Journal of Automata, Languages and Combinatorics **2** (1997) 3, 177–186

NFA TO DFA TRANSFORMATION FOR FINITE LANGUAGES OVER ARBITRARY ALPHABETS [1]

Kai Salomaa and Sheng Yu
Department of Computer Science, University of Western Ontario
London, Ontario, Canada N6A 5B7
e-mail: {ksalomaa,syu}@csd.uwo.ca

ABSTRACT
We consider the number of states of a DFA that is equivalent to an n-state NFA accepting a ***finite language*** over an arbitrary alphabet. We show that, for any n-state NFA accepting a finite language over a k-letter alphabet, $n, k > 1$, there is an equivalent DFA of $O(k^{n/(\log_2 k+1)})$ states, and show that this bound is optimal in the worst case.

Keywords: formal languages, finite automata, state complexity.

1. Introduction

It is well-known that for each positive integer n, there exists a regular language L such that L is accepted by an n-state NFA and any complete DFA accepting L requires at least 2^n states [3]. However, the same statement is not true if L is required to be finite. In [2], Mandl showed that for each n-state NFA accepting a finite language over *a two-letter alphabet*, there exists an equivalent DFA which has $O(2^{\frac{n}{2}})$ states; more specifically, no more than $2^{\frac{n}{2}+1} - 1$ states if n is even and $3 \cdot 2^{\lfloor \frac{n}{2} \rfloor} - 1$ states if n is odd. In [2], it was also shown that these bounds are optimal in the worst case. However, there have been no corresponding results concerning finite languages over an arbitrary k-letter alphabet, $k \geq 2$. Also, the proofs in [2] for the two-letter alphabet case are rather sketchy.

In this paper, we first give detailed proofs for the two-letter alphabet case. Then, as the main result of this paper, we give the optimal upper-bounds for the general cases of the problem, i. e., for the cases where finite languages are over an arbitrary k-letter alphabet, $k \geq 2$. Specifically, we show that for any n-state NFA accepting a finite language over a k-letter alphabet, $k \geq 2$, we can construct an equivalent DFA of $O(k^{n/(\log_2 k+1)})$ states; and we show that for each k-letter alphabet, $k \geq 2$, and for each $n \geq 2$, there exists a finite language L accepted by an n-state NFA such that the number of states of any DFA accepting L is $\Omega(k^{n/(\log_2 k+1)})$. One may observe that this bound, as a function of k, approaches to 2^n when k becomes larger.

[1]This research is supported by the Natural Sciences and Engineering Research Council of Canada grants OGP0041630.