Information Processing Letters 115 (2015) 155–158

Contents lists available at ScienceDirect

Information Processing Letters

www.elsevier.com/locate/ipl

Reachability problems for Markov chains

S. Akshay [a], Timos Antonopoulos [b,*], Joël Ouaknine [b], James Worrell [b]

[a] Department of Computer Science and Engineering, Indian Institute of Technology Bombay, India
[b] Department of Computer Science, University of Oxford, UK

ARTICLE INFO

Article history:
Received 3 March 2014
Received in revised form 29 August 2014
Accepted 31 August 2014
Available online 16 September 2014
Communicated by A. Muscholl

Keywords:
Formal methods
Skolem problem
Positivity problem
Probabilistic model checking
Markov chains

ABSTRACT

We consider the following decision problem: given a finite Markov chain with distinguished source and target states, and given a rational number r, does there exist an integer n such that the probability to reach the target from the source in n steps is r? This problem, which is not known to be decidable, lies at the heart of many model checking questions on Markov chains. We provide evidence of the hardness of the problem by giving a reduction from the Skolem Problem: a number-theoretic decision problem whose decidability has been open for many decades.

1. Introduction

By now there is a large body of work on model checking Markov chains; see [3] for references. Most of this work focuses on verifying linear- and branching-time properties of trajectories, typically by solving systems of linear equations or by linear programming. An alternative approach [1,2,4–6] considers specifications on the state distribution of the Markov chain at each time step, e.g., whether the probability to be in a given state is always at least 1/3. With this shift in view the associated algorithmic questions become surprisingly subtle, with not even decidability assured. Strikingly the works [1,2,4] only present incomplete or approximate verification algorithms. Similarly, in [5,6], the authors make additional assumptions (e.g., contraction properties, boundary assumptions) to obtain model-checking procedures.

The paper [4] highlights the following fundamental decision problem on Markov chains:

Markov Reachability. Given a finite stochastic matrix M with rational entries and a rational number r, does there exist $n \in \mathbb{N}$ such that $(M^n)_{1,2} = r$?

This problem asks whether there exists n such that the probability to go from State 1 to State 2 in n steps is exactly r. This is quite different from asking for the probability to go from State 1 to State 2 in any number of steps. Whereas the latter quantity can be computed in polynomial time by solving a system of linear equations, the Markov Reachability Problem is not known even to be decidable.

In Section 3 we observe that the Markov Reachability Problem can be encoded in the model checking frameworks of [1,2,4]. An inequality variant of the problem, asking for n such that $(M^n)_{1,2} > r$, is essentially the threshold problem for unary probabilistic automata [9], whose decidability is also open.

The paper [4] notes the close resemblance of the Markov Reachability Problem with the Skolem Problem in number theory and raises the question of whether the latter can be reduced to the Markov Reachability Problem.

* Corresponding author.

http://dx.doi.org/10.1016/j.ipl.2014.08.013
0020-0190/

Skolem Problem. Given a $k \times k$ integer matrix M, does there exist n such that $(M^n)_{1,2} = 0$?

The closely related *Positivity Problem* [8] asks whether there exists n such that $(M^n)_{1,2} > 0$.[1] There is a straightforward reduction of the Skolem Problem to the Positivity Problem (which however does not preserve the dimension of the matrices involved).

The Skolem and Positivity Problems have been the subject of much study, and their decidability has been open for several decades. Currently the Skolem Problem is only known to be decidable for matrices of dimension at most 4 (see, e.g., [7,11]) while the Positivity Problem is known only to be decidable up to dimension 5 (cf. [8]). Moreover for matrices of dimension 6 a decision procedure for the Positivity Problem would necessarily entail significant new results in Diophantine approximation—specifically the computability of the Lagrange constants of a general class of transcendental numbers [8].

While the Markov Reachability Problem and the Skolem Problem are very similar in form, the well-behaved spectral theory of stochastic matrices might lead one to conjecture that the former is more tractable. However in this note we give a reduction of the Skolem Problem to the Markov Reachability Problem. The same reduction transforms the Positivity Problem to the inequality version of the Markov Reachability Problem. In conjunction with the above-mentioned results of [8], this entails that the computability of some of the most basic problems in probabilistic verification will require significant advances in number theory.

2. Main result

In this section we give a polynomial-time reduction of the Skolem Problem to the Markov Reachability Problem. This is accomplished in two steps via the following intermediate problem:

Problem A. Given a $k \times k$ stochastic matrix M and column vector $\boldsymbol{y} \in \{0,1,2\}^k$, does there exist n such that $\boldsymbol{e}^T M^n \boldsymbol{y} = 1$, where $\boldsymbol{e} = (1, 0, \ldots, 0)^T$.

Thinking of M as the transition matrix of a Markov chain, Problem A asks if there exists n such that, starting from state 1, the state distribution $\boldsymbol{w}$ after n steps satisfies $\boldsymbol{w}^T \boldsymbol{y} = 1$.

Proposition 1. *The Skolem Problem can be reduced in polynomial time to Problem A.*

Proof. Given a $k \times k$ integer matrix $M = (m_{ij})$, we construct a stochastic $(2k+1) \times (2k+1)$ matrix $\widetilde{P}$ and a vector $\widetilde{\boldsymbol{v}} \in \{0,1,2\}^{2k+1}$, such that for all $n \in \mathbb{N}$, $(M^n)_{1,2} = 0$ if and only if $\widetilde{\boldsymbol{e}}^T \widetilde{P}^n \widetilde{\boldsymbol{v}} = 1$, where $\widetilde{\boldsymbol{e}}$ is the $(2k+1)$-dimensional coordinate vector $(1, 0, \ldots, 0)^T$.

