Yet Another Canonical Nondeterministic Automaton

Hendrik Maarand(✉) and Hellis Tamm

Department of Software Science, Tallinn University of Technology, Tallinn, Estonia
{hendrik,hellis}@cs.ioc.ee

Abstract. Several canonical forms of finite automata have been introduced over the decades. In particular, if one considers the minimal deterministic finite automaton (DFA), the canonical residual finite state automaton (RFSA), and the átomaton of a language, then the átomaton can be seen as the dual automaton of the minimal DFA, but no such dual has been presented for the canonical RFSA so far. We fill this gap by introducing a new canonical automaton that we call the maximized prime átomaton, and study its properties. We also describe how these four automata can be extracted from suitable observation tables used in the automata learning context.

Keywords: Canonical automaton · regular language · atoms of regular languages · automata learning

1 Introduction

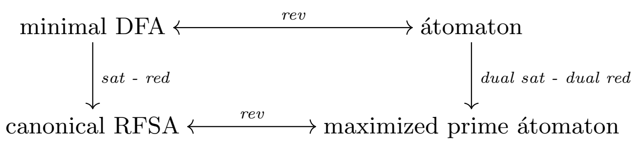
It is well known that every regular language has a unique minimal deterministic finite automaton (DFA) accepting the language. However, this nice property does not hold for the class of nondeterministic finite automata (NFAs), because a language may have several non-isomorphic NFAs with a minimum number of states. Nevertheless, several canonical forms of NFAs have been introduced over the decades: the *universal automaton* [9], the *canonical residual finite state automaton* (*canonical RFSA*) [6] (also known as *jiromaton* [10]), the *átomaton* [5], and the *maximized átomaton* [12] (same as *distromaton* [10]). We note that none of these NFAs are necessarily minimal NFAs.

While the states of the minimal DFA of a language L correspond to the *(left) quotients* of L, the canonical RFSA of L may have less states, since it is based on the *prime quotients* [6] of L, that is, non-empty quotients that are not unions of other quotients. The states of the átomaton of L correspond to the *atoms* [5] of L, which are non-empty intersections of complemented and uncomplemented quotients. Also, the notion of a *prime atom* was defined in [14], however, no automaton based on prime atoms has been presented so far.

This work was supported by the Estonian Research Council grant PRG1210. H. Maarand was also supported by the ERDF funded Estonian CoE project EXCITE (project 2014-2020.4.01.15-0018).

Published by Springer Nature Switzerland AG 2022
Y.-S. Han and G. Vaszil (Eds.): DCFS 2022, LNCS 13439, pp. 184–196, 2022.
https://doi.org/10.1007/978-3-031-13257-5_14

We fill this gap by introducing a new canonical NFA that we call the *maximized prime átomaton*, because it is a subautomaton of the maximized átomaton and its states correspond to the prime atoms of a language. While the átomaton of L is isomorphic to the reverse NFA of the minimal DFA of L^R [5], we show that the maximized prime átomaton of L is the reverse of the canonical RFSA of L^R. An informal description of the relationship between these automata is presented in the picture below. By applying *saturation* and *reduction* operations [6] to the minimal DFA, the canonical RFSA is obtained. By applying corresponding dual operations to the átomaton, we get the maximized prime átomaton.



Another way to construct a canonical RFSA is by using a modified subset construction operation C [6,12]. We define a dual operation of C and show how to use this operation to obtain the maximized prime átomaton.

We also describe how the four automata in the above picture can be extracted from suitable observation tables used in the automata learning context [1]. If an observation table is closed and consistent both for rows and columns (Definition 7), then its proper part forms the quotient-atom matrix [8,13] of the language. We believe that it can be helpful to think of these automata in terms of such matrices where the row and column indices are the right and left congruence classes of the language, respectively.

