Skolem and Positivity Completeness of Ergodic Markov Chains

Mihir Vahanwala

[a] *Max Planck Institute for Software Systems, Saarland Informatics Campus, Germany*

Abstract

We consider the following decision problems: given a finite, rational Markov Chain, source and target states, and a rational threshold, does there exist an n such that the probability of reaching the target from the source at the n^{th} step is equal to the threshold (resp. crosses the threshold)? These problems are known to be equivalent to the Skolem (resp. Positivity) problems for Linear Recurrence Sequences (LRS). These are number-theoretic problems whose decidability has been open for decades. We present a short, self-contained, and elementary reduction from LRS to Markov Chains that improves the state of the art as follows: (a) We reduce to ergodic Markov Chains, a class widely used in Model Checking. (b) We reduce LRS to Markov Chains of significantly lower order than before. We thus get sharper hardness results for a more ubiquitous class of Markov Chains. Immediate applications include problems in modeling biological systems, and regular automata-based counting problems.

Keywords: Ergodic Markov Chains, Reachability, Model checking, Linear Recurrence Sequences

1. Introduction

Markov Chains are a natural mathematical framework to describe probabilistic systems, such as those arising in computational biology. There is an extensive body of work on model checking Markov Chains: see [3] for a comprehensive set of references. Most of the focus has been on the verification of linear- and branching-time properties of Markov Chains through solving systems of linear equations, or linear programs. An alternative approach [1, 4, 7, 8] is to consider specifications on the state distribution at each time step, e.g., whether the probability of being in a given state at the n^{th} step is at least 1/4. Decidability in this setting is a lot more inaccessible: [1, 4] only present incomplete or approximate verification procedures, while [7, 8] owe their model-checking procedures to additional mathematical assumptions. The inherent difficulty of precisely solving decision problems in this fundamental setting is established in [2]: it is formally shown that verifying such specifications is tantamount to solving the Skolem/Positivity Problem for Linear Recurrence Sequences (LRS). The reduction therein is from LRS of order k to periodic Markov Chains of order

$2k{+}4$. However, it is the ergodic Markov Chains (irreducible and *aperiodic*) that are widely assumed in practice, and it may well be that the hardness is somehow mitigated by the additional spectral structure. On the automata-theoretic front, [6] reduces LRS to counting words of length n in a regular language. The nature of the reduction means that the resulting automaton is necessarily periodic. *Aperiodicity* in this context relates to LTL-definability, and the logical restriction could lead to a combinatorial breakthrough.

In this paper, we show that the such breakthroughs that circumvent the original reduction are unlikely without significant progress in the underlying number-theory itself. **Our novelty** is a reduction from order k LRS to ergodic Markov Chains of order $k+1$. An interesting feature of our reduction is that it shows that hard instances exist for *every* stationary distribution. In doing so, the translation of number-theoretic hardness for LRS (cf. [11]) to Markov Chains also becomes much sharper.

2. Markov Chain Preliminaries

Notation: Distributions are assumed to be column vectors. We use $\mathbf{1}$ to denote the column vector whose entries are all 1, and $\mathbf{I}$ to denote the identity matrix. We use $\mathbf{0}$ to denote the zero column vector, and $\mathbf{O}$ to denote the zero matrix. Superscript T denotes transposition. We use $\mathbf{e_i}$ to denote the elementary column vector, i.e. the vector whose i^{th} entry is 1 and all other entries are 0, e.g. $\mathbf{e_1} = \begin{bmatrix}1 & 0 & \dots & 0\end{bmatrix}^T$. We use $m_{ij}^{(n)}$ as shorthand to denote the entry in the i^{th} row and j^{th} column of $\mathbf{M}^n$, i.e. $m_{ij}^{(n)} = \mathbf{e_i}^T\mathbf{M}^n\mathbf{e_j}$. When not specified, $n = 1$.

