either $\bar{A} \rightsquigarrow_k B\bar{C}$ or $\bar{A} \rightsquigarrow_k \widetilde{B}$. In the first case we obtain $f(B\bar{C}\delta) = B\bar{C}f(\delta)$ and hence $f(\gamma_j) = \bar{A}f(\delta) \mapsto_k B\bar{C}f(\delta) = f(\gamma_{j+1})$. In the latter case, $\bar{C} \notin \bar{\Gamma}_k$ and $\widetilde{B} \in \bar{\Gamma}_k$, which means that $f(B\bar{C}\delta) = \widetilde{B}f(\delta)$ and $f(\gamma_j) = \bar{A}f(\delta) \mapsto_k \widetilde{B}f(\delta) = f(\gamma_{j+1})$.

  If $\delta \in \mathcal{S}$, then $\hat{A} = \widetilde{A}$. Further, we have that either $\widetilde{A} \rightsquigarrow_k B\widetilde{C}$ or $\widetilde{A} \rightsquigarrow_k \widetilde{B}$. In the first case we obtain $f(B\widetilde{C}\delta) = B\widetilde{C}f(\delta)$ and hence $f(\gamma_j) = \widetilde{A}f(\delta) \mapsto_k B\widetilde{C}f(\delta) = f(\gamma_{j+1})$. In the latter case, $\widetilde{C} \notin \bar{\Gamma}_k$ and $\widetilde{B} \in \bar{\Gamma}_k$, which means that $f(B\widetilde{C}\delta) = \widetilde{B}f(\delta)$ and $f(\gamma_j) = \widetilde{A}f(\delta) \mapsto_k \widetilde{B}f(\delta) = f(\gamma_{j+1})$.

□

A trivial corollary of Lemma 7.7 is the "$\subseteq$" direction of Proposition 6.9 (note that $\mathcal{W}_K \subseteq \mathcal{U}_K^* \mathcal{W}_K \bar{\Gamma}^*$ and $\mathcal{U}_K \subseteq \mathcal{U}_K^*$). Now we show the "$\supseteq$" direction.

Lemma 7.9. *We have that $\mathcal{W}_K \supseteq \mathscr{A}_0$ and $\mathcal{U}_K \supseteq \mathscr{C}_0$.*

Proof. Since $W_K = \emptyset$, due to Proposition 7.2 there is a regular MR strategy $\sigma_K \in \Sigma_K$ which is $(T_\varepsilon,{=}1)$-winning in every $\alpha \in \bar{\Gamma}_K^*$. Moreover, the strategy σ_K is computable in time which is polynomial in the size of Θ_K (assuming that Θ_K has already been computed). Let $\mathscr{B}_K = \bar{\Gamma} \setminus \mathcal{W}_K = \bar{\Gamma}_K$ and $\mathscr{D}_K = \bar{\Gamma} \setminus \mathcal{U}_K$. One can easily check that $\mathcal{D}_K \subseteq \mathscr{B}_K$ and

(i) if $\widetilde{X} \in \mathscr{B}_K$, then $X, \widetilde{X} \in \mathcal{D}_K$;

(ii) if $\widetilde{X} \notin \mathscr{B}_K$, then $X \notin \mathcal{D}_K$.

We show that the strategy σ_K can be efficiently transformed into regular MR strategies $\sigma_0, \hat{\sigma}_0 \in \Sigma_0$ such that

- σ_0 is $(T_\varepsilon,{=}1)$-winning in every configuration of $\mathscr{B}_K^* \cup \mathscr{B}_K^* \mathcal{D}_K \bar{\Gamma}^*$;

- $\hat{\sigma}_0$ is $(T,{=}1)$-winning in every configuration of $\mathscr{B}_K^* \mathcal{D}_K \bar{\Gamma}^*$.

In particular, $\mathscr{B}_K \subseteq [T_\varepsilon]_\Box^{=1}$ and $\mathscr{D}_K \subseteq [T]_\Box^{=1}$, hence $\mathcal{W}_K = \bar{\Gamma} \smallsetminus \mathscr{B}_K \supseteq \bar{\Gamma} \smallsetminus [T_\varepsilon]_\Box^{=1} = \mathscr{A}_0$ and $\mathcal{U}_K = \bar{\Gamma} \smallsetminus \mathscr{D}_K \supseteq \bar{\Gamma} \smallsetminus [T]_\Box^{=1} = \mathscr{C}_0$ as needed.

