I F I G

RESEARCH

REPORT

INSTITUT FÜR INFORMATIK

PROPERTIES OF RIGHT ONE-WAY JUMPING FINITE AUTOMATA

Simon Beier   Markus Holzer

IFIG RESEARCH REPORT 1802

MARCH 2018

Institut für Informatik
JLU Gießen
Arndtstraße 2
35392 Giessen, Germany
Tel: +49-641-99-32141
Fax: +49-641-99-32149
mail@informatik.uni-giessen.de
www.informatik.uni-giessen.de

JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

IFIG Research Report
IFIG Research Report 1802, March 2018

Properties of Right One-Way Jumping Finite Automata

Simon Beier and Markus Holzer[1]

Institut für Informatik, Universität Giessen
Arndtstraße 2, 35392 Giessen, Germany

Abstract. Right one-way jumping finite automata (ROWJFAs), were recently introduced in [H. Chigahara, S. Z. Fazekas, A. Yamamura: One-Way Jumping Finite Automata, *Internat. J. Found. Comput. Sci.*, 27(3), 2016] and are jumping automata that process the input in a discontinuous way with the restriction that the input head reads deterministically from left-to-right starting from the leftmost letter in the input and when it reaches the end of the input word, it returns to the beginning and continues the computation. We solve most of the open problems of these devices. In particular, we characterize the family of permutation closed languages accepted by ROWJFAs in terms of Myhill-Nerode equivalence classes. Using this, we investigate closure and non-closure properties as well as inclusion relations to other language families. We also give more characterizations of languages accepted by ROWJFAs for some interesting cases.

Categories and Subject Descriptors according to ACM Computing Classification System: F.1.1 [**Computation by Abstract Devices**]: Models of Computation—*Automata*; F.4.3 [**Mathematical Logic and Formal Languages**]: Formal Languages—*Classes Defined by Grammars or Automata*; F.4.3 [**Mathematical Logic and Formal Languages**]: Formal Languages—*Operations on Languages*

MSC Classification: 68Q45 Formal languages and automata

Additional Key Words and Phrases: Jumping Finite Automata, One-Way Restriction, Characterizations, Inclusion Relations, Closure Properties

[1] E-mail: {simon.beier,markus.holzer }@informatik.uni-giessen.de

1 Introduction

Jumping finite automata [11] are a machine model for discontinuous information processing. Roughly speaking, a jumping finite automaton is an ordinary finite automaton, which is allowed to read letters from anywhere in the input string, not necessarily only from the left of the remaining input. In a series of papers [1, 6, 7, 14] different aspects of jumping finite automata were investigated, such as, e.g., inclusion relations, closure and non-closure results, decision problems, computational complexity of jumping finite automata problems, etc. Shortly after the introduction of jumping automata a variant of this machine model was defined, namely (right) one-way jumping finite automata [3]. There the device moves the input head deterministically from left-to-right starting from the leftmost letter in the input and when it reaches the end of the input word, it returns to the beginning and continues the computation. As in the case of ordinary jumping finite automata inclusion relations to well-known formal language families, closure and non-closure results under standard formal language operations were investigated. Nevertheless, a series of problems on right one-way jumping automata (ROWJFAs) remained open in [3]. This is the starting point of our investigation.

First we develop a characterization of (permutation closed) languages that are accepted by ROWJFAs in terms of the Myhill-Nerode relation. It is shown that the permutation closed language L belongs to **ROWJ**, the family of all languages accepted by ROWJFAs, if and only if L can be written as the *finite union* of Myhill-Nerode equivalence classes. Observe, that the overall number of equivalence classes can be infinite. This result nicely contrasts the characterization of regular languages, which requires that the overall number of equivalence classes is finite. The characterization allows us to identify languages that are *not* accepted by ROWJFAs, which are useful to prove non-closure results on standard formal language operations. In this way we solve all of the open problems from [3] on the inclusion relations of ROWJFAs languages to other language families and on their closure properties. It is shown that the family **ROWJ** is an anti-abstract family of languages (anti-AFL), that is, it is not closed under any of the operations λ-free homomorphism, inverse homomorphism, intersection with regular sets, union, concatenation, or Kleene star. This is a little bit surprising for a language family defined by a deterministic automaton model. Although anti-AFLs are sometimes referred to an "unfortunate family of languages" there is linguistical evidence that such language families might be of crucial importance, since in [4] it was shown that the family of natural languages is an anti-AFL. On the other hand, the family **pROWJ**, of all permutation closed languages in **ROWJ**, almost form an anti-AFL, since this language family is closed under inverse homomorphism. Moreover, we obtain further characterizations of languages accepted by ROWJFAs. For instance, we show that

1. language wL is in **ROWJ** if and only if L is in **ROWJ**,
2. language Lw is in **ROWJ** if and only if L is regular, and
3. language L_1L_2 is in **ROWJ** if and only if L_1 is regular and L_2 is in **ROWJ**, where L_1 and L_2 have to fulfil some further easy pre-conditions.

The latter result is in similar vein as a result in [9] on linear context-free languages, where it was shown that L_1L_2 is a linear context-free language if and only if L_1 is regular and L_2 at most linear context free. Finally another characterization is given for letter bounded ROWJFA languages, namely, the language $L \subseteq a_1^*a_2^*\ldots a_n^*$ is in **ROWJ** if and only if L is regular. This result nicely generalizes the fact that every unary language accepted by an ROWJFA is regular.

The paper is organized as follows: in the next section we introduce the necessary notations on (one-way) jumping finite automata. Then we prove a characterization of the language family **ROWJ** in terms of the Myhill-Nerode equivalence relation in Section 3. Then Section 4 is devoted to inclusion relations between **ROWJ** and standard language families from formal language theory. There it is shown that the language family **ROWJ** is incomparable to the family **JFA**, of all languages accepted by jumping finite automata, solving an open problem from [3]. Closure properties of the language family in question and their permutation closed variant are investigated in Section 5. Finally, in Section 6 more characterizations of languages accepted by ROWJFA*s* are developed.

2 Preliminaries

We assume the reader to be familiar with the basics in automata and formal language theory as contained, for example, in [10]. Let $\mathbb{N} = \{0, 1, 2, \ldots\}$ be the set of non-negative integers. We use $\subseteq$ for inclusion, and $\subset$ for proper inclusion. Let Σ be an alphabet. Then Σ^* is the set of all words over Σ, including the empty word λ. For a language $L \subseteq \Sigma^*$ define the set $\mathsf{perm}(L) = \cup_{w\in L} \mathsf{perm}(w)$, where $\mathsf{perm}(w) = \{\, v \in \Sigma^* \mid v \text{ is a permutation of } w \,\}$. Then a language L is called *permutation closed* if $L = \mathsf{perm}(L)$. The length of a word $w \in \Sigma^*$ is denoted by $|w|$. For the number of occurrences of a symbol a in w we use the notation $|w|_a$. We denote the powerset of a set S by 2^S. For $\Sigma = \{a_1, a_2, \ldots, a_k\}$, the *Parikh-mapping* $\psi : \Sigma^* \to \mathbb{N}^k$ is the function $w \mapsto (|w|_{a_1}, |w|_{a_2}, \ldots, |w|_{a_k})$. A language $L \subseteq \Sigma^*$ is called *semilinear* if its *Parikh-image* $\psi(L)$ is a semilinear subset of $\mathbb{N}^k$, a definition of those can be found in [8].

The elements of $\mathbb{N}^k$ can be partially ordered by the $\leq$-relation on vectors. For $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{N}^k$ we write $\boldsymbol{x} \leq \boldsymbol{y}$ if all components of $\boldsymbol{x}$ are less or equal to the corresponding components of $\boldsymbol{y}$. The value $||\boldsymbol{x}||$ is the maximum norm of $\boldsymbol{x}$, that is, $||(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k)|| = \max\{\, |\boldsymbol{x}_i| \mid 1 \leq i \leq k \,\}$.

Let Σ be an alphabet and $v, w \in \Sigma^*$. We say that word v is a prefix of w if there is an $x \in \Sigma^*$ with $w = vx$ and v is a sub-word of w if there are $x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_{n+1} \in \Sigma^*$ with $v = x_1 x_2 \cdots x_n$ and $w = y_1 x_1 y_2 x_2 \cdots y_n x_n y_{n+1}$, for some $n \geq 0$. A language $L \subseteq \Sigma^*$ is called prefix-free if and only if there are no words $v, w \in L$ such that $v \neq w$ and v is a prefix of w.

For an alphabet Σ and a language $L \subseteq \Sigma^*$, let $\sim_L$ be the *Myhill-Nerode equivalence relation* on Σ^*. So, for $v, w \in \Sigma^*$, we have $v \sim_L w$ if and only if $vu \in L \Leftrightarrow wu \in L$, for all $u \in \Sigma^*$, holds. For $w \in \Sigma^*$, we call the equivalence class $[w]_{\sim_L}$ positive if and only if $w \in L$. Otherwise, the equivalence class $[w]_{\sim_L}$ is called negative.

A *deterministic finite automaton* (DFA) is defined as a tuple $A = (Q, \Sigma, R, s, F)$, where Q is the finite set of states, Σ is the finite input alphabet, $\Sigma \cap Q = \emptyset$, R is a partial function from $Q \times \Sigma$ to Q, $s \in Q$ is the start state, and $F \subseteq Q$ is the set of final states. Elements of R are referred to a rules of A and we write $py \to q \in R$ instead of $R(p, y) = q$. A configuration of A is a string in $Q\Sigma^*$. A DFA makes a transition from configuration paw to configuration qw if $pa \to q \in R$, where $p, q \in Q$, $a \in \Sigma$, and $w \in \Sigma^*$. We denote this by $paw \vdash_A qw$ or just $paw \vdash qw$ if it is clear which DFA we are referring to. In the standard manner, we extend $\vdash$ to $\vdash^n$, where $n \geq 0$. Let $\vdash^+$ and $\vdash^*$ denote the transitive closure of $\vdash$ and the transitive-reflexive closure of $\vdash$, respectively. Then, the language accepted by A is $L(A) = \{\, w \in \Sigma^* \mid \exists f \in F : sw \vdash^* f \,\}$. We say that A accepts $w \in \Sigma^*$ if $w \in L(A)$ and that A rejects w otherwise. The family of languages accepted by DFAs is referred to as **REG**.

A *jumping finite automaton* (JFA) is a tuple $A = (Q, \Sigma, R, s, F)$, where Q, Σ, R, s, and F are the same as in the case of DFAs. A configuration of A is a string in $\Sigma^* Q \Sigma^*$. The binary jumping relation, symbolically denoted by $\curvearrowright_A$, over $\Sigma^* Q \Sigma^*$ is defined as follows. Let x, z, x', z' be strings in Σ^* such that $xz = x'z'$ and $py \to q \in R$. Then, the automaton A makes a jump from $xpyz$ to $x'qz'$, symbolically written as $xpyz \curvearrowright_A x'qz'$ or just $xpyz \curvearrowright x'qz'$ if it is clear which JFA we are referring to. In the standard manner, we extend $\curvearrowright$ to $\curvearrowright^n$, where $n \geq 0$.