2 Automata, Quotients, and Atoms of Regular Languages

A *nondeterministic finite automaton (NFA)* is a quintuple $\mathcal{N} = (Q, \Sigma, \delta, I, F)$, where Q is a finite, non-empty set of *states*, Σ is a finite non-empty *alphabet*, $\delta : Q \times \Sigma \to 2^Q$ is the *transition function*, $I \subseteq Q$ is the set of *initial states*, and $F \subseteq Q$ is the set of *final states*. We extend the transition function to functions $\delta' : Q \times \Sigma^* \to 2^Q$ and $\delta'' : 2^Q \times \Sigma^* \to 2^Q$, using δ for all these functions. The *left language* of a state q of $\mathcal{N}$ is $L_{I,q}(\mathcal{N}) = \{w \in \Sigma^* \mid q \in \delta(I, w)\}$, and the *right language* of q is $L_{q,F}(\mathcal{N}) = \{w \in \Sigma^* \mid \delta(q, w) \cap F \neq \emptyset\}$. A state q of $\mathcal{N}$ is *reachable* if $L_{I,q}(\mathcal{N}) \neq \emptyset$, and it is *empty* if $L_{q,F}(\mathcal{N}) = \emptyset$. The *language accepted* by an NFA $\mathcal{N}$ is $L(\mathcal{N}) = \{w \in \Sigma^* \mid \delta(I, w) \cap F \neq \emptyset\}$. Two NFAs are *equivalent* if they accept the same language. An NFA is *minimal* if it has a minimum number of states among all equivalent NFAs. The *reverse* of an NFA $\mathcal{N} = (Q, \Sigma, \delta, I, F)$ is the NFA $\mathcal{N}^R = (Q, \Sigma, \delta^R, F, I)$, where $q \in \delta^R(p, a)$ if and only if $p \in \delta(q, a)$ for $p, q \in Q$ and $a \in \Sigma$.

A *deterministic finite automaton (DFA)* is a quintuple $\mathcal{D} = (Q, \Sigma, \delta, q_0, F)$, where Q, Σ, and F are as in an NFA, $\delta : Q \times \Sigma \to Q$ is the transition function, and q_0 is the initial state. The *left quotient*, or simply *quotient*, of a language L by a word $w \in \Sigma^*$ is the language $w^{-1}L = \{x \in \Sigma^* \mid wx \in L\}$. It is well known that the left quotients of L are the right languages of the states of the minimal DFA of L. Any NFA $\mathcal{N}$ can be *determinized* by the well-known subset construction, yielding a DFA $\mathcal{N}^D$ that has only reachable states.

Let L be a non-empty regular language with quotients $K_0, \dots, K_{n-1}$. An *atom* of L is any non-empty language of the form $\widetilde{K_0} \cap \cdots \cap \widetilde{K_{n-1}}$, where $\widetilde{K_i}$ is either K_i or $\overline{K_i}$, and $\overline{K_i}$ is the complement of K_i with respect to Σ^* [5]. An atom is *initial* if it has L (rather than $\overline{L}$) as a term; it is *final* if it contains ε. There is exactly one final atom, the atom $\widehat{K_0} \cap \cdots \cap \widehat{K_{n-1}}$, where $\widehat{K_i} = K_i$ if $\varepsilon \in K_i$, and $\widehat{K_i} = \overline{K_i}$ otherwise. If $\overline{K_0} \cap \cdots \cap \overline{K_{n-1}}$ is an atom, then it is called the *negative* atom, all the other atoms are *positive*. Thus atoms of L are pairwise disjoint languages uniquely determined by L; they define a partition of Σ^*. Every quotient K_i (including L) is a (possibly empty) union of atoms. An NFA $\mathcal{N}$ is *atomic* if the right languages of its states are unions of atoms of $L(\mathcal{N})$.

It is well known that quotients of L are in a one-one correspondence with the equivalence classes of the *Nerode right congruence* $\equiv_L$ of L [11] defined as follows: for $x, y \in \Sigma^*$, $x \equiv_L y$ if for every $v \in \Sigma^*$, $xv \in L$ if and only if $yv \in L$. Atoms of L are the classes of the *left congruence* ${}_L\!\equiv$ of L: for $x, y \in \Sigma^*$, $x \;{}_L\!\equiv y$ if for every $u \in \Sigma^*$, $ux \in L$ if and only if $uy \in L$ [7].

Let $A = \{A_0, A_1, \dots, A_{m-1}\}$ be the set of atoms of L, let A_I be the set of initial atoms, and let A_{m-1} be the final atom.

The *átomaton* of L is the NFA $\mathcal{A} = (S_A, \Sigma, \alpha, I_A, \{s_{m-1}\})$ where $S_A = \{s_0, s_1, \dots, s_{m-1}\}$, $I_A = \{s_i \in S_A \mid A_i \in A_I\}$, and $s_j \in \alpha(s_i, a)$ if and only if $A_j \subseteq a^{-1}A_i$, for all $i, j \in \{0, \dots, m-1\}$ and $a \in \Sigma$. It was shown in [5] that the atoms of L are the right languages of the states of the átomaton, and that the reverse NFA of the átomaton is the minimal DFA of the reverse language L^R.

