Power of Randomization in Automata on Infinite Strings

Rohit Chadha, A. Prasad Sistla, and Mahesh Viswanathan

[1] Univ. of Illinois, Urbana-Champaign
[2] Univ. of Illinois, Chicago

Abstract. Probabilistic Büchi Automata (PBA) are randomized, finite state automata that process input strings of infinite length. Based on the threshold chosen for the acceptance probability, different classes of languages can be defined. In this paper, we present a number of results that clarify the power of such machines and properties of the languages they define. The broad themes we focus on are as follows. We precisely characterize the complexity of the emptiness, universality, and language containment problems for such machines, answering canonical questions central to the use of these models in formal verification. Next, we characterize the languages recognized by PBAs topologically, demonstrating that though general PBAs can recognize languages that are not regular, topologically the languages are as simple as ω-regular languages. Finally, we introduce Hierarchical PBAs, which are syntactically restricted forms of PBAs that are tractable and capture exactly the class of ω-regular languages.

1 Introduction

Automata on infinite (length) strings have played a central role in the specification, modeling and verification of non-terminating, reactive and concurrent systems [8, 10, 17, 20, 21]. However, there are classes of systems whose behavior is probabilistic in nature; the probabilistic behavior being either due to the employment of randomization in the algorithms executed by the system or due to other uncertainties in the system, such as failures, that are modeled probabilistically. While Markov Chains and Markov Decision Processes have been used to model such behavior in the formal verification community [15], both these models do not adequately capture *open*, *reactive* probabilistic systems that continuously accept inputs from an environment. The most appropriate model for such systems are probabilistic automata on infinite strings, which are the focus of study in this paper.

Probabilistic Büchi Automata (PBA) have been introduced in [3] to capture such computational devices. These automata generalize probabilistic finite automata (PFA) [12, 14, 16] from finite length inputs to infinite length inputs. Informally, PBA's are like finite-state automata except that they differ in two respects. First, from each state and on each input symbol, the PBA may roll a dice to determine the next state. Second, the notion of acceptance is different because PBAs are probabilistic in nature and have infinite length input strings. The behavior of a PBA on a given infinite input string can be captured by an infinite Markov chain that defines a probability measure on the space of runs/executions of the machine on the given input. Like Büchi automata, a run is considered to be accepting if some accepting state occurs infinitely often, and

M. Bravetti and G. Zavattaro (Eds.): CONCUR 2009, LNCS 5710, pp. 229–243, 2009.

therefore, the probability of acceptance of the input is defined to be the measure of all accepting runs on the given input. There are two possible languages that one can associate with a PBA $\mathcal{B}$ [2, 3] — $\mathcal{L}_{>0}(\mathcal{B})$ (called *probable semantics*) consisting of all strings whose probability of acceptance is non-zero, and $\mathcal{L}_{=1}(\mathcal{B})$ (called *almost sure semantics*) consisting all strings whose probability of acceptance is 1. Based on these two languages, one can define two classes of languages — $\mathbb{L}(\mathrm{PBA}^{>0})$, and $\mathbb{L}(\mathrm{PBA}^{=1})$ which are the collection of all languages (of infinite length strings) that can be accepted by some PBA with respect to probable, and almost sure semantics, respectively. In this paper we study the expressive power of, and decision problems for these classes of languages.

We present a number of new results that highlight three broad themes. First, we establish the precise complexity of the canonical decision problems in verification, namely, emptiness, universality, and language containment, for the classes $\mathbb{L}(\mathrm{PBA}^{>0})$ and $\mathbb{L}(\mathrm{PBA}^{=1})$. For the decision problems, we focus our attention on RatPBAs which are PBAs in which all transition probabilities are rational. First we show the problem of checking emptiness of the language $\mathcal{L}_{=1}(\mathcal{B})$ for a RatPBA $\mathcal{B}$ is **PSPACE**-complete, which substantially improves the result of [2] where it was shown to be decidable in **EXPTIME** and conjectured to be **EXPTIME**-hard. This upper bound is established by observing that the complement of the language $\mathcal{L}_{=1}(\mathcal{B})$ is recognized by a special PBA $\mathcal{M}$ (with probable semantics) called a *finite state probabilistic monitor* (FPM) [4, 6] and then exploiting a result in [6] that shows that the language of an FPM is non-empty if and only if there is an *ultimately periodic word* in the language. This observation of the existence of ultimately periodic words does not carry over to the class $\mathbb{L}(\mathrm{PBA}^{>0})$. However, we show that $\mathcal{L}_{>0}(\mathcal{B})$, for a RatPBA $\mathcal{B}$, is non-empty iff it contains a *strongly asymptotic word*, which is a generalization of ultimately periodic word. This allows us to show that the emptiness problem for $\mathbb{L}(\mathrm{PBA}^{>0})$, though undecidable as originally shown in [2], is Σ_2^0-complete, where Σ_2^0 is a set in the second level of the arithmetic hierarchy. Next we show that the universality problems for $\mathbb{L}(\mathrm{PBA}^{>0})$ and $\mathbb{L}(\mathrm{PBA}^{=1})$ are also Σ_2^0-complete and **PSPACE**-complete, respectively. Finally, we show that for both $\mathbb{L}(\mathrm{PBA}^{>0})$ and $\mathbb{L}(\mathrm{PBA}^{=1})$, the language containment problems are Σ_2^0-complete. This is a surprising observation — given that emptiness and universality are both in **PSPACE** for $\mathbb{L}(\mathrm{PBA}^{=1})$, one would expect language containment to be at least decidable.

The second theme brings to sharper focus the correspondence between nondeterminism and probable semantics, and between determinism and almost sure semantics, in the context of automata on infinite words. This correspondence was hinted at in [2]. There it was observed that $\mathbb{L}(\mathrm{PBA}^{=1})$ is a strict subset of $\mathbb{L}(\mathrm{PBA}^{>0})$ and that while Büchi, Rabin and Streett acceptance conditions all yield the same class of languages under the probable semantics, they yield different classes of languages under the almost sure semantics. These observations mirror the situation in non-probabilistic automata — languages recognized by deterministic Büchi automata are a strict subset of the class of languages recognized by nondeterministic Büchi automata, and while Büchi, Rabin and Streett acceptances are equivalent for nondeterministic machines, Büchi acceptance is strictly weaker than Rabin and Streett for deterministic machines. In this paper we further strenghten this correspondence through a number of results on the closure properties as well as the topological structure of $\mathbb{L}(\mathrm{PBA}^{>0})$ and $\mathbb{L}(\mathrm{PBA}^{=1})$.

First we consider closure properties. It was shown in [2] that the class $\mathbb{L}(\mathrm{PBA}^{>0})$ is closed under all the Boolean operations (like the class of languages recognized by nondeterministic Büchi automata) and that $\mathbb{L}(\mathrm{PBA}^{=1})$ is not closed under complementation. We show that $\mathbb{L}(\mathrm{PBA}^{=1})$ is, however, closed under intersection and union, just like the class of languages recognized by deterministic Büchi automata. We also show that every language in $\mathbb{L}(\mathrm{PBA}^{>0})$ is a Boolean combination of languages in $\mathbb{L}(\mathrm{PBA}^{=1})$, exactly like every ω-regular language (or languages recognized by nondeterministic Büchi machines) is a Boolean combination of languages recognized by deterministic Büchi machines. Next, we characterize the classes topologically. There is natural topological space on infinite length strings called the *Cantor topology* [18]. We show that, like ω-regular languages, all the classes of languages defined by PBAs lie in very low levels of this Borel hierarchy. We show that $\mathbb{L}(\mathrm{PBA}^{=1})$ is strictly contained in $\mathcal{G}_\delta$, just like the class of languages recognized by deterministic Büchi is strictly contained in $\mathcal{G}_\delta$. From these results, it follows that $\mathbb{L}(\mathrm{PBA}^{>0})$ is in the Boolean closure of $\mathcal{G}_\delta$ much like the case for ω-regular languages.