Let $\curvearrowright^+$ and $\curvearrowright^*$ denote the transitive closure of $\curvearrowright$ and the transitive-reflexive closure of $\curvearrowright$, respectively. Then, the language accepted by A is $L(A) = \{\, uv \mid u, v \in \Sigma^*, \exists f \in F : usv \curvearrowright^* f \,\}$. We say that A accepts $w \in \Sigma^*$ if $w \in L(A)$ and that A rejects w otherwise. Let **JFA** be the family of all languages that are accepted by JFAs.

A *right one-way jumping finite automaton* (ROWJFA) is a tuple $A = (Q, \Sigma, R, s, F)$, where the elements Q, Σ, R, s, and F are defined as in a DFA. A configuration of A is a string in $Q\Sigma^*$. The right one-way jumping relation, symbolically denoted by $\circlearrowright_A$, over $Q\Sigma^*$ is defined as follows. For $p \in Q$ we set

$$\Sigma_p = \Sigma_{R,p} = \{\, b \in \Sigma \mid pb \to q \in R \text{ for some } q \in Q \,\}.$$

Now, let $pa \to q \in R$, $x \in (\Sigma \setminus \Sigma_p)^*$, and $y \in \Sigma^*$. Then, the ROWJFA A makes a jump from the configuration $pxay$ to the configuration qyx, symbolically written as $pxay \circlearrowright_A qyx$. We simply write $pxay \circlearrowright qyx$ if it is clear which ROWJFA we are referring to. In the standard manner, we extend $\circlearrowright$ to $\circlearrowright^n$, where $n \geq 0$. Let $\circlearrowright^+$ and $\circlearrowright^*$ denote the transitive closure of $\circlearrowright$ and the transitive-reflexive closure of $\circlearrowright$, respectively. The language accepted by A is the set $L(A) = \{\, w \in \Sigma^* \mid \exists f \in F : sw \circlearrowright^* f \,\}$. We say that A accepts $w \in \Sigma^*$ if $w \in L(A)$ and that A rejects w otherwise. Let **ROWJ** be the family of all languages that are accepted by ROWJFAs. Furthermore, for $n \geq 0$, be the class of all languages accepted by ROWJFAs with at most n accepting states is referred to as $\mathbf{ROWJ_n}$.

Besides the above mentioned language families let **FIN**, **DCF**, **CF**, and **CS** be the families of finite, deterministic context-free, context-free, and context-sensitive languages. Moreover, we are interested in permutation closed language families. These language families are referred to by a prefix **p**. E.,g., **pROWJ** denotes the language family of all permutation closed **ROWJ** languages.

Sometimes, for a DFA A, we will also consider the relations $\curvearrowright$ and $\circlearrowright$, that we get by interpreting A as a JFA or a ROWJFA. The following three languages are associated to A:

- $L_D(A)$ is the language accepted by A, interpreted as an ordinary DFA.
- $L_J(A)$ is the language accepted by A, interpreted as an JFA.
- $L_R(A)$ is the language accepted by A, interpreted as an ROWJFA.

From a result in [12] and from [3, Theorem 10], we get

$$L_D(A) \subseteq L_R(A) \subseteq L_J(A) = \mathsf{perm}(L_D(A)). \tag{1}$$

As a consequence, we have **JFA** = **pJFA**. We give an example of a DFA A with the property that $L_D(A) \subset L_R(A) \subset L_J(A)$:

Example 1. Let A be the DFA

$$A = (\{q_0, q_1, q_2, q_3\}, \{a, b\}, R, q_0, \{q_3\}),$$

where R consists of the rules $q_0 b \to q_1$, $q_0 a \to q_2$, $q_2 b \to q_3$, and $q_3 a \to q_2$. The automaton A is depicted in Figure 1.

It holds $L_D(A) = (ab)^+$ and

$$L_J(A) = \mathsf{perm}\left((ab)^+\right) = \left\{\, w \in \{a, b\}^+ \;\middle|\; |w|_a = |w|_b \,\right\}.$$

Then again, it is not hard to see that $L_R(A) = \{\, w \in a\{a, b\}^* \mid |w|_a = |w|_b \,\}$. Notice that this language is non-regular and not closed under permutation.

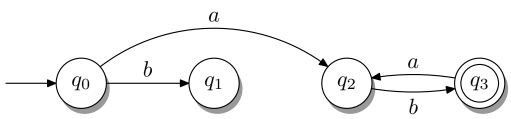

Fig. 1. The automaton A with $L_D(A) \subset L_R(A) \subset L_J(A)$.

The following basic property will be used later on.

Lemma 2. *Let $A = (Q, \Sigma, R, s, F)$ be a DFA. Consider two words $v, w \in \Sigma^*$, states $p, q \in Q$, and an $n \geq 0$ with $pv \circlearrowright^n qw$. Then, there is a word $x \in \Sigma^*$ such that xw is a permutation of v, and $px \vdash^n q$.*

Proof. We prove this by induction on n. If $n = 0$, we have $pv = qw$ and just set $x = \lambda$. Now, assume $n > 0$ and that the lemma is true for the relation $\circlearrowright^{n-1}$. We get a state $r \in Q$, a symbol $a \in \Sigma_r$, and words $y \in (\Sigma \setminus \Sigma_r)^*$ and $z \in \Sigma^*$ such that $w = zy$ and $pv \circlearrowright^{n-1} ryaz \circlearrowright qw$. By the induction hypothesis, there is a word $x' \in \Sigma^*$ such that $x'yaz$ is a permutation of v, and $px' \vdash^{n-1} r$. Set $x = x'a$. Then, the word $xw = x'azy$ is a permutation of $x'yaz$, which is a permutation of v. Furthermore, we get $px = px'a \vdash^{n-1} ra \vdash q$. This proves the lemma. ☐

3 A Characterization of Permutation Closed Languages Accepted by ROWJFAs

By the Myhill-Nerode theorem, a language L is regular if and only if the Myhill-Nerode equivalence relation $\sim_L$ has only a finite number of equivalence classes. Moreover, the number of equivalence classes equals the number of states of the minimal DFA accepting L, see for example [10]. We can give a similar characterization for permutation closed languages that are accepted by an ROWJFA.

Theorem 3. *Let L be a permutation closed language and $n \geq 0$. Then, the language L is in $\mathbf{ROWJ_n}$ if and only if the Myhill-Nerode equivalence relation $\sim_L$ has at most n positive equivalence classes.*

Proof. First, assume that L is in $\mathbf{ROWJ_n}$ and let $A = (Q, \Sigma, R, s, F)$ be a DFA with $|F| \leq n$ and $L_R(A) = L$. Consider $v, w \in L$ and $f \in F$ with $sv \circlearrowright^* f$ and $sw \circlearrowright^* f$. Lemma 2 shows that there are permutations v' and w' of v and w with $sv' \vdash^* f$ and $sw' \vdash^* f$. Because language L is closed under permutation we have $v \sim_L v'$ and $w \sim_L w'$. Now, let $u \in \Sigma^*$. Thus $sv'u \circlearrowright^* fu$ and $sw'u \circlearrowright^* fu$. That gives us

$$v'u \in L \Leftrightarrow (\exists g \in F : fu \circlearrowright^* g) \Leftrightarrow w'u \in L.$$

We have shown $v \sim_L v' \sim_L w' \sim_L w$. From $L = \bigcup_{f \in F} \{\, w \in \Sigma^* \mid sw \circlearrowright^* f \,\}$, we conclude that $|L/\sim_L| \leq |F| \leq n$, which means that $\sim_L$ has at most n positive equivalence classes.

Assume now that $\sim_L$ has at most n positive equivalence classes and let $\Sigma = \{a_1, a_2, \ldots, a_k\}$ be an alphabet with $L \subseteq \Sigma^*$. Set $L_\lambda = L \cup \{\lambda\}$. Define the map $S : L_\lambda/\sim_L \to 2^{\mathbb{N}^k}$ through

$$[w] \mapsto \left\{ \boldsymbol{x} \in \mathbb{N}^k \setminus \mathbf{0} \;\middle|\; \psi^{-1}(\psi(w) + \boldsymbol{x}) \subseteq L \right\}.$$

The definition of $\sim_L$ and the fact that L is closed under permutation make the map S well-defined. Consider the relation $\leq$ on $\mathbb{N}^k$. For each $[w] \in L_\lambda/\sim_L$, let $M([w])$ be the set of minimal elements of $S([w])$. So, for every $[w] \in L_\lambda/\sim_L$ and $\boldsymbol{x} \in S([w])$, there is an $\boldsymbol{x_0} \in M([w])$ such that $\boldsymbol{x_0} \leq \boldsymbol{x}$. Due to [5] each subset of $\mathbb{N}^k$ has only a finite number of minimal elements, so the sets $M([w])$ are finite. For $i \in \{1, 2, \ldots, k\}$, let $\pi_i : \mathbb{N}^k \to \mathbb{N}$ be the canonical projection on the ith factor and set

$$m_i = \max \left(\bigcup_{[w] \in L_\lambda/\sim_L} \{ \pi_i(\boldsymbol{x}) \mid \boldsymbol{x} \in M([w]) \} \right),$$

where $\max(\emptyset)$ should be 0. We have $m_i < \infty$, for all $i \in \{1, 2, \ldots, k\}$, because of $|L_\lambda/\sim_L| \leq n+1$. Let

$$Q = \left\{ q_{[wv]_{\sim_L}} \;\middle|\; w \in L_\lambda,\, v \in \Sigma^* \text{ with } |v|_{a_i} \leq m_i, \text{ for all } i \in \{1, 2, \ldots, k\} \right\}$$

be a set of states. The finiteness of $L_\lambda/\sim_L$ implies that Q is also finite. Set

$$F = \left\{ q_{[w]_{\sim_L}} \;\middle|\; w \in L \right\} \subseteq Q.$$

We get $|F| = |L/\sim_L| \leq n$. Define the partial mapping $R : Q \times \Sigma \to Q$ by $R(q_{[y]_{\sim_L}}, a) = q_{[ya]_{\sim_L}}$, if $q_{[ya]_{\sim_L}} \in Q$, and $R(q_{[y]_{\sim_L}}, a)$ be undefined otherwise, for $a \in \Sigma$ and $y \in \Sigma^*$ with $q_{[y]_{\sim_L}} \in Q$. Consider the DFA $A = (Q, \Sigma, R, q_{[\lambda]_{\sim_L}}, F)$. We will show that $L_R(A) = L$.

First, let $y \in L_R(A)$. Then, there exists $w \in L$ with $q_{[\lambda]_{\sim_L}} y \circlearrowright^* q_{[w]_{\sim_L}}$. From Lemma 2 it follows that there is a permutation y' of y with $q_{[\lambda]_{\sim_L}} y' \vdash^* q_{[w]_{\sim_L}}$. Now, the definition of R tells us $y' \sim_L w$. We get $y' \in L$ and also $y \in L$, because L is closed under permutation. That shows the inclusion $L_R(A) \subseteq L$.