First we show how to construct the strategy σ_0. Let $g : \bar{\Gamma}^* \to \bar{\Gamma}_K^*$ be a partial function defined as follows:

- $g(\varepsilon) = \varepsilon$

- $g(\bar{Y}\beta) = \begin{cases} Y g(\beta) & \text{if } \bar{Y} = Y \in \mathscr{B}_K \text{ and } \beta \in \mathscr{B}_K^* \cup \mathscr{B}_K^* \mathcal{D}_K \bar{\Gamma}^*; \\ \widetilde{Y} & \text{if } \widetilde{Y} \in \mathscr{B}_K \text{ and either } \bar{Y} = \widetilde{Y} \text{ or } \beta \notin \mathscr{B}_K^* \cup \mathscr{B}_K^* \mathcal{D}_K \bar{\Gamma}^*; \\ \bot & \text{otherwise.} \end{cases}$

A configuration $\alpha \in \bar{\Gamma}^*$ is *g-eligible* if $g(\alpha) \neq \bot$. Observe that $\alpha \in \bar{\Gamma}_T \bar{\Gamma}^*$ iff $g(\alpha) \in \bar{\Gamma}_T \bar{\Gamma}^*$, and $\alpha = \varepsilon$ iff $g(\alpha) = \varepsilon$. Further, for every configuration $\alpha \in \bar{\Gamma}_K^*$ which belongs to player $\Box$ we have that $\sigma_K(\alpha) = \sigma_K(g(\alpha))$, which follows immediately from the definition of g and Remark 7.4. These simple observations are used later when constructing the isomorphism $\mathcal{G}$.

The strategy σ_0 is constructed so that for every g-eligible $\bar{A}\alpha \in \bar{\Gamma}_\Box \bar{\Gamma}^*$, the following conditions are satisfied:

- If $\bar{A} \in \bar{\Gamma}[1]$ and $\sigma_K(g(\bar{A}\alpha))$ selects a rule $\hat{A} \rightsquigarrow_K \hat{B}$ with probability x, then $\sigma_0(\bar{A}\alpha)$ selects the rule $\bar{A} \rightsquigarrow_0 \bar{B}$ with probability x.

- If $\bar{A} \in \bar{\Gamma}[2] \cup \bar{\Gamma}[3]$, then $\sigma_0(\bar{A}\alpha)$ selects the only available rule with probability 1.

Note that the definition of σ_0 is effective in the sense that if the finite-state automaton $\mathscr{A}_{\sigma_K}$ associated with the regular MR strategy σ_K (see Definition 4.3) has already been computed, then the finite-state automaton $\mathscr{A}_{\sigma_0}$ associated with σ_0 simply "simulates" the execution of $\mathscr{A}_{\sigma_K}$ on the reverse of $g(\alpha)$ for every g-eligible $\alpha \in \bar{\Gamma}^*$. Hence, the automaton $\mathscr{A}_{\sigma_0}$ is constructible in polynomial time assuming that the BPA game Θ_K has already been computed (cf. Proposition 7.2).

We show that for every g-eligible initial configuration $\gamma \in \bar{\Gamma}^*$ and every $\pi_0 \in \Pi_0$ there is a strategy $\pi_K \in \Pi_K$ such that

$$\mathcal{P}_\gamma^{\sigma_0,\pi_0}(\mathit{Reach}(T_\varepsilon), G_{\Theta_0}) \quad = \quad \mathcal{P}_{g(\gamma)}^{\sigma_K,\pi_K}(\mathit{Reach}(T_\varepsilon), G_{\Theta_K}) \qquad (7)$$

Since $\mathcal{P}^{\sigma_K,\pi_K}_{g(\gamma)}(\mathit{Reach}(T_\varepsilon), G_{\Theta_K}) = 1$, we obtain that σ_0 is $(T_\varepsilon,=1)$-winning in every g-eligible γ.

