

Maximally Atomic Languages*

Janusz Brzozowski

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

Gareth Davies

Department of Pure Mathematics, University of Waterloo
Waterloo, ON, Canada N2L 3G1
gdavies@uwaterloo.ca

The atoms of a regular language are non-empty intersections of complemented and uncomplemented quotients of the language. Tight upper bounds on the number of atoms of a language and on the quotient complexities of atoms are known. We introduce a new class of regular languages, called the *maximally atomic languages*, consisting of all languages meeting these bounds. We prove the following result: If L is a regular language of quotient complexity n and G is the subgroup of permutations in the transition semigroup T of the minimal DFA of L , then L is maximally atomic if and only if G is transitive on k -subsets of $\{1, \dots, n\}$ for $0 \leq k \leq n$ and T contains a transformation of rank $n - 1$.

Keywords: atom, átomaton, finite automaton, quotient complexity, regular language, set-transitive group, state complexity, transition semigroup

1 Introduction

The *state/quotient complexity* of a regular language is the number of states in the minimal deterministic finite automaton (DFA) of the language, or equivalently, the number of left quotients of the language. An *atom* of a regular language is a non-empty intersection of the language's left quotients, some of which may be complemented. Brzozowski and Tamm have found tight upper bounds on the number of atoms of a language [4] and on the quotient complexities of atoms [3]. This lets us define a new class of regular languages which we call *maximally atomic*: these are regular languages whose atoms meet these bounds.

The *transition semigroup* of a DFA is the semigroup of transformations induced by the transition function of the DFA on its set of states. Our main result (stated formally in Section 3) is the following relationship between maximally atomic languages and transition semigroups:

A regular language with quotient complexity n is maximally atomic if and only if the transition semigroup of its minimal DFA contains permutations that can map any subset of $\{1, \dots, n\}$ to any other subset of the same size, as well as at least one transformation with an image of size $n - 1$.

In the process of proving this, we establish several other relationships between transition semigroups and atoms; in particular, we give sufficient conditions for a language to have the maximal number of atoms, and necessary and sufficient conditions for certain individual atoms to have maximal complexity. We also derive a general formula for the transition functions of “átomata” (nondeterministic finite automata whose states correspond to the atoms of the language they recognize).

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

2 Definitions and Terminology

2.1 Partially Ordered Sets

A *partially ordered set* (poset) is a pair (S, \leq) where S is a set and \leq is a partial order on S . A *subposet* of (S, \leq) is a poset (T, \leq) such that $T \subseteq S$. We often abbreviate (S, \leq) to simply S .

If T is a subposet of S , then for $a, b \in S$, the *interval* of T between a and b , denoted $[a, b]_T$, is the set of all $t \in T$ such that $a \leq t$ and $t \leq b$. Note that if $b < a$, then the interval $[a, b]_T$ is empty.

Let $Q_n = \{1, 2, \dots, n\}$, let $P = (2^{Q_n}, \subseteq)$, and let X be a subposet of P . For each non-empty interval $[V, U]_X$, define the *type* of $[V, U]_X$ to be the pair of integers $(|\cap[V, U]_X|, |\cup[V, U]_X|)$. Let the type of the empty interval be $(-1, -1)$.

2.2 Transformations

A *transformation* of a set X is a mapping $t: X \rightarrow X$. Since we deal only with finite sets, we assume without loss of generality that $X = Q_n$ for some n . A *permutation* is an invertible (one-to-one and onto) transformation. A *singular transformation* is a non-invertible transformation.

If a transformation t maps i to j , we say the *image* of i under t is j and write $t(i) = j$. The image of $S \subseteq Q_n$ is $t(S) = \{t(i) \mid i \in S\}$. The image of t itself is $\text{im } t = t(Q_n)$. The *coimage* of t is $\text{coim } t = \overline{\text{im } t}$, where $\overline{S} = Q_n \setminus S$. The *preimage* of an element i under t is $t^{-1}(i) = \{j \mid t(j) = i\}$. The preimage of $S \subseteq Q_n$ under t is $t^{-1}(S) = \bigcup_{i \in S} t^{-1}(i)$. The *rank* of a transformation is $|\text{im } t|$. The *composition* or *product* of two transformations s and t is $s \circ t$, defined by $(s \circ t)(i) = s(t(i))$.

A *transposition* (i, j) for $i \neq j$ is a transformation such that $t(i) = j$, $t(j) = i$, and $t(\ell) = \ell$ for all $\ell \notin \{i, j\}$. A permutation is *even* if it can be written as a product of an even number of transpositions and it is *odd* otherwise. A *unitary transformation*, denoted by $(i \rightarrow j)$ (with $i \neq j$), is a transformation such that $t(i) = j$ and $t(\ell) = \ell$ for all $\ell \neq i$.

2.3 Semigroups, Monoids, and Groups

A *semigroup* is a pair (S, \cdot) , where S is a non-empty set and \cdot is an associative binary operation. We often abbreviate (S, \cdot) to S . A *monoid* $M = (M, \cdot, e)$ is a semigroup with identity e , and a *group* $G = (G, \cdot, e)$ is a monoid in which each element has an inverse. A *subsemigroup* of (S, \cdot) is a semigroup (T, \cdot) where $T \subseteq S$. If (S, \cdot, e) and (M, \cdot, e) are monoids with $M \subseteq S$, then M is a *submonoid* of S . A *subgroup* of S is a submonoid G of S such that G is a group.

The *full transformation semigroup* of degree n , denoted T_n , is the set of all transformations $t: Q_n \rightarrow Q_n$ under the binary operation \circ . Note that T_n is a monoid, since the identity transformation of Q_n acts as the identity element. The *symmetric group* of degree n , denoted by S_n , is the subgroup of permutations in T_n . A *transformation semigroup* of degree n is a subsemigroup of T_n , and a *permutation group* of degree n is a subgroup of S_n . A *conjugate* of a permutation group G of degree n is a group of the form $\{p \circ g \circ p^{-1} \mid g \in G\}$, where $p \in S_n$.

Let G be a permutation group of degree n and let X be a set. For $x \in X$, the *orbit* of x under G is the set $\{g(x) \mid g \in G\}$. We say that G *acts transitively* or *is transitive* on a set X if for all $x, y \in X$ there exists $g \in G$ such that $g(x) = y$, or equivalently, if G has only one orbit when it acts on X . We say G is *k-set-transitive* if it is transitive on the set of k -subsets (subsets of cardinality k) of Q_n . If G is k -set-transitive for $0 \leq k \leq n$, we say G is *set-transitive*.