Now, let $y \in \Sigma^* \setminus L_R(A)$. There are two possibilities:

1. There is $w \in \Sigma^* \setminus L$ with $q_{[w]_{\sim_L}} \in Q$ such that $q_{[\lambda]_{\sim_L}} y \circlearrowright^* q_{[w]_{\sim_L}}$. Then, there is a permutation y' of y with $q_{[\lambda]_{\sim_L}} y' \vdash^* q_{[w]_{\sim_L}}$. We get $y' \sim_L w$. It follows $y' \notin L$, which gives us $y \notin L$.
2. There are a $w \in L_\lambda$, a $v \in \Sigma^*$ with $|v|_{a_i} \leq m_i$, for all $i \in \{1, 2, \ldots, k\}$, and a word z with $z \in (\Sigma \setminus \Sigma_{q_{[wv]_{\sim_L}}})^+$ such that $q_{[\lambda]_{\sim_L}} y \circlearrowright^* q_{[wv]_{\sim_L}} z$. By Lemma 2 there is a $y' \in \Sigma^*$ such that $y'z$ is a permutation of y and satisfies $q_{[\lambda]_{\sim_L}} y' \vdash^* q_{[wv]_{\sim_L}}$. We get $y' \sim_L wv$. Set

$$U = \bigcup_{t \in \Sigma^*} \{ u \in \Sigma^* \mid ut \in \mathsf{perm}(v) \text{ and } wu \in L_\lambda \}.$$

We have $\lambda \in U$. Let $u_0 \in U$ such that $|u_0| = \max(\{ |u| \mid u \in U \})$ and let $t_0 \in \Sigma^*$ such that $u_0 t_0 \in \mathsf{perm}(v)$. It follows that $|t_0|_{a_i} \leq |v|_{a_i} \leq m_i$, for all $i \in \{1, 2, \ldots, k\}$, and that there exists *no* $\boldsymbol{x} \in M([wu_0]_{\sim_L})$ with $\boldsymbol{x} \leq \psi(t_0)$. Otherwise, we would have an $x' \in \psi^{-1}(\boldsymbol{x})$ which is a non-empty sub-word of t_0 such that $wu_0x' \in L$, which implies $u_0x' \in U$. However, this is a contradiction to the maximality of $|u_0|$. That shows that there is no $\boldsymbol{x} \in M([wu_0]_{\sim_L})$ with $\boldsymbol{x} \leq \psi(t_0)$. Let now $\boldsymbol{x_0} \in M([wu_0]_{\sim_L})$. Then $|t_0|_{a_j} < \pi_j(\boldsymbol{x_0}) \leq m_j$, for some element $j \in \{1, 2, \ldots, k\}$. Because of $|t_0|_{a_i} \leq m_i$, for all i with $i \in \{1, 2, \ldots, k\}$, and since we have the equality $z \in (\Sigma \setminus \Sigma_{q_{[wv]_{\sim_L}}})^+ = (\Sigma \setminus \Sigma_{q_{[wu_0t_0]_{\sim_L}}})^+$, we get $|z|_{a_j} = 0$. That gives $|t_0z|_{a_j} < \pi_j(\boldsymbol{x_0})$ and that $\psi(t_0z) \geq \boldsymbol{x_0}$ is false. So, we have shown $\psi(t_0z) \notin S([wu_0]_{\sim_L})$, which immediatley implies $wu_0t_0z \notin L$. From $wu_0t_0z \sim_L wvz \sim_L y'z \sim_L y$, it follows that $y \notin L$.

We have seen $L_R(A) = L$. This shows that L is in $\mathbf{ROWJ_n}$. □

The previous theorem allows us to determine for a lot of interesting languages whether they belong to **ROWJ** or not.

Corollary 4. *Let L be a permutation closed language. Then, the language L is in* **ROWJ** *if and only if the Myhill-Nerode equivalence relation $\sim_L$ has only a finite number of positive equivalence classes.*

An application of the last corollary is the following.

Lemma 5. *The language $L = \{\, w \in \{a,b\}^* \mid |w|_b = 0 \vee |w|_b = |w|_a \,\}$ is not included in* **ROWJ**.

Proof. Obviously, the language L is closed under permutation. For $\sim_L$, the positive equivalence classes $[a^0], [a^1], \ldots$ are pairwise different, since $a^n b^m \in L$ if and only if $m \in \{0, n\}$. Corollary 4 tells us that L is not in **ROWJ**. □

There are counterexamples for both implications of Corollary 4, if we do not assume that the language L is closed under permutation. For instance, set $L = \{\, a^n b^n \mid n \geq 0 \,\}$, which was shown to be not in **ROWJ** in [3]. Then, the positive equivalence classes of $\sim_L$ are $[\lambda]$ and $[ab]$. On the other hand, we have:

Lemma 6. *There is a language L in* **ROWJ** *such that $\sim_L$ has an infinite number of positive equivalence classes.*

Proof. Let A be the ROWJFA

$$(\{q_0, q_1, q_2, q_3, q_4\}, \{a, b\}, R, q_0, \{q_2, q_3\}),$$

where R consists of the rules $q_0 b \to q_1$, $q_1 a \to q_2$, $q_2 a \to q_2$, $q_1 b \to q_3$, $q_3 a \to q_4$, and $q_4 b \to q_3$. The ROWJFA A is depicted in Figure 2. Let $n > 0$. Then, we have $q_0 a^n b \circlearrowright q_1 a^n \circlearrowright^+ q_2$, which

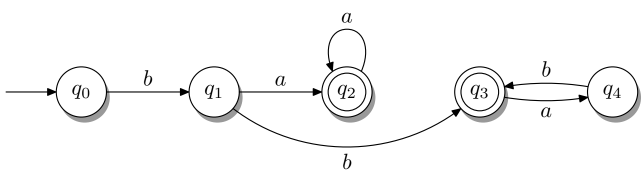

Fig. 2. The ROWJFA A accepting a language that has an infinite number of positive equivalence classes w.r.t. $\sim_{L(A)}$.

gives $a^n b \in L(A)$. We also have

$$q_0 a^n bbb^n \circlearrowright q_1 bb^n a^n \circlearrowright q_3 b^n a^n \circlearrowright^2 q_3 b^{n-1} a^{n-1} \circlearrowright^2 \cdots \circlearrowright^2 q_3 b^0 a^0.$$

It follows $a^n bbb^n \in L(A)$. Whenever A is in state q_3, the number of read b's equals the number of read a's plus 2. That implies $a^n bbb^m \notin L(A)$, for all $m \geq 0$ with $m \neq n$. So, the positive equivalence classes $[a^1 b]_{\sim_{L(A)}}, [a^2 b]_{\sim_{L(A)}}, \ldots$ are pairwise different. This proves the lemma. □

From Corollary 4 we conclude the following equivalence.

Corollary 7. *Let L be a permutation closed* **ROWJ** *language over the alphabet Σ. Then, the language L is regular if and only if $\Sigma^* \setminus L$ is in* **ROWJ**.

Proof. By Corollary 4, the Myhill-Nerode equivalence relation $\sim_L$ has only a finite number of positive equivalence classes. So, L is regular if and only if $\sim_L$ has only a finite number of negative equivalence classes, by the Myhill-Nerode theorem. The latter condition holds if and only if $\sim_{\Sigma^* \setminus L}$ has only a finite number of positive equivalence classes. Again by Corollary 4, this is equivalent to the condition that $\Sigma^* \setminus L$ is in **ROWJ**, because the complement of a permutation closed language is also permutation closed. □

The previous corollary gives us:

Lemma 8. *The language* $\{ w \in \{a,b\}^* \mid |w|_a \neq |w|_b \}$ *is not in* **ROWJ**.

Proof. Consider the permutation closed non-regular language $L = \{ w \in \{a,b\}^* \mid |w|_a = |w|_b \}$ over the alpbabet $\Sigma = \{a,b\}$. In [3], it was shown that L is in **ROWJ**. Now, by Corollary 7, the language

$$\Sigma^* \setminus L = \{ w \in \{a,b\}^* \mid |w|_a \neq |w|_b \}$$

is not in **ROWJ**. □

Having the statement of Theorem 3, it is natural to ask, which numbers arise as the number of positive equivalence classes of the Myhill-Nerode equivalence relation $\sim_L$ of a permutation closed language L. The answer is, that all natural numbers arise this way, even if we restrict ourselves to some special families:

Theorem 9. *For each $n > 0$, there is a permutation closed language which is (1) finite, (2) regular, but infinite, (3) context-free, but non-regular, (4) non-context-free such that the corresponding Myhill-Nerode equivalence relation has exactly n positive equivalence classes.*

Proof. For each $n > 0$, set

$$\begin{aligned}
L_n &= \{ a^m \mid m < n \} , \\
M_n &= \{ a^m \mid m \mod (n+1) \neq n \} , \\
N_n &= \left\{ w \in \{a,b\}^+ \;\middle|\; |w|_a = |w|_b \right\} \cup \{ c^m \mid 0 < m < n \} , \\
O_n &= \left\{ w \in \{a,b,c\}^+ \;\middle|\; |w|_a = |w|_b = |w|_c \right\} \cup \{ d^m \mid 0 < m < n \} .
\end{aligned}$$

All these languages are closed under permutation. Obviously, the language L_n is finite. The positive equivalence classes of $\sim_{L_n}$ are $[a^0], [a^1], \ldots, [a^{n-1}]$.

While M_n is infinite, the equivalence classes of $\sim_{M_n}$ are $[a^0], [a^1], \ldots, [a^n]$. So, the language M_n is regular, by the Myhill-Nerode Theorem. Only the last mentioned equivalence class is negative. Therefore, there are exactly n positive equivalence classes of $\sim_{M_n}$.

For $\sim_{N_n}$, the equivalence classes $[a^0], [a^1], \ldots$ are pairwise different. So, the language N_n is non-regular, by the Myhill-Nerode Theorem. It is context-free, because it is the union of a well known context-free language and a finite language. The positive equivalence classes of the relation $\sim_{N_n}$ are $[ab], [c^1], [c^2], \ldots, [c^{n-1}]$.