For the rest of this proof, we fix a g-eligible configuration γ and a strategy $\pi_0 \in \Pi_0$. We inductively construct a strategy $\pi_K \in \Pi_K$ and an isomorphism $\mathcal{G}$ between the reachable parts of $G_{\Theta_0}(\sigma_0, \pi_0)$ and $G_{\Theta_K}(\sigma_K, \pi_K)$ so that whenever $\mathcal{G}(\alpha_0, \ldots, \alpha_j) = \xi_0, \ldots, \xi_j$, then $g(\alpha_j) = g(\xi_j)$ (for technical reasons, we also maintain the invariant $\alpha_j \in \widetilde{\Gamma}\bar{\Gamma}^* \Rightarrow \xi_j \in \widetilde{\Gamma}\bar{\Gamma}^*$). From this we immediately obtain (7).

We proceed by induction on j. In the base case we put $\mathcal{G}(\gamma) = g(\gamma)$. Now assume that $\alpha_0, \ldots, \alpha_j$ is a reachable state of $G_{\Theta_0}(\sigma_0, \pi_0)$ such that $\alpha_0, \ldots, \alpha_{j-1} \xrightarrow{x} \alpha_0, \ldots, \alpha_j$, where $x > 0$. Then $\alpha_{j-1} \mapsto_0 \alpha_j$ is an edge in G_{Θ_0}, which is assigned the probability x either by *Prob*, σ_0, or π_0, depending on whether the first symbol of α_{j-1} belongs to $\bar{\Gamma}_\bigcirc$, $\bar{\Gamma}_\Box$, or $\bar{\Gamma}_\Diamond$, respectively. By induction hypothesis, $\mathcal{G}(\alpha_0, \ldots, \alpha_{j-1}) = \xi_0, \ldots, \xi_{j-1}$ where $g(\alpha_{j-1}) = g(\xi_{j-1})$, and $\alpha_{j-1} \in \widetilde{\Gamma}\bar{\Gamma}^*$ implies $\xi_{j-1} \in \widetilde{\Gamma}\bar{\Gamma}^*$. Hence, it suffices to show that there is an edge $\xi_{j-1} \mapsto_K \xi_j$ in G_{Θ_K} such that

- $g(\alpha_j) = g(\xi_j)$, and if $\alpha_j \in \widetilde{\Gamma}\bar{\Gamma}^*$ then also $\xi_j \in \widetilde{\Gamma}\bar{\Gamma}^*$;
- the edge $\xi_{j-1} \mapsto_K \xi_j$ is assigned the same probability x by *Prob*, σ_K, or the newly constructed π_k, respectively.

Let $\alpha_{j-1} = \bar{A}\beta$ and $\xi_{j-1} = \hat{A}\varrho$. We distinguish two possibilities:

- $g(\bar{A}\beta) = Ag(\beta)$. Then $\beta \in \mathscr{B}_K^* \cup \mathscr{B}_K^*\mathcal{D}_K\bar{\Gamma}^*$, $\hat{A} = A$, and $g(\beta) = g(\varrho)$. Let us first consider the case when $A \in \bar{\Gamma}_\bigcirc \cup \bar{\Gamma}_\Diamond \cup \bar{\Gamma}[1]$. Then $\alpha_{j-1} = A\beta \mapsto_0 B\beta = \alpha_j$ where $A \rightsquigarrow_0 B$. Since $A \in \mathscr{B}_K$, we also have $A \rightsquigarrow_K B$, and hence $\xi_{j-1} = A\varrho \mapsto_K B\varrho = \xi_j$. Obviously, $g(\alpha_j) = g(B\beta) = Bg(\beta) = Bg(\varrho) = g(B\varrho) = g(\xi_j)$. We show that the constructed edge $\xi_{j-1} \mapsto_K \xi_j$ is assigned the probability x. If $A \in \bar{\Gamma}_\bigcirc$, then x is the (fixed) probability of $A \rightsquigarrow_0 B$. Since $A \rightsquigarrow_K B$ has the same probability as $A \rightsquigarrow_0 B$, we are done. If $A \in \bar{\Gamma}_\Diamond$, we simply *define* the strategy π_K so that is selects the edge $\xi_{j-1} \mapsto_K \xi_j$ with probability x (note that π_K is also defined inductively). Finally, if $A \in \bar{\Gamma}[1]$, we observe that if the strategy σ_0 selects the edge $\alpha_{j-1} \mapsto_0 \alpha_j$ with probability x, then σ_K selects the edge $\xi_{j-1} \mapsto_0 \xi_j$ also with probability x, because $\sigma_K(\xi_{j-1}) = \sigma_K(g(\xi_{j-1}))$ (see above).