The last theme identifies syntactic restrictions on PBAs that capture regularity. Much like PFAs for finite word languages, PBAs, though finite state, allow one to recognize non-regular languages. It has been shown [2, 3] that both $\mathbb{L}(\mathrm{PBA}^{>0})$ and $\mathbb{L}(\mathrm{PBA}^{=1})$ contain non-ω-regular languages. A question initiated in [3] was to identify restrictions on PBAs that ensure that PBAs have the same expressive power as finite-state (non-probabilistic) machines. One such restriction was identified in [3], where it was shown that *uniform* PBAs with respect to the probable semantics capture exactly the class of ω-regular languages. However, the uniformity condition identified by Baier et. al. was semantic in nature. In this paper, we identify one simple syntactic restriction that capture regularity both for probable semantics and almost sure semantics. The restriction we consider is that of a hierarchical structure. A *Hierarchical* PBA *(*HPBA*)* is a PBA whose states are partitioned into different levels such that, from any state q, on an input symbol a, at most one transition with non-zero probability goes to a state at the same level as q and all others go to states at higher level. We show that HPBA with respect to probable semantics define exactly the class of ω-regular languages, and with respect to almost sure semantics define exactly the class of ω-regular languages in $\mathbb{L}(\mathrm{PBA}^{=1})$, namely, those recognized by deterministic Büchi automata. Next, HPBAs not only capture the notion of regularity, they are also very tractable. We show that the emptiness and universality problems for HPBA with probable semantics are **NL**-complete and **PSPACE**-complete, respectively; for almost sure semantics, emptiness is **PSPACE**-complete and universality is **NL**-complete. This is interesting because this is the exact same complexity as that for (non-probabilistic) Büchi automata. In contrast, the emptiness problem for uniform PBA has been shown to be in **EXPTIME** and co-**NP**-hard [3]; thus, they seem to be less tractable than HPBA.

The rest of the paper is organized as follows. After discussing closely related work, we start with some preliminaries (in Section 2) before introducing PBAs. We present our results about the probable semantics in Section 3, and almost sure semantics in

Section 4. Hierarchical PBAs are introduced in Section 5, and conclusions are presented in Section 6. In the interest of space, some proofs have been deferred to [5].

Related Work. Probabilistic Büchi automata (PBA), introduced in [3], generalize the model of Probabilistic Finite Automata [12, 14, 16] to consider inputs of infinite length. In [3], Baier and Größer only considered the probable semantics for PBA. They also introduced the model of uniform PBAs to capture ω-regular languages and showed that the emptiness problem for such machines is in **EXPTIME** and co-**NP**-hard. The almost sure semantics for PBA was first considered in [2] where a number of results were established. It was shown that $\mathbb{L}(\mathrm{PBA}^{>0})$ are closed under all Boolean operations, $\mathbb{L}(\mathrm{PBA}^{=1})$ is strictly contained in $\mathbb{L}(\mathrm{PBA}^{>0})$, the emptiness problem for $\mathbb{L}(\mathrm{PBA}^{>0})$ is undecidable, and the emptiness problem of $\mathbb{L}(\mathrm{PBA}^{=1})$ is in **EXPTIME**. We extend and sharpen the results of this paper. In a series of previous papers [4, 6], we considered a special class of PBAs called FPMs (Finite state Probabilistic Monitors) whose accepting set of states consists of all states excepting a rejecting state which is also absorbing. There we proved a number of results on the expressiveness and decidability/complexity of problems for FPMs. We draw on many of these observations to establish new results for the more general model of PBAs.

2 Preliminaries

We assume that the reader is familiar with arithmetical hierarchy. The set of natural numbers will be denoted by $\mathbb{N}$, the closed unit interval by $[0,1]$ and the open unit interval by $(0,1)$. The power-set of a set X will be denoted by 2^X.

Sequences. Given a finite set S, $|S|$ denotes the cardinality of S. Given a sequence (finite or infinite) $\kappa = s_0, s_1, \ldots$ over S, $|\kappa|$ will denote the length of the sequence (for infinite sequence $|\kappa|$ will be ω), and $\kappa[i]$ will denote the ith element s_i of the sequence. As usual S^* will denote the set of all finite sequences/strings/words over S, S^+ will denote the set of all finite non-empty sequences/strings/words over S and S^ω will denote the set of all infinite sequences/strings/words over S. Given $\eta \in S^*$ and $\kappa \in S^* \cup S^\omega$, $\eta\kappa$ is the sequence obtained by concatenating the two sequences in order. Given $\mathsf{L}_1 \subseteq \Sigma^*$ and $\mathsf{L}_2 \subseteq \Sigma^\omega$, the set $\mathsf{L}_1\mathsf{L}_2$ is defined to be $\{\eta\kappa \mid \eta \in \mathsf{L}_1 \text{ and } \kappa \in \mathsf{L}_2\}$. Given natural numbers $i, j \leq |\kappa|$, $\kappa[i:j]$ is the finite sequence $s_i, \ldots s_j$, where $s_k = \kappa[k]$. The set of *finite prefixes* of κ is the set $\mathit{Pref}(\kappa) = \{\kappa[0,j] \mid j \in \mathbb{N}, j \leq |\kappa|\}$.

Languages of infinite words. A language L of infinite words over a finite alphabet Σ is a subset of Σ^ω. (Please note we restrict only to finite alphabets). A set of languages of infinite words over Σ is said to be a class of languages of infinite words over Σ. Given a class $\mathcal{L}$, the Boolean closure of $\mathcal{L}$, denoted $\mathsf{BCl}(\mathcal{L})$, is the smallest class containing $\mathcal{L}$ that is closed under the Boolean operations of complementation, union and intersection.

Automata and ω-regular Languages. A *finite automaton on infinite words*, $\mathcal{A}$, over a (finite) alphabet Σ is a tuple (Q, q_0, F, Δ), where Q is a finite set of states, $\Delta \subseteq Q \times \Sigma \times Q$ is the transition relation, $q_0 \in Q$ is the initial state, and F defines the accepting condition. The nature of F depends on the type of automaton we are considering; for a *Büchi automaton* $F \subseteq Q$, while for a *Rabin automaton* F is a finite subset of $2^Q \times 2^Q$.

If for every $q \in Q$ and $a \in \Sigma$, there is exactly one q' such that $(q, a, q') \in \Delta$ then $\mathcal{A}$ is called a *deterministic* automaton. Let $\alpha = a_0, a_1, \ldots$ be an infinite string over Σ. A *run r of $\mathcal{A}$* on α is an infinite sequence $s_0, s_1, \ldots$ over Q such that $s_0 = q_0$ and for every $i \geq 0$, $(s_i, a_i, s_{i+1}) \in \Delta$. The notion of an *accepting run* depends on the type of automaton we consider. For a Büchi automaton r is accepting if some state in F appears infinitely often in r. On the other hand for a Rabin automaton, r is accepting if it satisfies the *Rabin acceptance condition*— there is some pair $(B_i, G_i) \in F$ such that all the states in B_i appear only finitely many times in r, while at least one state in G_i appears infinitely many times. The automaton $\mathcal{A}$ *accepts* the string α if it has an accepting run on α. The *language accepted (recognized) by $\mathcal{A}$*, denoted by $\mathcal{L}(\mathcal{A})$, is the set of strings that $\mathcal{A}$ accepts. A language $\mathsf{L} \subseteq \Sigma^\omega$ is called *ω-regular* iff there is some Büchi automata $\mathcal{A}$ such that $\mathcal{L}(\mathcal{A}) = \mathsf{L}$. In this paper, given a fixed alphabet Σ, we will denote the class of ω-regular languages by $\mathsf{Regular}$. It is well-known that unlike the case of finite automata on finite strings, deterministic Büchi automata are less powerful than nondeterministic Büchi automata. On the other hand, nondeterministic Rabin automata and deterministic Rabin automata have the expressive power and they recognize exactly the class $\mathsf{Regular}$. Finally, we will sometimes find it convenient to consider automata $\mathcal{A}$ that do not have finitely many states. We will say that a language L is *deterministic* iff it can be accepted by a deterministic Büchi automaton that does not necessarily have finitely many states. We denote by $\mathrm{Deterministic}$ the collection of all deterministic languages. Please note that the class $\mathrm{Deterministic}$ strictly contains the class of languages recognized by finite state deterministic Büchi automata. The following are well-known results [13, 18].