If O_n was context-free, then

$$O_n \cap \{a,b,c\}^* = \left\{ w \in \{a,b,c\}^+ \;\middle|\; |w|_a = |w|_b = |w|_c \right\}$$

would also be context-free, as the intersection of a context-free and a regular language. However, the language $\{ w \in \{a,b,c\}^+ \mid |w|_a = |w|_b = |w|_c \}$ is a well known non-context-free language. It follows, that O_n is also non-context-free. The positive equivalence classes of the relation $\sim_{O_n}$ are $[abc], [d^1], [d^2], \ldots, [d^{n-1}]$. This proves the theorem. □

The previous theorem, together with Theorem 3, implies that the language families $\mathbf{ROWJ_n}$ form a proper hierarchy, even if we only consider languages out of special language families:

Corollary 10. *For all $n \geq 0$, we have*

$$\begin{gathered}
\mathbf{ROWJ_n} \subset \mathbf{ROWJ_{n+1}}, \\
\mathbf{ROWJ_n} \cap \mathbf{FIN} \subset \mathbf{ROWJ_{n+1}} \cap \mathbf{FIN}, \\
\mathbf{ROWJ_n} \cap (\mathbf{REG} \setminus \mathbf{FIN}) \subset \mathbf{ROWJ_{n+1}} \cap (\mathbf{REG} \setminus \mathbf{FIN}), \\
\mathbf{ROWJ_n} \cap (\mathbf{CF} \setminus \mathbf{REG}) \subset \mathbf{ROWJ_{n+1}} \cap (\mathbf{CF} \setminus \mathbf{REG}), \\
\mathbf{ROWJ_n} \cap (\mathbf{CS} \setminus \mathbf{CF}) \subset \mathbf{ROWJ_{n+1}} \cap (\mathbf{CS} \setminus \mathbf{CF}).
\end{gathered}$$

The statement remains valid if restricted to permutation closed languages.

4 Inclusion Relations Between Language Families

We investigate inclusion relations between **ROWJ** and other important languages families. The following inclusion relations were given in [3]:

- $\mathbf{REG} \subset \mathbf{ROWJ}$,
- **ROWJ** and **CF** are incomparable,
- $\mathbf{ROWJ} \nsubseteq \mathbf{JFA}$.

It was stated as an open problem if $\mathbf{JFA} \subset \mathbf{ROWJ}$. We can answer this:

Theorem 11. *The language families* **ROWJ** *and* **JFA** *are incomparable.*

Proof. The language $\{\, w \in \{a,b\}^* \mid |w|_b = 0 \vee |w|_b = |w|_a \,\}$ is not included in the family **ROWJ**, by Lemma 5, but it belongs to **JFA**, because it is the permutation closure of the regular language $a^* \cup (ab)^*$. So, we get $\mathbf{JFA} \nsubseteq \mathbf{ROWJ}$. Together with the result $\mathbf{ROWJ} \nsubseteq \mathbf{JFA}$ from [3] the incomparability of the language families **ROWJ** and **JFA** follows. □

For the complexity of **ROWJ**, we get:

Theorem 12. *The language family* **ROWJ** *is included in* $\mathbf{DTIME}(n^2)$ *and* $\mathbf{DSPACE}(n)$.

Proof. Right revolving automata were described in [2]. It was shown that every language accepted by a deterministic right revolving automaton belongs to bothclasses $\mathbf{DTIME}(n^2)$ and $\mathbf{DSPACE}(n)$. In [3] it was proven that **ROWJ** is properly included in the family of languages accepted by deterministic right revolving automata. □

This implies that **ROWJ** is properly included in **CS**:

Theorem 13. *We have* $\mathbf{ROWJ} \subset \mathbf{CS}$.

Proof. From Theorem 12 we get $\mathbf{ROWJ} \subseteq \mathbf{CS}$. On the other hand, **ROWJ** and **CF** are incomparable, which proves the theorem. □

We also get a result for the inclusion relation between **ROWJ** and the family of deterministic context-free languages:

Theorem 14. *The language families* **ROWJ** *and* **DCF** *are incomparable.*

Proof. The families **ROWJ** and **CF** are incomparable, so there are non context-free languages in **ROWJ**. Morever, it was shown in [3] that the deterministic context-free language

$$\{\, a^n b^n \mid n \geq 0 \,\}$$

is not accepted by any ROWJFA. □

By the famous result in [13], every context-free language is semilinear. In [3] it was proven that every language in **JFA** is also semilinear. This holds for **ROWJ**, too:

Theorem 15. *Every language in* **ROWJ** *is semilinear.*

Proof. For every language L in **ROWJ**, there exists a DFA A such that $L = L_R(A)$. From (1) we get

$$\psi(L_D(A)) \subseteq \psi(L_R(A)) \subseteq \psi(L_J(A)).$$

Because of $L_J(A) = \mathsf{perm}(L_D(A))$, we have $\psi(L_J(A)) = \psi(L_D(A))$. So,

$$\psi(L) = \psi(L_R(A)) = \psi(L_D(A)),$$

which is a semilinear set, because $L_D(A)$ is regular. □

We now consider inclusion relations between families of permutation closed languages. It holds

$$\mathbf{pFIN} \subset \mathbf{pREG} \subset \mathbf{pDCF} \subseteq \mathbf{pCF} \subset \mathbf{pCS}, \tag{2}$$

witness languages are a^*, $\{\, w \in \{a,b\}^* \mid |w|_a = |w|_b \,\}$, and $\left\{\, a^{2^n} \;\middle|\; n \geq 0 \,\right\}$. There is also a language that distinguishes **pDCF** and **pCF**:

Theorem 16. *We have* **pDCF** $\subset$ **pCF**.

Proof. Consider the permutation closed language

$$L = \{\, w \in \{a,b,c\}^* \mid |w|_a = |w|_b \vee |w|_b = |w|_c \,\}.$$

It is context-free as the union of two context-free languages. If L was deterministic context-free, then

$$L' = L \cap a^*b^*c^* = \left\{\, a^ib^jc^k \;\middle|\; (i,j,k \geq 0) \wedge (i = j \vee j = k) \,\right\}$$

was also deterministic context-free as the intersection of a deterministic context-free and a regular language. However, language L' is not deterministic context-free, as shown in [10]. Hence language L is also not deterministic context-free, which proves the theorem. □

The next theorem places **JFA** in the hierarchy (2).

Theorem 17. *We have* **pCF** $\subset$ **JFA** $\subset$ **pCS**.

Proof. The first strict inclusion is seen as follows: it was shown that every context-free language is semilinear in [13], while in [3] it was proven that **JFA** is the family of all permutation closed semilinear languages. So, we get **pCF** $\subseteq$ **JFA**. On the other hand, the non-context free language

$$\{\, w \in \{a,b,c\}^* \mid |w|_a = |w|_b = |w|_c \,\}$$

is in **JFA**, which was shown in [11]. This proves the first inclusion.

For the second strict inclusion we argue as follows: in [11] it was proven that **JFA** $\subset$ **CS** and that all languages in **JFA** are closed under permutation. This gives us **JFA** $\subseteq$ **pCS**. The permutation closed context-sensitive language $\left\{\, a^{2^n} \;\middle|\; n \geq 0 \,\right\}$ is not in **JFA**, because it is not semilinear. □

So, we get

$$\mathbf{pFIN} \subset \mathbf{pREG} \subset \mathbf{pDCF} \subset \mathbf{pCF} \subset \mathbf{JFA} \subset \mathbf{pCS}.$$

We investigate the inclusion relations of **pROWJ**, now.

Theorem 18. *We have* $\mathbf{pREG} \subset \mathbf{pROWJ} \subset \mathbf{JFA}$*.*

Proof. Since $\mathbf{REG} \subset \mathbf{ROWJ}$, we have $\mathbf{pREG} \subseteq \mathbf{pROWJ}$. The permutation closed, non-regular language $\{\, w \in \{a,b\}^* \mid |w|_a = |w|_b \,\}$ was shown to be included in **ROWJ** in [3].

Theorem 15 implies $\mathbf{pROWJ} \subseteq \mathbf{JFA}$, because **JFA** is the family of all permutation closed semilinear languages. On the other hand, Theorem 11 tells us that there is a language in **JFA**, which is not in **pROWJ**. □

Next, we consider the inclusion relations between **pROWJ** and the language families **pDCF** and **pCF**.

Theorem 19. *The language family* **pROWJ** *is incomparable to* **pDCF** *and to* **pCF***.*

Proof. From [3] we know that the permutation closed, non-context-free language

$$\{\, w \in \{a,b,c\}^* \mid |w|_a = |w|_b = |w|_c \,\}$$

is in **ROWJ**. The language $L = \{\, w \in \{a,b\}^* \mid |w|_a = |w|_b \,\}$ is in **pDCF** and so is $\{a,b\}^* \setminus L$, because **pDCF** is closed under complementation. Lemma 8 gives us that $\{a,b\}^* \setminus L$ is not in **ROWJ**, which proves the theorem. □

Finally, we get from Theorems 11 and 18:

Theorem 20. *We have* $\mathbf{pROWJ} \subset \mathbf{ROWJ}$*.*

5 Closure Properties of ROWJ and pROWJ

We consider closure properties of the language families **ROWJ** and **pROWJ**. Our results are summarized in Table 1.

The language family **ROWJ** is not closed under the operations of intersection, intersection with regular languages, reversal, concatenation, concatenation with regular languages from the right, Kleene star, Kleene plus, and substitution. All these properties were proven in [3]. In the following we will show that **ROWJ** is also not closed under the operations of union, union with regular languages, complement, concatenation with regular languages from the left, homomorphism, λ-free homomorphism, inverse homomorphism and permutation closure. However, we will prove one positive closure result: the family **ROWJ** is closed under concatenation with prefix-free regular languages from the left.

Theorem 21. *The family* **ROWJ** *is not closed under union and under union with regular languages.*

Proof. Consider $L_1 = a^*$ and $L_2 = \{\, w \in \{a,b\}^* \mid |w|_a = |w|_b \,\}$. The language L_1 is in **ROWJ**, because it is regular, while L_2 was shown to be in **ROWJ** in [3]. In Lemma 5 it was shown that the union $L_1 \cup L_2$ is not in **ROWJ**. □

Next we consider the complementation operation.

Theorem 22. *The family* **ROWJ** *is not closed under complement.*

Proof. While $\{\, w \in \{a,b\}^* \mid |w|_a = |w|_b \,\}$ is in **ROWJ**, its complement is not, which was shown in Lemma 8. □

Closed under	Language family REG	pROWJ	ROWJ	JFA
Union	yes	no	no	yes
Union with reg. lang.	yes	no	no	no
Intersection	yes	yes	no	yes
Intersection with reg. lang.	yes	no	no	no
Complementation	yes	no	no	yes
Reversal	yes	yes	no	yes
Concatenation	yes	no	no	no
Right conc. with reg. lang.	yes	no	no	no
Left conc. with reg. lang.	yes	no	no	no
Left conc. with prefix-free reg. lang.	yes	no	yes	no
Kleene star	yes	no	no	no
Kleene plus	yes	no	no	no
Homomorphism	yes	no	no	no
λ-free homomorphism	yes	no	no	no
Inv. homomorphism	yes	yes	no	yes
Substitution	yes	no	no	no
Permutation	no	yes	no	yes

Table 1. Closure properties of **ROWJ** and **pROWJ**. The gray shaded results are proven in this paper. The non-shaded closure properties for **REG** are folklore. For **ROWJ** the closure/non-closure results can be found in [3] and that for the language family **JFA** in [1, 6, 7, 12].

