Journal of Automata, Languages and Combinatorics **7** (2002) 1, 109–125

VALID IDENTITY PROBLEM FOR SHUFFLE REGULAR EXPRESSIONS

Albert R. Meyer
MIT Laboratory for Computer Science
Cambridge, MA 02139, USA
e-mail: `meyer@lcs.mit.edu`

and

Alexander Rabinovich
Department of Computer Science, The Sackler School of Exact Sciences
Tel Aviv University, Israel 69978
e-mail: `rabino@math.tau.ac.il`

ABSTRACT

We show that the following problem is decidable: given are expressions E_1 and E_2 constructed from variables by the regular operations and shuffle. Is the identity $E_1 = E_2$ true for all instantiations of its variables by formal languages?

Keywords: Decidability, valid equations, shuffle regular expressions

1. Introduction

We consider valid identity decision problem for shuffle regular expressions. Namely, given are expressions E_1 and E_2 constructed from variables by the regular operations and shuffle. Is the identity $E_1 = E_2$ true for all instantiations of its variables by formal languages?

For example, the identity $(X^*Y^*)^* = (X+Y)^*$ is true because for all languages L_1 and L_2, the languages $(L_1^*L_2^*)^*$ and $(L_1+L_2)^*$ are the same.

The above identity contains only regular operations: concatenation, union and iteration. An easy 'folk' theorem (see e.g., [10] Exercise 14, Chapter 3) shows that the validity of an identity over regular operations can be verified by instantiating the language variables as single letters. For example, in order to check the validity of $(X^*Y^*)^* = (X+Y)^*$ we instantiate the variables X and Y by single letters a and b and verify that $(a^*b^*)^* = (a+b)^*$. Checking this variable-free identity is a routine matter of checking equivalence of finite state automata.

In concurrency a very important role is played by parallel composition operators. The simplest of these operators is the non-communicating parallel connective $||$, corresponding to shuffle of languages. The above folk theorem fails for the expressions