The set-transitive permutation groups have been fully classified by Beaumont and Peterson [1]. In general, a set-transitive group is either the symmetric group S_n or the alternating group A_n (the subgroup of even permutations in S_n). When n is small there are four exceptions (up to conjugation):

Proposition 1. *A set-transitive permutation group of degree n is S_n or A_n or a conjugate of one of the following permutation groups:*

1. For $n = 5$, the affine general linear group $\text{AGL}(1,5)$.
2. For $n = 6$, the projective general linear group $\text{PGL}(2,5)$.
3. For $n = 9$, the projective special linear group $\text{PSL}(2,8)$.
4. For $n = 9$, the projective semilinear group $\text{P}\Gamma\text{L}(2,8)$.

2.4 Finite Automata

A *nondeterministic finite automaton (NFA)* is a tuple $\mathcal{N} = (Q, \Sigma, \eta, I, F)$, where Q is a finite, non-empty set of *states*, Σ is a finite, non-empty *alphabet*, $\eta: Q \times \Sigma \rightarrow 2^Q$ is a *transition function*, $I \subseteq Q$ is a set of *initial states*, and $F \subseteq Q$ is a set of *final states*. We extend η to $\eta: 2^Q \times \Sigma^* \rightarrow 2^Q$ as follows: for $S \subseteq Q$ and $w = xa$, $x \in \Sigma^*$, $a \in \Sigma$, we define $\eta(S, w)$ inductively by $\eta(S, \varepsilon) = S$ and $\eta(S, xa) = \eta(\eta(S, x), a) = \bigcup_{s \in \eta(S, x)} \eta(s, a)$. We define $\eta_w: 2^Q \rightarrow 2^Q$ by $\eta_w(S) = \eta(S, w)$.

A word w is *accepted* by \mathcal{N} if $\eta_w(I) \cap F \neq \emptyset$. The *language accepted* by \mathcal{N} is the set of all words accepted by \mathcal{N} . The *language of a state* $q \in Q$ is the language accepted by the modified NFA $\mathcal{N}_q = (Q, \Sigma, \eta, \{q\}, F)$. For $S, T \subseteq Q$, we say S is *reachable from* T in \mathcal{N} if there exists $w \in \Sigma^*$ such that $\eta_w(T) = S$. If S is reachable from I , we simply say S is *reachable* in \mathcal{N} . An NFA that accepts a language L is *minimal* if the number of states is minimal among all NFAs that accept L .

A *deterministic finite automaton (DFA)* is a tuple $\mathcal{D} = (Q, \Sigma, \delta, q_1, F)$, where Q , Σ and F have the same meaning as in an NFA, $\delta: Q \times \Sigma \rightarrow Q$ is a transition function, and $q_1 \in Q$ is an initial state. Since DFAs are special cases of NFAs, all the definitions above apply also to DFAs. While minimal NFAs need not be unique, there is a unique (up to isomorphism) minimal DFA for each regular language.

For all $w \in \Sigma^*$, $\delta_w: Q \rightarrow Q$ is a transformation of the set of states of \mathcal{D} ; we call this the *transformation induced by* w in \mathcal{D} . The *transition semigroup* of \mathcal{D} is the semigroup (T, \circ) , where $T = \{\delta_w \mid w \in \Sigma^+\}$. This is the semigroup of transformations of Q induced by non-empty words over Σ in \mathcal{D} .

For an NFA $\mathcal{N} = (Q, \Sigma, \eta, I, F)$, define the *reverse* of \mathcal{N} to be the NFA $\mathcal{N}^R = (Q, \Sigma, \eta^R, F, I)$, where $\eta^R(q, a) = \{p \in Q \mid q \in \eta(p, a)\}$. Note that if $\mathcal{N} = \mathcal{D}$ is a DFA with transition function δ , then δ_w is a transformation and we have $\delta_w^R = \delta_w^{-1}$. Define the *determinization* of an NFA \mathcal{N} to be the DFA $\mathcal{N}^D = (Q', \Sigma, \eta^D, I, F')$, where $Q' = \{S \in 2^Q \mid S \text{ is reachable in } \mathcal{N}\}$, $F' = \{S \in Q' \mid S \cap F \neq \emptyset\}$, and $\eta^D(S, a) = \bigcup_{s \in S} \eta(s, a)$.

2.5 Languages, Quotients, and Atoms

Let L be a regular language over the alphabet Σ and let $\mathcal{D} = (Q_n, \Sigma, \delta, q_1, F)$ be the minimal DFA of L . The *left quotient* (or simply *quotient*) of L by the word $w \in \Sigma^*$ is $w^{-1}L = \{x \mid wx \in L\}$. There is a one-to-one correspondence between quotients of L and states of the minimal DFA of L : the languages of distinct states of \mathcal{D} are distinct quotients of L . We use the following convention when discussing quotients of L : the set of quotients is $\{K_1, K_2, \dots, K_n\}$, where K_i is the language of state i of \mathcal{D} . Due to the one-to-one correspondence between states and quotients, the *complexity* of L can be equivalently defined as the number of states in the minimal DFA of L (*state complexity*) or the number of distinct quotients of L (*quotient complexity*).

From now on we deal with non-empty languages only. Denote the complement of a language L by $\bar{L} = \Sigma^* \setminus L$. For $S \subseteq Q_n$, let A_S denote the intersection $\bigcap_{i \in S} K_i \cap \bigcap_{i \in \bar{S}} \bar{K}_i$. If A_S is non-empty, then A_S is called an *atom* of L . Let \mathbf{A} be the set of all atoms of L . The *atom map* $\phi: \mathbf{A} \rightarrow 2^{Q_n}$ is defined by $\phi(A_S) = S$. This map is well-defined, since for each atom A there is precisely one subset S of Q_n such that $A_S = A$. The *basis* of an atom A is $\mathcal{B}(A) = \{K_i \mid i \in \phi(A)\}$.

The *átomaton* of L is the NFA $\mathcal{A} = (\mathbf{A}, \Sigma, \eta, I, F)$, where $\eta(A_i, a) = \{A_j \mid aA_j \subseteq A_i\}$, $I = \{A \in \mathbf{A} \mid q_1 \in \phi(A)\}$, and $F = \{A \in \mathbf{A} \mid \varepsilon \in A\}$. Note that the initial atoms are those that contain L in their bases. Also, there is precisely one final atom: the atom for which all the quotients in its basis contain ε and all other quotients do not. The language of state A of \mathcal{A} is the atom A [4].