From [3] we know that **ROWJ** is not closed under concatenation, not even under concatenation with regular languages from the right. Also, under concatenation with regular languages from the left, the family **ROWJ** is not closed:

Theorem 23. *The family* **ROWJ** *is not closed under concatenation with regular languages from the left.*

Proof. Consider the regular language $L_1 = a^*$ and the **ROWJ** language

$$L_2 = \{\, w \in \{a,b\}^* \mid |w|_a = |w|_b \,\}.$$

Assume that there is a DFA $A = (Q, \{a,b\}, R, s, F)$ with $L_R(A) = L_1L_2$. For each $n \geq 0$, there is exactly one $q_n \in F$ with $sa^n \vdash^* q_n$. Because of $|F| \leq \infty$, there are $0 \leq n < m$ with $q_n = q_m$. Since the word $\lambda a^m b^m$ belongs to L_1L_2, there exists $q \in F$ with $sa^m b^m \vdash^* q_m b^m \vdash^* q$. This implies that $sa^n b^m \vdash^* q_m b^m \vdash^* q$, which gives us $a^n b^m \in L_1L_2$. That is a contradiction, because of $m > n$. Thus, the language L_1L_2 is not in **ROWJ**. $\square$

If we add the condition that the regular language has to be prefix-free, we get a positive closure result:

Theorem 24. *The family* **ROWJ** *is closed under concatenation with prefix-free regular languages from the left.*

Proof. For an alphabet Σ, let $L_1 \subseteq \Sigma^*$ be a prefix-free regular language and moreover $L_2 \subseteq \Sigma^*$ be a **ROWJ** language. If $\lambda \in L_1$, we have $L_1 = \{\lambda\}$ and therefore $L_1L_2 = L_2$. Thus, assume form now on that $\lambda \notin L_1$. Let $A_1 = (Q_1, \Sigma, R_1, s_1, F_1)$ be a DFA with total transition function R_1

and $L_D(A_1) = L_1$. Moreover, let $A_2 = (Q_2, \Sigma, R_2, s_2, F_2)$ be a DFA with $L_R(A_2) = L_2$ and assume $Q_1 \cap Q_2 = \emptyset$ without loss of generality. Consider the DFA

$$B = ((Q_1 \setminus F_1) \cup Q_2, \Sigma, S, s_1, F_2),$$

where S is defined as follows: for $(q, a) \in (Q_1 \setminus F_1) \times \Sigma$, let $S(q, a) = R_1(q, a)$, if $R_1(q, a) \notin F_1$, and $S(q, a) = s_2$, otherwise. For $(q, a) \in Q_2 \times \Sigma$, the value $S(q, a)$ is defined if and only if $R_2(q, a)$ is defined. In this case we have $S(q, a) = R_2(q, a)$. We will show that $L_R(B) = L_1L_2$.

First, let $v \in L_1$ and $w \in L_2$. So, there is a symbol $a \in \Sigma$ and states $p \in Q_1$, $q \in F_1$, and $r \in F_2$ such that $s_1v \vdash_{A_1}^* pa \vdash_{A_1} q$ and $s_2w \circlearrowright_{A_2}^* r$. Because L_1 is prefix-free, there are no word $x \in \Sigma^+$ and $q' \in F_1$ such that $s_1v \vdash_{A_1}^* q'x$. This gives us $s_1vw \vdash_B^* paw \vdash_B s_2w \circlearrowright_B^* r$, which implies $vw \in L_R(B)$.

Let now $v \in L_R(B)$. Since R_1 is a total function, there are a symbol $a \in \Sigma$, words $w, x \in \Sigma^*$, and states $p \in Q_1 \setminus F_1$ and $q \in F_2$ such that $v = wax$ and $s_1wax \vdash_B^* pax \vdash_B s_2x \circlearrowright_B^* q$. So, there is an $r \in F_1$ with $s_1wa \vdash_{A_1}^* pa \vdash_{A_1} r$ and $s_2x \circlearrowright_{A_2}^* q$. This gives us $wa \in L_1$ and $x \in L_2$, which proves the theorem. □

The previous theorem allows us for a large family of languages to show that they belong to **ROWJ**. From Corollary 4 and Theorem 24 it follows that:

Corollary 25. *Let Σ be an alphabet and $w \in \Sigma^*$. Furthermore, let $L \subseteq \Sigma^*$ be a permutation closed language such that the Myhill-Nerode equivalence relation $\sim_L$ has only a finite number of positive equivalence classes. Then, the language wL is in* **ROWJ**.

For marked concatenation we find a similar result, which can be deduced from Corollary 4 and Theorem 24, too, because L_1a is a prefix-free regular language.

Corollary 26. *Let Σ be an alphabet and $a \in \Sigma$. Moreover, let $L_1 \subseteq (\Sigma \setminus \{a\})^*$ be a regular language and $L_2 \subseteq \Sigma^*$ be a permutation closed language such that the Myhill-Nerode equivalence relation $\sim_{L_2}$ has only a finite number of positive equivalence classes. Then, the language L_1aL_2 is in* **ROWJ**.

Now, we turn back to the closure properties of **ROWJ**.

Theorem 27. *The family* **ROWJ** *is not closed under λ-free homomorphism nor under homomorphism.*

Proof. Consider the permutation closed language

$$L = a^* \cup \{\, w \in \{b, c\}^* \mid |w|_b = |w|_c \,\}.$$

The positive equivalence classes of $\sim_L$ are $[\lambda]$, $[a]$, and $[bc]$. So, the language L is in **ROWJ**, by Corollary 4. Let the λ-free homomorphism $h : \{a, b, c\}^* \to \{a, b\}^*$ be defined by $h(a) = a$, $h(b) = b$, and $h(c) = a$. Then, we get

$$h(L) = \{\, w \in \{a, b\}^* \mid |w|_b = 0 \vee |w|_b = |w|_a \,\},$$

which was shown to be not in **ROWJ** in Lemma 5. □

We also consider the operation of inverse homomorphism:

Theorem 28. *The family* **ROWJ** *is not closed under inverse homomorphism.*

Proof. Let A be the ROWJFA $A = (\{q_0, q_1, q_2\}, \{a, b, c\}, R, q_0, \{q_0, q_2\})$, where R consists of the rules $q_0c \to q_0$, $q_0b \to q_1$, $q_1a \to q_2$, and $q_2b \to q_1$. The ROWJFA A is depicted in Figure 3.

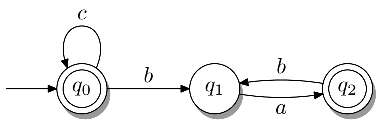

Fig. 3. The ROWJFA A satisfying $L(A) \cap \{ac, b\}^* = \{\, (ac)^n b^n \mid n \geq 0 \,\}$.

Let $h : \{a, b\}^* \to \{a, b, c\}^*$ be the homomorphism, given by $h(a) = ac$ and $h(b) = b$. It is not hard to see that $h(\{a, b\}^*) = \{ac, b\}^*$.

Let now $\lambda \neq w \in L(A) \cap \{ac, b\}^*$, which implies $|w|_b > 0$. When A reads w, it reaches the first occurrence of the symbol b in state q_0. After reading this b, the automaton is in state q_1. Now, no more c can be read. So, we get $w \in (ac)^+ b^+$. Whenever A is in state q_2, it has read the same number of a's and b's. This gives us $w \in \{\, (ac)^n b^n \mid n > 0 \,\}$. That shows the invlusion of $L(A) \cap \{ac, b\}^*$ within $\{\, (ac)^n b^n \mid n \geq 0 \,\}$

On the other hand, for $n > 0$, we have

$$q_0 (ac)^n b^n \circlearrowright^n q_0 b^n a^n \circlearrowright^2 q_2 a^{n-1} b^{n-1} \circlearrowright^2 q_2 a^{n-2} b^{n-2} \circlearrowright^2 \cdots \circlearrowright^2 q_2 ab \circlearrowright^2 q_2.$$

This implies $L(A) \cap \{ac, b\}^* = \{\, (ac)^n b^n \mid n \geq 0 \,\}$. We get

$$\begin{aligned} h^{-1}(L(A)) &= h^{-1}(L(A) \cap h(\{a, b\}^*)) \\ &= h^{-1}(L(A) \cap \{ac, b\}^*) \\ &= h^{-1}(\{\, (ac)^n b^n \mid n \geq 0 \,\}) = \{\, a^n b^n \mid n \geq 0 \,\}. \end{aligned}$$

In [3] it was shown that this language is not in **ROWJ**. □

Finally, we take a look at the permutation closure of **ROWJ**.

Theorem 29. *The family* **ROWJ** *is not closed under permutation closure.*

Proof. By Theorem 11, there is a language L, that is in **JFA**, but not in **ROWJ**. There exists a DFA A with $L_J(A) = L$. Because of (1), we have $\mathsf{perm}(L_R(A)) = L_J(A) = L$. □

Next, we consider the language family **pROWJ** in more detail. One can easily find witness languages to see that **pROWJ** is not closed under union with regular languages, intersection with regular languages, concatenation, concatenation with regular languages (from both sides), Kleene star, Kleene plus, substitution, homomorphism, and λ-free homomorphism. For all these operations, the witness languages can be chosen in a way such that the resulting language is not even permutation closed. On the other hand, it is not hard to see that the family of permutation closed languages is closed under union, intersection, complement, and inverse homomorphism. We investigate how the language family **pROWJ** behaves under the latter four operations. From the proofs of the Theorems 21 and 22 we get:

Theorem 30. *The family* **pROWJ** *is not closed under union and under complement.*

The next theorem shows that **pROWJ** is closed under intersection.

Theorem 31. *Let* $L_1 \in \mathbf{pROWJ_m}$ *and* $L_2 \in \mathbf{pROWJ_n}$*, for some* $n, m \geq 0$*. Then, the language* $L_1 \cap L_2 \in \mathbf{pROWJ_{mn}}$*.*

Proof. Let Σ be an alphabet such that L_1, $L_2 \subseteq \Sigma^*$. The set Σ^* is enumerable, so, there is a total order on Σ^* such that each non-empty subset of Σ^* has exactly one minimal element. Set

$$X = \bigcup_{\substack{(S,T)\in(L_1/\sim_{L_1})\times(L_2/\sim_{L_2}) \\ \text{with } S\cap T\neq\emptyset}} \{\min(S \cap T)\}.$$

Because of Theorem 3 we have $|L_1/\sim_{L_1}| \leq m$ and $|L_2/\sim_{L_2}| \leq n$. That gives us $|X| \leq mn$. Now, let $w \in L_1 \cap L_2$. There exists exactly one

$$(S, T) \in (L_1/\sim_{L_1}) \times (L_2/\sim_{L_2})$$

such that $w \in S \cap T$. Let $v = \min(S \cap T)$ and u be an arbitrary word in Σ^*. For $i \in \{1, 2\}$, we have $wu \in L_i$ if and only if $vu \in L_i$, because of $w \sim_{L_i} v$. This implies that $wu \in L_1 \cap L_2$ if and only if $vu \in L_1 \cap L_2$. We get $w \sim_{L_1\cap L_2} v$. So, for $\sim_{L_1\cap L_2}$, we have shown that each element of the intersection $L_1 \cap L_2$ is equivalent to an element out of X. It follows that $\sim_{L_1\cap L_2}$ has at most mn positive equivalence classes. By using Theorem 3 again, we get $L_1 \cap L_2 \in \mathbf{pROWJ_{mn}}$. $\square$

As an immediate consequence we get:

Corollary 32. *The family* **pROWJ** *is closed under intersection.*

Our next result implies that **pROWJ** is closed under inverse homomorphism.