The next theorem is a slightly modified version of the result by Brzozowski [4]:

Theorem 1. *If an NFA $\mathcal{N}$ has no empty states and $\mathcal{N}^R$ is deterministic, then $\mathcal{N}^D$ is minimal.*

By Theorem 1, for any NFA $\mathcal{N}$, the DFA $\mathcal{N}^{RDRD}$ is the minimal DFA equivalent to $\mathcal{N}$. This result is known as Brzozowski's double-reversal method for DFA minimization. In [5], a generalization of Theorem 1 was presented, providing a characterization of the class of NFAs for which applying determinization procedure produces a minimal DFA:

Theorem 2. *For any NFA $\mathcal{N}$, the DFA $\mathcal{N}^D$ is minimal if and only if $\mathcal{N}^R$ is atomic.*

3 Residual Finite State Automata

Residual finite state automata (RFSAs) were introduced by Denis, Lemay, and Terlutte in [6]. In this section, we state some basic properties of RFSAs. However, we note here that we usually prefer to use the term "quotient" over "residual".

An NFA $\mathcal{N} = (Q, \Sigma, \delta, I, F)$ is a *residual finite state automaton* (RFSA) if for every state $q \in Q$, $L_{q,F}(\mathcal{N})$ is a quotient of $L(\mathcal{N})$. Clearly, any DFA having only reachable states, is an RFSA.

Let L be a regular language over Σ. A non-empty quotient of L is *prime* if it is not a union of other quotients. Let $K' = \{K_0, \ldots, K_{n'-1}\}$ be the set of prime quotients of L.

The *canonical RFSA* of L is the NFA $\mathcal{R} = (Q_{K'}, \Sigma, \delta, I_{K'}, F_{K'})$, where $Q_{K'} = \{q_0, \ldots, q_{n'-1}\}$, $I_{K'} = \{q_i \in Q_{K'} \mid K_i \subseteq L\}$, $F_{K'} = \{q_i \in Q_{K'} \mid \varepsilon \in K_i\}$, and $\delta(q_i, a) = \{q_j \in Q_{K'} \mid K_j \subseteq a^{-1}K_i\}$ for every $q_i \in Q_{K'}$ and $a \in \Sigma$.

The canonical RFSA is a state-minimal RFSA with a maximal number of transitions. One way to build a canonical RFSA is to use the *saturation* and *reduction* operations defined in the following.

Let $\mathcal{N} = (Q, \Sigma, \delta, I, F)$ be an NFA. The *saturation* operation S, applied to $\mathcal{N}$, produces the NFA $\mathcal{N}^S = (Q, \Sigma, \delta_S, I_S, F)$, where $\delta_S(q, a) = \{q' \in Q \mid aL_{q',F}(\mathcal{N}) \subseteq L_{q,F}(\mathcal{N})\}$ for all $q \in Q$ and $a \in \Sigma$, and $I_S = \{q \in Q \mid L_{q,F}(\mathcal{N}) \subseteq L(\mathcal{N})\}$. An NFA $\mathcal{N}$ is saturated if $\mathcal{N}^S = \mathcal{N}$. Saturation may add transitions and initial states to an NFA, without changing its language. If $\mathcal{N}$ is an RFSA, then $\mathcal{N}^S$ is an RFSA.

For any state q of $\mathcal{N}$, let $R(q)$ be the set $\{q' \in Q\backslash\{q\} \mid L_{q',F}(\mathcal{N}) \subseteq L_{q,F}(\mathcal{N})\}$. A state q is *erasable* if $L_{q,F}(\mathcal{N}) = \bigcup_{q' \in R(q)} L_{q',F}(\mathcal{N})$. If q is erasable, a *reduction* operator ϕ is defined as follows: $\phi(\mathcal{N}, q) = (Q', \Sigma, \delta', I', F')$ where $Q' = Q\backslash\{q\}$, $I' = I$ if $q \notin I$, and $I' = (I\backslash\{q\}) \cup R(q)$ otherwise, $F' = F \cap Q'$, $\delta'(q', a) = \delta(q', a)$ if $q \notin \delta(q', a)$, and $\delta'(q', a) = (\delta(q', a)\backslash\{q\}) \cup R(q)$ otherwise, for every $q' \in Q'$ and every $a \in \Sigma$. If q is not erasable, let $\phi(\mathcal{N}, q) = \mathcal{N}$.

