Probabilistic Automata with Isolated Cut-points

Rohit Chadha[1], A. Prasad Sistla[2], and Mahesh Viswanathan*[3]

[1] University of Missouri
[2] University of Illinois, Chicago
[3] University of Illinois, Urbana-Champaign

Abstract. We consider various decision problems for probabilistic finite automata (PFA)s with isolated cut-points. Recall that a cut-point x is said to be isolated for a PFA if the acceptance probability of all finite strings is bounded away from x. First we establish the exact level of undecidability of the problem of determining if a cut-point is isolated; we show this problem to be $\mathbf{\Sigma^0_2}$-complete. Next we introduce a new class of PFAs called eventually weakly ergodic PFAs that generalize ergodic and weakly ergodic PFAs. We show that the emptiness and universality problem for these PFAs is decidable provided the cut-point is isolated.

1 Introduction

A probabilistic finite automaton (PFA) [21, 20] is like a deterministic finite automaton except that after reading an input symbol the automaton rolls a dice to determine the next state. Thus the transition function of a PFA associates a probability distribution on next states with each state and input symbol. Given an acceptance threshold or cut-point x and an initial distribution μ, the language recognized by a PFA $\mathcal{B}$ (denoted as $\mathsf{L}_{>x}(\mathcal{B}, \mu)$) is the collection of all finite words u that reach a final state with probability $> x$ when $\mathcal{B}$ is started with initial distribution μ. Surprisingly, even though PFAs have only finitely many states, they are known to recognize non-regular languages [21].

One semantic restriction that has been extensively studied is that of cut-points being isolated [21, 4, 5, 15] — a cut-point x is isolated for PFA $\mathcal{B}$ and initial distribution μ, if there is an $\epsilon > 0$ such that any input word u is either accepted with probability at most $x - \epsilon$ or with probability at least $x + \epsilon$. Thus, the acceptance probability of any input word is bounded away from the cut-point x. Isolated cut-points are important because algorithms described by PFAs are useful mainly when there is a separation between the probability of accepting the good inputs from the probability of accepting the bad inputs. Isolation allows one to use standard algorithmic techniques like amplification by running multiple copies of the algorithm to drive down the probability of error. PFAs with isolated cut-points are the constant space analogues of probabilistic polynomial time complexity classes like **BPP** and **RP**.

In this paper, we consider various decision problems for PFAs with isolated cut-points. The first problem we consider is that of determining if a cut-point x is isolated for a PFA $\mathcal{B}$ and initial distribution μ. The problem was shown to be undecidable (in fact, **r.e.**-hard) by Bertoni [4, 5] when $x \in (0, 1)$. Recently, the problem was shown to

* Mahesh Viswanathan was supported by NSF CCF 1016989 and NSF CNS 1016791.

be undecidable (in fact, **co-r.e.**-hard) even when x is either 0 or 1 [15]. Determining the exact level of undecidability was posed as a problem in Bertoni's original paper and has remained open until now. We show that this problem is $\mathbf{\Sigma_2^0}$-complete (Theorem 1).

Next, we consider the emptiness and universality problems for PFAs with isolated cut-points. The decidability of these problems is still open. We conjecture that these problems are undecidable. Our belief in this result stems from the undecidability of the problem of determining if a cut-point is isolated. Also, Condon-Lipton's [13] undecidability proof for the emptiness problem for PFAs (without the cut-point being necessarily isolated) can be modified to establish the undecidability of the emptiness problem for PFAs with semi-isolated cut-points — x is a semi-isolated cut-point for PFA $\mathcal{B}$ and initial distribution μ if there is an $\epsilon > 0$ such that every input is either accepted with probability at most x or with probability at least $x + \epsilon$. Thus, if x is semi-isolated then $x + \epsilon/2$ is an isolated cut-point (for some unknown ϵ) .

Given our belief in the undecidability of the emptiness and universality problems for general PFAs with isolated cut-points, we consider restricted classes of PFAs. Ergodicity and weak ergodicity have played an important role in the study of Markov Chains and non-homogeneous Markov Chains, and have been considered in the context of PFAs in the past [22, 19, 17, 7]. Recall that a Markov Chain is ergodic if its transition graph forms an aperiodic, strongly connected component. Weak ergodicity for non-homogeneous Markov Chains means that any sequence of input symbols has only one terminal strongly connected component and this component is aperiodic. In this paper we generalize both ergodic and weakly ergodic PFAs to define a new class that we call *eventually weakly ergodic* (see Definition 4). Informally, these are PFAs such that the states can be partitioned into sets $Q_T, Q_1, \ldots Q_r$ and there is an ℓ such that in the transition graph on any word of length ℓ, $Q_1, \ldots Q_r$ are the terminal strongly connected components, and these are aperiodic. Any state in Q_T has a non-zero probability of reaching some state in $\cup_i Q_i$ on any word of length ℓ. Note that any Markov chain is eventually weakly ergodic (see Proposition 1).

There are several natural classes of systems that can be modeled as eventually weakly ergodic PFAs. One such class of protocols is randomized leadership election protocols in which a leader is elected amongst a set of "equally" likely candidates. Such a protocol usually proceeds in rounds until a leader is elected. Once a leader is elected the protocol stops and each "elected" choice forms a closed communicating class. Furthermore, leadership election protocols normally ensure that there is a constant number k, such that in every k rounds the probability that a leader is elected is > 0.

