Available online at www.sciencedirect.com
SCIENCE DIRECT®

Theoretical Computer Science 330 (2005) 349–360

Theoretical Computer Science

www.elsevier.com/locate/tcs

Complementing unary nondeterministic automata☆

Filippo Mera, Giovanni Pighizzini*

Dipartimento di Informatica e Comunicazione, Università degli Studi di Milano, via Comelico 39, 20135 Milano, Italy

Received 14 November 2003; received in revised form 7 April 2004; accepted 23 April 2004

Abstract

We compare the nondeterministic state complexity of unary regular languages and that of their complements: if a unary language $\mathscr{L}$ has a succinct nondeterministic finite automaton, then nondeterminism is useless in order to recognize its complement, namely, the smallest nondeterministic automaton accepting the complement of $\mathscr{L}$ has as many states as the minimum deterministic automaton accepting it. The same property does not hold in the case of automata and languages defined over larger alphabets. We also show the existence of infinitely many unary regular languages for which nondeterminism is useless in their recognition and in the recognition of their complements.

Keywords: Formal languages; Finite state automata; Unary languages; Descriptional complexity; Complementation

1. Introduction

In the last few years, we observed a renewed interest for research in automata theory (for a discussion, we address the reader to [9,23]).

Some aspects of this field that were only partially considered in the early developments of the theory are now extensively and deeply investigated. Two such aspects are the descriptional complexity of automata and the analysis of unary languages.

☆ Partially supported by MIUR, under the projects “Complessità descrizionale di automi e strutture correlate” and “Linguaggi formali: metodi, modelli e applicazioni”.

* Corresponding author.
E-mail addresses: mera@dico.unimi.it (F. Mera), pighizzi@dico.unimi.it (G. Pighizzini).

0304-3975/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.tcs.2004.04.015

Descriptional complexity studies the costs of the description of objects (e.g., languages) by different formal systems (e.g., deterministic or nondeterministic automata, grammars, etc.). Furthermore, in this area formal systems with the same expressive power are compared with respect to their conciseness. Probably the first and most widely known result of this kind is the simulation of nondeterministic finite automata (nfa) by deterministic ones (dfa). While these two models share the same computational power, i.e., they characterize the class of regular languages [21], from the point of view of descriptional complexity they are very different. In fact, nfa's can be exponentially more succinct than dfa's, i.e., for any nonnegative integer n there is a language accepted by an n-state nfa that *requires* 2^n states to be accepted by a dfa [17,18]. For a recent survey on descriptional complexity, we address the reader to [4], while for descriptional complexity of regular languages some recent works are [24,3,7].

Languages and automata are called *unary* when they are defined over a one-letter alphabet. Some interesting properties of unary automata were studied by Chrobak [2], showing many important differences with respect to the general, i.e., "nonunary" case. These studies have been recently deepened, mainly in connection with descriptional complexity issues (see, e.g., [3,6,16,19,20]).

This paper continues this stream of research. In particular, we study some questions related to the complement operation. While it is trivial to show that any regular language and its complement can be recognized by dfa's with the same number of states, in the case of nfa's the situation can be totally different: the only known standard way to achieve complementation is to convert an nfa accepting a given language to a dfa, and then complement the set of final states. This may lead to an increment from n to 2^n in the number of states. This gap was originally proved to be optimal for a four-letter alphabet by Birget [1] and, recently, for a two-letter alphabet by Jirásková [11], who showed that for each integer n there exists a language $\mathcal{L}$ accepted by an nfa with n states such that *each nfa* accepting the complement of $\mathcal{L}$ has at least 2^n states. In this paper, we consider the same problem for unary languages. Our main result states that if a unary regular language $\mathcal{L}$ has a succinct nfa (i.e., an nfa with the minimum possible number of states with respect to the periodicity of the accepted language [10]), then nondeterminism is useless in order to recognize its complement, i.e., each nfa accepting $\mathcal{L}^c$ has at least the same number of states as the minimum dfa accepting it. Roughly speaking, succinct nfa's witness the optimality of the simulation of n-state unary nfa's by $e^{\Theta(\sqrt{n \ln n})}$-state dfa's proved in [2]. A similar result does not hold in the general case. In fact, we prove that, without the unary restriction, for each integer $n \geqslant 2$, there is a language $\mathcal{L}$ accepted by an nfa with n states such that the minimum dfa accepting it must have 2^n states (i.e., $\mathcal{L}$ is a witness of the state gap between nfa's and dfa's in the general case) and the complement of $\mathcal{L}$ is accepted by an nfa with $n+1$ states. In other words, in the general case there are languages with small nfa's accepting them and their complements.

We further deepen this investigation by proving the existence of unary languages for which nondeterminism is useless in their recognition and in the recognition of their complements. In particular, we show that for each integer $n \geqslant 2$, there exists a language $\mathcal{L}$ such that all nfa's accepting either $\mathcal{L}$ or its complement must have at least the same number of states n as the minimum dfa's accepting them.