The *atomic poset* of L is $\phi(\mathbf{A}) = (\phi(\mathbf{A}), \subseteq)$; this is the set of all subsets S of Q_n such that A_S is an atom. An *atomic interval* of L is an interval in $\phi(\mathbf{A})$, that is, an interval of the form $[V, U]_{\phi(\mathbf{A})}$. We denote an atomic interval using double brackets, since this makes the notation cleaner: we write $[[V, U]]$ instead of $[V, U]_{\phi(\mathbf{A})}$. Since $\phi(\mathbf{A})$ is a subposet of $(2^{Q_n}, \subseteq)$, any two subsets of Q_n can act as endpoints of an atomic interval. Furthermore, every atomic interval $[[V, U]]$ has an associated type (v, u) , as defined in the section on posets.

Note that, if $[[V, U]]$ contains both of its endpoints (i.e., $V, U \in [[V, U]]$), then the type of $[[V, U]]$ is $(|V|, |U|)$. However, we cannot always use the sizes of the endpoints to determine the type of an interval, since there may be multiple ways to choose the endpoints of an interval. For example, if $A_{\{1\}}$ is an atom but A_\emptyset and $A_{\{1,2\}}$ are not, then $[[\{1\}, \{1\}]] = [[\emptyset, \{1\}]] = [[\{1\}, \{1,2\}]] = \{\{1\}\}$. But this interval has type $(1, 1)$, not $(0, 1)$ or $(1, 2)$.

Some basic facts about atoms and átomata follow. The following proposition, proved in [3], shows that we may view the states of \mathcal{A} as subsets of Q_n :

Proposition 2. *Let L be a regular language with átomaton \mathcal{A} and minimal DFA \mathcal{D} . Then the atom map ϕ is an NFA isomorphism between \mathcal{A} and \mathcal{D}^{RDR} .*

The next proposition relates the number of atoms of L to the complexity of the reverse L^R . The proof follows easily from Proposition 2.

Proposition 3 (Number of Atoms). *Let L be a regular language with complexity n , and let the minimal DFA of L be $\mathcal{D} = (Q_n, \Sigma, \delta, q_1, F)$. Then for $S \subseteq Q_n$, the intersection A_S is an atom of L if and only if S is reachable in \mathcal{D}^R , i.e. if and only if there exists $w \in \Sigma^*$ such that $\delta_w^{-1}(F) = S$. Thus there is a bijection between atoms of L and states of \mathcal{D}^{RD} , the minimal DFA of L^R .*

It is well-known that if the complexity of L is n , then the complexity of L^R is at most 2^n , and for $n \geq 2$ this bound is tight. Thus 2^n is also a tight bound on the number of atoms of a regular language when $n \geq 2$.

In [3], a tight upper bound on the complexity of individual atoms was derived and a formula for the bound was given. We give a different (but equivalent) formula below:

Proposition 4 (Complexity of Atoms). *Let L be a regular language with complexity n . Define the function Ψ as follows:*

$$\Psi(n, k) = \begin{cases} 2^n - 1, & \text{if } k = 0 \text{ or } k = n; \\ 1 + \sum_{v=1}^k \sum_{u=k}^{n-1} \binom{n}{u} \binom{u}{v}, & \text{if } 1 \leq k \leq n-1. \end{cases}$$

If A_S is an atom of L , $\Psi(n, |S|)$ is a tight upper bound on the complexity of A_S .

With these bounds established, we can formally define the class of maximally atomic languages. A non-empty regular language L of complexity n is *maximally atomic* if it has the maximal number of atoms (1 if $n = 1$, 2^n if $n \geq 2$) and if for each atom A_S of L , A_S has the maximal complexity $\Psi(n, |S|)$.

3 Main Results

Note that when $n = 1$, the only nonempty language over Σ is Σ^* , and it is maximally atomic. The following proposition characterizes the maximally atomic languages of complexity $n = 2$:

Proposition 5. *Let L be a regular language of complexity 2 and let \mathcal{D} be its minimal DFA with state set Q_2 . Let T be the transition semigroup of \mathcal{D} . Then:*

- *There are four transformations of Q_2 : the identity transformation, the transposition $(1, 2)$, and the unitary transformations $(1 \rightarrow 2)$ and $(2 \rightarrow 1)$.*
- *T contains all four transformations of Q_2 if and only if T contains $(1, 2)$ and at least one unitary transformation.*
- *All subsets of Q_2 are reachable in \mathcal{D}^R (and hence L has all 2^2 atoms) if and only if T contains all four transformations of Q_2 .*
- *Each atom of L has maximal complexity if and only if T contains all four transformations of Q_2 .*
- *Thus, L is maximally atomic if and only if T contains all four transformations of Q_2 .*

The computations required to prove this proposition can be easily done by hand. Henceforth we will be concerned only with languages of complexity $n \geq 3$.

Our main theorem is the following:

Theorem 1. *Let L be a regular language over Σ with complexity $n \geq 3$, and let T be the transition semigroup of the minimal DFA of L . Then L is maximally atomic if and only if the subgroup of permutations in T is set-transitive and T contains a transformation of rank $n - 1$.*

In view of this, let us consider how the class of maximally atomic languages relates to other language classes. Let **FTS** denote the class of languages whose minimal DFAs have the *full transformation semigroup* as their transition semigroup, let **STS** denote the class whose minimal DFAs have transition semigroups with a *set-transitive subgroup* of permutations and a transformation of rank $n - 1$, let **MAL** denote the class of *maximally atomic languages*, let **MNA** denote the class of languages with the *maximal number of atoms*, and let **MCR** denote the class of languages with a *maximally complex reverse*.

1. **FTS** is properly contained in **STS**, by Proposition 1.
2. **STS** is equal to **MAL**, by Theorem 1.
3. **MAL** is contained in **MNA**. Figure 1 in [2] shows the containment is proper.
4. **MNA** is equal to **MCR**, by Proposition 3.

To summarize, we have: **FTS** \subset **STS** = **MAL** \subset **MNA** = **MCR**.

The proof of Theorem 1 relies on two intermediate results. The first gives a condition that is sufficient (but not necessary) for L to have 2^n atoms:

Theorem 2. *Let L be a regular language over Σ with complexity $n \geq 3$, and let T be the transition semigroup of the minimal DFA of L . If T contains all unitary transformations, then L has 2^n atoms.*

The second result establishes Theorem 1 in all but a few cases; it gives necessary and sufficient conditions for individual atoms of L to have maximal complexity, but only when the bases of the atoms are in a certain size range.