Another class of systems relates to "Dolev-Yao" modeling of probabilistic security protocols such as probabilistic anonymity protocols. In this setting, protocol participants are modeled as processes that can send and receive messages, and the communication is mediated through an attacker than can intercept messages, inject and modify messages. The attacker keeps track of the messages exchanged and *nondeterministically* chooses to send new messages to protocol participants. An "attack" is a particular resolution of the nondeterministic choices of the attacker, and protocol analysis checks for security under *every* possible attack. For a faithful analysis [14, 12, 6, 16, 11, 10], we have to consider *view-consistent* [10] attack strategies in which, at any instance, the attacker must do the same actions in all computations in which its view is the same

upto that instance. Under suitable bounds (memory of the attacker, message size and number of sessions), we can model the resulting system as a PFA: an input letter being a "view-consistent" function from the (bounded) view of the attacker to the set of its possible choices. The resulting PFA is also likely to satisfy eventual weak ergodicity because there will often be a constant number k such that each session finishes within k-steps with probability > 0.

We establish the following two results for eventually weakly ergodic PFAs: (a) the problem of determining if x is isolated is **r.e.**-complete (as opposed to $\mathbf{\Sigma_2^0}$-complete for general PFAs) (Theorem 3), and (b) $\mathsf{L}_{>x}(\mathcal{B}, \mu)$ is regular and can be computed, if x is isolated (Theorem 2). The second observation allows us to conclude that the emptiness and universality problems for eventually weakly ergodic PFAs with isolated cut-points is decidable. These results are useful when we know that x is an isolated cut point and we want to know whether at least one string is accepted with probability $> x$. Note that if the cut-point is not isolated, the emptiness and universality problems for such special PFAs is undecidable (Proposition 3).

Related work. As already mentioned above, the problem of checking emptiness of PFAs is undecidable [13]. For weakly ergodic PFAs, [7] shows that the problem of checking emptiness/universality is decidable under the assumption that the cut-point is isolated. The class of eventually weakly ergodic PFAs is a strict superset of weakly ergodic PFAs (See Example 2). Hence that result does not apply to our setting. Furthermore, the proof of that result relies on the existence of a unique compact non-empty set of distributions W which is invariant on the set of inputs (that is $W = \{\mu\delta_a \mid \mu \in W, \delta_a$ is the transition matrix on the input $a\}$). Eventually weakly ergodic matrices do not enjoy these properties and we have to appeal to different proof methods.

A decidability result under the assumption of isolation is also obtained in [18]. However, our results are incomparable to the results in [18]. They consider contracting PFAs which are different from eventually weakly ergodic matrices (see Remark 1 on Page 7). Furthermore, they only consider emptiness/universality problem relative to restricted sets of inputs (and not over the whole language).

2 Preliminaries

We assume that the reader is familiar with regular languages and basic measure theory. We will also assume that the reader is familiar with the basic theory of Markov Chains. The set of natural numbers will be denoted by $\mathbb{N}$. The powerset of any set A will be denoted by 2^A. Given any set Σ, Σ^*(Σ^+ respectively) will denote the set of finite words (nonempty finite words respectively) over Σ. A set $\mathsf{L} \subseteq \Sigma^*$ is said to be a language over Σ. Given $\rho \in \Sigma^*$, $|\rho|$ will denote the length of ρ. Given $\ell \in \mathbb{N}$, Σ^ℓ will denote the set $\{u \in \Sigma^* \mid |u| = \ell\}$ and $\Sigma^{<\ell}$ will denote the set $\{u \in \Sigma^* \mid |u| < \ell\}$.

2.1 Arithmetical Hierarchy

Let Δ be a finite alphabet. A language L over Δ is a set of finite strings over Δ. Arithmetical hierarchy consists of classes of languages $\mathbf{\Sigma}_n^0$, $\mathbf{\Pi}_n^0$ for each integer $n > 0$. Fix

an $n > 0$. A language $L \in \mathbf{\Sigma}^0_n$ iff there exists a recursive predicate $\phi(u, \boldsymbol{x}_1, ..., \boldsymbol{x}_n)$ where u is a variable ranging over Δ^*, and for each $i, 0 < i \leq n$, $\boldsymbol{x}_i$ is a finite sequence of variables ranging over integers such that

$$L \;=\; \{u \in \Delta^* \mid \exists \boldsymbol{x}_1, \forall \boldsymbol{x}_2, \ldots, Q_n \boldsymbol{x}_n \; \phi(u, \boldsymbol{x}_1, ..., \boldsymbol{x}_n)\}$$

where Q_n is an existential quantifier if n is odd, else it is a universal quantifier. Note that the quantifiers in the above equation are alternating starting with an existential quantifier. The class $\mathbf{\Pi}^0_n$ is exactly the class of languages that are complements of languages in $\mathbf{\Sigma}^0_n$. $\mathbf{\Sigma}^0_1$, $\mathbf{\Pi}^0_1$ are exactly the class of **R.E.**-sets and **co-R.E.**-sets. Let $\mathcal{C}$ be a class in the arithmetic hierarchy. $L \in \mathcal{C}$ is said to be $\mathcal{C}$-complete if for every $L' \in \mathcal{C}$ there is a computable function f such that $x \in L'$ iff $f(x) \in L$. A well known $\mathbf{\Sigma}^0_2$-complete language is the set of deterministic Turing machine encodings that halt on finitely many inputs.

2.2 Distributions and Stochastic Matrices

Distributions. A *probability distribution* over a *finite* set Q is a map $\mu : Q \to [0, 1]$ s.t. $\sum_{q \in Q} \mu(q) = 1$. For $Q' \subseteq Q$, we shall write $\mu(Q')$ for $\sum_{q \in Q'} \mu(q)$. $\mathrm{Dist}(Q)$ will denote the set of all distributions over Q. The map $d : \mathrm{Dist}(Q) \times \mathrm{Dist}(Q) \to [0, 1]$ defined as

$$d(\mu, \nu) = \frac{\sum_{q \in Q} |\mu(q) - \nu(q)|}{2} = \max_{Q' \subseteq Q} |\mu(Q') - \nu(Q')|$$

defines a metric on the set $\mathrm{Dist}(Q)$. Note that $d(\mu, \nu) \leq 1$. Unless otherwise stated, we assume that $\mu(q)$ is a rational number.