  If $A \in \bar{\Gamma}[2]$, then $\alpha_{j-1} = A\beta \mapsto_0 \beta = \alpha_j$ where $A \rightsquigarrow_0 \varepsilon$. Since $A \in \mathscr{B}_K$, we also have $A \rightsquigarrow_K \varepsilon$, and hence $\xi_{j-1} = A\varrho \mapsto_K \varrho = \xi_j$. Note that

$g(\alpha_j) = g(\beta) = g(\varrho) = g(\xi_j)$ and the edges $\alpha_{j-1} \mapsto_0 \alpha_j$ and $\xi_{j-1} \mapsto_K \xi_j$ are selected with probability 1 by σ_0 and σ_K, respectively.

Finally, if $A \in \bar{\Gamma}[2]$, then $\alpha_{j-1} = A\beta \mapsto_0 BC\beta = \alpha_j$ where $A \rightsquigarrow_0 BC$. Since $A \in \mathscr{B}_K$, there are two possibilities:

- $A \rightsquigarrow_K BC$. Then $\xi_{j-1} = A\varrho \mapsto_K BC\varrho = \xi_j$, and $g(\alpha_j) = g(BC\beta) = BCg(\beta) = BCg(\varrho) = g(BC\varrho) = g(\xi_j)$.
- $A \rightsquigarrow_K \widetilde{B}$. Then $C \notin \mathscr{B}_K$ and $\xi_{j-1} = A\varrho \mapsto_K \widetilde{B}\varrho = \xi_j$. Since $C \notin \mathscr{B}_K$, we obtain $g(\alpha_j) = g(BC\beta) = \widetilde{B} = g(\widetilde{B}\varrho) = g(\xi_j)$.

Also note that if $A \in \bar{\Gamma}[2]$, then the edges $\alpha_{j-1} \mapsto_0 \alpha_j$ and $\xi_{j-1} \mapsto_K \xi_j$ are selected with probability 1 by σ_0 and σ_K, respectively.

- $g(\bar{A}\beta) = \widetilde{A}$. If $\bar{A} = \hat{A} = \widetilde{A}$, the arguments are very similar to the ones above. Now let us consider the case when $\bar{A} = A$ and $\hat{A} = \widetilde{A}$. Then $\beta \notin \mathscr{B}_K^* \cup \mathscr{B}_K^* \mathcal{D}_K \bar{\Gamma}^*$. If $A \in \bar{\Gamma}_{\bigcirc} \cup \bar{\Gamma}_{\Diamond} \cup \bar{\Gamma}[1]$, we can again argue similarly as above. If $A \in \bar{\Gamma}[2]$, then $\widetilde{A} \mapsto_K \widetilde{A}$ is the only rule with $\widetilde{A}$ on the left hand side, which means that $\widetilde{A} \in W_K$, and we have a contradiction with $W_K = \emptyset$. So, it remains to discuss the case when $A \in \bar{\Gamma}[3]$. Then $\alpha_{j-1} = A\beta \mapsto_0 BC\beta = \alpha_j$ where $A \rightsquigarrow_0 BC$. Since $\widetilde{A} \in \mathscr{B}_K$, there are two possibilities:

  - $\widetilde{A} \rightsquigarrow_K B\widetilde{C}$. Then $\xi_{j-1} = \widetilde{A}\varrho \mapsto_K B\widetilde{C}\varrho = \xi_j$. Since $\widetilde{C} \in \mathscr{B}_K$, we have that $C \in \mathcal{D}_K$ (see Observation (i) above). Further, $\beta \notin \mathscr{B}_K^* \cup \mathscr{B}_K^* \mathcal{D}_K \bar{\Gamma}^*$ and hence $g(\alpha_j) = g(BC\beta) = B\widetilde{C} = g(B\widetilde{C}\varrho) = g(\xi_j)$.
  - $\widetilde{A} \rightsquigarrow_K \widetilde{B}$. Then $\widetilde{C} \notin \mathscr{B}_K$ and $\xi_{j-1} = \widetilde{A}\varrho \mapsto_K \widetilde{B}\varrho = \xi_j$. Since $\widetilde{C} \notin \mathscr{B}_K$, we have that $C \notin \mathcal{D}_K$ (see Observation (ii) above), hence $C\beta \notin \mathscr{B}^* \cup \mathscr{B}_K^* \mathcal{D}_K \bar{\Gamma}^*$, which implies $g(\alpha_j) = g(BC\beta) = \widetilde{B} = g(\widetilde{B}\varrho) = g(\xi_j)$.