Theorem 3. *Let L be a regular language over Σ with complexity $n \geq 3$, and let T be the transition semigroup of the minimal DFA of L . Let A_S be an atom of L and suppose that either $n \geq 4$ and $2 \leq |S| \leq n - 2$, or $n = 3$ and $1 \leq |S| \leq 2$. Then A_S has maximal complexity if and only if the subgroup of permutations in T is $|S|$ -set-transitive and T contains a transformation of rank $n - 1$.*

The rest of the paper consists of the proofs of these three theorems. Shortly before the deadline for this paper, we were informed that the proof of our main result can be simplified by replacing the átomaton with a different construction [5]. Below we present our original proofs, which use the átomaton.

4 Proof of Theorem 2

Let L be a language of complexity $n \geq 3$ and let $\mathcal{D} = (Q_n, \Sigma, \delta, q_1, F)$ be its minimal DFA. Let T be the transition semigroup of \mathcal{D} and assume it contains all unitary transformations. By Proposition 3, L has 2^n atoms if and only if for all $S \subseteq Q_n$, S is reachable in \mathcal{D}^R , i.e., there exists $w \in \Sigma^*$ such that $\delta_w^R(F) = \delta_w^{-1}(F) = S$.

Suppose $X \subseteq Q_n$, with $1 \leq |X| \leq n-1$. Let $t = (i \rightarrow j)$ and $s = (i \rightarrow k)$ for $i \in Q_n$, $j \in X$, $k \notin X$; then $t^{-1}(X) = X \cup \{i\}$ and $s^{-1}(X) = X \setminus \{i\}$. Since T contains all unitary transformations, it contains t and s . Thus for every non-empty $X \subseteq Q_n$ and every $i \in Q_n$, there are words $w, x \in \Sigma^*$ such that $\delta_w^{-1}(X) = X \cup \{i\}$ and $\delta_x^{-1}(X) = X \setminus \{i\}$.

In other words, from any non-empty proper subset X of Q_n , we can reach (in \mathcal{D}^R) all subsets that differ from X by the addition or removal of a single element. Repeatedly applying this fact, we see that from X we can reach any subset S of Q_n : shrink X to a singleton $\{i\} \subseteq X$, expand $\{i\}$ to $\{i, j\}$ for $j \in S$, shrink again to $\{j\} \subseteq S$, and then expand to S (or shrink to \emptyset for $S = \emptyset$).

Now, if $|F| = 0$ then $L = \emptyset$, and if $|F| = n$ then $L = \Sigma^*$; since \mathcal{D} is minimal, $n = 1$ in either case. Since $n \geq 3$, we have that F is a non-empty proper subset of Q_n . Thus by the argument above, we can reach all subsets of Q_n in \mathcal{D}^R ; hence L has 2^n atoms. \square

5 Proof of Theorem 3

5.1 The Átomaton and Minimal DFAs of Atoms

In this section we prove the \Rightarrow direction of Theorem 3. Two results on átomata and atoms are needed for this. We first describe the transition function of the átomaton, in the case where the states are viewed as subsets of Q_n . Define $\Delta_w: 2^{Q_n} \rightarrow 2^{Q_n}$ by $\Delta_w(S) = \delta_w(\overline{S}) = Q_n \setminus \delta_w(Q_n \setminus S)$.

Lemma 1. *Let L be a regular language over Σ . Let $\mathcal{D} = (Q_n, \Sigma, \delta, q_1, F)$ be the minimal DFA of L . Let \mathcal{A} be the átomaton of L with transition function η . If $[[V, U]]$ is an atomic interval of L and a set of states of \mathcal{A} , then for all $w \in \Sigma^*$, we have $\eta_w([[V, U]]) = [[\delta_w(V), \Delta_w(U)]]$.*

Proof. It was shown in [3] that $\eta_a(S) = \{T \mid A_T \text{ is an atom of } L, T \supseteq \delta_a(S) \text{ and } \delta_a(\overline{S}) \cap T = \emptyset\}$. If $\delta_a(\overline{S}) \cap T = \emptyset$, then $T \subseteq \overline{\delta_a(\overline{S})} = \Delta_a(S)$. Thus $\eta_a(S)$ is the set of $T \subseteq Q_n$ such that A_T is an atom of L and $\delta_a(S) \subseteq T \subseteq \Delta_a(S)$, which is precisely $[[\delta_a(S), \Delta_a(S)]]$. One verifies that this can be extended to words, giving $\eta_w(S) = [[\delta_w(S), \Delta_w(S)]]$.

Next, we want to show $\eta_w([[V, U]]) = [[\delta_w(V), \Delta_w(U)]]$. For $T \in [[V, U]]$, consider $\eta_w(T)$. Since $V \subseteq T$, $\delta_w(V) \subseteq \delta_w(T)$. Since $T \subseteq U$, we have $\overline{T} \supseteq \overline{U}$, and thus $\delta_w(\overline{T}) \supseteq \delta_w(\overline{U})$. It follows that $\Delta_w(T) \subseteq \Delta_w(U)$. Hence $\eta_w(T) = [[\delta_w(T), \Delta_w(T)]] \subseteq [[\delta_w(V), \Delta_w(U)]]$, and $\eta_w([[V, U]]) \subseteq [[\delta_w(V), \Delta_w(U)]]$.

For containment in the other direction, suppose that T is in $[[\delta_w(V), \Delta_w(U)]]$; then $\delta_w(V) \subseteq T \subseteq \Delta_w(U)$ and A_T is an atom. Let $S = \delta_w^{-1}(T)$; then we claim $S \in [[V, U]]$. Since $T \supseteq \delta_w(V)$, we have $\delta_w^{-1}(T) = S \supseteq V$. If $i \in T \subseteq \Delta_w(U)$, then $i \notin \delta_w(\overline{U})$. Hence $\delta_w^{-1}(i)$ is disjoint from \overline{U} for all $i \in T$, and so $\delta_w^{-1}(T)$ is disjoint from \overline{U} . It follows that $\delta_w^{-1}(T) = S \subseteq U$. It remains to show A_S is an atom; but

since A_T is an atom, by Proposition 3, there exists $x \in \Sigma^*$ such that $\delta_x^{-1}(F) = T$. Thus $S = \delta_w^{-1}(T) = \delta_w^{-1}(\delta_x^{-1}(F)) = \delta_{xw}^{-1}(F)$, so by Proposition 3, A_S is also an atom.

