Journal of Computer and System Sciences 80 (2014) 744–765

Contents lists available at ScienceDirect

Journal of Computer and System Sciences

www.elsevier.com/locate/jcss

Infinite vs. finite size-bounded randomized computations ☆

Richard Královič [a,b,*]

[a] Department of Computer Science, ETH Zurich, Switzerland
[b] Google Zurich, Switzerland

ARTICLE INFO

Article history:
Received 4 August 2011
Received in revised form 5 October 2013
Accepted 22 November 2013
Available online 6 December 2013

Keywords:
Probabilistic computations
Las Vegas randomization
Time complexity
Size complexity
Rotating finite automata
Sweeping finite automata

ABSTRACT

Randomized computations can be very powerful with respect to space complexity, e.g., for logarithmic space, LasVegas is equivalent to nondeterminism. This power depends on the possibility of infinite computations, however, it is an open question if they are necessary. We answer this question for rotating finite automata (RFAS) and sweeping finite automata (SFAS). We show that LasVegas RFAS (SFAS) allowing infinite computations, although only with probability 0, can be exponentially smaller than LasVegas RFAS (SFAS) forbidding them. In particular, we show that even RFAS (SFAS) with linear expected running time may require exponentially more states than RFAS (SFAS) running in exponential time. We also strengthen this result, showing that the restriction on time cannot be traded for the more powerful bounded-error randomization. To prove our results, we introduce a technique for proving lower bounds on size of RFAS (SFAS) that generalizes the notion of generic strings discovered by M. Sipser.

1. Introduction

The comparative study of probabilistic complexity classes is an important part of theoretical computer science. Results of such kind are important to obtain a deeper understanding of the power of randomness. One of the most natural questions about randomization is to compare probabilistic complexity classes of Turing machines to deterministic and nondeterministic ones. Several prominent open problems, such as *"Does ZPP* (*zero probability of error, polynomial time*) *equal P*?", *"Does BPP* (*bounded probability of error, polynomial time*) *equal NP*?" belong here.

Known results about probabilistic complexity classes suggest that randomized computations can be very powerful with respect to space complexity. For example, zero probability of error (LasVegas) Turing machines working in logarithmic space are as powerful as nondeterministic ones [13,18]. The core of this power is the possibility of infinite computations, even if their probability tends to 0. The main open problem posed in this context is whether restricting computations to finite ones restricts the power of randomized machines or not [16].

In this paper, we analyze the simplest models of LasVegas machines where the possibility of infinite computations makes sense – restricted versions of two-way finite automata. In this context, we prove an exponential gap in the size complexity between machines with and without infinite computations.

Various models of finite automata (FAS) are often used to analyze space-bounded computations. The space complexity of the analyzed computation directly translates into the size of the finite automaton, measured by its number of states. On one hand, the *one-way* FAS, which are the simplest models of FAS, are well understood. The exponential gap in the number of states between deterministic and nondeterministic one-way FAS was discovered long ago. Randomized FAS have been

☆ This work was supported by the Swiss National Science Foundation grants 200020-120073 and 200021_132510. A preliminary version of this paper was published as [11] and were part of [12].
* Correspondence to: ETH Zürich, Information Technology and Education, CAB F 15.1, Universitätstrasse 6, 8092 Zürich, Switzerland.

0022-0000/$ – see front matter
http://dx.doi.org/10.1016/j.jcss.2013.11.004

introduced in [14] and more recently studied in [6,7]. In [6], it was proven that deterministic and LasVegas one-way FAS are equivalent within a quadratic blow-up in their number of states. Since one-way FAS do not possess the ability of infinite computations, this fact does not contradict our intuition that infinite computations are crucial for the efficiency of LasVegas machines.

On the other hand, *two-way* FAS are a quite complex model – e.g., the relationship between determinism and nondeterminism in two-way FAS (raised in [15]) is a long-standing open problem. To fill the gap between the one-way and two-way FAS, the *sweeping automata* (SFAS) were introduced in [17]: an SFA is a restricted form of a two-way FA that can change the direction of the head motion only at the end-markers. It was proven in [17] that nondeterministic SFAS can be exponentially more succinct than deterministic SFAS; more recently, an exponential gap between nondeterministic and co-nondeterministic SFAS was proven in [8].