The strategy $\hat{\sigma}_0$ is constructed similarly as σ_0, using the following partial function $\hat{g} : \bar{\Gamma}^* \to \bar{\Gamma}_K^*$ instead of g:

- $\hat{g}(\varepsilon) = \bot$
- $\hat{g}(\bar{Y}\beta) = \begin{cases} Y\hat{g}(\beta) & \text{if } \bar{Y} = Y \in \mathscr{B}_K \text{ and } \beta \in \mathscr{B}_K^* \mathcal{D}_K \bar{\Gamma}^*; \\ \widetilde{Y} & \text{if } \widetilde{Y} \in \mathscr{B}_K \text{ and either } \bar{Y} = \widetilde{Y} \text{ or } \beta \notin \mathscr{B}_K^* \mathcal{D}_K \bar{\Gamma}^*; \\ \bot & \text{otherwise.} \end{cases}$

We show that for every $\hat{g}$-eligible initial configuration $\gamma \in \bar{\Gamma}^*$ and every $\pi_0 \in \Pi_0$ there is a strategy $\hat{\pi}_K \in \Pi_K$ such that

$$\mathcal{P}_{\gamma}^{\hat{\sigma}_0,\pi_0}(\mathit{Reach}(T), G_{\Theta_0}) \;=\; \mathcal{P}_{\hat{g}(\gamma)}^{\sigma_K,\hat{\pi}_K}(\mathit{Reach}(T), G_{\Theta_K}) \qquad (8)$$

Since $\mathcal{P}_{\hat{g}(\gamma)}^{\sigma_K,\hat{\pi}_K}(\mathit{Reach}(T_\varepsilon), G_{\Theta_K}) = 1$ and $\hat{g}(\gamma)$ inevitably contains a symbol of $\widetilde{\Gamma}$ which cannot by removed from the stack, we obtain $\mathcal{P}_{\hat{g}(\gamma)}^{\sigma_K,\hat{\pi}_K}(\mathit{Reach}(T), G_{\Theta_K}) = 1$, which means that $\hat{\sigma}_0$ is $(T,{=}1)$-winning in every $\hat{g}$-eligible γ. Equality (8) is proven by constructing the strategy $\hat{\pi}_K$ inductively, together with an isomorphism $\hat{\mathcal{G}}$ between the reachable parts of the corresponding plays. Technically, this construction is very similar to the one above. □

Lemma 7.7 and Lemma 7.9 together imply Proposition 6.9. It remains to prove Theorem 6.11. The strategy $\hat{\sigma}_0$ constructed in the proof of Lemma 7.9 is $(T,{=}1)$-winning in every configuration of $\mathscr{B}_K^* \mathscr{D}_K \bar{\Gamma}^*$. Since $\mathscr{B}_K = \mathscr{B}$ and $\mathscr{D}_K = \mathscr{D}$ by Proposition 6.9, the strategy $\hat{\sigma}_0$ is $(T,{=}1)$-winning in every configuration of $[T]_{\Box}^{=1}$. As it was noted in the proof of Lemma 7.9, the strategy $\hat{\sigma}_0$ is constructible in polynomial time assuming that the BPA game Θ_K has already been computed. Since Θ_K is computable by a deterministic polynomial-time algorithm with **NP** $\cap$ **co-NP** oracle, the first part of Theorem 6.11 is proven. The second part of Theorem 6.11 follows immediately from Lemma 7.7 because $\mathcal{A} = \mathcal{W}_K$ and $\mathscr{C} = \mathcal{U}_K$.

8. Conclusions

We have solved the qualitative reachability problem for stochastic BPA games, retaining the same upper complexity bounds that have previously been established for termination [15]. One interesting question which remains unsolved is the decidability of the problem whether $\mathit{val}(\alpha) = 1$ for a given BPA configuration α (we can only decide whether player □ has a $({=}1)$-winning strategy, which is sufficient but not necessary for $\mathit{val}(\alpha) = 1$). Another open problem is quantitative reachability for stochastic BPA games, where the methods presented in this paper seem insufficient.

References

[1] P. Abdulla, N.B. Henda, L. de Alfaro, R. Mayr, and S. Sandberg. Stochastic games with lossy channels. In *Proceedings of FoSSaCS*

2008, volume 4962 of *Lecture Notes in Computer Science*, pages 35–49. Springer, 2008.