Stochastic Matrices. A stochastic matrix over a *finite* set Q is a matrix $\delta : Q \times Q \to [0, 1]$ s.t. $\forall q \in Q$. $\sum_{q' \in Q} \delta(q, q') = 1$. $\mathsf{Mat}_{=1}(Q)$ will denote the set of all stochastic matrices over the set Q. For $\delta \in \mathsf{Mat}_{=1}(Q)$ and $\mu \in \mathrm{Dist}(Q)$, $\mu\delta$ denotes the distribution, given by $\mu\delta(q) = \sum_{q' \in Q} \mu(q')\delta(q', q)$. Unless otherwise stated, we assume that $\delta(q, q')$ is a rational number. Given $\delta_1, \delta_2 \in \mathsf{Mat}_{=1}(Q)$, we write $\delta_1\delta_2$ to denote the matrix product of δ_1 and δ_2 and we write $(\delta_1)^\ell$ to denote the ℓ-times product of δ_1.

Given a state $q \in Q$, and a matrix $\delta \in \mathsf{Mat}_{=1}(Q)$, we write $\mathsf{post}(q, \delta) = \{q' \mid \delta(q, q') > 0\}$. Given $Q' \subseteq Q$, we write $\mathsf{post}(Q', \delta) = \cup_{q \in Q'} \mathsf{post}(q, \delta)$. $Q' \subseteq Q$ is said to be *closed for* δ if $\mathsf{post}(Q', \delta) \subseteq Q'$. It is easy to see that if Q' is closed for δ, then the matrix $\delta|_{Q'}$ obtained by restricting δ to $Q' \times Q'$ is a stochastic matrix over Q'. Given $\Delta \subseteq \mathsf{Mat}_{=1}(Q)$, $Q' \subseteq Q$ is said to be *closed for* Δ if Q' is closed for each $\delta \in \Delta$. If Q' is closed for Δ, we let $\Delta|_{Q'} = \{\delta|_{Q'} \mid \delta \in \Delta\}$.

$\delta \in \mathsf{Mat}_{=1}(Q)$ is said to be *irreducible* if for each $q, q' \in Q$, there is an $\ell > 0$ s.t. $\delta^\ell(q, q') > 0$. The *period* of q, written $period_\delta(q)$, is defined to be the the greatest common divisor of $\{j \mid \delta^j(q, q) > 0\}$. δ is said to be *aperiodic* if for every $q \in Q$, $period_\delta(q) = 1$. δ is said to be *ergodic* if it is aperiodic and irreducible.

Markov chains. A Markov chain $\mathcal{M}$ is a tuple (Q, δ, μ) s.t. Q is a finite set of *states*, $\delta \in \mathsf{Mat}_{=1}(Q)$ and an *initial distribution* $\mu \in \mathrm{Dist}(Q)$. A Markov chain defines a sequence of distributions $\mu_0, \mu_1, \cdots$ where $\mu_i = \mu\delta^i$.

2.3 Probabilistic finite automata

A probabilistic finite automaton [20, 21] is like a deterministic automaton except that the transition function from a state on a given input is described as a probability distribution that determines the probability of transitioning to the next state.

Definition 1. *A Probabilistic Automaton (*PFA*) is a tuple* $\mathcal{B} = (\Sigma, Q, Q_f, \Delta = \{\delta_a\}_{a\in\Sigma})$ *where* Σ *is a finite nonempty set of input symbols and is called the input alphabet,* Q *is a finite set of states,* $Q_f \subseteq Q$ *is the set of accepting/final states and* $\Delta = \{\delta_a\}_{a\in\Sigma}$ *is a collection of stochastic matrices, one each for each input letter* a*.*

Notation: Given a PFA $\mathcal{B} = (\Sigma, Q, Q_f, \Delta = \{\delta_a\}_{a\in\Sigma})$ and a word $u = a_0 \cdots a_m \in \Sigma^*$, we denote the matrix $\delta_{a_0} \cdots \delta_{a_m}$ by δ_u. If u is the empty word then δ_u shall denote the identity matrix. Given a nonempty set $\Sigma_1 \subseteq \Sigma$, we denote the set of matrices $\{\delta_a\}_{a\in\Sigma_1}$ by Δ_{Σ_1}.

Language of a PFA**.** Language of a PFA is defined relative to an initial distribution and a cut-point. Formally, given a PFA $\mathcal{B} = (\Sigma, Q, Q_f, \Delta = \{\delta_a\}_{a\in\Sigma})$, an *initial distribution* $\mu \in \mathrm{Dist}(Q)$ and $v \in \Sigma^*$, the quantity $\mu\delta_v(Q_f)$ is called the *probability of* $\mathcal{B}$ *accepting* v *when started in* μ and shall be denoted by $\mathsf{Pr}^{acc}_{\mathcal{B},\mu}(v)$. For $x \in [0, 1]$, the set of words

$$\mathsf{L}_{>x}(\mathcal{B}, \mu) = \{v \in \Sigma^* \mid \mathsf{Pr}^{acc}_{\mathcal{B},\mu}(v) > x\}$$

is said to be the *language accepted by* $\mathcal{B}$ *with initial distribution* μ *and cut-point* x.

PFAs can recognize non-regular languages [21]. Furthermore, the problem of deciding emptiness and universality respectively for PFAs is undecidable [20, 13].