Theorem 33. *Let Γ and Σ be alphabets and $h : \Gamma^* \to \Sigma^*$ be a homomorphism. Furthermore let $L \subseteq \Sigma^*$ be in $\mathbf{pROWJ_n}$, for some $n \geq 0$. Then, the language $h^{-1}(L)$ is also in $\mathbf{pROWJ_n}$.*

Proof. Theorem 3 gives us $|L/\sim_L| \leq n$. From $L = \bigcup_{S\in L/\sim_L} S$, we get

$$h^{-1}(L) = \bigcup_{S\in L/\sim_L} h^{-1}(S).$$

Consider now an element $S \in L/\sim_L$, two words $v, w \in h^{-1}(S)$, and an arbitrary $u \in \Gamma^*$. Because of $h(v), h(w) \in S$, we have $h(v) \sim_L h(w)$. It follows that

$$vu \in h^{-1}(L) \Leftrightarrow h(v)h(u) \in L \Leftrightarrow h(w)h(u) \in L \Leftrightarrow wu \in h^{-1}(L).$$

We have shown $v \sim_{h^{-1}(L)} w$. So, we get $\left|h^{-1}(L)/\sim_{h^{-1}(L)}\right| \leq |L/\sim_L| \leq n$, which implies that $h^{-1}(L)$ is in $\mathbf{pROWJ_n}$, by Theorem 3. $\square$

Thus we immediately get:

Corollary 34. *The family* **pROWJ** *is closed under inverse homomorphism.*

6 More on Languages Accepted by ROWJFAs

In Corollary 4 a characterization of the permutation closed languages that are in **ROWJ** was given. In this section, we characterize languages in **ROWJ** for some cases where the considered language does not need to be permutation closed.

Theorem 35. *For an alphabet Σ, let $w \in \Sigma^*$ and $L \subseteq \Sigma^*$. Then, the language wL is in* **ROWJ** *if and only if L is in* **ROWJ**.

Proof. If L is in **ROWJ**, then wL is also in **ROWJ**, because the language family **ROWJ** is closed under concatenation with prefix-free languages from the left. Now assume that wL is in **ROWJ** and $L \neq \emptyset$. We may also assume that $|w| = 1$. The general case follows from this special case *via* a trivial induction over the length of w. Thus, let $w = a$ for an $a \in \Sigma$ and let $A = (Q, \Sigma, R, s, F)$ be a DFA with $L_R(A) = aL$. In the following, we will show *via* a contradiction that the value $R(s, a)$ is defined. Assume that $R(s, a)$ is undefined and let v be an arbitrary word out of L. Because $av \in L_R(A)$, there is a symbol $b \in \Sigma_s$, two words $x \in (\Sigma \setminus \Sigma_s)^*$ and $y \in \Sigma^*$, and a state $p \in F$ such that $v = xby$ and $saxby \circlearrowleft R(s, b)yax \circlearrowleft^* p$. This gives us $sbyax \vdash R(s, b)yax \circlearrowleft^* p$, which implies $byax \in L_R(A) = aL$. However, this is a contradiction, because $b \neq a$. So, the value $R(s, a)$ is defined.

Consider the DFA $B = (Q, \Sigma, R, R(s, a), F)$. For a word $z \in \Sigma^*$, we have $z \in L_R(B)$ if and only if $az \in L_R(A) = aL$, because of $saz \vdash R(s, a)z$. That gives us $L_R(B) = L$ and we have shown that L is in **ROWJ**. □

From the previous theorem and Corollary 4 we get a generalization of the latter corollary.

Corollary 36. *For an alphabet Σ, let $w \in \Sigma^*$ and let $L \subseteq \Sigma^*$ be a permutation closed language. Then, the language wL is in* **ROWJ** *if and only if the Myhill-Nerode equivalence relation $\sim_L$ has only a finite number of positive equivalence classes.*

Next, we will give a characterization for the concatenation Lw of a language L and a word w. To do so, we need the following lemma. It treats the case of an ROWJFA that is only allowed to jump over one of the input symbols.

Lemma 37. *Let $A = (Q, \Sigma, R, s, F)$ be a DFA with a symbol $a \in \Sigma$ such that the value $R(q, b)$ is defined for all $(q, b) \in Q \times (\Sigma \setminus \{a\})$. Then, the language $L_R(A)$ is regular.*

Proof. Consider the DFA

$$B = (Q \times (Q \cup \{d\})^Q, \Sigma, S, (s, \mathrm{id}_{Q \to Q \cup \{d\}}), G),$$

where $d \notin Q \cup \Sigma$, the map $\mathrm{id}_{Q \to Q \cup \{d\}}$ is the identity map and

$$G = \left\{ (q, f) \in Q \times (Q \cup \{d\})^Q \;\middle|\; f(q) \in F \right\}.$$

The total map S is defined as follows:

$$S((q, f), b) = \begin{cases} (R(q, b), f) & \text{if } (q, f) \in Q \times (Q \cup \{d\})^Q \text{ and } b \in \Sigma_{R,q} \\ (q, g \circ f) & \text{if } (q, f) \in Q \times (Q \cup \{d\})^Q \text{ and } b \notin \Sigma_{R,q}. \end{cases}$$

The map $g : (Q \cup \{d\}) \to (Q \cup \{d\})$ is defined in the following way: let

$$g(p) = \begin{cases} R(p, a) & \text{if } p \in Q \text{ and } a \in \Sigma_{R,p} \\ d & \text{if } p \in Q \text{ and } a \notin \Sigma_{R,p} \\ d & \text{otherwise.} \end{cases}$$

This completes the description of B. We will show that $L_D(B) = L_R(A)$.

Let $w \in \Sigma^*$. We decompose the word w into factors that are consumed by A and factors that are jumped over by the automaton in question: there exists a number $m > 0$, words $w_1, w_2, \ldots, w_m \in \Sigma^*$, symbols $b_2, b_3, \ldots, b_m \in \Sigma \setminus \{a\}$, numbers $n_1, \ldots, n_m \in \mathbb{N}$ such that $n_1 n_2 \cdots n_{m-1} > 0$, and states $p_1, p_2, \ldots, p_m \in Q$ and $q_2, q_3, \ldots, q_m \in Q$ with

$$\begin{aligned}
sw = sw_1 a^{n_1} \prod_{i=2}^{m} b_i w_i a^{n_i} &\vdash_A^{|w_1|} p_1 a^{n_1} \prod_{i=2}^{m} b_i w_i a^{n_i} \\
&\circlearrowleft_A q_2 w_2 a^{n_2} \left(\prod_{i=3}^{m} b_i w_i a^{n_i} \right) a^{n_1} \\
&\vdash_A^{|w_2|} p_2 a^{n_2} \left(\prod_{i=3}^{m} b_i w_i a^{n_i} \right) a^{n_1} \\
&\circlearrowleft_A q_3 w_3 a^{n_3} \left(\prod_{i=4}^{m} b_i w_i a^{n_i} \right) a^{n_1 + n_2} \\
&\ldots \\
&\vdash_A^{|w_{m-1}|} p_{m-1} a^{n_{m-1}} b_m w_m a^{n_m + \sum_{i=1}^{m-2} n_i} \\
&\circlearrowleft_A q_m w_m a^{\sum_{i=1}^{m} n_i} \\
&\vdash_A^{|w_m|} p_m a^{\sum_{i=1}^{m} n_i}.
\end{aligned}$$

We have $w \in L_R(A)$ if and only if $g^{\sum_{i=1}^{m} n_i}(p_m) \in F$. On the other hand, we get the following computation

$$\begin{aligned}
(s, \mathrm{id}_{Q \to Q \cup \{d\}}) w = (s, \mathrm{id}_{Q \to Q \cup \{d\}}) w_1 a^{n_1} \prod_{i=2}^{m} b_i w_i a^{n_i} \\
&\vdash_B^{|w_1|} (p_1, \mathrm{id}_{Q \to Q \cup \{d\}}) a^{n_1} \prod_{i=2}^{m} b_i w_i a^{n_i} \\
&\vdash_B^{n_1 + 1} (q_2, g^{n_1}|_Q) w_2 a^{n_2} \left(\prod_{i=3}^{m} b_i w_i a^{n_i} \right) \\
&\vdash_B^{|w_2|} (p_2, g^{n_1}|_Q) a^{n_2} \left(\prod_{i=3}^{m} b_i w_i a^{n_i} \right) \\
&\vdash_B^{n_2 + 1} (q_3, g^{n_1 + n_2}|_Q) w_3 a^{n_3} \left(\prod_{i=4}^{m} b_i w_i a^{n_i} \right) \\
&\ldots \\
&\vdash_B^{|w_{m-1}|} (p_{m-1}, g^{\sum_{i=1}^{m-2} n_i}|_Q) a^{n_{m-1}} b_m w_m a^{n_m} \\
&\vdash_B^{n_{m-1}+1} (q_m, g^{\sum_{i=1}^{m-1} n_i}|_Q) w_m a^{n_m} \\
&\vdash_B^{|w_m|} (p_m, g^{\sum_{i=1}^{m-1} n_i}|_Q) a^{n_m}.
\end{aligned}$$

Set $k = \max\{\, r \in \{0, 1, \ldots, n_m\} \mid g^r(p_m) \in Q \,\}$. That gives

$$\begin{aligned}
(p_m, g^{\sum_{i=1}^{m-1} n_i}|_Q) a^{n_m} &\vdash_B^{k} (g^k(p_m), g^{\sum_{i=1}^{m-1} n_i}|_Q) a^{n_m - k} \\
&\vdash_B^{n_m - k} (g^k(p_m), g^{\sum_{i=1}^{m} n_i - k}|_Q).
\end{aligned}$$

Thus, we have $w \in L_D(B)$ if and only if

$$g^{\sum_{i=1}^{m} n_i}(p_m) = g^{\sum_{i=1}^{m} n_i - k}\left(g^k(p_m)\right) \in F,$$

which holds if and only if $w \in L_R(A)$. That shows $L_D(B) = L_R(A)$ and that $L_R(A)$ is a regular language. This proves the lemma. $\square$

Our characterization for languages of the form Lw generalizes a result from [3], which says that the language $\{ va \mid v \in \{a,b\}^*, |v|_a = |v|_b \}$ is not in **ROWJ**:

Theorem 38. *For an alphabet Σ, let $w \in \Sigma^*$ be a non-empty word and $L \subseteq \Sigma^*$. Then, the language Lw is in* **ROWJ** *if and only if L is regular.*

Proof. If L is regular, then Lw is also regular, which means that Lw is in **ROWJ**. Assume now, that Lw is in **ROWJ**. As in the proof of Theorem 35, we can assume that $w = a$ for an $a \in \Sigma$. Let $A = (Q, \Sigma, R, s, F)$ be a DFA with $L_R(A) = La$. Consider the DFA $B = (Q \cup \{d\}, \Sigma, S, s, F)$ and let d be a new symbol with $d \notin Q \cup \Sigma$. The map S is defined as follows: for $(q, b) \in Q \times \Sigma$, we set $S(q, b) = R(q, b)$, if $R(q, b)$ is defined. If $R(q, b)$ is undefined and $b \neq a$, we define $S(q, b) = d$. For all $q \in Q$, the value $S(q, a)$ is undefined, if $R(q, a)$ is undefined. Finally, for all $b \in \Sigma$, it holds $S(d, b) = d$. By Lemma 37, the language $L_R(B)$ is regular. We will show that

$$L_R(A) = L_R(B).$$

Then, the regularity of $La = L_R(A)$ implies the regularity of L, because regular languages are closed under the operation of quotient with a regular language.