[2] K.B. Athreya and P.E. Ney. *Branching Processes*. Springer, 1972.

[3] J.C.M. Baeten and W.P. Weijland. *Process Algebra*. Number 18 in Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 1990.

[4] C. Baier, N. Bertrand, and Ph. Schnoebelen. Verifying nondeterministic probabilistic channel systems against ω-regular linear-time properties. *ACM Transactions on Computational Logic*, 9(1), 2007.

[5] T. Brázdil, V. Brožek, V. Forejt, and A. Kučera. Reachability in recursive Markov decision processes. *Information and Computation*, 206(5):520–537, 2008.

[6] V. Brožek. *Basic Model Checking Problems for Stochastic Games*. PhD thesis, Masaryk University, Faculty of Informatics, 2009.

[7] K. Chatterjee, L. de Alfaro, and T. Henzinger. The complexity of stochastic Rabin and Streett games. In *Proceedings of ICALP 2005*, volume 3580 of *Lecture Notes in Computer Science*, pages 878–890. Springer, 2005.

[8] K. Chatterjee, M. Jurdziński, and T. Henzinger. Simple stochastic parity games. In *Proceedings of CSL'93*, volume 832 of *Lecture Notes in Computer Science*, pages 100–113. Springer, 1994.

[9] K. Chatterjee, M. Jurdziński, and T. Henzinger. Quantitative stochastic parity games. In *Proceedings of SODA 2004*, pages 121–130. SIAM, 2004.

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

[11] L. de Alfaro and R. Majumdar. Quantitative solution of omega-regular games. *Journal of Computer and System Sciences*, 68:374–397, 2004.

[12] J. Esparza, A. Kučera, and S. Schwoon. Model-checking LTL with regular valuations for pushdown systems. *Information and Computation*, 186(2):355–376, 2003.

[13] K. Etessami, D. Wojtczak, and M. Yannakakis. Recursive stochastic games with positive rewards. In *Proceedings of ICALP 2008, Part I*, volume 5125 of *Lecture Notes in Computer Science*, pages 711–723. Springer, 2008.

[14] K. Etessami and M. Yannakakis. Recursive Markov decision processes and recursive stochastic games. In *Proceedings of ICALP 2005*, volume 3580 of *Lecture Notes in Computer Science*, pages 891–903. Springer, 2005.

[15] K. Etessami and M. Yannakakis. Efficient qualitative analysis of classes of recursive Markov decision processes and simple stochastic games. In *Proceedings of STACS 2006*, volume 3884 of *Lecture Notes in Computer Science*, pages 634–645. Springer, 2006.

[16] K. Etessami and M. Yannakakis. Recursive concurrent stochastic games. In *Proceedings of ICALP 2006*, volume 4052 of *Lecture Notes in Computer Science*, pages 324–335. Springer, 2006.

[17] R. Fagin, A.R. Karlin, J. Kleinberg P. Raghavan, S. Rajagopalan, R. Rubinfeld, M. Sudan, and A. Tomkins. Random walks with "back buttons". In *Proceedings of STOC 2000*, pages 484–493. ACM Press, 2000.

[18] T.E. Harris. *The Theory of Branching Processes.* Springer, 1963.

[19] J.E. Hopcroft and J.D. Ullman. *Introduction to Automata Theory, Languages, and Computation.* Addison-Wesley, 1979.

[20] A. Maitra and W. Sudderth. Finitely additive stochastic games with Borel measurable payoffs. *International Journal of Game Theory*, 27:257–267, 1998.

[21] C. Manning and H. Schütze. *Foundations of Statistical Natural Language Processing.* The MIT Press, 1999.

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

[23] A. Tarski. A lattice-theoretical fixpoint theorem and its applications. *Pacific Journal of Mathematics*, 5(2):285–309, 1955.

[24] W. Thomas. Infinite games and verification. In *Proceedings of CAV 2003*, volume 2725 of *Lecture Notes in Computer Science*, pages 58–64. Springer, 2003.

[25] N. Vieille. Stochastic games: Recent results. *Handbook of Game Theory*, pages 1833–1850, 2002.

[26] I. Walukiewicz. A landscape with games in the background. In *Proceedings of LICS 2004*, pages 356–366. IEEE Computer Society Press, 2004.