2.4 Isolated cut-points.

A much celebrated result of PFAs concerns isolated cut-points. A cut-point x is said to be isolated if there is an ϵ such that every word is either accepted with probability at least $x + \epsilon$ or accepted with probability at most $x - \epsilon$. Formally,

Definition 2. *Given a* PFA $\mathcal{B} = (\Sigma, Q, Q_f, \Delta = \{\delta_a\}_{a\in\Sigma})$ *and an initial distribution* μ*,* x *is said to be an isolated cut-point for* $(\mathcal{B}, \mu)$ *with a degree of isolation* $\epsilon > 0$ *if for each* $v \in \Sigma^*$*,* $|\mathsf{Pr}^{acc}_{\mathcal{B},\mu}(v) - x| > \epsilon$*.* x *is said to be an isolated cut-point for* $(\mathcal{B}, \mu)$ *if there is an* $\epsilon > 0$ *s.t.* x *is an isolated cut-point for* $(\mathcal{B}, \mu)$ *with a degree of isolation* ϵ*.*

A famous result of Rabin [21] says that if x is an isolated cut-point for $(\mathcal{B}, \mu)$ then $\mathsf{L}_{>x}(\mathcal{B}, \mu)$ is a regular language. This fact raises two interesting questions.

The first one asks if there is an algorithm that decides given a PFA $\mathcal{B}$, an initial distribution μ and a cut-point $x \in [0, 1]$, whether x is an isolated cut-point for $(\mathcal{B}, \mu)$ or not. Bertoni [4, 5] showed that the problem is undecidable when $x \in (0, 1)$. A close examination of the proof reveals that this problem is **r.e.**-hard. Recently, Gimbert and Oualhadj [15] showed that the problem remains undecidable even when x is 0 or 1. A close examination of their proof reveals that this problem is **co-r.e.**-hard also. However, the exact level of undecidability of this problem remained open.

The second question asks if there is a decision procedure that given a PFA $\mathcal{B}$, an initial distribution μ and a cut-point $x \in [0,1]$ isolated for $(\mathcal{B}, \mu)$ decides whether the language $\mathsf{L}_{>x}(\mathcal{B}, \mu)$ is empty or not. This problem seems to be less studied in literature. A close examination of Rabin's proof shows that if a degree of isolation ϵ is known, then $\mathsf{L}_{>x}(\mathcal{B}, \mu)$ can be computed as the proof computes an upper bound on the number of states of the deterministic automaton recognizing $\mathsf{L}_{>x}(\mathcal{B}, \mu)$ in terms of ϵ and the number of states of $\mathcal{B}$. However, the status of the problem when the degree of isolation is not known, remains open.

3 Checking isolation in PFAs

The following theorem states that the problem of checking whether a given x is isolated for a given PFA A is $\mathbf{\Sigma_2^0}$-complete, thus settling the open problem posed in [4, 5] (please note that the results also apply when x is 0 or 1).

Theorem 1. *Given a* PFA $\mathcal{B} = (\Sigma, Q, Q_f, \Delta = \{\delta_a\}_{a \in \Sigma})$*, an initial distribution* $\mu \in \mathrm{Dist}(Q)$ *and rational* $x \in [0,1]$*, the problem of checking if* x *is an isolated cut-point for* $(\mathcal{B}, \mu)$ *is* $\mathbf{\Sigma_2^0}$*-complete.*

Proof. Please note that x is isolated for $(\mathcal{B}, \mu)$ iff

$$\exists n \in \mathbb{N}, n > 0.\ \forall u \in \Sigma^*. |\mathsf{Pr}^{acc}_{\mathcal{B},\mu}(u) - x| > \frac{1}{n}.$$

This demonstrates that the problem of checking if x is isolated for $(\mathcal{B}, \mu)$ is in $\mathbf{\Sigma_2^0}$.

For the lower bound, please observe that it suffices to show that the problem of checking if 1 is isolated for $(\mathcal{B}, \mu)$ is $\mathbf{\Sigma_2^0}$-hard. We will demonstrate this by a reduction from emptiness checking problem of Probabilistic Büchi Automata (PBA)s [2]. A PBA $\mathcal{B}' = (\Sigma', Q', Q'_f, \Delta' = \{\delta'_a\}_{a \in \Sigma'})$ is like a PFA except that it is used to define languages over infinite words. Given a PBA $\mathcal{B}'$ and an initial distribution μ', $\mathcal{L}_{>0}(\mathcal{B}', \mu') \subseteq \Sigma^\omega$ denotes the set of infinite words accepted by $\mathcal{B}'$ with probability > 0. Intuitively, $\mathcal{B}$ accepts an infinite word α with probability > 0 if on input α, the measure of all (infinite) paths that visit Q'_f infinitely often is > 0. The exact definition of what it means for a PBA to accept an infinite word α with probability > 0 is beyond the scope of the paper and the interested reader is referred to [1]. We recall the necessary results.

We had shown the following problem to be $\mathbf{\Sigma_2^0}$-complete in [9]: Given a PBA $\mathcal{B}'$ and an initial distribution μ', check whether $\mathcal{L}_{>0}(\mathcal{B}', \mu') = \emptyset$. We will use this decision problem to establish the lower bound result.

Given a PBA, $\mathcal{B}' = (\Sigma', Q', Q'_f, \Delta' = \{\delta'_a\}_{a \in \Sigma'})$, an initial distribution μ' and $Q'' \subseteq Q'$, let $reachable(Q'')$ be the predicate $(\exists u \in \Sigma^*.\ (\mu' \delta_u)(Q'') > 0)$. Given $q \in Q$ and $v \in \Sigma^+$, let $\mathsf{Pr}^{Q'_f}_{q,v}(Q'')$ be the probability that the PFA $\mathcal{B}'$, on input v, when started in q reaches Q'' after passing through a state in Q'_f.