Hence $S \in [[V, U]]$, and it follows that $\eta_w(S) = [[T, \Delta_w(S)]] \subseteq \eta_w([[V, U]])$. To complete the proof, we must show $\eta_w(S)$ is non-empty (and thus contains T) by showing that $T \subseteq \Delta_w(S) = \delta_w(\delta_w^{-1}(T))$. Observe that if $i \in T$, then $\delta_w^{-1}(i) \subseteq \delta_w^{-1}(T)$. Thus $\delta_w^{-1}(i) \cap \delta_w^{-1}(T) = \emptyset$, and so $i \notin \delta_w(\delta_w^{-1}(T))$, which gives $i \in \Delta_w(S)$ as required. Thus $T \in \eta_w(S) = [[T, \Delta_w(S)]]$, and it follows that if $T \in [[\delta_w(V), \Delta_w(U)]]$, then $T \in \eta_w([[V, U]])$. This proves that the two intervals must be equal. \square

	δ	a
\rightarrow	1	2
\leftarrow	2	3
\leftarrow	3	4
	4	4

	δ^R	a
\leftarrow	1	
\rightarrow	2	{1}
\rightarrow	3	{2}
	4	{3,4}

	δ^{RD}	a
\rightarrow	{2,3}	{1,2}
\leftarrow	{1,2}	{1}
\leftarrow	{1}	\emptyset
	\emptyset	\emptyset

	η	a
\leftarrow	{2,3}	
\rightarrow	{1,2}	{2,3}
\rightarrow	{1}	{1,2}
	\emptyset	{ \emptyset , {1}}

Example 1. The minimal DFA \mathcal{D} of Table 1 accepts the language $\{a, aa\}$. The NFA \mathcal{D}^R is in Table 2 and the DFA \mathcal{D}^{RD} , in Table 3. The átomaton \mathcal{A} is in Table 4. In NFAs \mathcal{D}^R and \mathcal{A} , a blank in an entry (q, a) indicates that there is no transition from q under a . However, when determinization is used in Table 3, the empty set of states of \mathcal{D}^R becomes a state of the resulting DFA \mathcal{D}^{RD} . A right arrow (\rightarrow) indicates an initial state and a left arrow (\leftarrow) indicates a final state.

Consider the atomic interval $[[\emptyset, \{1, 2\}]] = \{\emptyset, \{1\}, \{1, 2\}\}$; we have $\delta_a(\emptyset) = \emptyset$, and $\Delta_a(\{1, 2\}) = \delta_a(\overline{\{1, 2\}}) = \delta_a(\overline{\{3, 4\}}) = \overline{\{4\}} = \{1, 2, 3\}$. Thus to determine the result of $\eta_a([[\emptyset, \{1, 2\}]])$, we take the interval $[\emptyset, \{1, 2, 3\}]_{2Q_4} = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$ and we remove the sets that do not represent atoms. After this removal, we get $\{\emptyset, \{1\}, \{1, 2\}, \{2, 3\}\}$. Hence $\eta_a(\{\emptyset, \{1\}, \{1, 2\}\}) = \eta_a([[\emptyset, \{1, 2\}]]) = [[\emptyset, \{1, 2, 3\}]] = \{\emptyset, \{1\}, \{1, 2\}, \{2, 3\}\}$. \blacksquare

Remark 1. If we treat the set of states of \mathcal{A} as a subset of 2^{Q_n} , then it is possible that the empty set is a state of \mathcal{A} , as in Example 1. Since we use the same symbol for η and its extension to subsets of states, an ambiguity arises when η is applied to the empty set. Specifically, $\eta_w(\emptyset)$ may mean “ η_w applied to the state $\emptyset \in 2^{Q_n}$ ”, in which case $\eta_w(\emptyset) = \eta_w([[\emptyset, \emptyset]]) = [[\emptyset, \text{coim } \delta_w]]$, or it may mean “ η_w applied to the empty subset of states $\emptyset \subseteq 2^{Q_n}$ ”, in which case $\eta_w(\emptyset) = \emptyset$. We avoid this ambiguity by adopting the convention that $\eta_w(\emptyset)$ always means “ η_w applied to $\emptyset \subseteq 2^{Q_n}$ ” and $\eta_w([[\emptyset, \emptyset]])$ has the other meaning.

A corollary of this is that every reachable subset of states in the átomaton $\mathcal{A} = (\mathbf{A}, \Sigma, \eta, I, F)$ is an atomic interval of L . The same holds for every reachable subset of states in the NFA $\mathcal{A}_S = (\mathbf{A}, \Sigma, \eta, \{S\}, F)$ recognizing the atom A_S . Since the determinization \mathcal{A}_S^D is the minimal DFA of A_S [3], it follows that *the states of minimal DFAs of atoms of L may be represented as atomic intervals of L .*

If A_S is an atom of L with maximal complexity, certain restrictions apply to the *types* of the atomic intervals in \mathcal{A}_S^D . For $S \subseteq Q_n$, define an *S-type* to be a pair of integers (v, u) satisfying:

1. If $|S| = 0$, then $v = 0$ and $0 \leq u \leq n - 1$.
2. If $|S| = n$, then $1 \leq v \leq n$ and $u = n$.
3. If $1 \leq |S| \leq n - 1$, then $1 \leq v \leq |S|$ and $|S| \leq u \leq n - 1$.

A non-empty interval that has an S -type is called an S -interval. The empty interval is a special case: it is an S -interval if and only if $1 \leq |S| \leq n-1$. The significance of S -types and S -intervals is as follows:

Lemma 2. *Let L be a regular language with complexity n and let A_S be an atom of L . If A_S has maximal complexity $\Psi(n, |S|)$, then the set of states of \mathcal{A}_S^D equals the set of atomic S -intervals of L .*

Proof. A simple counting argument shows that the number of intervals of type (v, u) in a subposet of $P = (2^{\mathcal{Q}_n}, \subseteq)$ is bounded from above by $\binom{n}{u} \binom{n}{v}$. Combining this fact with the definition of an S -type gives that $\Psi(n, |S|)$ is an upper bound on the number of S -intervals in a subposet of P . Now, we know that \mathcal{A}_S^D has exactly $\Psi(n, |S|)$ states; if we show that these states are all atomic S -intervals of L , then this implies the state set of \mathcal{A}_S^D contains exactly $\Psi(n, |S|)$ distinct atomic S -intervals of L , and nothing else. Since the atomic poset of L is a subposet of P , there can be no more than $\Psi(n, |S|)$ atomic S -intervals of L , and this proves the result. Thus we just need to show that every state of \mathcal{A}_S^D is an S -interval.