Definition 1 (Markov Chain)**.** *A k-state Markov Chain over $\mathbb{Q}$ is a matrix $\mathbf{M} \in \mathbb{Q}^{k\times k}$, such that $m_{ij} = \mathbf{e_i}^T\mathbf{M}\mathbf{e_j}$ denotes the probability of moving from state j to state i. We have $\mathbf{1}^T\mathbf{M} = \mathbf{1}^T$.*

Definition 2 (Irreducible Markov Chain)**.** *A Markov Chain is called irreducible if every state has a path to every other state.*

Definition 3 (Periodicity)**.** *The period d_i of a state i of a Markov chain $\mathbf{M}$ is defined as*

$$\mathsf{gcd}\{n \geq 1 : \mathbf{e_i}^T\mathbf{M}^n\mathbf{e_i} > 0\}$$

State i is called aperiodic if $d_i = 1$. $\mathbf{M}$ is said to be aperiodic iff all its states are aperiodic.

Definition 4 (Stationary distribution)**.** *A distribution $\mathbf{s}$ is said to be a stationary distribution of a Markov Chain $\mathbf{M}$, if $\mathbf{M}\mathbf{s} = \mathbf{s}$.*

Theorem 1 (Fundamental Theorem of (Ergodic) Markov Chains)**.** *A Markov chain $\mathbf{M}$ is called ergodic if it is irreducible and aperiodic. An ergodic Markov Chain has a unique stationary distribution.*

The following technical lemma will help us construct an Ergodic Markov Chain in our reduction.

Lemma 2. *Let* $\mathbf{s}$ *be a distribution with all entries strictly positive, and let* $\mathbf{S} = \begin{bmatrix} \mathbf{s} & \mathbf{s} & \ldots & \mathbf{s} \end{bmatrix}$*. A stochastic matrix* $\mathbf{M}$ *is an ergodic Markov Chain with stationary distribution* $\mathbf{s}$ *if and only if* $\mathbf{Ms} = \mathbf{s}$ *and there exists* $\mathbf{D}$ *such that*

- $\mathbf{M} = \mathbf{S} + \mathbf{D}$
- $\mathbf{DS} = \mathbf{SD} = \mathbf{O}$
- $\mathbf{D}$ *has spectral radius less than* 1*, i.e.* $\lim_{n\to\infty} \mathbf{D}^n = \mathbf{O}$

In particular, we observe that the first two properties of $\mathbf{D}$ *imply that* $\mathbf{M}^n = \mathbf{S} + \mathbf{D}^n$ *for* $n \geq 1$*.*

Proof. **Only If**:
Let $\mathbf{M}$ be an ergodic Markov Chain with stationary distribution $\mathbf{s}$. Then, by definition, $\mathbf{Ms} = \mathbf{s}$, $\mathbf{MS} = \mathbf{S}$, and $\lim_{n\to\infty} \mathbf{M}^n = \mathbf{S}$. We note that both $\mathbf{M}$ and $\mathbf{S}$ are stochastic matrices, and thus $\mathbf{1}^T(\mathbf{M} - \mathbf{S}) = \mathbf{0}^T$. Denote $\mathbf{M} - \mathbf{S}$ by $\mathbf{D}$. From the previous observation, it is clear that $\mathbf{SD} = \mathbf{O}$, since all the rows of $\mathbf{S}$ are scaled multiples of $\mathbf{1}^T$. We also have that $\mathbf{S} = \mathbf{MS} = \mathbf{S}^2 + \mathbf{DS} = \mathbf{S} + \mathbf{DS}$, which means that $\mathbf{DS} = \mathbf{O}$.

We use the fact that $\mathbf{DS} = \mathbf{SD} = \mathbf{O}$ and that $\mathbf{S}^n = \mathbf{S}$ to observe

$$\mathbf{M}^n = (\mathbf{S} + \mathbf{D})^n = \mathbf{S} + \mathbf{D}^n$$

because all other terms in the binomial expansion are nullified. Now, since $\lim_{n\to\infty} \mathbf{M}^n = \mathbf{S}$, it forces $\lim_{n\to\infty} \mathbf{D}^n = \mathbf{O}$.