If $\mathcal{N}$ is saturated and if q is an erasable state of $\mathcal{N}$, then $\phi(\mathcal{N}, q)$ is obtained by deleting q and its associated transitions from $\mathcal{N}$. An NFA $\mathcal{N}$ is *reduced* if there is no erasable state in $\mathcal{N}$. Applying ϕ to $\mathcal{N}$ does not change its language. If $\mathcal{N}$ is an RFSA, then $\phi(\mathcal{N}, q)$ is an RFSA. The following proposition is from [6]:

Proposition 1. *If an NFA $\mathcal{N}$ is a reduced saturated RFSA of L, then $\mathcal{N}$ is the canonical RFSA for L.*

The canonical RFSA can be obtained from a DFA having only reachable states, by using saturation and reduction operations.

Next we will discuss another method to compute the canonical RFSA, suggested in [6]. In Sect. 2, we recalled the result that for any NFA $\mathcal{N}$, the DFA $\mathcal{N}^{RDRD}$ is the minimal DFA equivalent to $\mathcal{N}$. In [6], a similar double-reversal method is proposed to obtain a canonical RFSA from a given NFA, using a modified subset construction operation C to be applied to an NFA as follows:

Definition 1. *Let $\mathcal{N} = (Q, \Sigma, \delta, I, F)$ be an NFA. Let Q_D be the set of states of the determinized version $\mathcal{N}^D$ of $\mathcal{N}$. A state $s \in Q_D$ is* coverable *if there is a set $Q_s \subseteq Q_D\backslash\{s\}$ such that $s = \bigcup_{s' \in Q_s} s'$. The NFA $\mathcal{N}^C$ is defined as $(Q_C, \Sigma, \delta_C, I_C, F_C)$, where $Q_C = \{s \in Q_D \mid s \text{ is not coverable }\}$, $I_C = \{s \in Q_C \mid s \subseteq I\}$, $F_C = \{s \in Q_C \mid s \cap F \neq \emptyset\}$, and $\delta_C(s, a) = \{s' \in Q_C \mid s' \subseteq \delta(s, a)\}$ for any $s \in Q_C$ and $a \in \Sigma$.*

Applying the operation C to any NFA $\mathcal{N}$ produces an RFSA $\mathcal{N}^C$. Denis et al. [6] have the following result:

Theorem 3. *If an NFA $\mathcal{N}$ has no empty states and $\mathcal{N}^R$ is an RFSA, then $\mathcal{N}^C$ is the canonical RFSA.*

By Theorem 3, for any NFA $\mathcal{N}$, the RFSA $\mathcal{N}^{RCRC}$ is the canonical RFSA equivalent to $\mathcal{N}$. Hence, the operation C has a similar role for RFSAs as determinization has for DFAs.

In Sect. 2, we recalled Theorem 2 from [5], a generalization of Theorem 1, characterizing the class of NFAs to which applying the determinization procedure produces a minimal DFA. Theorem 3 was generalized in [12] in a similar way:

Theorem 4. *For any NFA $\mathcal{N}$ of L, $\mathcal{N}^C$ is a canonical RFSA if and only if the left languages of $\mathcal{N}$ are unions of left languages of the canonical RFSA of L.*

4 Maximized Átomaton

Let L be a non-empty regular language, $K = \{K_0, \ldots, K_{n-1}\}$ be the set of quotients, and $A = \{A_0, \ldots, A_{m-1}\}$ be the set of atoms of L, with the set of initial atoms $A_I \subseteq A$, and the final atom A_{m-1}.

In [12], the notions of a *maximized atom* and the *maximized átomaton* of a regular language L were introduced. For every atom A_i of L, the corresponding *maximized atom* M_i is the union of all the atoms which occur in every quotient containing A_i:

Definition 2. *The* maximized atom M_i *of an atom* A_i *is the union of atoms* $M_i = \bigcup\{A_h \mid A_h \subseteq \bigcap_{A_i \subseteq K_k} K_k\}$.