Let $[[V, U]]$ be a state of \mathcal{A}_S^D and suppose $[[\delta_w(S), \Delta_w(S)]] = [[V, U]]$. If $[[V, U]]$ is the empty interval, then it is automatically an S -interval for $1 \leq |S| \leq n-1$, by definition. For $|S| = 0$ or $|S| = n$, the empty interval is not an S -interval, but this does not matter since it is not reachable in \mathcal{A}_S^D . In the $|S| = 0$ case, a state of \mathcal{A}_S^D has the form $[[\delta_w(\emptyset), \Delta_w(\emptyset)]] = [[\emptyset, \text{coim } \delta_w]]$, which always contains \emptyset (since $A_S = A_\emptyset$ is an atom); thus every state of \mathcal{A}_S^D is a non-empty interval. For $|S| = n$, a similar argument works.

Next, suppose $[[V, U]]$ is non-empty. For this case, some setup is needed. Define the set $X_S = \{[\delta_w(S), \Delta_w(S)]_P \mid w \in \Sigma^*\}$ of intervals in P . One can verify that $(|\delta_w(S)|, |\Delta_w(S)|)$ is an S -type for all S and w , and thus X_S is a set of S -intervals of P . This means $|X_S|$ is bounded from above by $\Psi(n, |S|)$. Now, let $Y_S = \{[[\delta_w(S), \Delta_w(S)]] \mid w \in \Sigma^*\}$; this is just the set of states of \mathcal{A}_S^D , and thus it has size $\Psi(n, |S|)$. Define $\alpha: X_S \rightarrow Y_S$ by $\alpha([X, Y]_P) = [[X, Y]]$; this is clearly a surjection, and thus $|X_S| \geq |Y_S| = \Psi(n, |S|)$. Since we also have $|X_S| \leq \Psi(n, |S|)$, we get $|X_S| = |Y_S| = \Psi(n, |S|)$ and hence α is a bijection.

Now, assume without loss of generality that the type of $[[V, U]]$ is $(|V|, |U|)$. Suppose for a contradiction that $|V| > |S|$. Then since $[[V, U]] = [[\delta_w(S), \Delta_w(S)]]$, we have $\delta_w(S) \subset V$. We can find a set X such that $|X| = |S|$ and $\delta_w(S) \subseteq X \subset V$. Now, since $\delta_w(S) \subseteq X \subseteq \Delta_w(S)$, we have $X \in [[\delta_w(S), \Delta_w(S)]]$ if and only if A_X is an atom. But $X \notin [[V, U]]$ since $X \subset V$, and thus A_X is not an atom. It follows that the interval $[[X, X]]$ is empty. If $|S| = 0$ or $|S| = n$, then in fact $|S| = |X|$ and $[[S, S]]$ is clearly non-empty, a contradiction. If $1 \leq |S| \leq n-1$, observe that $\alpha([X, X]_P) = [[X, X]] = \emptyset$. But $\emptyset \in X_S$ since \emptyset is an S -interval of P for $1 \leq |S| \leq n-1$, so also $\alpha(\emptyset) = \emptyset$. This is a contradiction, since α is a bijection. Thus for S of any size, we always have $|V| \leq |S|$. A similar argument to the above shows that $|U| \geq |S|$.

Thus, if $|S| = 0$ we have $|V| = 0$ and $0 \leq |U| \leq |\Delta_w(S)| = n-1$. If $|S| = n$ we have $1 \leq |\delta_w(S)| \leq |V| \leq n$ and $|U| = n$. If $1 \leq |S| \leq n-1$, then $1 \leq |V| \leq |S|$ and $|S| \leq |U| \leq n-1$. Thus we have proved $(|V|, |U|)$ is an S -type. Hence every state $[[V, U]]$ of \mathcal{A}_S^D is an atomic S -interval of L , and the number of states equals the upper bound $\Psi(n, |S|)$ on the number of atomic S -intervals of L , proving the lemma. \square

Lemma 2 has two particularly useful consequences. Let A_S be an atom of maximal complexity, and suppose $V, U \subseteq \mathcal{Q}_n$ are sets such that $(|V|, |U|)$ is an S -type. Then:

1. $[[V, U]]$ has type $(|V|, |U|)$. In particular, this means $[[V, U]]$ contains its endpoints V and U .
2. $[[V, U]]$ is a state of \mathcal{A}_S^D .

(1) follows since $(|V|, |U|)$ is an S -type, and so if $[[V, U]]$ does not have type $(|V|, |U|)$, the number of atomic S -intervals of type $(|V|, |U|)$ is not maximal and hence A_S is not maximally complex. (2) follows since if $[[V, U]]$ has the S -type $(|V|, |U|)$, it is an atomic S -interval and thus a state of \mathcal{A}_S^D .

These facts are sufficient to prove one direction of Theorem 3:

Theorem 3 (\Rightarrow Direction). Let L be a language of complexity $n \geq 3$, let T be the transition semigroup of the minimal DFA of L , and let A_S be an atom of L . Suppose either $n = 3$ and $1 \leq |S| \leq 2$, or $n \geq 4$ and $2 \leq |S| \leq n - 2$. We prove that if A_S has maximal complexity, then the subgroup of permutations in T is $|S|$ -set-transitive and T contains a transformation of rank $n - 1$.

The minimal DFA of A_S is \mathcal{A}_S^D , and its initial state is $[[S, S]]$. For all $X \subseteq Q_n$ with $|X| = |S|$, $(|X|, |X|)$ is an S -type. Thus by Lemma 2, $[[X, X]]$ is a state of \mathcal{A}_S^D of type $(|X|, |X|)$. Thus $\eta_w([[S, S]]) = [[\delta_w(S), \Delta_w(S)]] = [[X, X]]$ for some $w \in \Sigma^*$. Applying Lemma 2 again gives $(|\delta_w(S)|, |\Delta_w(S)|) = (|X|, |X|)$. Hence $|X| = |\delta_w(S)| = |S|$, and so $\delta_w \in T$ is a permutation. It follows for all $X \subseteq Q_n$ with $|X| = |S|$, there is a permutation that sends S to X ; thus the subgroup of permutations in T is $|S|$ -set-transitive.

Now, let $\delta_w \in T$ have rank $n - k$ and consider $[[\delta_w(S), \Delta_w(S)]]$. By Lemma 2 this interval has type $(|\delta_w(S)|, |\Delta_w(S)|)$, so it is a non-empty interval. This implies $\delta_w(S)$ and $\delta_w(\bar{S})$ are disjoint. It follows that $|\text{im } \delta_w| = |\delta_w(Q_n)| = |\delta_w(S)| + |\delta_w(\bar{S})|$. Since the rank of δ_w is $n - k$, $|\delta_w(\bar{S})| = (n - k) - |\delta_w(S)|$. Thus $|\Delta_w(S)| = n - (n - k - |\delta_w(S)|) = |\delta_w(S)| + k$, which gives $|\Delta_w(S)| - |\delta_w(S)| = k$.

