Journal of Automata, Languages and Combinatorics **23** (2018) 1–3, 67 – 101

TOWARDS A THEORY OF COMPLEXITY OF REGULAR LANGUAGES

Janusz A. Brzozowski

David R. Cheriton School of Computer Science, University of Waterloo
Waterloo, ON, Canada N2L 3G1
brzozo@uwaterloo.ca

ABSTRACT

We survey recent results concerning the complexity of regular languages represented by their minimal deterministic finite automata. In addition to the quotient complexity of the language – which is the number of its (left) quotients, and is the same as its state complexity – we also consider the size of its syntactic semigroup and the quotient complexity of its atoms – basic components of every regular language. We then turn to the study of the quotient/state complexity of common operations on regular languages: reversal, (Kleene) star, product (concatenation) and boolean operations. We examine relations among these complexity measures. We discuss several subclasses of regular languages defined by convexity. In many, but not all, cases there exist "most complex" languages, languages satisfying all these complexity measures.

Keywords: atom, boolean operation, complexity measure, concatenation, convex language, most complex language, quotient complexity, regular language, reversal, star, state complexity, syntactic semigroup, unrestricted complexity

1. Introduction

We assume the reader is familiar with basic properties of regular languages and finite automata, as discussed in [58, 66], for example; formal definitions are given later.

We study the complexity of regular languages represented by their minimal deterministic finite automata (DFAs). The number of states in the minimal DFA of a language is its *state complexity* [54, 67]; this number is used as a first measure of complexity. But languages having the same state complexity can be quite simple or very complex. How do we decide whether one language is more complex than another? In this respect, the size of the syntactic semigroup of the language – which is isomorphic to the transition semigroup of its minimal DFA – appears to be a good measure.

Another way to distinguish two regular languages of the same state complexity is by comparing how difficult it is to perform operations on these languages. The state complexity of a regularity preserving unary operation on a language is defined

This work was supported by the Natural Sciences and Engineering Research Council of Canada grant No. OGP0000871.