If:
The above argument is reversible. If instead we are given the three properties of $\mathbf{D}$ to begin with, we can conclude that $\mathbf{MS} = \mathbf{S}$ and $\lim_{n\to\infty} \mathbf{M}^n = \mathbf{S}$, which is precisely the definition of $\mathbf{M}$ being an ergodic Markov Chain with stationary distribution $\mathbf{s}$. □

3. Overview of Problems

Problem 1 (Ergodic Markov Chain Reachability)**.** *Given an ergodic Markov Chain* $\mathbf{M} \in \mathbb{Q}^{k\times k}$ *and* $r \in \mathbb{Q}$*, the Ergodic Markov Chain Reachability problem asks whether there exists an* $n \in \mathbb{N}$ *such that the probability of returning to state* 1 *at the* n^{th} *step is exactly* r*, i.e.* $\mathbf{e_1}^T\mathbf{M}^n\mathbf{e_1} = r$*.*

Problem 2 (Threshold Ergodic Markov Chain Reachability)**.** *Given an ergodic Markov Chain* $\mathbf{M} \in \mathbb{Q}^{k\times k}$ *and* $r \in \mathbb{Q}$*, the Threshold Ergodic Markov Chain Reachability problem asks whether for all* $n \in \mathbb{N}$*, the probability of returning to state* 1 *at the* n^{th} *step is at least* r*, i.e.* $\mathbf{e_1}^T\mathbf{M}^n\mathbf{e_1} \geq r$*.*

We relate these problems to long-standing open problems on Linear Recurrence Sequences.

Definition 5 (Linear Recurrence Sequence a.k.a. LRS)**.** *An LRS of order k over $\mathbb{Q}$ is an infinite sequence $\langle u_n \rangle_{n=0}^{\infty}$ satisfying a recurrence relation*

$$u_{n+k} = \sum_{i=0}^{k-1} a_i u_{n+i}$$

for all $n \in \mathbb{N}$. The recurrence relation is given by the tuple $(a_0, \ldots, a_{k-1}) \in \mathbb{Q}^k$ with $a_0 \neq 0$. The sequence is uniquely determined by the starting values $(u_0, \ldots, u_{k-1}) \in \mathbb{Q}_k$.

Problem 3 (Skolem Problem for LRS)**.** *Given an LRS $\langle u_n \rangle_{n=0}^{\infty}$ (via the recurrence relation and starting values), the Skolem problem asks whether there exists an $n \in \mathbb{N}$ such that $u_n = 0$.*

Problem 4 (Positivity Problem for LRS)**.** *Given an LRS $\langle u_n \rangle_{n=0}^{\infty}$ (via the recurrence relation and starting values), the Positivity problem asks whether for all $n \in \mathbb{N}$, $u_n \geq 0$.*

The Skolem Problem is known to be decidable for LRS of order up to 4, see [9, 12]. Very recently, there have been conditional decidability results for LRS of order 5 [5]. The Positivity Problem is decidable up to order 5, decidability at order 6 would entail significant number-theoretic breakthroughs [11]. If we restrict ourselves to the class of *simple* LRS (no repeated characteristic roots), then Positivity is decidable up to order 9, see [10].

We are now ready to state our main reduction, which is agnostic to the spectral nature of the LRS, and minimal with respect to the order.

Theorem 3 (Main Result)**.** *Problem 3 reduces to Problem 1, while Problem 4 reduces to Problem 2. Moreover, applying the reduction to an LRS of order k results in an ergodic Markov Chain of order $k+1$.*

It is well known that for any matrix $\mathbf{M}$, $\langle m_{ij}^{(n)} \rangle_{n=1}^{\infty}$, i.e. the entries in the i^{th} row and j^{th} column in the powers of $\mathbf{M}$ form an LRS. This is most easily seen through the Cayley-Hamilton Theorem: any matrix $\mathbf{M}$ satisfies its characteristic polynomial equation, i.e. if $p(\lambda) = \det(\mathbf{M} - \lambda \mathbf{I})$, then $p(\mathbf{M}) = \mathbf{O}$.