Proposition 1. $\mathsf{L} \in \mathsf{Regular} \cap \mathrm{Deterministic}$ iff there is a finite state deterministic Büchi automaton $\mathcal{A}$ such that $\mathcal{L}(\mathcal{A}) = \mathsf{L}$. Furthermore, $\mathsf{Regular} \cap \mathrm{Deterministic} \subsetneq \mathsf{Regular}$ and $\mathsf{Regular} = \mathsf{BCl}(\mathsf{Regular} \cap \mathrm{Deterministic})$.

Topology on infinite strings. The set Σ^ω comes equipped with a natural topology called the *Cantor topology*. The collection of open sets is the collection $\mathcal{G} = \{\mathsf{L}\Sigma^\omega \mid \mathsf{L} \subseteq \Sigma^+\}$[1]. The collection of closed sets, $\mathcal{F}$, is the collection of *prefix-closed sets* — L is prefix-closed if for every infinite string α, if every prefix of α is a prefix of some string in L, then α itself is in L. In the context of verification of reactive systems, closed sets are also called *safety languages* [1, 11]. One remarkable result in automata theory is that the class of languages $\mathcal{G}_\delta$ coincides exactly with the class of languages recognized by infinite-state deterministic Büchi automata [13, 18]. This combined with the fact that the class of ω-regular languages is the Boolean closure of ω-regular deterministic Büchi automata yields that the class of ω-regular languages is strictly contained in $\mathsf{BCl}(\mathcal{G}_\delta)$ which itself is strictly contained in $\mathcal{G}_{\delta\sigma} \cap \mathcal{F}_{\sigma\delta}$ [13, 18].

Proposition 2. $\mathcal{G}_\delta = \mathrm{Deterministic}$, and $\mathsf{Regular} \subsetneq \mathsf{BCl}(\mathcal{G}_\delta) \subsetneq \mathcal{G}_{\delta\sigma} \cap \mathcal{F}_{\sigma\delta}$.

2.1 Probabilistic Büchi Automata

We shall now recall the definition of probabilistic Büchi automata given in [3]. Informally, PBA's are like finite-state deterministic Büchi automata except that the transition

[1] This topology is also generated by the metric $d : \Sigma^\omega \times \Sigma^\omega \to [0, 1]$ where $d(\alpha, \beta)$ is 0 iff $\alpha = \beta$; otherwise it is $\frac{1}{2^i}$ where i is the smallest integer such that $\alpha[i] \neq \beta[i]$.

function from a state on a given input is described as a probability distribution that determines the probability of the next state. PBAs generalize the probabilistic finite automata (PFA) [12, 14, 16] on finite input strings to infinite input strings. Formally,

Definition: A *finite state probabilistic Büchi automata (*PBA*)* over a finite alphabet Σ is a tuple $\mathcal{B} = (Q, q_s, Q_f, \delta)$ where Q is a finite set of *states*, $q_s \in Q$ is the *initial state*, $Q_f \subseteq Q$ is the set of *accepting/final states*, and $\delta : Q \times \Sigma \times Q \to [0, 1]$ is the *transition relation* such that for all $q \in Q$ and $a \in \Sigma$, $\sum_{q' \in Q} \delta(q, a, q') = 1$. In addition, if $\delta(q, a, q')$ is a rational number for all $q, q' \in Q, a \in \Sigma$, then we say that $\mathcal{M}$ is a rational probabilistic Büchi automata (RatPBA).

Notation: The transition function δ of PBA $\mathcal{B}$ on input a can be seen as a square matrix δ_a of order $|Q|$ with the rows labeled by "current" state, columns labeled by "next state" and the entry $\delta_a(q, q')$ equal to $\delta(q, a, q')$. Given a word $u = a_0 a_1 \dots a_n \in \Sigma^+$, δ_u is the matrix product $\delta_{a_0}\delta_{a_1}\dots\delta_{a_n}$. For an empty word $\epsilon \in \Sigma^*$ we take δ_ϵ to be the identity matrix. Finally for any $Q_0 \subseteq Q$, we say that $\delta_u(q, Q_0) = \sum_{q' \in Q_0} \delta_u(q, q')$. Given a state $q \in Q$ and a word $u \in \Sigma^+$, $\mathsf{post}(q, u) = \{q' \mid \delta_u(q, q') > 0\}$.

Intuitively, the PBA starts in the initial state q_s and if after reading $a_0, a_1 \dots, a_n$ results in state q, then it moves to state q' with probability $\delta_{a_{i+1}}(q, q')$ on symbol a_{i+1}. Given a word $\alpha \in \Sigma^\omega$, the PBA $\mathcal{B}$ can be thought of as a infinite state Markov chain which gives rise to the standard σ-algebra defined using cylinders and the standard probability measure on Markov chains [9, 19]. We denote this measure by $\mu_{\mathcal{B},\alpha}$. A *run* of the PBA $\mathcal{B}$ is an infinite sequence $\rho \in Q^\omega$. A run ρ is *accepting* if $\rho[i] \in Q_f$ for infinitely many i. A run ρ is said to be *rejecting* if it is not accepting. The set of accepting runs and the set of rejecting runs are measurable [19]. Given a word α, the measure of the set of accepting runs is said to be the *probability of accepting* α and is henceforth denoted by $\mu^{acc}_{\mathcal{B},\alpha}$; and the measure of the set of rejecting runs is said to be the *probability of rejecting* α and is henceforth denoted by $\mu^{rej}_{\mathcal{B},\alpha}$. Clearly $\mu^{acc}_{\mathcal{B},\alpha} + \mu^{rej}_{\mathcal{B},\alpha} = 1$. Following, [2, 3], a PBA $\mathcal{B}$ on alphabet Σ defines two *semantics*:

- $\mathcal{L}_{>0}(\mathcal{B}) = \{\alpha \in \Sigma^\omega \mid \mu^{acc}_{\mathcal{B},\alpha} > 0\}$, henceforth referred to as the *probable semantics* of $\mathcal{B}$, and
- $\mathcal{L}_{=1}(\mathcal{B}) = \{\alpha \in \Sigma^\omega \mid \mu^{acc}_{\mathcal{B},\alpha} = 1\}$, henceforth referred to as the *almost-sure semantics* of $\mathcal{B}$.

This gives rise to the following classes of languages of infinite words.

Definition: Given a finite alphabet Σ, $\mathbb{L}(\text{PBA}^{>0}) = \{\mathsf{L} \subseteq \Sigma^\omega \mid \exists \text{PBA } \mathcal{B}.\ \mathsf{L} = \mathcal{L}_{>0}(\mathcal{B})\}$ and $\mathbb{L}(\text{PBA}^{=1}) = \{\mathsf{L} \subseteq \Sigma^\omega \mid \exists \text{PBA } \mathcal{B}.\ \mathsf{L} = \mathcal{L}_{=1}(\mathcal{B})\}$.

Remark: Given $x \in [0, 1]$, one can of course, also define the languages $\mathcal{L}_{>x}(\mathcal{B}) = \{\alpha \in \Sigma^\omega \mid \mu^{acc}_{\mathcal{B},\alpha} > x\}$ and $\mathcal{L}_{\geq x}(\mathcal{B}) = \{\alpha \in \Sigma^\omega \mid \mu^{acc}_{\mathcal{B},\alpha} \geq x\}$. The exact value of x is not important and thus one can also define classes $\mathbb{L}(\text{PBA}^{>\frac{1}{2}})$ and $\mathbb{L}(\text{PBA}^{\geq\frac{1}{2}})$.