We had shown in [9] that $\mathcal{L}_{>0}(\mathcal{B}', \mu') \neq \emptyset$ iff

$$\exists Q'' \subseteq Q'.\ (reachable(Q'') \text{ and } (\forall n \in \mathbb{N}, n > 0.\ \exists v \in \Sigma^+.\ \forall q \in Q''.\ \mathsf{Pr}^{Q'_f}_{q,v}(Q'') > 1 - \tfrac{1}{2^n})).$$

Observe first that the predicate $reachable(\cdot)$ is a recursive predicate. Now, pick a new element † not in Q' and for all subsets $Q'' \subseteq Q'$ s.t. $reachable(Q'')$ is true, construct a PFA $\mathcal{B}_{Q''}$ and an initial distribution $\mu_{Q''}$ as follows. The input alphabet is Σ'. The states of $\mathcal{B}_{Q''}$ are $Q' \cup (Q' \times \{\dagger\})$. The set of final states of $\mathcal{B}_{Q''}$ are $Q'' \times \{\dagger\}$. The set of transitions $\Delta'' = \{\delta''_a\}_{a \in \Sigma'}$ is as follows. For each $a \in \Sigma'$:

- $\delta''_a(q_1, q_2) = \delta'_a(q_1, q_2)$ if $q_1 \in Q'$ and $q_2 \in Q' \setminus Q'_f$.
- $\delta''_a(q_1, (q_2, \dagger)) = \delta'_a(q_1, q_2)$ if $q_1 \in Q'$ and $q_2 \in Q'_f$.
- $\delta''_a((q_1, \dagger), (q_2, \dagger)) = \delta'_a(q_1, q_2)$ if $q_1, q_2 \in Q'$.

Let $\mu_{Q''}$ be the distribution assigns probability $\frac{1}{|Q''|}$ to each $q \in Q''$ where $|Q''|$ is the number of elements of Q''. It can be easily shown that $\mathcal{L}_{>0}(\mathcal{B}', \mu') = \emptyset$ iff $\forall Q'' \subseteq Q$, $reachable(Q'')$ implies that 1 is an isolated cut-point for $(\mathcal{B}_{Q''}, \mu_{Q''})$.

While the reduction above is a truth-table reduction, note that by taking "disjoint" union of the PFAs $\mathcal{B}_{Q''}$, adding a new initial state, a new reject state and new input symbols, we can easily construct a many-to-one reduction. The result now follows. □

Remark 1. We can conclude from the proof of Theorem 1 that the problem of checking whether 1 is an isolated cut-point for a PFA A is equivalent to the problem of checking whether a PBA $\mathcal{B}$ accepts an infinite word with probability > 0. The proof of Theorem 1 establishes one side of this equivalence and the converse is established in [3] (see Remark 5.7 on Page 41).

4 Weak ergodicity and eventually weak ergodicity

Ergodicity is an important concept that is useful in the study of stochastic matrices. We will recall this notion shortly and its extension to sets of stochastic matrices. We shall need one notation: Given a nonempty finite set $\Delta \subseteq \mathsf{Mat}_{=1}(Q)$ of stochastic matrices and $\ell > 0$, let $\Delta^\ell = \{\delta_1 \delta_2 \cdots \delta_\ell \mid \delta_i \in \Delta\}$.

Recall that a stochastic matrix is *ergodic* if it is irreducible and aperiodic. A Markov chain $\mathcal{M} = (Q, \delta, \mu)$ is ergodic if the matrix δ is ergodic. Ergodic chains have a special property that they converge to a *unique stationary distribution* in the limit irrespective of the starting distribution. More generally, this fact generalizes to Markov chains that a) have a single closed communicating class and b) this class is aperiodic. The notion of ergodicity has been extended to sets of stochastic matrices [22, 19, 17] and such sets are called *weakly ergodic sets*. Analogous to the convergence to the stationary distribution, if Δ is weakly ergodic then for any "long enough sequence" $\delta_1, \cdots, \delta_\ell$, any two distributions $\mu_1 \delta_1 \cdots \delta_\ell, \mu_2 \delta_1 \cdots \delta_\ell$ are "very close." We recall the formal definition of weakly ergodic matrices introduced in [7].

Definition 3. *A finite set of stochastic matrices Δ over a finite state space Q is said to be strongly semi-regular if for each $\delta \in \Delta$ there is a state q_δ s.t. for each $q \in Q$ $\delta(q, q_\delta) > 0$. Δ is weakly ergodic if there is an $\ell > 0$ s.t. Δ^ℓ is strongly semi-regular. A* PFA $\mathcal{B} = (\Sigma, Q, Q_f, \Delta = \{\delta_a\}_{a \in \Sigma})$ *is weakly ergodic if Δ is weakly ergodic.*

Example 1. Consider the 2-element set Δ shown in Figure 1.a). Δ can be seen to be strongly semi-regular and hence weakly ergodic as follows: the transition represented

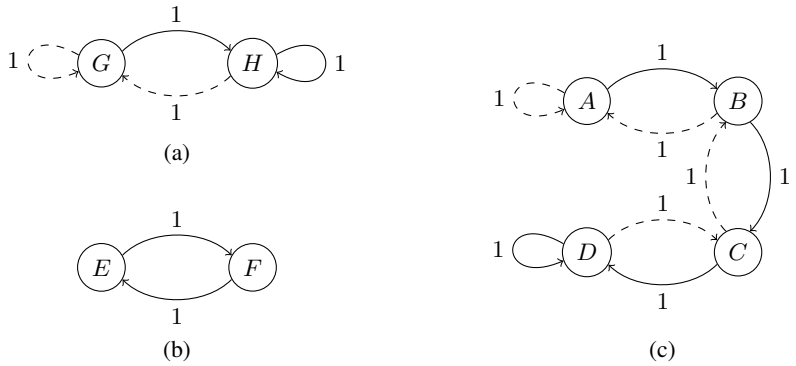

Fig. 1. a) Δ shown is weakly ergodic. There are two matrices in Δ: transitions of first one are shown as solid edges, while transitions of second one are shown as dashed edges. b) Δ is singleton and hence is eventually weakly ergodic. Δ is not weakly ergodic. c) Δ shown is not eventually weakly ergodic. There are two matrices in Δ: transitions of first one are shown as solid edges, while transitions of second one are shown as dashed edges.

by solid edges always "hits" the state H and the transition corresponding to the dashed edges always "hits" the state G.