The model of sweeping automata is one of the simplest models of randomized machines that allows to use infinite computations. (More precisely, we should talk about computations of arbitrary length, since the probability of any infinite computation is 0, provided that the expected running time of the machine is finite.) In [9], it was proven that LasVegas SFAS can be exponentially more succinct than deterministic SFAS. However, this result relies heavily on LasVegas SFAS with *exponential* expected running time. It is not difficult to see that if computations of superlinear length are possible for some SFA, there is a non-zero probability that this SFA runs in a loop. Hence arbitrarily long computations are possible for any SFA with exponential expected running time. The natural question is if randomization can be of any help if we forbid arbitrarily long computations.

In this paper, we give a partial negative answer to this question. In particular, we show that LasVegas SFAS running in exponential expected time (i.e., those with arbitrarily long computations allowed) can be exponentially more succinct than LasVegas SFAS running in linear expected time. Since LasVegas SFAS without arbitrarily long computations always work in linear time, our result confirms the conjecture that the possibility of infinite computations is essential to the efficiency of LasVegas randomization, at least in the restricted case of sweeping automata.

The time complexity of randomized finite automata has been analyzed e.g., in [1]. There, it has been proven that exponential expected running time is necessary for two-way bounded-error finite automata to accept non-regular languages, i.e., avoiding the possibility of infinite computations reduces the power of the automata. To the best of our knowledge, no analogous results are known for LasVegas randomization. Since LasVegas finite automata always accept regular languages, we take size complexity into consideration. Our results show that avoiding infinite computations introduces exponential blow-up in the size complexity for LasVegas automata.

Apart from sweeping automata, we consider also their simplified version – the *rotating automata* (RFAS) [15,10]. RFAS differ from SFAS in that they are not able to read the input from right to left. After reaching the right end-marker, their head is positioned again at the first symbol of the input. Our motivation to consider RFAS is twofold: On one hand, they allow us to present the core ideas in a simpler way and often it is rather straightforward to generalize arguments about RFAS to SFAS. On the other hand, we show that the power of randomized RFAS is related to one-way randomized automata, which are a very natural model. To have an analogous relation for SFAS, we need to introduce a rather artificial modification of one-way machines.

To obtain lower bounds on size complexity of randomized FAS with restricted running time, we use a technique that generalizes the ideas of [17] in such a way that makes it possible to prove such bounds by purely algebraic arguments. Hence, the contribution of the paper is also in introducing proof methods whose further development could help the original problem in the general setting.

Preliminary version of this paper was published as [11]. Results presented here are part of [12] as well.

1.1. Organization of the paper

The paper is organized as follows: We introduce the models of finite automata, as well as other notation needed in our proofs, in Section 2. In Section 3, we show how is it possible to construct languages that are hard for one-way Monte-Carlo automata out of languages that are hard for one-way deterministic automata. In Section 4, we show how to use languages that are hard for one-way Monte-Carlo automata to construct languages that are hard for rotating bounded-error automata running in linear time. Later we extend this result to SFAS in Section 5. In Section 6, we complement our lower bounds on size complexity of RFAS and SFAS running in linear time by upper bounds for automata without restriction on running time. Finally, we apply the presented results to conclude an exponential gap in the size complexity between randomized RFAS (SFAS) running in linear time and those running in exponential time.

2. Preliminaries

We say that a matrix is *substochastic* if and only if all its values are nonnegative and every row has sum at most 1. A matrix is *stochastic* if and only if it is nonnegative and every row has sum equal to 1. A $k \times k$ matrix is *a permutation matrix* if and only if all its values are either 0 or 1 and in every row and every column there is exactly one value 1.

If Σ is an alphabet, then Σ^* is the set of all finite strings over Σ. If $z \in \Sigma^*$, then $|z|$ is its length, and z^{R} is its reverse. A *language* over Σ is any $L \subseteq \Sigma^*$; $\bar{L}$ is its complement. An automaton *solves* (*accepts*, or *recognizes*) a language if and only