Probabilistic Rabin automaton. Analogous to the definition of a PBA and RatPBA, one can define a Probabilistic Rabin automaton PRA and RatPRA [2, 7]; where instead of using a set of final states, a set of pairs of subsets of states is used. A run in that case is said to be accepting if it satisfies the Rabin acceptance condition. It is shown in [2, 7] that PRAs have the same expressive power under both probable and almost-sure semantics. Furthermore, it is shown in [2, 7] that for any PBA $\mathcal{B}$, there is PRA $\mathcal{R}$ such that a word α is accepted by $\mathcal{R}$ with probability 1 iff α is accepted by $\mathcal{B}$ with probability > 0. All other words are accepted with probability 0 by $\mathcal{R}$.

Proposition 3 ([2]). For any PBA $\mathcal{B}$ there is a PRA $\mathcal{R}$ such that $\mathcal{L}_{>0}(\mathcal{B}) = \mathcal{L}_{>0}(\mathcal{R}) = \mathcal{L}_{=1}(\mathcal{R})$ and $\mathcal{L}_{=0}(\mathcal{B}) = \mathcal{L}_{=0}(\mathcal{R})$. Furthermore, if $\mathcal{B}$ is a RatPBA then $\mathcal{R}$ can be taken to be RatPRA and the construction of $\mathcal{R}$ is recursive in this case.

Finite probabilistic monitors (FPM)s. We identify one useful syntactic restriction of PBAs, called *finite probabilistic monitors* (FPM)s. In an FPM, all the states are accepting except a special absorbing *reject* state. We studied them extensively in [4, 6].

Definition: A PBA $\mathcal{M} = (Q, q_s, Q_f, \delta)$ on Σ is said to be an FPM if there is a state $q_r \in Q$ such that $q_r \neq q_s$, $Q_f = Q \setminus \{q_r\}$ and $\delta(q_r, a, q_r) = 1$ for each $a \in \Sigma$. The state q_r said to be the *reject* state of $\mathcal{M}$. If in addition $\mathcal{M}$ is a RatPBA, we say that $\mathcal{M}$ is a rational finite probabilistic monitor (RatFPM).

3 Probable Semantics

In this section, we shall study the expressiveness of the languages contained in $\mathbb{L}(\mathrm{PBA}^{>0})$ as well as the complexity of deciding emptiness and universality of $\mathcal{L}_{>0}(\mathcal{B})$ for a given RatPBA $\mathcal{B}$. We assume that the alphabet Σ is fixed and contains at least two letters.

3.1 Expressiveness

It was already shown in [3] that the class of ω-regular languages is strictly contained in the class $\mathbb{L}(\mathrm{PBA}^{>0})$ and that $\mathbb{L}(\mathrm{PBA}^{>0})$ is closed under the Boolean operations of complementation, finite intersection and finite union. We will now show that even though the class $\mathbb{L}(\mathrm{PBA}^{>0})$ strictly contains ω-regular languages, it is not topologically harder. More precisely, we will show that for any PBA $\mathcal{B}$, $\mathcal{L}_{>0}(\mathcal{B})$ is a $\mathsf{BCl}(\mathcal{G}_\delta)$-set. The proof of this fact relies on two facts. The first is that just as the class of ω-regular languages is the Boolean closure of the class of ω-regular recognized by deterministic Büchi automata, the class $\mathbb{L}(\mathrm{PBA}^{>0})$ coincides with the Boolean closure of the class $\mathbb{L}(\mathrm{PBA}^{=1})$. This is the content of the following theorem whose proof is of independent interest and shall be used later in establishing that the containment of languages of two PBAs under almost-sure semantics is undecidable (see Theorem 4).

Theorem 1. $\mathbb{L}(\mathrm{PBA}^{>0}) = \mathsf{BCl}(\mathbb{L}(\mathrm{PBA}^{=1}))$.

Proof. First observe that it was already shown in [2] that $\mathbb{L}(\mathrm{PBA}^{=1}) \subseteq \mathbb{L}(\mathrm{PBA}^{>0})$. Since $\mathbb{L}(\mathrm{PBA}^{>0})$ is closed under Boolean operations, we get that $\mathsf{BCl}(\mathbb{L}(\mathrm{PBA}^{=1})) \subseteq \mathbb{L}(\mathrm{PBA}^{>0})$. We have to show the reverse inclusion.

It suffices to show that given a PBA $\mathcal{B}$, the language $\mathcal{L}_{>0}(\mathcal{B}) \in \mathsf{BCl}(\mathbb{L}(\mathrm{PBA}^{=1}))$. Fix $\mathcal{B}$. Recall that results of [2, 7] (see Proposition 3) imply that there is a probabilistic Rabin automaton (PRA) $\mathcal{R}$ such that 1) $\mathcal{L}_{>0}(\mathcal{B}) = \mathcal{L}_{=1}(\mathcal{R}) = \mathcal{L}_{>0}(\mathcal{R})$ and 2)

$\mathcal{L}_{=0}(\mathcal{B}) = \mathcal{L}_{=0}(\mathcal{R})$. Let $\mathcal{R} = (Q, q_s, F, \delta)$ where $F \subseteq 2^Q \times 2^Q$ is the set of the Rabin pairs. Assuming that F consists of n-pairs, let $F = ((B_1, G_1), \ldots, (B_n, G_n))$.

Given an index set $\mathcal{I} \subseteq \{1, \ldots, n\}$, let $\mathsf{Good}_{\mathcal{I}} = \cup_{r \in \mathcal{I}} G_r$. Let $\mathcal{R}_{\mathcal{I}}$ be the PBA obtained from $\mathcal{R}$ by taking the set of final states to be $\mathsf{Good}_{\mathcal{I}}$. In other words, $\mathcal{R}_{\mathcal{I}} = (Q, q_s, \mathsf{Good}_{\mathcal{I}}, \delta)$. Given $\mathcal{I} \subseteq \{1, \ldots, n\}$ and an index $j \in \mathcal{I}$, let $\mathsf{Bad}_{\mathcal{I},j} = B_j \cup \cup_{r \in \mathcal{I}, r \neq j} G_r$. Let $\mathcal{R}^j_{\mathcal{I}}$ be the PBA obtained from $\mathcal{R}$ by taking the set of final states to be $\mathsf{Bad}_{\mathcal{I},j}$, *i.e.*, $\mathcal{R}^j_{\mathcal{I}} = (Q, q_s, \mathsf{Bad}_{\mathcal{I},j}, \delta)$. The result follows from the following claim.

Claim:

$$\mathcal{L}_{>0}(\mathcal{B}) = \bigcup_{\mathcal{I} \subseteq \{1,\ldots,n\}, j \in \mathcal{I}} \mathcal{L}_{=1}(\mathcal{R}_{\mathcal{I}}) \cap (\Sigma^\omega \setminus \mathcal{L}_{=1}(\mathcal{R}^j_{\mathcal{I}})).$$

The proof of the claim is detailed in [5]. □

The second component needed for showing that $\mathbb{L}(\mathrm{PBA}^{>0}) \subseteq \mathsf{BCl}(\mathcal{G}_\delta)$ is the fact that for any PBA $\mathcal{B}$ and $x \in [0,1]$, the language $\mathcal{L}_{\geq x}(\mathcal{B})$ is a $\mathcal{G}_\delta$-set.

Lemma 1. For any PBA $\mathcal{B}$ and $x \in [0,1]$, $\mathcal{L}_{\geq x}(\mathcal{B})$ is a $\mathcal{G}_\delta$ set.

Using Lemma 1, one immediately gets that $\mathbb{L}(\mathrm{PBA}^{>0}) \subseteq \mathsf{BCl}(\mathcal{G}_\delta)$. Even though PBAs accept non-ω-regular languages, they cannot accept all the languages in $\mathsf{BCl}(\mathcal{G}_\delta)$.

Lemma 2. $\mathsf{Regular} \subsetneq \mathbb{L}(\mathrm{PBA}^{>0}) \subsetneq \mathsf{BCl}(\mathcal{G}_\delta)$.