Thus, we trivially have that Problems 3 and 4 respectively reduce to Problems 1 and 2.

One can define the "off-diagonal" variants of Problems 1 and 2, i.e. queries on the probability of reaching state 1 from state 2. The above equivalences hold for the off-diagonal variants with an almost identical proof. We discuss the difference after presenting the reduction.

4. The Reduction

The key idea is to use Lemma 2 to construct an ergodic Markov Chain via the decomposition $\mathbf{M} = \mathbf{S} + \mathbf{D}$. Given an LRS $\langle u_n \rangle_{n=0}^{\infty}$ of order k over $\mathbb{Q}$, we will choose $\mathbf{S}, \mathbf{D} \in \mathbb{Q}^{(k+1)\times(k+1)}$, $r = s_{11} = s_1$, a rational η and a large rational ρ in such a way that for all $n \geq 1$, $d_{11}^{(n)} = \eta u_n / \rho^n$.

Since $m_{11}^{(n)} = s_1 + d_{11}^{(n)}$, deciding the Skolem (resp. Positivity) problem reduces to checking whether there exists n such that $m_{11}^{(n)} = r = s_1$ (resp. for all n, $m_{11}^{(n)} \geq r = s_1$), which is precisely the reduction we want.

We assume without loss of generality that none of the initial terms of the LRS are 0, and that $u_0 > 0$.

To begin with, we choose some arbitrary probability distribution

$$\mathbf{s} = \begin{bmatrix} s_1 & s_2 & \dots & s_{k+1} \end{bmatrix}^T \in \mathbb{Q}^{k+1}$$

such that all the entries of $\mathbf{s}$ are strictly positive. $\mathbf{S}$ denotes the square matrix, each of whose columns is $\mathbf{s}$.

Let $\mathbf{A} \in \mathbb{Q}^{k\times k}$ be the companion matrix of the given LRS, i.e.

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ a_0 & a_1 & a_2 & \dots & a_{k-1} \end{bmatrix}$$

and let $\mathbf{u} = \begin{bmatrix} u_0 & u_1 & \dots & u_{k-1} \end{bmatrix}^T$ We have that $u_n = \mathbf{e_1}^T \mathbf{A}^n \mathbf{u}$

Now, we choose $\eta \in \mathbb{Q}$, $\eta > 0$ such that $\eta u_0 = 1 - s_1$. Let $\mathbf{F} \in \mathbb{Q}^{k\times k}$ be the invertible diagonal matrix such that

$$\mathbf{F} \begin{bmatrix} 1 - s_1 \\ -s_2 \\ \vdots \\ -s_k \end{bmatrix} = \eta \mathbf{u}$$

i.e. $\mathbf{F} = \mathrm{diag}(1, -\eta u_1/s_2, \dots, -\eta u_{k-1}/s_k)$. Observe that the top left entry in both $\mathbf{F}$ and $\mathbf{F}^{-1}$ is 1. Now, let $\mathbf{B} = \mathbf{F}^{-1}\mathbf{A}\mathbf{F}$.

Let $\mathbf{C} \in \mathbb{Q}^{(k+1)\times(k+1)}$ be the matrix

$$\begin{bmatrix} \mathbf{B} & \mathbf{0} \\ -\mathbf{1}^T\mathbf{B} & 0 \end{bmatrix}$$

We note, by a simple induction, that for $n \geq 1$

$$\mathbf{C}^n = \begin{bmatrix} \mathbf{B}^n & \mathbf{0} \\ -\mathbf{1}^T\mathbf{B}^n & 0 \end{bmatrix}$$

By construction $\mathbf{1}^T\mathbf{C} = \mathbf{0}^T$, and hence $\mathbf{S}\mathbf{C} = \mathbf{O}$. Let $\mathbf{D} = \frac{1}{\rho}(\mathbf{C} - \mathbf{C}\mathbf{S})$. The choice of ρ is large enough to ensure that:

- The entries of $\mathbf{S} + \mathbf{D}$ are non-negative.

