LENGTH-REDUCING AUTOMATA (ALMOST) WITHOUT AUXILIARY SYMBOLS[1]

ARTUR JEŻ and TOMASZ JURDZIŃSKI

Institute of Computer Science, University of Wrocław
Joliot-Curie 15, PL-50-383 Wrocław, Poland
e-mail: aje@cs.uni.wroc.pl, tju@ii.uni.wroc.pl

ABSTRACT

Length-reducing two-pushdown automaton (lr-tpda) is a machine model for growing context-sensitive languages and Church-Rosser languages, the language classes complementing and refining the Chomsky hierarchy. These automata are closely related to restarting automata, an analytic model for some natural language processing techniques. While the latter model was considered in numerous variants depending on usage of the auxiliary symbols, lr-tpdas were always assumed to use the auxiliary symbols in non-limited way.

We study lr-tpdas with limited usage of auxiliary symbols. We show that, the deterministic automata without auxiliary symbols can recognize all deterministic context-free languages. Moreover, non-deterministic automata using one extra alphabet symbol can recognize all context-free languages. Next, we inspect automata over one-letter alphabet, without auxiliary symbol. We show that even this restricted variant of lr-tpdas is quite powerful. Finally, we show that most of our constructions are stateless.

1. Introduction

Context-free languages (CFL) form possibly the most intensively studied class in formal language theory. Among various reasons of its popularity are very natural characterizations by grammars and automata. Since expressiveness of these formalisms is too weak for many applications, their extensions were investigated over the years. One of such extensions is based on the machine model for CFL, i.e., the pushdown automata. An appealing generalization is an automaton with two stacks. Since this model is equivalent to a Turing machine (TM) in its full generality, it is natural to impose some additional restrictions on it. A length-reducing two-pushdown automaton (length-reducing tpda or lr-tpda) is such a restricted variant possessing several advantageous properties: for instance, lr-tpdas accept exactly the growing context-sensitive languages (GCSL), and their deterministic counterparts characterize

---