Remark: Please note that Lemma 1 can be used to show that the classes $\mathbb{L}(\mathrm{PBA}^{>\frac{1}{2}})$ and $\mathbb{L}(\mathrm{PBA}^{\geq\frac{1}{2}})$ are also contained within the first few levels of Borel hierarchy. However, we can show that no version of Theorem 1 holds for those classes. More precisely, $\mathbb{L}(\mathrm{PBA}^{>\frac{1}{2}}) \not\subseteq \mathbb{L}(\mathrm{PBA}^{\geq\frac{1}{2}})$ and $\mathbb{L}(\mathrm{PBA}^{\geq\frac{1}{2}}) \not\subseteq \mathbb{L}(\mathrm{PBA}^{>\frac{1}{2}})$. These results are out of the scope of this paper.

3.2 Decision Problems

Given a RatPBA $\mathcal{B}$, the problems of emptiness and universality of $\mathcal{L}_{>0}(\mathcal{B})$ are known to be undecidable [2]. We sharpen this result by showing that the problem is Σ^0_2-complete. This is interesting in the light of the fact that problems on infinite string automata that are undecidable tend to typically lie in the analytical hierarchy, and not in the arithmetic hierarchy.

Before we proceed with the proof of the upper bound, let us recall an important property of finite-state Büchi automata [13, 18]. The language recognized by a finite-state Büchi automata $\mathcal{A}$ is non-empty iff there is a final state q_f of $\mathcal{A}$, and finite words u and v such that q_f is reachable from the initial state on input u, and q_f is reachable from the state q_f on input v. This is equivalent to saying that any non-empty ω-regular language contains an ultimately periodic word. We had extended this observation to FPMs in [4, 6]. In particular, we had shown that the language $\mathcal{L}_{>x}(\mathcal{M}) \neq \emptyset$ for a given $\mathcal{M}$ iff there exists a set of final states C of $\mathcal{M}$ and words u and v such that the probability of reaching C from the initial state on input u is $> x$ and for each state $q \in \mathsf{C}$ the probability of reaching C from q on input v is 1. This immediately implies that if $\mathcal{L}_{>x}(\mathcal{M})$ is non-empty then $\mathcal{L}_{>x}(\mathcal{M})$ must contain an ultimately periodic word. In contrast, this fact does not hold for non-empty languages in $\mathbb{L}(\mathrm{PBA}^{>0})$. In fact, Baier and Größer [3], construct a PBA $\mathcal{B}$ such that $\mathcal{L}_{>0}(\mathcal{B})$ does not contain any ultimately periodic word.

However, we will show that even though the probable semantics may not contain an ultimately periodic, they nevertheless are restrained in the sense that they must contain *strongly asymptotic* words. Given a PBA $\mathcal{B} = (Q, q_s, Q_f, \delta)$ and a set C of states of $\mathcal{B}$, a word $\alpha \in \Sigma^\omega$ is said to be *strongly asymptotic with respect to $\mathcal{B}$ and C* if there is an infinite sequence $i_1 < i_2 < ....$ such that 1) $\delta_{\alpha[0:i_1]}(q_s, C) > 0$ and 2) all $j > 0$, for all $q \in C$, the probability of being in C from q after passing through a final state on the finite input string $\alpha[i_j, i_{j+1}]$ is strictly greater than $1 - \frac{1}{2^j}$. A word α is said to be *strongly asymptotic with respect to $\mathcal{B}$* if there is some C such that α is strongly asymptotic with respect to $\mathcal{B}$ and C . The following notations shall be useful.

Notation: Let $\mathcal{B} = (Q, q_s, Q_f, \delta)$. Given $C \subseteq Q$, $q \in C$ and a finite word $u \in \Sigma^+$, let $\delta_u^{Q_f}(q, C)$ be the probability that the PBA $\mathcal{B}$, when started in state q, on the input string u, is in some state in C at the end of u after passing through a final state. Let $Reach(\mathcal{B}, C, x)$ denote the predicate that for some finite non-empty input string u, the probability of being in C having started from the initial state q_s is $> x$, i.e., $Reach(\mathcal{B}, C, x) = \exists u \in \Sigma^+ . \delta_u(q_s, C) > x$.

The asymptotic sequence property is an immediate consequence of the following Lemma.

Lemma 3. Let $\mathcal{B} = (Q, q_s, Q_f, \delta)$. For any $x \in [0, 1)$, $\mathcal{L}_{>x}(\mathcal{B}) \neq \emptyset$ iff $\exists C \subseteq Q$ such that $Reach(\mathcal{B}, C, x)$ is true and $\forall j > 0$ there is a finite non-empty word u_j such that $\forall q \in C.\ \delta_{u_j}^{Q_f}(q, C) > (1 - \frac{1}{2^j})$.

Proof. The ($\Leftarrow$)-direction is proved in [5]. We outline here the proof of ($\Rightarrow$)-direction. The missing parts of the proof will be cast in terms of claims which are proved in [5].

Assume that $\mathcal{L}_{>x}(\mathcal{B}) \neq \emptyset$. Fix an infinite input string $\gamma \in \mathcal{L}_{>x}(\mathcal{B})$. Recall that the probability measure generated by γ and $\mathcal{B}$ is denoted by $\mu_{\mathcal{B},\gamma}$. For the rest of this proof we will just write μ for $\mu_{\mathcal{B},\gamma}$.

We will call a non-empty set of states C *good* if there is an $\epsilon > 0$, a measurable set $\mathsf{Paths} \subseteq Q^\omega$ of runs, and an infinite sequence of natural numbers $i_1 < i_2 < i_3 < \ldots$ such that following conditions hold.

- $\mu(\mathsf{Paths}) \geq x + \epsilon$;
- For each $j > 0$ and each run ρ in Paths, we have that
  1. $\rho[0] = q_s$, $\rho[i_j] \in C$ and
  2. at least one state in the finite sequence $\rho[i_j, i_{j+1}]$ is a final state.

We say that a good set C is *minimal* if C is good but for each $q \in C$, the set $C \setminus \{q\}$ is not good. Clearly if there is a good set of states then there is also a minimal good set of states.

Claim:

- There is a good set of states C.

– Let C be a minimal good set of states. Fix ϵ, Paths and the sequence $i_1 < i_2 < \ldots$ which witness the fact that C is good set of states. For each $q \in C$ and each $j > 0$, let $\mathsf{Paths}_{j,q}$ be the subset of Paths such that each run in Paths passes through q at point i_j, *i.e.*, $\mathsf{Paths}_{j,q} = \{\rho \in \mathsf{Paths} \mid \rho[i_j] = q\}$. Then there exists a $p > 0$ such that $\mu(\mathsf{Paths}_{j,q}) \geq p$ for each $q \in C$ and each $j > 0$.

Now, fix a minimal set of good states C. Fix ϵ, Paths and the sequence $i_1 < i_2 < \ldots$ which witness the fact that C is a good set of states. We claim that C is the required set of states. As $\mu(\mathsf{Paths}) \geq x + \epsilon$ and for each $\rho \in \mathsf{Paths}$, $\rho[i_1] \in C$, it follows immediately that $Reach(\mathcal{B}, C, x)$. Assume now, by way of contradiction, that there exists a $j_0 > 0$ such that for each finite word u, there exists a $q \in C$ such that $\delta_u^{Q_f}(q, C) \leq 1 - \frac{1}{2^{j_0}}$. Fix j_0. Also fix $p > 0$ be such that $\mu(\mathsf{Paths}_{j,q}) \geq p$ for each j and $q \in C$, where $\mathsf{Paths}_{j,q}$ is the subset of Paths such that each run in $\mathsf{Paths}_{j,q}$ passes through q at point i_j; the existence of p is guaranteed by the above claim.