- The spectral radius of $\mathbf{D}$ is less than 1.

By Lemma 2, this makes the stochastic $\mathbf{M} = \mathbf{S} + \mathbf{D}$ an ergodic Markov Chain with stationary distribution $\mathbf{s}$, since, indeed, $\mathbf{DS} = \mathbf{SD} = \mathbf{O}$.

We now observe that for $n \geq 1$, $\mathbf{D}^n = \frac{1}{\rho^n}\mathbf{C}^n(\mathbf{I} - \mathbf{S})$. We see this inductively: $(\mathbf{C} - \mathbf{CS})(\mathbf{C}^n - \mathbf{C}^n\mathbf{S}) = \mathbf{C}^{n+1} - \mathbf{C}^{n+1}\mathbf{S} - \mathbf{CSC}^n + \mathbf{CSC}^n\mathbf{S} = \mathbf{C}^{n+1} - \mathbf{C}^{n+1}\mathbf{S}$, since $\mathbf{SC} = \mathbf{O}$.

To complete the proof, we now compute the top-left entry of $\mathbf{D}^n$, $n \geq 1$:

$$
\begin{aligned}
\mathbf{e_1}^T\mathbf{D^n}\mathbf{e_1} &= \frac{1}{\rho^n}\mathbf{e_1}^T\mathbf{C}^n(\mathbf{I} - \mathbf{S})\mathbf{e_1} \\
&= \frac{1}{\rho^n}\mathbf{e_1}^T \begin{bmatrix} \mathbf{B}^n & \mathbf{0} \\ -\mathbf{1}^T\mathbf{B}^n & 0 \end{bmatrix} \begin{bmatrix} 1 - s_1 \\ -s_2 \\ \vdots \\ -s_k \\ -s_{k+1} \end{bmatrix} \\
&= \frac{1}{\rho^n}\mathbf{e_1}^T\mathbf{B}^n \begin{bmatrix} 1 - s_1 \\ -s_2 \\ \vdots \\ -s_k \end{bmatrix} \quad \text{(note the change in dimension of } \mathbf{e_1}) \\
&= \frac{1}{\rho^n}(\mathbf{e_1}^T\mathbf{F}^{-1})\mathbf{A}^n \left(\mathbf{F} \begin{bmatrix} 1 - s_1 \\ -s_2 \\ \vdots \\ -s_k \end{bmatrix} \right) \\
&= \frac{\eta}{\rho^n}\mathbf{e_1}^T\mathbf{A}^n\mathbf{u} \\
&= \frac{\eta u_n}{\rho^n}
\end{aligned}
$$

which is precisely a scaled version of our LRS.

We have thus established that

$$m_{11}^{(n)} = s_1 + \frac{\eta u_n}{\rho^n}$$

Let $r = s_1$. The original LRS is a YES instance of the Skolem Problem (resp. the Positivity Problem) if and only if there exists an n such that $m_{11}^{(n)} = r$ (resp. for all n, $m_{11}^{(n)} \geq r$). These are precisely the YES instances of the reachability problems we defined.

5. Discussion: The off-diagonal variants

In this variant, we query $m_{12}^{(n)}$ instead of $m_{11}^{(n)}$. The proof proceeds identically, except for the choice of η and diagonal matrix $\mathbf{F}$. Here

$$\mathbf{F}\begin{bmatrix} -s_1 \\ 1-s_2 \\ \vdots \\ -s_k \end{bmatrix} = \eta\mathbf{u}$$

We choose $r = s_1$, but $\eta = -u_0/s_0 < 0$, and thus

$$\mathbf{F} = \mathrm{diag}(1, \eta u_1/(1-s_2), -\eta u_2/s_3, \ldots, -\eta u_{k-1}/s_k)$$

In the same way as above, we get $d_{12}^{(n)} = \frac{\eta u_n}{\rho^n}$.

$$m_{12}^{(n)} = s_1 + \frac{\eta u_n}{\rho^n}$$

In the previous case, $d_{11}^{(n)}$ and u_n had the same sign, however, $d_{12}^{(n)}$ and u_n have opposite signs here. Thus, Positivity is equivalent to $m_{12}^{(n)} \leq s_1$ for all n, whereas Skolem is still equivalent to $m_{12}^{(n)} = s_1$ for some n.