Let P be a $2k \times 2k$ matrix of non-negative integers obtained by replacing each entry m_{ij} of M by the symmetric matrix $\left(\begin{smallmatrix} p_{ij} & q_{ij} \\ q_{ij} & p_{ij} \end{smallmatrix}\right)$, where $p_{ij} := \max(m_{ij}, 0)$ and $q_{ij} := \max(-m_{ij}, 0)$ satisfy $p_{ij} - q_{ij} = m_{i,j}$.

The map φ sending $\left(\begin{smallmatrix} a & b \\ b & a \end{smallmatrix}\right)$ to $a - b$ is a homomorphism from the ring of 2×2 symmetric integer matrices to $\mathbb{Z}$. By definition of P, partitioning P into 2×2 blocks and applying φ to each block one obtains M. Since matrix products can be computed block-wise and φ is a homomorphism, it follows that applying φ to each 2×2 sub-block of P^n one obtains the matrix M^n. Thus $(M^n)_{1,2} = \boldsymbol{e}^T P^n \boldsymbol{v}$, where $\boldsymbol{e} = (1, 0, \ldots, 0)^T$ and $\boldsymbol{v} = (0, 0, 1, -1, 0, \ldots, 0)^T$ are $2k$-dimensional vectors.

Since P is non-negative, there exists a non-negative scalar $s \in \mathbb{Q}$ such that sP is *sub-stochastic*, i.e., the sum of the entries in each row is at most one. Now define a $(2k+1)$-dimensional matrix $\widetilde{P}$ and vectors $\widetilde{\boldsymbol{e}}, \widetilde{\boldsymbol{v}}$ by

$$\widetilde{\boldsymbol{e}} = \begin{pmatrix} \boldsymbol{e} \\ 0 \end{pmatrix} \qquad \widetilde{P} = \begin{pmatrix} sP & \mathbf{1} - sP\mathbf{1} \\ 0 & 1 \end{pmatrix} \qquad \widetilde{\boldsymbol{v}} = \begin{pmatrix} \boldsymbol{v} \\ 0 \end{pmatrix} + \mathbf{1},$$

where $\mathbf{1} = (1, \ldots, 1)^T$ denotes a column vector of 1's of the appropriate dimension. The rightmost column of $\widetilde{P}$ is defined to make $\widetilde{P}$ a stochastic matrix.

Since $\widetilde{P}^n$ is stochastic for each $n \in \mathbb{N}$, we have that

$$\begin{aligned} \widetilde{\boldsymbol{e}}^T \widetilde{P}^n \widetilde{\boldsymbol{v}} &= \widetilde{\boldsymbol{e}}^T \widetilde{P}^n \begin{pmatrix} \boldsymbol{v} \\ 0 \end{pmatrix} + \widetilde{\boldsymbol{e}}^T \widetilde{P}^n \mathbf{1} = \boldsymbol{e}^T (sP)^n \boldsymbol{v} + 1 \\ &= s^n \left(M^n\right)_{1,2} + 1. \end{aligned}$$

From this we conclude that $(M^n)_{1,2} = 0$ iff $\widetilde{\boldsymbol{e}}^T \widetilde{P}^n \widetilde{\boldsymbol{v}} = 1$. □

The next step shows how the vector $\widetilde{\boldsymbol{v}}$ can be made into a coordinate vector.

Proposition 2. *Problem A can be reduced in polynomial time to the Markov Reachability Problem.*

Proof. Given k-dimensional vectors $\boldsymbol{e} = (1, 0, \ldots, 0)^T$ and $\boldsymbol{y} \in \{0,1,2\}^k$, and a $k \times k$ stochastic matrix Q, we construct a $2k+3$-dimensional stochastic matrix $\widetilde{Q}$ such that $\boldsymbol{e}^T Q^n \boldsymbol{y} = 1$ if and only if $(\widetilde{Q}^{2n+1})_{1,2k+1} = \frac{1}{4}$ for all $n \in \mathbb{N}$. In addition, the construction of $\widetilde{Q}$ is such that for all n, $(\widetilde{Q}^{2n})_{1,2k+1} = 0$, and thus by rearranging the rows and columns of $\widetilde{Q}$ we get an instance of the Markov Reachability Problem.

We first give an informal description of $\widetilde{Q}$, making reference to the example in Fig. 1. Thinking of Q as the transition matrix of a Markov chain, the idea is that $\widetilde{Q}$ contains two copies of each state of Q (the circle and square states in Fig. 1). Each transition of Q is split into a length-two path in $\widetilde{Q}$ connecting two circle states via an intermediate square state. Thus the underlying transition graph of $\widetilde{Q}$ is bipartite. We also create a new bottom strongly connected component in $\widetilde{Q}$ with three states (states a, b and c in Fig. 1). The transition weights from Q are halved in $\widetilde{Q}$, with half of the mass in each transition redirected to the new bottom strongly connected component according to the final-state vector $\boldsymbol{y}$. Looking at Fig. 1, the total mass entering state a from the shaded region in

[1] Strictly speaking [8] defines the Positivity Problem to be the complement of the problem stated here. Since we are interested in questions of decidability the difference is inconsequential.