We first construct a sequence of sets $L_i \subseteq Q^+$ as follows. Let $L_1 \subseteq Q^+$ be the set of finite words on states of Q of length i_1+1 such that each word in L_1 starts with the state q_s and ends in a state in C. Formally $L_1 = \{\eta \subseteq Q^+ \,||\eta| = i_1+1, \eta[0] = q_s \text{ and } \eta[i_1] \in C\}$. Assume that L_r has been constructed. Let $L_{r+1} \subseteq Q^+$ be the set of finite words on states of Q of length $i_{r+1} + 1$ such that each word in L_{r+1} has a prefix in L_r, passes through a final state in between i_r and i_{r+1}, and ends in a state in C. Formally, $L_{r+1} = \{\eta \subseteq Q^+ \mid |\eta| = i_{r+1} + 1, \eta[0 : i_r] \in L_r, \exists i.(i_r < i < i_{r+1} \ \wedge \ \rho[i] \in Q_f)\}$.

Note that $L_r\Sigma^\omega$ is a decreasing sequence of measurable subsets and $\mathsf{Paths} \subseteq \cap_{r>1} L_r\Sigma^\omega$. Now, it is easy to see from the choice of j_0 and p that $\mu(L_{r+1}\Sigma^\omega) \leq \mu(L_r\Sigma^\omega) - \frac{p}{2^{j_0}}$. This, however, implies that there is a r_0 such that $\mu(L_{r_0}\Sigma^\omega) < 0$. A contradiction. □

Lemma 3 implies that checking the non-emptiness of $\mathcal{L}_{>0}(\mathcal{B})$ for a given a RatPBA $\mathcal{B}$ is in Π_2^0. We can exhibit that non-emptiness checking is Π_2^0-hard also. Since the class $\mathbb{L}(\text{PBA}^{>0})$ is closed under complementation and the complementation procedure is recursive [2] for RatPBAs, we can conclude that checking universality of $\mathcal{L}_{>0}(\mathcal{B})$ is also Σ_2^0-complete. The same bounds also apply to checking language containment under probable semantics. Note that these problems were already shown to undecidable in [2], but the exact complexity was not computed therein.

Theorem 2. Given a RatPBA, $\mathcal{B}$, the problems 1) deciding whether $\mathcal{L}_{>0}(\mathcal{B}) = \emptyset$ and 2) deciding whether $\mathcal{L}_{>0}(\mathcal{B}) = \Sigma^\omega$, are Σ_2^0-complete. Given another RatPBA, $\mathcal{B}'$, the problem of deciding whether $\mathcal{L}_{>0}(\mathcal{B}) \subseteq \mathcal{L}_{>0}(\mathcal{B}')$ is also Σ_2^0-complete.

Remark: Lemma 3 can be used to show that emptiness-checking of $\mathcal{L}_{>\frac{1}{2}}(\mathcal{B})$ for a given RatPBA $\mathcal{B}$ is in Σ_2^0. In contrast, we had shown in [6] that the problem of deciding whether $\mathcal{L}_{>\frac{1}{2}}(\mathcal{M}) = \Sigma^\omega$ for a given FPM $\mathcal{M}$ lies beyond the arithmetical hierarchy.

4 Almost-Sure Semantics

The class $\mathbb{L}(\text{PBA}^{=1})$ was first studied in [2], although they were not characterized topologically. In this section, we study the expressiveness and complexity of the class

$\mathbb{L}(\mathrm{PBA}^{=1})$. We will also demonstrate that the class $\mathbb{L}(\mathrm{PBA}^{=1})$ is closed under finite unions and intersections. As in the case of probable semantics, we assume that the alphabet Σ is fixed and contains at least two letters.

4.1 Expressiveness

Lemma 1 already implies that topologically, the class $\mathbb{L}(\mathrm{PBA}^{=1}) \subseteq \mathcal{G}_\delta$. Recall that $\mathcal{G}_\delta$ coincides exactly with the class of languages recognizable with infinite-state deterministic Büchi automata (see Section 2). Thanks to Theorem 1 and Lemma 2, it also follows immediately that the inclusion $\mathbb{L}(\mathrm{PBA}^{=1}) \subseteq \mathcal{G}_\delta$ is strict (otherwise we will have $\mathbb{L}(\mathrm{PBA}^{>0}) = \mathsf{BCl}(\mathbb{L}(\mathrm{PBA}^{=1})) = \mathsf{BCl}(\mathcal{G}_\delta))$. The fact that every language $\mathbb{L}(\mathrm{PBA}^{=1})$ is contained in $\mathcal{G}_\delta$ implies immediately that there are ω-regular languages not in $\mathbb{L}(\mathrm{PBA}^{=1})$. That there are ω-regular languages not in $\mathbb{L}(\mathrm{PBA}^{=1})$ was also proved in [2], although the proof therein is by explicit construction of an ω-regular language which is then shown to be not in $\mathbb{L}(\mathrm{PBA}^{=1})$. Our topological characterization of the class $\mathbb{L}(\mathrm{PBA}^{=1})$ has the advantage that we can characterize the intersection $\mathsf{Regular} \cap \mathbb{L}(\mathrm{PBA}^{=1})$ exactly: $\mathsf{Regular} \cap \mathbb{L}(\mathrm{PBA}^{=1})$ is the class of ω-regular languages that can be recognized by a finite-state deterministic Büchi automaton.

Proposition 4. For any PBA $\mathcal{B}$, $\mathcal{L}_{=1}(\mathcal{B})$ is a $\mathcal{G}_\delta$ set. Furthermore, $\mathsf{Regular} \cap \mathbb{L}(\mathrm{PBA}^{=1}) = \mathsf{Regular} \cap \mathrm{Deterministic}$ and $\mathsf{Regular} \cap \mathrm{Deterministic} \subsetneq \mathbb{L}(\mathrm{PBA}^{=1}) \subsetneq \mathcal{G}_\delta = \mathrm{Deterministic}$.

An immediate consequence of the characterization of the intersection $\mathsf{Regular} \cap \mathrm{Deterministic}$ is that the class $\mathbb{L}(\mathrm{PBA}^{=1})$ is not closed under complementation as the class of ω-regular languages recognized by deterministic Büchi automata is not closed under complementation. That the class $\mathbb{L}(\mathrm{PBA}^{=1})$ is not closed under complementation is also observed in [2], and is proved by constructing an explicit example. However, even though the class $\mathbb{L}(\mathrm{PBA}^{=1})$ is not closed under complementation, we have a "partial" complementation operation— for any PBA $\mathcal{B}$ there is another PBA $\mathcal{B}'$ such that $\mathcal{L}_{>0}(\mathcal{B}')$ is the complement of $\mathcal{L}_{=1}(\mathcal{B})$. This also follows from the results of [2] as they showed that $\mathbb{L}(\mathrm{PBA}^{=1}) \subseteq \mathbb{L}(\mathrm{PBA}^{>0})$ and $\mathbb{L}(\mathrm{PBA}^{>0})$ is closed under complementation. However our construction has two advantages: 1) it is much simpler than the one obtained by the constructions in [2], and 2) the PBA $\mathcal{B}'$ belongs to the restricted class of finite probabilistic monitors FPMs (see Section 2 for definition of FPMs). This construction plays a critical role in our complexity analysis of decision problems.

Lemma 4. For any PBA $\mathcal{B}$, there is an FPM $\mathcal{M}$ such that $\mathcal{L}_{=1}(\mathcal{B}) = \Sigma^\omega \setminus \mathcal{L}_{>0}(\mathcal{M})$.