Remark 2. There are several other equivalent definitions of weakly ergodic set of matrices. For example, one formulation [22] says that a finite set of matrices is weakly ergodic if every finite product of matrices has only one closed communicating class that is irreducible and aperiodic. There is an algorithm [22, 19] that given a nonempty, finite set $\Delta \subseteq \mathsf{Mat}_{=1}(Q)$ checks if Δ is weakly ergodic or not.

Eventually weakly ergodic sets. Even when stochastic matrices are not ergodic, the notion of ergodicity proves useful for analysis of Markov chains. This is because for any stochastic matrix δ, there is an $\ell > 0$ such that Q can be written as a disjoint sum $Q = Q_T \cup Q'_1 \cdots Q'_m$ where Q'_j is an aperiodic, closed communicating class for δ^ℓ and Q_T is the set of transient states for δ^ℓ. This observation motivates the following:

Definition 4. *$\Delta \subseteq \mathsf{Mat}_{=1}(Q)$ is said to be eventually weakly ergodic if there is a partition $Q_T, Q_1, \ldots Q_r$ of Q and a natural number $\ell > 0$ s.t. for each $1 \leq i \leq r$ the following conditions hold–*

- *Q_i is closed for Δ^ℓ.*
- *$\Delta^\ell|_{Q_i}$ is strongly semi-regular.*
- *For each $q \in Q_T$ and each $\delta \in \Delta^\ell$, $\mathsf{post}(q, \delta) \cap (\cup_{1 \leq j \leq r} Q_j) \neq \emptyset$.*

The tuple $(\ell, Q_T, (Q_0, \ldots, Q_r))$ is said to be a witness of eventual weak ergodicity. A PFA *$\mathcal{B} = (\Sigma, Q, Q_f, \Delta = \{\delta_a\}_{a \in \Sigma})$ is eventually weakly ergodic if Δ is eventually weakly ergodic.*

Remark 3. Please note that we are requiring that each Q_i be closed for Δ^ℓ and not for Δ. This means that Q_i is closed for all $\Delta^{\ell'}$ s.t. ℓ' is a multiple of ℓ. For ℓ' which is not a multiple of ℓ, a $\delta \in \Delta^{\ell'}$ may take a state in Q_i to a state not in Q_i. For example, consider the singleton Δ in Figure 1.b). Note that Δ is eventually weakly ergodic (but not weakly ergodic) with witness $(2, \emptyset, (\{E\}, \{F\}))$. Note that the sets $\{E\}$ and $\{F\}$ are closed for $\Delta^{\ell'}$ iff ℓ' is even.

As expected, each singleton turns out to be eventually weakly ergodic.

Proposition 1. *If $\Delta \subseteq \mathsf{Mat}_{=1}(Q)$ and $|\Delta| = 1$ then Δ is eventually weakly ergodic.*

Example 2. Observe that the Δ shown in Figure 1.c) is not eventually weakly ergodic. This can be seen as follows. For each ℓ, the only closed class of Δ^ℓ is the set of all states. Let δ_{solid} and δ_{dashed} be the matrices shown with solid lines and dashed lines respectively. For any $\delta \in \delta^\ell$ of the form $\delta_{solid}\delta_{dashed}\delta_{solid}\delta_{dashed}\cdots$, $\mathsf{post}(\delta, \{B\}) \cap \mathsf{post}(\delta, \{C\}) = \emptyset$. On the other hand, Δ shown in Figure 1.b) is eventually weakly ergodic with $(2, \emptyset, (\{E\}, \{F\}))$ as a witness of eventual weak ergodicity.

Remark 4. The *contracting* PFAs considered in [18] are different from eventually weakly ergodic matrices. Contracting PFAs are PFAs in which each transition matrix has only one closed communicating class which is aperiodic. The set Δ in Figure (1.c) is not eventually weakly ergodic but is contracting and the set Δ in Figure (1.b) is eventually weakly ergodic but not contracting.

The algorithm for checking whether a finite set of matrices are weakly ergodic can be extended to checking whether a finite set of matrices is eventually weakly ergodic.

Proposition 2. *The problem of checking given nonempty, finite set $\Delta \subseteq \mathsf{Mat}_{=1}(Q)$, whether Δ is eventually weakly ergodic is decidable. Furthermore, if Δ is eventually weakly ergodic then a witness of eventual weak ergodicity can be computed.*

5 Decision problems for eventually weakly ergodic PFAs

We shall now study the problems of deciding emptiness and isolation for eventually weakly ergodic PFAs. We start by discussing the emptiness problem.

5.1 Emptiness/universality checking for eventually weakly ergodic PFAs

The problem of checking emptiness of a PFA is undecidable [20, 13]. The problem continues to remain undecidable for eventually weakly ergodic PFAs. The proof of undecidability is similar to the one used in [8] to show that the problem of checking emptiness of *finite probabilistic monitors* is undecidable.

Proposition 3. *The following problems are undecidable: Given an eventually weakly ergodic* PFA $\mathcal{B} = (\Sigma, Q, Q_f, \Delta = \{\delta_a\}_{a\in\Sigma})$*, an initial distribution $\mu \in \mathrm{Dist}(Q)$ and rational $x \in (0, 1)$ check whether a) $\mathsf{L}_{>x}(\mathcal{B}, \mu) = \emptyset$ and whether b) $\mathsf{L}_{>x}(\mathcal{B}, \mu) = \Sigma^*$.*

Therefore, it follows that the *syntactic restriction* of eventually weak ergodicity is not enough for deciding emptiness of probabilistic automata. However, we will show that the problem of checking emptiness becomes decidable under the promise that the cut-point is isolated. In order to establish this result, we shall first establish a useful lemma (Lemma 1). In order to state this lemma, we need one auxiliary definition. Note that by the notation $m|n$ we mean that m divides n.