Consider $[[S, S \cup \{i\}]]$ for $i \notin S$. Since $(|S|, |S| + 1)$ is an S -type, by Lemma 2 this interval is reachable in \mathcal{A}_S^D . Thus there is a $\delta_w \in T$ such that $(|\delta_w(S)|, |\Delta_w(S)|) = (|S|, |S| + 1)$. By the argument above, this δ_w must have rank $n - (|\Delta_w(S)| - |\delta_w(S)|) = n - 1$. Hence T contains a transformation of rank $n - 1$. \square

5.2 Semigroups and Groups

To prove the other direction of Theorem 3, we use some results from semigroup and group theory. The first is a result of Livingstone and Wagner [6]:

Proposition 6. *Let G be a permutation group of degree $n \geq 4$. If $2 \leq k \leq \frac{n}{2}$, then the number of orbits when G acts on k -subsets of Q_n is at least the number of orbits when G acts on $(k - 1)$ -subsets of Q_n .*

Using this proposition, we can easily prove

Lemma 3. *Let G be a k -set-transitive permutation group of degree $n \geq 4$ and suppose $2 \leq k \leq \frac{n}{2}$. Then:*

1. G is $(n - k)$ -set-transitive.
2. G is ℓ -set-transitive for each ℓ such that $0 \leq \ell \leq k$ or $n - k \leq \ell \leq n$.

Proof. (1): Suppose G is k -set-transitive. If U and V are $(n - k)$ -subsets of Q_n , then \bar{U} and \bar{V} are k -subsets, and there exists a permutation $p \in G$ mapping \bar{U} to \bar{V} . But if p maps \bar{U} to \bar{V} , then it maps U to V ; thus G can map any $(n - k)$ -subset to any other $(n - k)$ -subset, and so is $(n - k)$ -set-transitive.

(2): Suppose G is k -set-transitive and $2 \leq k \leq \frac{n}{2}$. Then there is one orbit when G acts on k -subsets. By Proposition 6, there is one orbit when G acts on $(k - 1)$ -subsets. This implies G is $(k - 1)$ -set-transitive. Repeating this argument we conclude that G is ℓ -set-transitive for $0 \leq \ell \leq k$. By (1), G is also ℓ -set-transitive for $n - k \leq \ell \leq n$. \square

Note that for $n = 3$, a permutation group of degree 3 is set-transitive if and only if it is transitive.

The second result we use is a theorem of Ruškuc, published by McAlister [7]:

Proposition 7. *Let G be a permutation group of degree $n \geq 3$ and let $t: Q_n \rightarrow Q_n$ be a unitary transformation. Let T be the transformation semigroup generated by $G \cup \{t\}$. Then T contains all singular transformations if and only if G is 2-set-transitive.*

We can use this to prove the following lemma:

Lemma 4. *Let G be a 2-set-transitive permutation group of degree $n \geq 3$ and let $t: Q_n \rightarrow Q_n$ be a transformation of rank $n - 1$. Then the transformation semigroup T generated by $G \cup \{t\}$ contains all singular transformations.*

Proof. By Proposition 7, if G is 2-set-transitive and t is a unitary transformation, then T contains all singular transformations. Thus it suffices to show that if t is any transformation of rank $n - 1$, then $G \cup \{t\}$ generates a unitary transformation.

For each transformation $s: Q_n \rightarrow Q_n$, we define a set of tuples called s -paths. For $k \geq 2$, a tuple (i_1, \dots, i_k) of distinct elements of Q_n is an s -path of length k if $s(i_j) = i_{j+1}$ for $1 \leq j < k$ and $s(i_k) = i_\ell$ for some $\ell < k$. An s -path (i_1, \dots, i_k) is *incomplete* if there exists a in Q_n such that (a, i_1, \dots, i_k) is an s -path, and *complete* otherwise. An s -path (i_1, \dots, i_k) is *cyclic* if $s(i_k) = i_1$ and *acyclic* otherwise. The element i_1 of the acyclic s -path (i_1, \dots, i_k) is called the *head*.

Let t be a transformation of rank $n - 1$, and consider the t -paths. If a t -path is complete and acyclic, its head must be an element of $\text{coim } t$. Since t has rank $n - 1$, $|\text{coim } t| = 1$, and so there is precisely one complete acyclic t -path. Let (a_1, \dots, a_k) be that complete acyclic t -path, and suppose $t(a_k) = a_\ell$.

Since G is 2-set-transitive, it is 1-set-transitive by Lemma 3. Thus there exists a permutation $p \in G$ with $p(a_1) = a_{\ell-1}$. Let $pt = p \circ t$, and consider pt -paths. Since pt has rank $n - 1$, there is only one complete acyclic pt -path; the head of this path must be $a_{\ell-1}$, since $\text{coim } pt = \{a_{\ell-1}\}$. Observe that $pt(a_k) = p(t(a_k)) = p(a_\ell) = p(t(a_{\ell-1})) = pt(a_{\ell-1})$; it follows that $(a_{\ell-1}, p(a_\ell), pt(p(a_\ell)), \dots, a_k)$ is the complete acyclic pt -path.

Now, let n be the product of the lengths of all the complete cyclic pt -paths and the incomplete cyclic pt -path $(p(a_\ell), pt(p(a_\ell)), \dots, a_k)$. Then we have $(pt)^n = (a_{\ell-1} \rightarrow (pt)^n(a_{\ell-1}))$, where $(pt)^n$ is pt composed with itself n times. This proves that T must contain all singular transformations, since it is 2-set-transitive and contains a unitary transformation. \square

These results are sufficient to prove the other direction of Theorem 3:

Theorem 3 (\Leftarrow Direction). Let L be a language of complexity $n \geq 3$, let T be the transition semigroup of the minimal DFA of L , and let A_S be an atom of L . Suppose either $n = 3$ and $1 \leq |S| \leq 2$, or $n \geq 4$ and $2 \leq |S| \leq n - 2$. We prove that if the subgroup of permutations in T is $|S|$ -set-transitive and T contains a transformation of rank $n - 1$, then A_S has maximal complexity.

