

POSITION AUTOMATA FOR SEMI-EXTENDED EXPRESSIONS

SABINE BRODA ANTONÍO MACHIAVELO NELMA MOREIRA ROGÉRIO REIS

CMUP & DCC & DM

Faculdade de Ciências da Universidade do Porto

Rua do Campo Alegre 1021, 4219-007, Portugal

`sbb@dcc.fc.up.pt, ajmachia@fc.up.pt, {nam,rvr}@dcc.fc.up.pt`

ABSTRACT

Positions and derivatives are two essential notions in the conversion methods from regular expressions to equivalent finite automata. Partial derivative based methods have recently been extended to regular expressions with intersection (semi-extended). In this paper, we present a position automaton construction for those expressions. This construction generalizes the notion of position, making it compatible with intersection. The resulting automaton is homogeneous and has the partial derivative automaton as a quotient.

Keywords: regular expressions, nondeterministic finite automata, position automaton, intersection

1. Introduction

The position automaton (\mathcal{A}_{pos}), introduced by Glushkov [14], permits the conversion of a simple regular expression (involving only the union, concatenation and star operations) into an equivalent nondeterministic finite automaton (NFA) without ε -transitions. The states in the position automaton correspond to the positions of letters in the corresponding regular expression plus an additional initial state. McNaughton and Yamada [17] also used the positions of a regular expression to define an automaton, however they directly computed a deterministic version of the position automaton. The position automaton has been well studied [4, 10] and it is considered the *standard* automaton simulation of a regular expression [18]. Some of its interesting properties are: homogeneity, i. e., for each state, all in-transitions have the same label (letter); whenever deterministic, these automata characterize certain families of unambiguous regular expressions; and can be computed in quadratic time [6]; other

This work was partially supported by CMUP (UID/MAT/00144/2013), which is funded by FCT (Portugal) with national (MEC) and European structural funds through the programs FEDER, under the partnership agreement PT2020.