Definition 5. *Given an alphabet Σ and natural numbers $\ell, \ell' > 0$ such that $\ell'|\ell$, let $\mathsf{c}_{(\ell,\ell')} : \Sigma^* \to \Sigma^*$ be defined as follows.*

$$\mathsf{c}_{(\ell,\ell')}(u) = \begin{cases} u & \text{if } |u| < \ell' + 2\ell; \\ u_0u_1v_1 & \text{if } u = u_0u_1wv_1, |u_0| < \ell', |u_1| = \ell, w \in (\Sigma^{\ell'})^+ \text{ and } |v_1| = \ell \end{cases}.$$

Informally, given ℓ, ℓ' and Σ such that $\ell'|\ell$, the function $\mathsf{c}_{(\ell,\ell')}(\cdot)$ works as follows. If u is a string whose length is a multiple of ℓ', then $\mathsf{c}_{(\ell,\ell')}(u)$ keeps the prefix of length ℓ of u and the suffix of the length ℓ of u and "cuts" away the rest of the string. If the length of the u is not a multiple of ℓ' then it selects the largest suffix whose length is a multiple of ℓ' and applies $\mathsf{c}_{(\ell,\ell')}(\cdot)$ to it.

The following lemma states that if $\mathcal{B}$ is eventually weakly ergodic then the distribution obtained by inputting a word in Σ^* is determined up to a given ϵ by the initial distribution and an appropriate "cut" (which depends only on $\mathcal{B}$ and ϵ).

Lemma 1. *Given an eventually weakly ergodic* PFA $\mathcal{B} = (\Sigma, Q, Q_f, \Delta = \{\delta_a\}_{a\in\Sigma})$ *and $\epsilon > 0$, there are $\ell > 0$ and $\ell' > 0$ s.t. $\ell'|\ell$ and*

$$\forall \mu_0 \in \mathrm{Dist}(Q) \,.\, \forall u \in \Sigma^* \,.\, d(\mu_0\delta_u, \mu_0\delta_{\mathsf{c}_{(\ell,\ell')}(u)}) < \epsilon.$$

Furthermore, if ϵ is rational then ℓ, ℓ' can be computed from $\mathcal{B}$ and ϵ.

We shall now show that if $\mathcal{B}$ is eventually weakly ergodic and x is an isolated cut point for $(\mathcal{B}, \mu)$ then there is an algorithm that computes the regular language $\mathsf{L}_{>x}(\mathcal{B}, \mu)$. Recall that regularity is a consequence of Rabin's theorem on isolated cut-points (see Section 2.4). Indeed, as observed in Section 2.4, Rabin's proof can also be used to compute the regular language, provided we can compute a degree of isolation. The eventually weak ergodicity condition allows us to compute this.

Lemma 2. *There is a procedure that given an eventually weakly ergodic* PFA $\mathcal{B} = (\Sigma, Q, Q_f, \Delta = \{\delta_a\}_{a\in\Sigma})$, *a distribution $\mu \in \mathrm{Dist}(Q)$ and a rational $x \in [0, 1]$ such that x is an isolated cut-point for $(\mathcal{B}, \mu)$ terminates and outputs $\epsilon > 0$ such that ϵ is a degree of isolation.*

Proof. Consider the procedure in Figure 2. Thanks to Lemma 1, if the procedure terminates then the ϵ returned is a degree of isolation. Hence, it suffices to show that the procedure terminates if x is an isolated cut-point for $(\mathcal{B}, \mu)$. Let ϵ_0 be a degree of isolation and fix it. Thus, for all $u \in \Sigma^*$, $|\mathsf{Pr}^{acc}_{\mathcal{B},\mu}(u) - x| > \epsilon_0$. Let $\epsilon^{(n)}$ be the value of variable ϵ at the beginning of the nth unrolling of the while loop. As long as $isolation_{found}$ is false, $\epsilon^{(n)} = \frac{1}{2^n}$. Let $N_0 = \lfloor \log_2 \epsilon_0 \rfloor + 1$. As $\epsilon^{(N_0)} < \epsilon_0$ and $\forall u \in \Sigma^*$, $|\mathsf{Pr}^{acc}_{\mathcal{B},\mu}(u) - x| > \epsilon_0$, $isolation_{found}$ must become true at the end of N_0th unrolling, if not before. □

Input: $\mathcal{B}, \mu, x$ where
$\mathcal{B} = (\Sigma, Q, Q_f, \Delta = \{\delta_a\}_{a\in\Sigma})$ is an eventually weakly ergodic PFA,
$\mu \in \mathrm{Dist}(Q)$ is the initial distribution and $x \in [0,1]$ is a rational number
{
$isolation_{found} := false$;
$\epsilon := \frac{1}{2}$;
while $not(isolation_{found})$
do
Compute $\ell, \ell' > 0$ such that
$\forall \mu_0 \in \mathrm{Dist}(Q).\ \forall u \in \Sigma^*.\ d(\mu_0\delta_u, \mu_0\delta_{c_{(\ell,\ell')}(u)}) < \epsilon$;
$curr_{isolation} := \min_{v\in(\Sigma)^{<2\ell+\ell'}} |\mu\delta_v(Q_f) - x|$;
If $curr_{isolation} \leq \epsilon$ then $\epsilon := \frac{\epsilon}{2}$
else $\{\epsilon := curr_{isolation} - \epsilon;\ isolation_{found} = true;\}$
od;
return(ϵ);
}

Fig. 2. Procedure for computing the degree of isolation

Lemma 2 along with the proof of Rabin's theorem on isolated cut-points can now be used to establish our main result.

