Syntactic Complexity of $\mathcal{R}$- and $\mathcal{J}$-Trivial Regular Languages*

Janusz Brzozowski and Baiyu Li**

David R. Cheriton School of Computer Science, University of Waterloo
Waterloo, ON, Canada N2L 3G1
{brzozo,b5li}@uwaterloo.ca

Abstract. The syntactic complexity of a subclass of the class of regular languages is the maximal cardinality of syntactic semigroups of languages in that class, taken as a function of the state complexity $n$ of these languages. We prove that $n!$ and $\lceil e(n-1)! \rceil$ are tight upper bounds for the syntactic complexity of $\mathcal{R}$- and $\mathcal{J}$-trivial regular languages, respectively.

Keywords: finite automaton, $\mathcal{J}$-trivial, monoid, regular language, $\mathcal{R}$-trivial, semigroup, syntactic complexity.

1 Introduction

The state complexity of a regular language $L$ is the number of states in the minimal deterministic finite automaton (DFA) accepting $L$. An equivalent notion is quotient complexity, which is the number of distinct left quotients of $L$. The syntactic complexity of $L$ is the cardinality of the syntactic semigroup of $L$. Since the syntactic semigroup of $L$ is isomorphic to the semigroup of transformations performed by the minimal DFA of $L$, it is natural to consider the relation between syntactic complexity and state complexity. The syntactic complexity of a subclass of regular languages is the maximal syntactic complexity of languages in that class, taken as a function of the state complexity of these languages.

Here we consider the classes of languages defined using the well-known Green equivalence relations on semigroups [13]. Let $M$ be a monoid, that is, a semigroup with an identity, and let $s, t \in M$ be any two elements of $M$. The Green relations on $M$, denoted by $\mathcal{L}, \mathcal{R}, \mathcal{J}$ and $\mathcal{H}$, are defined as follows: $s \mathcal{L} t \iff Ms = Mt$, $s \mathcal{R} t \iff sM = tM$, $s \mathcal{J} t \iff MsM = MtM$, and $s \mathcal{H} t \iff s \mathcal{L} t$ and $s \mathcal{R} t$. For $\rho \in \{\mathcal{L}, \mathcal{R}, \mathcal{J}, \mathcal{H}\}$, $M$ is $\rho$-trivial if and only if $(s, t) \in \rho$ implies $s = t$ for all $s, t \in M$. A language is $\rho$-trivial if and only if its syntactic monoid is $\rho$-trivial. In this paper we consider only regular $\rho$-trivial languages. $\mathcal{H}$-trivial regular languages are exactly the star-free languages [13], and $\mathcal{L}$-, $\mathcal{R}$-, and $\mathcal{J}$-trivial regular languages are exactly the trivalent languages [13].

* This work was supported by the Natural Sciences and Engineering Research Council of Canada under grant No. OGP0000871 and a Postgraduate Scholarship.
** Present address: Optumsoft, Inc., 275 Middlefield Rd, Suite 210, Menlo Park, CA 94025, USA.

© Springer-Verlag Berlin Heidelberg 2013
languages are all subclasses of the class of star-free languages. The class of $J$-trivial languages is the intersection of the classes of $R$- and $L$-trivial languages.

A language $L \subseteq \Sigma^*$ is piecewise-testable if it is a finite boolean combination of languages of the form $\Sigma^*a_1\Sigma^* \cdots \Sigma^*a_l\Sigma^*$, where $a_i \in \Sigma$. Simon [15,16] proved in 1972 that a language is piecewise-testable if and only if it is $J$-trivial. A biautomaton is a finite automaton which can read the input word alternatively from left and right. In 2011 Klíma and Polák [9] showed that a language is piecewise-testable if and only if it is accepted by an acyclic biautomaton; here self-loops are allowed, as they are not considered cycles.

In 1979 Brzozowski and Fich [1] proved that a regular language is $R$-trivial if and only if its minimal DFA is partially ordered, that is, it is acyclic as above. They also showed that $R$-trivial regular languages are finite boolean combinations of languages $\Sigma_i^*a_1\Sigma^* \cdots \Sigma_i^*a_l\Sigma^*$, where $a_i \in \Sigma$ and $\Sigma_i \subseteq \Sigma \setminus \{a_i\}$. Recently Jirásková and Masopust proved a tight upper bound on the state complexity of reversal of $R$-trivial languages [8].

In the past, the syntactic complexity of the following subclasses of regular languages was considered: In 1970 Maslov [11] noted that $n^n$ was a tight upper bound on the number of transformations performed by a DFA of $n$ states. In 2003–2004, Holzer and König [7], and Krawetz, Lawrence and Shallit [10] studied unary and binary languages. In 2010 Brzozowski and Ye [5] examined ideal and closed regular languages. In 2012 Brzozowski, Li and Ye studied prefix-, suffix-, bifix-, and factor-free regular languages [3], Brzozowski and Li [2] considered the class of star-free languages and three of its subclasses, and Brzozowski and Liu [4] studied finite/cofinite, definite, and reverse definite languages, where $L$ is definite (reverse-definite) if it can be decided whether a word $w$ belongs to $L$ by examining the suffix (prefix) of $w$ of some fixed length.

We state basic definitions and facts in Section 2. In Sections 3 and 4 we prove tight upper bounds on the syntactic complexities of $R$- and $J$-trivial regular languages, respectively. Section 5 concludes the paper. Omitted proofs can be found at http://arxiv.org/abs/1208.4650.

2 Preliminaries

Let $Q$ be a non-empty finite set with $n$ elements, and assume without loss of generality that $Q = \{1,2,\ldots,n\}$. There is a linear order on $Q$, namely the natural order $<$ on integers. If $X$ is a non-empty subset of $Q$, then the maximal element in $X$ is denoted by $\max(X)$. A partition $\pi$ of $Q$ is a collection $\pi = \{X_1,X_2,\ldots,X_m\}$ of non-empty subsets of $Q$ such that $Q = X_1 \cup X_2 \cup \cdots \cup X_m$, and $X_i \cap X_j = \emptyset$ for all $1 \leq i < j \leq m$. We call each subset $X_i$ a block in $\pi$. For any partition $\pi$ of $Q$, let $\operatorname{Max}(\pi) = \{\max(X) \mid X \in \pi\}$. The set of all partitions of $Q$ is denoted by $\Pi_Q$. We define a partial order $\preceq$ on $\Pi_Q$ such that, for any $\pi_1,\pi_2 \in \Pi_Q$, $\pi_1 \preceq \pi_2$ if and only if each block of $\pi_1$ is contained in some block of $\pi_2$. We say $\pi_1$ refines $\pi_2$ if $\pi_1 \preceq \pi_2$. The poset $(\Pi_Q,\preceq)$ is a finite lattice: For any $\pi_1,\pi_2 \in \Pi_Q$, the meet $\pi_1 \land \pi_2$ is the $\preceq$-largest partition that refines both $\pi_1$ and $\pi_2$, and the join $\pi_1 \lor \pi_2$ is the $\preceq$-smallest partition that is refined by both $\pi_1$ and $\pi_2$. From now on, we refer to the lattice $(\Pi_Q,\preceq)$ simply as $\Pi_Q$. 

Syntactic Complexity of $R$- and $J$-Trivial Regular Languages 161