Note the difference in the inequalities in the diagonal and off-diagonal cases. The diagonal and off-diagonal variants seem to have some inherent structural differences. The trivial justification for making this choice of inequalities is that regardless of choice of r, for $n = 0$, the probability of being in the source state is 1, and can never be less than r. Similarly, for $n = 0$ the probability of being in a state different from the starting state is 0. On a philosophical note, the diagonal variant can be thought of as a safety property (e.g. a fraction of the population will invariably be in the desirable state we started off in), whereas the violation of the off-diagonal variant can be thought of as a liveness property (e.g. at some point, a large fraction of the population will be in an active state). We do not have more technical insights (e.g. what if the problem were defined for only $n \geq 1$?): if any, they are likely to be beyond the scope of the simple reduction we present here.

References

[1] Agrawal, M., Akshay, S., Genest, B., Thiagarajan, P.S.: Approximate verification of the symbolic dynamics of markov chains. Journal of the ACM (JACM) **62**, 1 – 34 (2015)

[2] Akshay, S., Antonopoulos, T., Ouaknine, J., Worrell, J.: Reachability problems for markov chains. Information Processing Letters **115**(2), 155–158 (2015). https://doi.org/https://doi.org/10.1016/j.ipl.2014.08.013, `https://www.sciencedirect.com/science/article/pii/S0020019014001781`

[3] Baier, C., Katoen, J.P.: Principles of model checking. MIT Press (2000)

[4] Beauquier, D., Rabinovich, A., Slissenko, A.: A logic of probability with decidable model checking. J. Log. Comput. **16**, 461–487 (08 2006). https://doi.org/10.1093/logcom/exl004

[5] Bilu, Y., Luca, F., Nieuwveld, J., Ouaknine, J., Purser, D., Worrell, J.: Skolem meets schanuel (2022)

[6] Fried, D., Legay, A., Ouaknine, J., Vardi, M.Y.: Sequential relational decomposition. In: Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science. p. 432–441. LICS '18, Association for Computing Machinery, New York, NY, USA (2018). https://doi.org/10.1145/3209108.3209203, `https://doi.org/10.1145/3209108.3209203`

[7] Korthikanti, V.A., Viswanathan, M., Agha, G., Kwon, Y.M.: Reasoning about mdps as transformers of probability distributions. In: QEST. pp. 199–208 (2010)

[8] Kwon, Y., Agha, G.: Linear inequality ltl (iltl): A model checker for discrete time markov chains. In: Davies, J., Schulte, W., Barnett, M. (eds.) Formal Methods and Software Engineering. pp. 194–208. Springer Berlin Heidelberg (2004)

[9] Mignotte, M., Shorey, T.N., Tijdeman, R.: The distance between terms of an algebraic recurrence sequence. Journal für die reine und angewandte Mathematik **349**, 63–76 (1984)

[10] Ouaknine, J., Worrell, J.: On the positivity problem for simple linear recurrence sequences. In: International Colloquium on Automata, Languages, and Programming. pp. 318–329. Springer (2014)

[11] Ouaknine, J., Worrell, J.: Positivity problems for low-order linear recurrence sequences. In: Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2014, Portland, Oregon, USA, January 5-7, 2014. pp. 366–379. SIAM (2014)

[12] Vereshchagin, N.K.: The problem of appearance of a zero in a linear recurrence sequence. Matematicheskie Zametki (in Russian) **38**, 177–189 (1985)