By Lemmas 3 and 4, T contains all singular transformations. By Theorem 2, L has 2^n atoms. From the proof of Lemma 2, $\Psi(n, |S|)$ is a tight bound on the number of S -intervals in the atomic poset of L . Since L has 2^n atoms (the maximal possible), the number of atomic S -intervals of L meets the bound $\Psi(n, |S|)$. It remains to show that all these intervals are reachable in the minimal DFA \mathcal{A}_S^D of A_S . From the initial state $[[S, S]]$ of \mathcal{A}_S^D , we can reach the empty interval by $(i \rightarrow j)$ where $i \in S$ and $j \notin S$; thus it suffices to consider non-empty intervals.

Let $[[V, U]]$ be a non-empty atomic S -interval of L with type $(|V|, |U|)$. By the definition of an atomic S -interval, $1 \leq |V| \leq |S|$ and $|S| \leq |U| \leq n - 1$ and $V \subseteq U$. Thus there exists a set X such that $|X| = |S|$ and $V \subseteq X \subseteq U$. Since the subgroup of permutations in T is $|S|$ -set-transitive, there is a permutation $\delta_w \in T$ that sends S to X ; thus $\eta_w([[S, S]]) = [[X, X]]$. If $V = X = U$, we are done, so assume that $V \subset X$ or $X \subset U$. If $V \subset X$ and $|V| \geq 2$, we can shrink the lower bound of $[[X, X]]$ as follows: select distinct $i, j \in Q_n$ such that $i \in X \setminus V$ and $j \in V$. Since T contains all unitary transformations, there is a $\delta_x \in T$ such that $\delta_x = (i \rightarrow j)$. Since $i \notin \overline{X}$, $\delta_x(\overline{X}) = \overline{X}$ and thus $\Delta_x(X) = X$. It follows that $\eta_x([[X, X]]) = [[X \setminus \{i\}, X]]$. Repeating this process, we can reach $[[V, X]]$ for all V with $1 \leq |V| \leq |S|$. By a similar process, we can repeatedly enlarge the upper bound of $[[V, X]]$ to reach $[[V, U]]$. Thus all $\Psi(n, |S|)$ atomic S -intervals of L are reachable in \mathcal{A}_S^D . By Lemma 2, A_S has maximal complexity. \square

Remark 2. The proof above works for the $|S| = 1$ and $|S| = n - 1$ cases if we assume that T contains all unitary transformations, rather than only assuming it contains some transformation of rank $n - 1$.

6 Proof of Theorem 1

Having proved Theorems 2 and 3, we need only a bit more work to prove our main theorem.

Let L be a language with complexity $n \geq 3$ and let T be the transition semigroup of the minimal DFA of L . If L is maximally atomic, then by Theorem 3 and Lemma 3, the subgroup of permutations in T is k -set-transitive for $1 \leq k \leq n-1$, and hence is set-transitive; also, by Theorem 3, T contains a transformation of rank $n-1$. This proves one direction of the theorem.

For the other direction, suppose the subgroup of permutations in T is set-transitive and contains a transformation of rank $n-1$. By Theorem 2, L has 2^n atoms. By Theorem 3, if $n \geq 4$ and $2 \leq |S| \leq n-2$ or $n = 3$ and $1 \leq |S| \leq 2$, then A_S has maximal complexity. By Lemma 4, T contains all singular transformations and hence all unitary transformations; so by Remark 2, A_S has maximal complexity if $|S| = 1$ or $|S| = n-1$. The only remaining cases are $|S| = 0$ and $|S| = n$.

Let \mathcal{A}_S^D be the minimal DFA of A_S . By Lemma 2, to show that A_S has maximal complexity, it suffices to show that all atomic S -intervals of L are reachable in \mathcal{A}_S^D . If $|S| = 0$, then $S = \emptyset$, and the atomic \emptyset -intervals of L are those with type $(0, i)$ where $0 \leq i \leq n-1$. The initial state of \mathcal{A}_\emptyset^D is $[[\emptyset, \emptyset]]$; thus a reachable state looks like $[[\delta_w(\emptyset), \Delta_w(\emptyset)]] = [[\emptyset, \text{coim } \delta_w]]$ for some $w \in \Sigma^*$.

Since T contains all singular transformations, for all $U \subset Q_n$, there exists $t \in T$ such that $\text{coim } t = U$. Hence for all $U \subset Q_n$, $[[\emptyset, U]]$ is reachable in \mathcal{A}_\emptyset^D . Thus all intervals of type $(0, i)$ are reachable, for $0 \leq i \leq n-1$. By Lemma 2, A_\emptyset has maximal complexity. By a similar argument, when $|S| = n$, the atom A_{Q_n} has maximal complexity. Thus all 2^n atoms have maximal complexity; this completes the proof. \square

7 Conclusions

We have defined a new class of regular languages – the maximally atomic languages – and proven that a language of complexity n is maximally atomic if and only if the transition semigroup of its minimal DFA is set-transitive and contains a transformation of rank $n-1$. Since the set-transitive groups have been fully classified, it is easy to construct examples of maximally atomic languages and study them. We have also derived a formula for the transition functions of átomata and minimal DFAs of atoms.

Acknowledgements: We thank a referee for giving many suggestions to improve our proofs.

References

- [1] Ross A. Beaumont & Raymond P. Peterson (1955): *Set-transitive permutation groups*. *Canadian Journal of Mathematics* 7, pp. 35–42, doi:10.4153/CJM-1955-005-x.
- [2] Janusz Brzozowski & Gareth Davies (2013): *Maximal Syntactic Complexity of Regular Languages Implies Maximal Quotient Complexities of Atoms*. Available at <http://arxiv.org/abs/1302.3906>.
- [3] Janusz Brzozowski & Hellis Tamm (2013): *Complexity of Atoms of Regular Languages*. *Int. J. Found. Comput. Sci.* 24(7), pp. 1009–1027, doi:10.1142/S0129054113400285.
- [4] Janusz Brzozowski & Hellis Tamm (2014): *Theory of Átomata*. *Theoret. Comput. Sci.*, doi:10.1016/j.tcs.2014.04.016. In press.
- [5] Szabolcs Iván (2014): *Handle Atoms with Care*. Available at <http://arxiv.org/abs/1404.6632>.
- [6] Donald Livingstone & Ascher Wagner (1965): *Transitivity of finite permutation groups on unordered sets*. *Mathematische Zeitschrift* 90(5), pp. 393–403, doi:10.1007/BF01112361.
- [7] Donald B. McAlister (1998): *Semigroups generated by a group and an idempotent*. *Communications in Algebra* 26(2), pp. 243–254, doi:10.1080/00927879808826145.