Theorem 2. *There is a procedure that given an eventually weakly ergodic* PFA $\mathcal{B} = (\Sigma, Q, Q_f, \Delta = \{\delta_a\}_{a\in\Sigma})$, *a distribution* $\mu \in \mathrm{Dist}(Q)$ *and a rational* $x \in [0,1]$ *such that* x *is an isolated cut-point for* $(\mathcal{B}, \mu)$ *terminates and outputs the regular language* $\mathsf{L}_{>x}(\mathcal{B}, \mu)$. *Therefore, the following problems are decidable: Given an eventually weakly ergodic* PFA $\mathcal{B} = (\Sigma, Q, Q_f, \Delta = \{\delta_a\}_{a\in\Sigma})$, *a distribution* $\mu \in \mathrm{Dist}(Q)$ *and rational* $x \in [0,1]$ *s.t.* x *is an isolated cut-point, a) check whether* $\mathsf{L}_{>x}(\mathcal{B}, \mu) = \emptyset$ *and b) check whether* $\mathsf{L}_{>x}(\mathcal{B}, \mu) = \Sigma^*$.

5.2 Checking Isolation for weakly ergodic PFAs

A close examination of the proof of undecidability of checking isolation in PFAs given in Bertoni [4, 5] reveals that the problem of checking isolation is **r.e.**-hard even for eventually weakly ergodic automata. Furthermore, if we run the procedure in Lemma 2 on an arbitrary (i.e., not necessarily isolated) cut-point x then the procedure terminates if and only if x is an isolated cut-point, implying that checking isolation is in **r.e.**.

Theorem 3. *The following problem is* **r.e.***-complete: Given an eventually weakly ergodic* PFA $\mathcal{B} = (\Sigma, Q, Q_f, \Delta = \{\delta_a\}_{a\in\Sigma})$, *a distribution* $\mu \in \mathrm{Dist}(Q)$ *and rational* x, *check if* x *is isolated for* $(\mathcal{B}, \mu)$.

6 Conclusions

We have established the exact level of undecidability of checking if a given threshold x is an isolated cut point for a given PFA A, showing it to be Σ_2^0-complete. We have also

proved decidability of non-emptiness (and universality) for eventually weakly ergodic automata, given that the automaton has an isolated cut-point. The problem of decidability/undecidability of checking non-emptiness for arbitrary PFAs with isolated cut points is still an open problem.

References

1. C. Baier, N. Bertrand, and M. Größer. On decision problems for probabilistic Büchi automata. In *Proceedings of FoSSaCS*, pages 287–301, 2008.
2. C. Baier and M. Größer. Recognizing ω-regular languages with probabilistic automata. In *Proceedings of LICS*, pages 137–146, 2005.
3. C. Baier, M. Größer, and N. Bertrand. Probabilistic ω-automata. *J. ACM*, 59(1):1, 2012.
4. A. Bertoni. The solution of problems relative to probabilistic automata in the frame of the formal languages theory. In *Gl-4.Jahrestagung*, volume 26 of *Lecture Notes in Computer Science*, pages 107–112. Springer Berlin / Heidelberg, 1975.
5. A. Bertoni, G. Mauri, and M. Torelli. Some recursive unsolvable problems relating to isolated cutpoints in probabilistic automata. In *Proceedings of ICALP*, pages 87–94, 1977.
6. R. Canetti, L. Cheung, D. Kaynar, M. Liskov, N. Lynch, P. Pereira, and R. Segala. Task-Structured Probabilistic I/O Automata. In *Workshop on Discrete Event Systems*, 2006.
7. R. Chadha, V. A. Korthikranthi, M. Viswanathan, G. Agha, and Y. Kwon. Model checking mdps with a unique compact invariant set of distributions. In *Proceedings of QEST*, pages 121–130, 2011.
8. R. Chadha, A. P. Sistla, and M. Viswanathan. On the expressiveness and complexity of randomization in finite state monitors. *Journal of the ACM*, 56(5), 2009.
9. R. Chadha, A. P. Sistla, and M. Viswanathan. Power of randomization in automata on infinite strings. In *Proceedings of CONCUR*, pages 229–243, 2009.
10. R. Chadha, A.P. Sistla, and M. Viswanathan. Model checking concurrent programs with nondeterminism and randomization. In *Proceedings of FSTTCS*, pages 364–375, 2010.
11. K. Chatzikokolakis and C. Palamidessi. Making Random Choices Invisible to the Scheduler. *Information and Computation*, 208(6):694–715, 2010.
12. L. Cheung. *Reconciling Nondeterministic and Probabilistic Choices*. PhD thesis, Radboud University of Nijmegen, 2006.
13. A. Condon and R. J. Lipton. On the complexity of space bounded interactive proofs (extended abstract). In *Proceedings of FOCS*, pages 462–467, 1989.
14. L. de Alfaro. The Verification of Probabilistic Systems under Memoryless Partial Information Policies is Hard. In *PROBMIV*, 1999.
15. H. Gimbert and Y. Oualhadj. Probabilistic automata on finite words: Decidable and undecidable problems. In *Proceedings of ICALP*, pages 527–538, 2010.
16. S. Giro and P.R. D'Argenio. Quantitative model checking revisited: Neither decidable nor approximable. In *Proceedings of FORMATS*, pages 179–194, 2007.
17. J. Hajnal and M. S. Bartlett. Weak ergodicity in non-homogeneous markov chains. *Mathematical Proceedings of the Cambridge Philosophical Society*, 54(02):233–246, 1958.
18. V. A. Korthikanti, M. Viswanathan, G. Agha, and Y. Kwon. Reasoning about mdps as transformers of probability distributions. In *Proceedings of QEST*, pages 199–208, 2010.
19. A. Paz. Definite and quasidefinite sets of stochastic matrices. *Proceedings of the American Mathematical Society*, 16(4):634–641, 1965.
20. A. Paz. *Introduction to Probabilistic Automata*. Academic Press, 1971.
21. M. O. Rabin. Probabilitic automata. *Information and Control*, 6(3):230–245, 1963.
22. J. Wolfowitz. Products of indecomposable, aperiodic, stochastic matrices. *Proceedings of the American Mathematical Society*, 14(5):733–737, 1963.