Proof. Let $\mathcal{B} = (Q, q_s, Q_f, \delta)$. We construct $\mathcal{M}$ as follows. First we pick a new state q_r, which will be the reject state of the FPM $\mathcal{M}$. The set of states of $\mathcal{M}$ would be $Q \cup \{q_r\}$. The initial state of $\mathcal{M}$ will be q_s, the initial state of $\mathcal{B}$. The set of final states of $\mathcal{M}$ will be Q, the set of states of $\mathcal{B}$. The transition relation of $\mathcal{M}$ would be defined as follows. If q is not a final state of $\mathcal{B}$ then the transition function would be the same as for $\mathcal{B}$. If q is an final state of $\mathcal{B}$ then $\mathcal{M}$ will transit to the reject state with probability $\frac{1}{2}$ and with probability $\frac{1}{2}$ continue as in $\mathcal{B}$. Formally, $\mathcal{M} = (Q \cup \{q_r\}, q_s, Q, \delta_\mathcal{M})$ where $\delta_\mathcal{M}$ is defined as follows. For each $a \in \Sigma$, $q, q' \in Q$,

- $\delta_{\mathcal{M}}(q,a,q_r) = \frac{1}{2}$ and $\delta_{\mathcal{M}}(q,a,q') = \frac{1}{2}\delta(q,a,q')$ if $q \in Q_f$,
- $\delta_{\mathcal{M}}(q,a,q_r) = 0$ and $\delta_{\mathcal{M}}(q,a,q') = \delta(q,a,q')$ if $q \in Q \setminus Q_f$,
- $\delta_{\mathcal{M}}(q_r,a,q_r) = 1$.

It is easy to see that a word $\alpha \in \Sigma^\omega$ is rejected with probability 1 by $\mathcal{M}$ iff it is accepted with probability 1 by $\mathcal{B}$. The result now follows. □

The "partial" complementation operation has many consequences. One consequence is that the class $\mathbb{L}(\mathrm{PBA}^{=1})$ is closed under union. The class $\mathbb{L}(\mathrm{PBA}^{=1})$ is easily shown to be closed under intersection. Hence for closure properties, $\mathbb{L}(\mathrm{PBA}^{=1})$ behave like deterministic Büchi automata. Please note that closure properties were not studied in [2].

Corollary 1. The class $\mathbb{L}(\mathrm{PBA}^{=1})$ is closed under finite union and finite intersection.

4.2 Decision Problems

The problem of checking whether $\mathcal{L}_{=1}(\mathcal{B}) = \emptyset$ for a given RatPBA $\mathcal{B}$ was shown to be decidable in **EXPTIME** in [2], where it was also conjectured to be **EXPTIME**-complete. The decidability of the universality problem was left open in [2]. We can leverage our "partial" complementation operation to show that a) the emptiness problem is in fact **PSPACE**-complete, thus tightening the bound in [2] and b) the universality problem is also **PSPACE**-complete.

Theorem 3. Given a RatPBA $\mathcal{B}$, the problem of deciding whether $\mathcal{L}_{=1}(\mathcal{B}) = \emptyset$ is **PSPACE**-complete. The problem of deciding whether $\mathcal{L}_{=1}(\mathcal{B}) = \Sigma^\omega$ is also **PSPACE**-complete.

Proof. (**Upper bounds**.) We first show the upper bounds. The proof of Lemma 4 shows that for any RatPBA $\mathcal{B}$, there is a RatFPM $\mathcal{M}$ constructed in polynomial time such that $\mathcal{L}_{=1}(\mathcal{B}) = \Sigma^\omega \setminus \mathcal{L}_{>0}(\mathcal{M})$. $\mathcal{L}_{=1}(\mathcal{B})$ is empty (universal) iff $\mathcal{L}_{>0}(\mathcal{M})$ is universal (empty respectively). Now, we had shown in [4, 6] that given a RatFPM $\mathcal{M}$, the problems of checking emptiness and universality of $\mathcal{L}_{>0}(\mathcal{M})$ are in **PSPACE**, thus giving us the desired upper bounds.
(**Lower bounds**.) We had shown in [4, 6] that given a RatFPM $\mathcal{M}$, the problems of deciding the emptiness and universality of $\mathcal{L}_{>0}(\mathcal{M})$ are **PSPACE**-hard respectively. Given a RatFPM $\mathcal{M} = (Q, q_s, Q_0, \delta)$ with q_r as the absorbing reject state, consider the PBA $\overline{\mathcal{M}} = (Q, q_s, \{q_r\}, \delta)$ obtained by considering the unique reject state of $\mathcal{M}$ as the only final state of $\overline{\mathcal{M}}$. Clearly we have that $\mathcal{L}_{>0}(\mathcal{M}) = \Sigma^\omega \setminus \mathcal{L}_{=1}(\overline{\mathcal{M}})$. Thus $\mathcal{L}_{>0}(\mathcal{M})$ is empty (universal) iff $\mathcal{L}_{=1}(\overline{\mathcal{M}})$ is universal (empty respectively). The result now follows. □

Even though the problems of checking emptiness and universality of almost-sure semantics of a RatPBA are decidable, the problem of deciding language containment under almost-sure semantics turns out to be undecidable, and is indeed as hard as the problem of deciding language containment under probable semantics (or, equivalently, checking emptiness under probable semantics).

Theorem 4. Given RatPBAs, $\mathcal{B}_1$ and $\mathcal{B}_2$, the problem of deciding whether $\mathcal{L}_{=1}(\mathcal{B}_1) \subseteq \mathcal{L}_{=1}(\mathcal{B}_2)$ is Σ^0_2-complete.

5 Hierarchical PBAs

We shall now identify a simple syntactic restriction on PBAs which under probable semantics coincide exactly with ω-regular languages and under almost-sure semantics coincide exactly with ω-regular deterministic languages. These restricted PBAs shall be called *hierarchical* PBAs.

Intuitively, a hierarchical PBA is a PBA such that the set of its states can be stratified into (totally) ordered levels. From a state q, for each letter a, the machine can transition with non-zero probability to at most one state in the same level as q, and all other probabilistic transitions go to states that belong to a higher level. Formally,

Definition: Given a natural number k, a PBA $\mathcal{B} = (Q, q_s, Q, \delta)$ over an alphabet Σ is said to be a *k-level hierarchical PBA (k-*PBA) if there is a function $\mathsf{rk} : Q \rightarrow \{0, 1, \ldots, k\}$ such that the following holds.

> Given $j \in \{0, 1, \ldots, k\}$, let $Q_j = \{q \in Q \mid \mathsf{rk}(Q) = j\}$. For every $q \in Q$ and $a \in \Sigma$, if $j_0 = \mathsf{rk}(q)$ then $\mathsf{post}(q, a) \subseteq \cup_{j_0 \leq \ell \leq k} Q_\ell$ and $|\mathsf{post}(q, a) \cap Q_{j_0}| \leq 1$.

The function rk is said to be a *compatible ranking function* of $\mathcal{B}$ and for $q \in Q$ the natural number $\mathsf{rk}(q)$ is said to be the *rank* or *level* of q. $\mathcal{B}$ is said to be a *hierarchical* PBA *(*HPBA*)* if $\mathcal{B}$ is k-hierarchical for some k. If $\mathcal{B}$ is also a RatPBA, we say that $\mathcal{B}$ is a rational hierarchical PBA (RatHPBA).

We can define classes analogous to $\mathbb{L}(\mathrm{PBA}^{>0})$ and $\mathbb{L}(\mathrm{PBA}^{=1})$; and we shall call them $\mathbb{L}(\mathrm{HPBA}^{>0})$ and $\mathbb{L}(\mathrm{HPBA}^{=1})$ respectively.

Before we proceed to discuss the probable and almost-sure semantics for HPBAs, we point out two interesting facts about hierarchical HPBAs. First is that for the class of ω-regular deterministic languages, HPBAs like non-deterministic Büchi automata can be exponentially more succinct.

Lemma 5. Let $\Sigma = \{a, b, c\}$. For each $n \in \mathbb{N}$, there is a ω-regular deterministic property $\mathsf{L}_n \subseteq \Sigma^\omega$ such that i) any deterministic Büchi automata for L_n has at least $O(2^n)$ number of states, and ii) there are HPBAs $\mathcal{B}_n$ s.t. $\mathcal{B}_n$ has $O(n)$ number of states and $\mathsf{L}_n = \mathcal{L}_{=1}(\mathcal{B}_n)$.