First, let $v \in L_R(B)$ and $f \in F$ with $sv \circlearrowleft_B^* f$. For a state $q \in Q$, a symbol $b \in \Sigma_{S,q}$, and words $x \in (\Sigma \setminus \Sigma_{S,q})^*$ and $y \in \Sigma^*$ with $sv \circlearrowleft_B^* qxby$, we have $x \in (\Sigma \setminus \Sigma_{R,q})^*$ and

$$qxby \circlearrowleft_B S(q, b)yx \circlearrowleft_B^* f.$$

This implies $S(q, b) \neq d$, which tells us $b \in \Sigma_{R,q}$ and $R(q, b) = S(q, b)$. We get $qxby \circlearrowleft_A S(q, b)yx$. By induction, we see that $sv \circlearrowleft_A^* f$. Therefore, we have $v \in L_R(A)$.

Now, let $v \in L_R(A)$ and $f \in F$ with $sv \circlearrowleft_A^* f$. Assume that $v \notin L_R(B)$. Then, there exists a symbol out of $\Sigma \setminus \{a\}$ that is jumped over during the processing of A, when the starting configuration is sv. The part of v that is visited by A before it jumps over the first symbol out of $\Sigma \setminus \{a\}$ will be decomposed into factors that are consumed by A and factors that are jumped over by the device under consideration: there is a natural number $m > 0$, words $w_1, w_2, \ldots, w_{m+2} \in \Sigma^*$, symbols $b_2, b_3, \ldots, b_{m+1} \in \Sigma \setminus \{a\}$ and $c \in \Sigma$, numbers $n_1, \ldots, n_m \in \mathbb{N}$ with $n_1 n_2 \cdots n_{m-1} > 0$, and states $p_1, p_2, \ldots, p_m, q_2, q_3, \ldots, q_{m+1} \in Q$ with symbols $b_{i+1} \in \Sigma_{p_i}$, for every i satisfying $i \in \{1, 2, \ldots, m-1\}$, such that word $a^{n_m} b_{m+1} w_{m+1} \in (\Sigma \setminus \Sigma_{R,p_m})^+$, $c \in \Sigma_{R,p_m}$, and

$$\begin{aligned}
sv &= sw_1 a^{n_1} \left(\prod_{i=2}^{m} b_i w_i a^{n_i} \right) b_{m+1} w_{m+1} c w_{m+2} \\
&\vdash_A^{|w_1|} p_1 a^{n_1} \left(\prod_{i=2}^{m} b_i w_i a^{n_i} \right) b_{m+1} w_{m+1} c w_{m+2} \\
&\circlearrowleft_A q_2 w_2 a^{n_2} \left(\prod_{i=3}^{m} b_i w_i a^{n_i} \right) b_{m+1} w_{m+1} c w_{m+2} a^{n_1}
\end{aligned}$$

continued by

$$
\begin{aligned}
q_2 w_2 a^{n_2} \left(\prod_{i=3}^{m} b_i w_i a^{n_i} \right) & b_{m+1} w_{m+1} c w_{m+2} a^{n_1} \\
& \vdash_A^{|w_2|} p_2 a^{n_2} \left(\prod_{i=3}^{m} b_i w_i a^{n_i} \right) b_{m+1} w_{m+1} c w_{m+2} a^{n_1} \\
& \circlearrowleft_A q_3 w_3 a^{n_3} \left(\prod_{i=4}^{m} b_i w_i a^{n_i} \right) b_{m+1} w_{m+1} c w_{m+2} a^{n_1+n_2} \\
& \dots \\
& \vdash_A^{|w_{m-1}|} p_{m-1} a^{n_{m-1}} b_m w_m a^{n_m} b_{m+1} w_{m+1} c w_{m+2} a^{\sum_{i=1}^{m-2} n_i} \\
& \circlearrowleft_A q_m w_m a^{n_m} b_{m+1} w_{m+1} c w_{m+2} a^{\sum_{i=1}^{m-1} n_i} \\
& \vdash_A^{|w_m|} p_m a^{n_m} b_{m+1} w_{m+1} c w_{m+2} a^{\sum_{i=1}^{m-1} n_i} \\
& \circlearrowleft_A q_{m+1} w_{m+2} a^{\sum_{i=1}^{m} n_i} b_{m+1} w_{m+1} \circlearrowleft_A^* f.
\end{aligned}
$$

We get

$$
\begin{aligned}
s w_1 \left(\prod_{i=2}^{m} b_i w_i \right) & w_{m+1} c w_{m+2} a^{\sum_{i=1}^{m} n_i} b_{m+1} \\
& \vdash_A^{\left| w_1 \prod_{i=2}^{m} b_i w_i \right|} p_m w_{m+1} c w_{m+2} a^{\sum_{i=1}^{m} n_i} b_{m+1} \\
& \qquad \circlearrowleft_A q_{m+1} w_{m+2} a^{\sum_{i=1}^{m} n_i} b_{m+1} w_{m+1} \circlearrowleft_A^* f.
\end{aligned}
$$

This implies

$$
w_1 \left(\prod_{i=2}^{m} b_i w_i \right) w_{m+1} c w_{m+2} a^{\sum_{i=1}^{m} n_i} b_{m+1} \in L_R(A) = La,
$$

a contradiction. This shows $v \in L_R(B)$, which proves the theorem. □

Now, we consider the case of two languages over disjoint alphabets.

Theorem 39. *For disjoint alphabets Σ_1 and Σ_2, let languages $L_1 \subseteq \Sigma_1^*$ and $L_2 \subseteq \Sigma_2^*$ satisfy $L_1 \neq \emptyset \neq L_2 \neq \{\lambda\}$ such that $L_1 L_2$ is in* **ROWJ**. *Then, the language L_1 is regular and L_2 is in* **ROWJ**.

Proof. The proof is similar as those of Theorem 35. Let $A = (Q, \Sigma = \Sigma_1 \cup \Sigma_2, R, s, F)$ be a DFA with $L_R(A) = L_1 L_2$. For an $m \geq 0$ and $a_1, a_2, \dots, a_m \in \Sigma_1$, let $w = a_1 a_2 \dots a_m \in L_1$. We will show by induction that for each $0 \leq n \leq m$, there is a state $q_n \in Q$ with $sw \vdash^n q_n a_{n+1} a_{n+2} \dots a_m$. For $n = 0$, we just set $q_0 = s$. Assume that, for a fixed k with $0 \leq k < m$, we already know that there is a state $q_k \in Q$ with

$$
sw \vdash^k q_k a_{k+1} a_{k+2} \dots a_m.
$$

If the value $R(q_k, a_{k+1})$ is defined, then we have

$$
sw \vdash^{k+1} R(q_k, a_{k+1}) a_{k+2} a_{k+3} \dots a_m.
$$

Therefore, now let $R(q_k, a_{k+1})$ be undefined and let v be an arbitrary non-empty word out of L_2. We get

$$
swv \vdash^k q_k a_{k+1} a_{k+2} \dots a_m v.
$$

Because of $wv \in L_R(A)$, there exist a symbol $b \in \Sigma_{q_k}$, words $x \in (\Sigma \setminus \Sigma_{q_k})^*$ and $y \in \Sigma^*$, and a state $p \in F$ such that $a_{k+2}a_{k+3} \dots a_m v = xby$ and

$$q_k a_{k+1} xby \circlearrowright R(q_k, b) y a_{k+1} x \circlearrowright^* p.$$

This implies

$$s a_1 a_2 \dots a_k xby a_{k+1} \vdash^k q_k xby a_{k+1} \circlearrowright R(q_k, b) y a_{k+1} x \circlearrowright^* p,$$

which gives us that $a_1 a_2 \dots a_k xby a_{k+1} \in L_R(A)$. However, this word is equal to

$$a_1 a_2 \dots a_k a_{k+2} a_{k+3} \dots a_m v a_{k+1}$$

and we have

$$a_1 a_2 \dots a_k a_{k+2} a_{k+3} \dots a_m v a_{k+1} \in \Sigma_1^* \Sigma_2^+ \Sigma_1 \subseteq \Sigma^* \setminus (\Sigma_1^* \Sigma_2^*) \subseteq \Sigma^* \setminus (L_1 L_2),$$

which is a contradiction. So, the value $R(q_k, a_{k+1})$ has to be defined and we have shown by induction that for each $0 \leq n \leq m$, there is a state $q_n \in Q$ with $sw \vdash^n q_n a_{n+1} a_{n+2} \dots a_m$. We set $q_w = q_m$ and get $sw \vdash^{|w|} q_w$.

For every $w \in L_1$, we consider the DFA $B_w = (Q, \Sigma_2, R|_{Q \times \Sigma_2}, q_w, F)$. For every $v \in \Sigma_2^*$, we have

$$\begin{aligned} v \in L_R(B_w) &\Leftrightarrow (\exists f \in F : q_w v \circlearrowright_{B_w}^* f) \\ &\Leftrightarrow (\exists f \in F : swv \circlearrowright_A^* f) \\ &\Leftrightarrow wv \in L_R(A) = L_1 L_2 \\ &\Leftrightarrow v \in L_2. \end{aligned}$$

This shows $L_R(B_w) = L_2$, so L_2 is in **ROWJ**.

For every $v \in L_2$, we define the set $Q_v = \{\, q \in Q \mid \exists f \in F : qv \circlearrowright_A^* f \,\}$ and the determinstic finite stace device $C_v = (Q, \Sigma_1, R|_{Q \times \Sigma_1}, s, Q_v)$. For every $w \in \Sigma_1^*$, we have

$$\begin{aligned} w \in L_D(C_v) &\Leftrightarrow (\exists q \in Q_v : sw \vdash_{C_v}^* q) \\ &\Leftrightarrow (\exists q \in Q,\ f \in F : swv \vdash_A^* qv \circlearrowright_A^* f) \\ &\Leftrightarrow wv \in L_R(A) = L_1 L_2 \\ &\Leftrightarrow w \in L_1. \end{aligned}$$

Therefore, we conclude $L_D(C_v) = L_1$, so L_1 is regular, which proves the theorem. ◻

Adding prefix-freeness for L_1, we get an equivalence, by Theorem 39 and the closure of **ROWJ** under left-concatenation with prefix-free regular sets.

Corollary 40. *For disjoint alphabets Σ_1 and Σ_2, let $L_1 \subseteq \Sigma_1^*$ be a prefix-free set and $L_2 \subseteq \Sigma_2^*$ be an arbitrary language with $L_1 \neq \emptyset \neq L_2 \neq \{\lambda\}$. Then, the language $L_1 L_2$ is in* **ROWJ** *if and only if L_1 is regular and L_2 is in* **ROWJ**.

The previous corollary directly implies the following characterization that is another generalization of Corollary 4.

Corollary 41. *For disjoint alphabets Σ_1 and Σ_2, let $L_1 \subseteq \Sigma_1^*$ be a prefix-free set and $L_2 \subseteq \Sigma_2^*$ be a permutation closed language with $L_1 \neq \emptyset \neq L_2 \neq \{\lambda\}$. Then, the language $L_1 L_2$ is in* **ROWJ** *if and only if L_1 is regular and the Myhill-Nerode equivalence relation $\sim_{L_2}$ has only a finite number of positive equivalence classes.*

If a non-empty language and a non-empty permutation closed language over disjoint alphabets are separated by a symbol, we get the following result:

Corollary 42. *For disjoint alphabets Σ_1 and Σ_2, let $L_1 \subseteq \Sigma_1^*$ be non empty and $L_2 \subseteq \Sigma_2^*$ be a non-empty permutation closed language. Furthermore, let $a \in \Sigma_2$. Then, the language $L_1 a L_2$ is in* **ROWJ** *if and only if L_1 is regular and the Myhill-Nerode equivalence relation $\sim_{L_2}$ has only a finite number of positive equivalence classes.*

Proof. If L_1 is regular and the Myhill-Nerode equivalence relation $\sim_{L_2}$ has only a finite number of positive equivalence classes, Corollary 26 tells us that $L_1 a L_2$ is in **ROWJ**.

Now, assume that $L_1 a L_2$ is in **ROWJ**. From Theorem 39 we get that L_1 is regular and aL_2 is in **ROWJ**. Corollary 36 gives us that $\sim_{L_2}$ has only a finite number of positive equivalence classes. □

For an alphabet $\Sigma = \{a_1, a_2, \ldots, a_n\}$, the family of subsets of $a_1^* a_2^* \ldots a_n^*$ is kind of a counterpart of the family of permutation closed languages over Σ. In a permutation closed language L, for each word $w \in L$, all permutations of w are also in L. In a language $M \subseteq a_1^* a_2^* \ldots a_n^*$, for each word $w \in M$, no other permutation of w is in M. We can characterize the subsets of $a_1^* a_2^* \ldots a_n^*$ that are in **ROWJ**. The following lemma helps us to do so.

Lemma 43. *Let A be a DFA with input alphabet $\{a_1, a_2, \ldots, a_n\}$ accepting a letter bounded language, i.e., $L_R(A) \subseteq a_1^* a_2^* \ldots a_n^*$. Then, $L_R(A) = L_D(A)$.*

Proof. Because of (1) we have $L_D(A) \subseteq L_R(A)$. Now assume that $w \in L_R(A)$. Again, because of this inclusion chain, there is a permutation v ofthe word w with $v \in L_D(A) \subseteq L_R(A)$. Since $L_R(A) \subseteq a_1^* a_2^* \ldots a_n^*$, we conclude that $w = v \in L_D(A)$. Thus, $L_R(A) = L_D(A)$. □

In [3] it was shown that the language $\{\, a^n b^n \mid n \geq 0 \,\}$ is not in **ROWJ**. Our characterization generalizes this result:

Theorem 44. *Let $\{a_1, a_2, \ldots, a_n\}$ be an alphabet and $L \subseteq a_1^* a_2^* \ldots a_n^*$. Then, the language L is in* **ROWJ** *if and only if L is regular.*

Proof. If L is regular, then L is also in **ROWJ**, because of **REG** $\subset$ **ROWJ**. If L is in **ROWJ**, then there exists a DFA A with $L = L_R(A)$. Because of Lemma 43, we get $L = L_R(A) = L_D(A)$, which is a regular language. This proves the stated claim. □

References

1. S. Beier, M. Holzer, and M. Kutrib. Operational state complexity and decidability of jumping finite automata. In É. Charlier, J. Leroy, and M. Rigo, editors, *Proceedings of the 21st International Conference on Developments in Language Theory*, number 10396 in LNCS, pages 96–108, Liège, Belgium, 2017. Springer.
2. S. Bensch, H. Bordihn, M. Holzer, and M. Kutrib. On input-revolving deterministic and nondeterministic finite automata. *Inform. Comput.*, 207(11):1140–1155, 2009.
3. H. Chigahara, S. Fazekas, and A. Yamamura. One-way jumping finite automata. *Internat. J. Found. Comput. Sci.*, 27(3):391–405, 2016.
4. C. Culy. Formal properties of natural language and linguistic theories. *Linguistics and Philosophy*, 19:599–617, 1996.
5. L. E. Dickson. Finiteness of the odd perfect and primitive abundant numbers with n distinct prime factors. *American J. Math.*, 35(4):413–422, 1913.
6. H. Fernau, M. Paramasivan, and M. L. Schmid. Jumping finite automata: Characterizations and compliexity. In F. Drewes, editor, *Proceedings of the 20th Conference on Implementation and Application of Automata*, number 9223 in LNCS, pages 89–101, Umeå, Sweden, 2015. Springer.

7. H. Fernau, M. Paramasivan, M. L. Schmid, and V. Vorel. Characterization and complexity results on jumping finite automata. http://arxiv.org/abs/1512.00482, 2015.
8. S. Ginsburg and E. H. Spanier. Bounded ALGOL-like languages. *Trans. AMS*, 113:333–368, 1964.
9. S. A. Greibach. An infinite hierarchy of context-free languages. *J. ACM*, 16(1):91–106, 1969.
10. J. E. Hopcroft and J. D. Ullman. *Introduction to Automata Theory, Languages and Computation.* Addison-Wesley, 1979.
11. A. Meduna and P. Zemek. Jumping finite automata. *Internat. J. Found. Comput. Sci.*, 23(7):1555–1578, 2012.
12. A. Meduna and P. Zemek. *Regulated Grammars and Automata*, chapter 17: Jumping Finite Automata, pages 567–585. Springer, 2014.
13. R. J. Parikh. On context-free languages. *J. ACM*, 13(4):570–581, 1966.
14. V. Vorel. Basic properties of jumping finite automata. http://arxiv.org/abs/1511.08396v2, 2015.

Institut für Informatik
Justus-Liebig-Universität Giessen
Arndtstr. 2, 35392 Giessen, Germany

Recent Reports

(Further reports are available at `www.informatik.uni-giessen.de`.)

B. Truthe, *Hierarchy of Subregular Language Families*, Report 1801, February 2018.

M. Holzer, M. Hospodár, *On the Magic Number Problem of the Cut Operation*, Report 1703, October 2017.

M. Holzer, S. Jakobi, *A Note on the Computational Complexity of Some Problems for Self-Verifying Finite Automata*, Report 1702, April 2017.

S. Beier, M. Holzer, M. Kutrib, *On the Descriptional Complexity of Operations on Semilinear Sets*, Report 1701, April 2017.

M. Holzer, S. Jakobi, M. Wendlandt, *On the Computational Complexity of Partial Word Automata Problems*, Report 1404, May 2014.

H. Gruber, M. Holzer, *Regular Expressions From Deterministic Finite Automata, Revisited*, Report 1403, May 2014.

M. Kutrib, A. Malcher, M. Wendlandt, *Deterministic Set Automata*, Report 1402, April 2014.

M. Holzer, S. Jakobi, *Minimal and Hyper-Minimal Biautomata*, Report 1401, March 2014.

J. Kari, M. Kutrib, A. Malcher (Eds.), *19th International Workshop on Cellular Automata and Discrete Complex Systems AUTOMATA 2013 Exploratory Papers*, Report 1302, September 2013.

M. Holzer, S. Jakobi, *Minimization, Characterizations, and Nondeterminism for Biautomata*, Report 1301, April 2013.

A. Malcher, K. Meckel, C. Mereghetti, B. Palano, *Descriptional Complexity of Pushdown Store Languages*, Report 1203, May 2012.

M. Holzer, S. Jakobi, *On the Complexity of Rolling Block and Alice Mazes*, Report 1202, March 2012.

M. Holzer, S. Jakobi, *Grid Graphs with Diagonal Edges and the Complexity of Xmas Mazes*, Report 1201, January 2012.

H. Gruber, S. Gulan, *Simplifying Regular Expressions: A Quantitative Perspective*, Report 0904, August 2009.

M. Kutrib, A. Malcher, *Cellular Automata with Sparse Communication*, Report 0903, May 2009.

M. Holzer, A. Maletti, *An $n \log n$ Algorithm for Hyper-Minimizing States in a (Minimized) Deterministic Automaton*, Report 0902, April 2009.

H. Gruber, M. Holzer, *Tight Bounds on the Descriptional Complexity of Regular Expressions*, Report 0901, February 2009.

M. Holzer, M. Kutrib, and A. Malcher (Eds.), *18. Theorietag Automaten und Formale Sprachen*, Report 0801, September 2008.

M. Holzer, M. Kutrib, *Flip-Pushdown Automata: Nondeterminism is Better than Determinism*, Report 0301, February 2003

M. Holzer, M. Kutrib, *Flip-Pushdown Automata: $k + 1$ Pushdown Reversals are Better Than k*, Report 0206, November 2002

M. Holzer, M. Kutrib, *Nondeterministic Descriptional Complexity of Regular Languages*, Report 0205, September 2002

H. Bordihn, M. Holzer, M. Kutrib, *Economy of Description for Basic Constructions on Rational Transductions*, Report 0204, July 2002

M. Kutrib, J.-T. Löwe, *String Transformation for n-dimensional Image Compression*, Report 0203, May 2002

A. Klein, M. Kutrib, *Grammars with Scattered Nonterminals*, Report 0202, February 2002

A. Klein, M. Kutrib, *Self-Assembling Finite Automata*, Report 0201, January 2002

M. Holzer, M. Kutrib, *Unary Language Operations and its Nondeterministic State Complexity*, Report 0107, November 2001

A. Klein, M. Kutrib, *Fast One-Way Cellular Automata*, Report 0106, September 2001

M. Holzer, M. Kutrib, *Improving Raster Image Run-Length Encoding Using Data Order*, Report 0105, July 2001

M. Kutrib, *Refining Nondeterminism Below Linear-Time*, Report 0104, June 2001

M. Holzer, M. Kutrib, *State Complexity of Basic Operations on Nondeterministic Finite Automata*, Report 0103, April 2001