The second thing is that even though HPBAs yield only ω-regular languages under both almost-sure semantics and probable semantics, we can recognize non-ω-regular languages with cut-points.

Lemma 6. There is a HPBA $\mathcal{B}$ such that both $\mathcal{L}_{\geq \frac{1}{2}}(\mathcal{B})$ and $\mathcal{L}_{> \frac{1}{2}}(\mathcal{B})$ are not ω-regular.

Remark: We will see shortly that the problems of deciding emptiness and universality for a HPBA turn out to be decidable under both probable and almost-sure semantics. However, with cut-points, they turn out to be undecidable. The latter, however, is out of scope of this paper.

5.1 Probable Semantics

We shall now show that the class $\mathbb{L}(\mathrm{HPBA}^{>0})$ coincides with the class of ω-regular languages under probable semantics. In [3], a restricted class of PBAs called uniform

PBAs was identified that also accept exactly the class of ω-regular languages. We make a couple of observations, contrasting our results here with theirs. First the definition of uniform PBA was semantic (i.e., the condition depends on the acceptance probability of infinitely many strings from different states of the automaton), whereas HPBA are a syntactic restriction on PBA. Second, we note that the definitions themselves are incomparable in some sense; in other words, there are HPBAs which are not uniform, and vice versa. Finally, HPBAs appear to be more tractable than uniform PBAs. We show that the emptiness problem for $\mathbb{L}(\mathrm{HPBA}^{>0})$ is **NL**-complete. In contrast, the same problem was demonstrated to be in **EXPTIME** and co-**NP**-hard [3].

Our main observation is the Hierarchical PBAs capture exactly the class of ω-regular languages.

Theorem 5. $\mathbb{L}(\mathrm{HPBA}^{>0}) = \mathsf{Regular}$.

We will show that the problem of deciding whether $\mathcal{L}_{>0}(\mathcal{B})$ is empty for hierarchical RatPBA's is **NL**-complete while the problem of deciding whether $\mathcal{L}_{>0}(\mathcal{B})$ is universal is **PSPACE**-complete. Thus "algorithmically", hierarchical PBAs are much "simpler" than both PBAs and uniform PBAs. Note that the emptiness and universality problem for finite state Büchi-automata are also **NL**-complete and **PSPACE**-complete respectively.

Theorem 6. Given a RatHPBA, $\mathcal{B}$, the problem of deciding whether $\mathcal{L}_{>0}(\mathcal{B}) = \emptyset$ is **NL**-complete. The problem of deciding whether $\mathcal{L}_{>0}(\mathcal{B}) = \Sigma^\omega$ is **PSPACE**-complete.

5.2 Almost-Sure Semantics

For a hierarchical PBA, the "partial" complementation operation for almost-sure semantics discussed in Section 4 yields a hierarchical PBA. Therefore using Theorem 5, we immediately get that a language $\mathcal{L} \in \mathbb{L}(\mathrm{HPBA}^{=1})$ is ω-regular. Thanks to the topological characterization of $\mathbb{L}(\mathrm{HPBA}^{=1})$ as a sub-collection of deterministic languages, we get that $\mathbb{L}(\mathrm{HPBA}^{=1})$ is exactly the class of languages recognized by deterministic finite-state Büchi automata.

Theorem 7. $\mathbb{L}(\mathrm{HPBA}^{=1}) = \mathsf{Regular} \cap \mathrm{Deterministic}$.

The "partial" complementation operation also yields the complexity of emptiness and universality problems.

Theorem 8. Given a RatHPBA, $\mathcal{B}$, the problem of deciding whether $\mathcal{L}_{=1}(\mathcal{B}) = \emptyset$ is **PSPACE**-complete. The problem of deciding whether $\mathcal{L}_{=1}(\mathcal{B}) = \Sigma^\omega$ is **NL**-complete.

6 Conclusions

In this paper, we investigated the power of randomization in finite state automata on infinite strings. We presented a number of results on the expressiveness and decidability problems under different notions of acceptance based on the probability of acceptance. In the case of decidability, we gave tight bounds for both the universality and emptiness problems. As part of future work, it will be interesting to investigate the power of randomization in other models of computations on infinite strings such as pushdown automata etc. Since the universality and emptiness problems are PSPACE-complete for almost-sure semantics, their application to practical systems needs further enquiry.

Acknowledgements

Rohit Chadha was supported in part by NSF grants CCF04-29639 and NSF CCF04-48178. A. Prasad Sistla was supported in part by NSF CCF-0742686. Mahesh Viswanathan was supported in part by NSF CCF04-48178 and NSF CCF05-09321.

References

1. Alpern, B., Schneider, F.: Defining liveness. Information Processing Letters 21, 181–185 (1985)
2. Baier, C., Bertrand, N., Größer, M.: On decision problems for probabilistic büchi automata. In: Amadio, R.M. (ed.) FOSSACS 2008. LNCS, vol. 4962, pp. 287–301. Springer, Heidelberg (2008)
3. Baier, C., Größer, M.: Recognizing ω-regular languages with probabilistic automata. In: Proceedings of the IEEE Symposium on Logic in Computer Science, pp. 137–146 (2005)
4. Chadha, R., Sistla, A.P., Viswanathan, M.: On the expressiveness and complexity of randomization in finite state monitors. In: LICS 2008- 23rd Annual IEEE Symposium on Logic in Computer Science, pp. 18–29. IEEE Computer Society, Los Alamitos (2008)
5. Chadha, R., Sistla, A.P., Viswanathan, M.: On the expressiveness and complexity of randomization in finite state monitors. Journal of the ACM (to appear), `http://www.cs.uiuc.edu/homes/rch/TechReports/MonitoringTechReport.pdf`
6. Condon, A., Lipton, R.J.: On the complexity of space bounded interactive proofs (extended abstract). In: 30th Annual Symposium on Foundations of Computer Science, pp. 462–467 (1989)
7. Groesser, M.: Reduction Methods for Probabilistic Model Checking. PhD thesis, TU Dresden (2008)
8. Holzmann, G.J., Peled, D.: The state of spin. In: Alur, R., Henzinger, T.A. (eds.) CAV 1996. LNCS, vol. 1102. Springer, Heidelberg (1996)
9. Kemeny, J., Snell, J.: Denumerable Markov Chains. Springer, Heidelberg (1976)
10. Kurshan, R.P.: Computer Aided Verification of the Coordinated Processes: The Automata Theoretic Approach. Princeton University Press, Princeton (1994)
11. Lamport, L.: Logical foundation, distributed systems- methods and tools for specification, vol. 190. Springer, Heidelberg (1985)
12. Paz, A.: Introduction to Probabilistic Automata. Academic Press, London (1971)
13. Perrin, D., Pin, J.-E.: Infinite Words: Automata, Semigroups, Logic and Games. Elsevier, Amsterdam (2004)
14. Rabin, M.O.: Probabilitic automata. Information and Control 6(3), 230–245 (1963)
15. Rutten, J.M., Kwiatkowska, M., Norman, G., Parker, D.: Mathematical Techniques for Analyzing Concurrent and Probabilistic Systems. AMS (2004)
16. Salomaa, A.: Formal Languages. Academic Press, London (1973)
17. Sistla, A.P.: Theoretical Issues in the Design and Verification of Distributed Systems. PhD thesis, Harvard University (1983)
18. Thomas, W.: Automata on infinite objects. In: Handbook of Theoretical Computer Science, vol. B, pp. 133–192 (1990)
19. Vardi, M.: Automatic verification of probabilistic concurrent systems. In: 26th annual Symposium on Foundations of Computer Science, pp. 327–338. IEEE Computer Society Press, Los Alamitos (1985)
20. Vardi, M., Wolper, P.: An automata theoretic approach to automatic program verification. In: Proceedings of the first IEEE Symposium on Logic in Computer Science (1986)
21. Vardi, M., Wolper, P., Sistla, A.P.: Reasoning about infinite computation paths. In: Proceedings of the 24th IEEE Symposium on Foundations of Computer Science (1983)