Clearly, since atoms are pairwise disjoint, and every quotient is a union of atoms, $M_i = \bigcap_{A_i \subseteq K_k} K_k$. In [12], the following properties of maximized atoms were shown:

Proposition 2. *Let A_i and A_j be some atoms of L. The following properties hold:*

1. $A_i \subseteq M_i$.
2. *If* $A_i \neq A_j$, *then* $M_i \neq M_j$.
3. $A_i \subseteq M_j$ *if and only if* $M_i \subseteq M_j$.
4. $A_j \subseteq a^{-1}M_i$ *if and only if* $M_j \subseteq a^{-1}M_i$.

Let $M = \{M_0, \ldots, M_{m-1}\}$ be the set of the maximized atoms of L. The *maximized átomaton* was defined in [12] as follows:

Definition 3. *The* maximized átomaton *of L is the NFA defined as* $\mathcal{M} = (Q_M, \Sigma, \mu, I_M, F_M)$, *where* $Q_M = \{q_0, q_1, \ldots, q_{m-1}\}$, $I_M = \{q_i \in Q_M \mid A_i \in A_I\}$, $F_M = \{q_i \in Q_M \mid A_{m-1} \subseteq M_i\}$, *and* $q_j \in \mu(q_i, a)$ *if and only if* $M_j \subseteq a^{-1}M_i$, *for all* $i, j \in \{0, \ldots, m-1\}$ *and* $a \in \Sigma$.

It was shown in [12] that the maximized átomaton $\mathcal{M}$ of L is isomorphic to the reverse NFA of the saturated version of the minimal DFA of L^R.

Using results from [13] and Proposition 2, we can see that the right language of any state of the maximized átomaton is the corresponding maximized atom:

Proposition 3. *For every state $q_i \in Q_M$ of the maximized átomaton $\mathcal{M} = (Q_M, \Sigma, \mu, I_M, F_M)$ of L, the equality $L_{q_i, F_M}(\mathcal{M}) = M_i$ holds.*

5 Maximized Prime Átomaton

We recall that a non-empty quotient is *prime* if it is not a union of other quotients.

The notion of a *prime atom* was defined in [14] as follows: any positive atom $A_i = \bigcap_{j \in S_i} K_j \cap \bigcap_{j \in \overline{S_i}} \overline{K_j}$, where $S_i \subseteq \{0, \dots, n-1\}$ and $\overline{S_i} = \{0, \dots, n-1\} \setminus S_i$, is *prime* if the set $\{K_j \mid j \in S_i\}$ of uncomplemented quotients in the intersection of A_i is not a union of such sets of quotients corresponding to other atoms.

By results in [5], it is known that the reverse of the átomaton $\mathcal{A}$ of L is the minimal DFA of L^R. Since the right language of any state of $\mathcal{A}$ is some atom of L, and the right language of any state of $\mathcal{A}^R$ is some quotient of L^R, there is a natural one-one-correspondence between the set of atoms of L and the set of quotients of L^R, based on the state set of $\mathcal{A}$ (and $\mathcal{A}^R$). Also, there is a one-one correspondence between the set of prime atoms of L and the set of prime quotients of L^R:

Proposition 4. *The right language of any state of the átomaton $\mathcal{A}$ of L is a prime atom of L if and only if the right language of the same state of $\mathcal{A}^R$ is a prime quotient of L^R.*

Now, let $A' \subseteq A$ be the set of prime atoms of L, and let $M' \subseteq M$ be the corresponding set of maximized prime atoms. We define the *maximized prime átomaton* of L as follows:

Definition 4. *The* maximized prime átomaton *of L is the NFA defined by $\mathcal{M}' = (Q_{M'}, \Sigma, \mu, I_{M'}, F_{M'})$, where $Q_{M'} = \{q_i \mid M_i \text{ is prime}\}$, $I_{M'} = Q_{M'} \cap I_M$, $F_{M'} = Q_{M'} \cap F_M$, and $q_j \in \mu(q_i, a)$ if and only if $M_j \subseteq a^{-1}M_i$, for $q_i, q_j \in Q_{M'}$ and $a \in \Sigma$.*

In [12], it was shown that the maximized átomaton $\mathcal{M}$ of L is isomorphic to $\mathcal{E}^{SR}$, where $\mathcal{E}$ is the minimal DFA of L^R. That is, $\mathcal{M}^R$ is isomorphic to $\mathcal{E}^S$. Now, the canonical RFSA of L^R is the reduced version of $\mathcal{E}^S$, where those states of $\mathcal{E}^S$ corresponding to non-prime quotients of L^R, have been removed. Since by Proposition 4, the states of $\mathcal{E}$ corresponding to prime quotients of L^R are exactly those states of $\mathcal{E}^R$ corresponding to prime atoms of L, the canonical RFSA of L^R is isomorphic to the subautomaton of $\mathcal{M}^R$, where the states corresponding to non-prime atoms, together with their in- and out-transitions, have been removed. We have the following result:

Proposition 5. *The maximized prime átomaton $\mathcal{M}'$ of L is isomorphic to the reverse NFA of the canonical RFSA of L^R.*

There is a one-one correspondence between the set M' of maximized prime atoms and the state set $Q_{M'}$ of the maximized prime átomaton of L. However, the right language of a state q_i of $\mathcal{M}'$ is not necessarily equal to the corresponding maximized prime atom M_i. By a result in [12], for the left language L_i of a state q_i of the canonical RFSA of L^R, the inclusions $A_i^R \subseteq L_i \subseteq M_i^R$ hold, where A_i and M_i are respectively the corresponding atom and the maximized atom of L. Since the right language of any state of the maximized prime átomaton of L is the reverse of the left language of the corresponding state of the canonical RFSA of L^R, we can state the following:

Proposition 6. *For any state q_i of the maximized prime átomaton $\mathcal{M}'$ of L, the inclusions $A_i \subseteq L_{q_i,F_{M'}}(\mathcal{M}') \subseteq M_i$ hold.*

By Proposition 5, we are able to obtain the maximized prime átomaton of L by finding the canonical RFSA of L^R, and then reversing it. Since by Theorem 3, for any NFA $\mathcal{N}$, the RFSA $\mathcal{N}^{RCRC}$ is the canonical RFSA equivalent to $\mathcal{N}$, it is clear that $\mathcal{N}^{CRCR}$ is the maximized prime átomaton of L.

We define an operation *coC* to be applied to an NFA as follows:

Definition 5. *Let $\mathcal{N} = (Q, \Sigma, \delta, I, F)$ be an NFA. Let Q_{coD} be the set of states of the determinized version $\mathcal{N}^{RD}$ of $\mathcal{N}^R$. A state $s \in Q_{coD}$ is* coverable *if there is a set $Q_s \subseteq Q_{coD} \setminus \{s\}$ such that $s = \bigcup_{s' \in Q_s} s'$. The NFA $\mathcal{N}^{coC} = (Q_{coC}, \Sigma, \delta_{coC}, I_{coC}, F_{coC})$ is defined as follows: $Q_{coC} = \{s \in Q_{coD} \mid s$ is not coverable$\}$, $I_{coC} = \{s \in Q_{coC} \mid s \cap I \neq \emptyset\}$, $F_{coC} = \{s \in Q_{coC} \mid s \subseteq F\}$, and for any $s, s' \in Q_C$ and $a \in \Sigma$, $s' \in \delta_{coC}(s, a)$ if and only if for every $q \in s$ there is some $q' \in s'$ such that $q' \in \delta(q, a)$.*

Clearly, $\mathcal{N}^{coC}$ is isomorphic to $\mathcal{N}^{RCR}$. Hence, given any NFA $\mathcal{N}$ of L, the maximized prime átomaton of L can be obtained by applying first the operation C to $\mathcal{N}$, yielding $\mathcal{N}^C$, and then applying *coC* to $\mathcal{N}^C$, resulting in the automaton $\mathcal{N}^{C(coC)}$. Also, the NFA $\mathcal{N}^{(coC)C}$ is the canonical RFSA of L. The following theorem holds:

Theorem 5. *For any NFA $\mathcal{N}$ of L, the NFA $\mathcal{N}^{coC}$ is the maximized prime átomaton of L if and only if the right language of every state of $\mathcal{N}$ is a union of right languages of the maximized prime átomaton of L.*

Example 1. We consider a modification of an example from [6], and define a family of NFAs $\mathcal{B}_n = (Q, \Sigma, \delta, I, F)$, $n \geqslant 1$, where $Q = \{q_0, \dots, q_{n-1}\}$, $\Sigma = \{a, b\}$, $I = \{q_i \mid 0 \leqslant i < n/2\}$, $F = \{q_0\}$, and $\delta(q_i, a) = \{q_{(i+1) \bmod n}\}$ for $i = 0, \dots, n-1$, and $\delta(q_0, b) = \{q_0, q_1\}$, $\delta(q_1, b) = \{q_{n-1}\}$, and $\delta(q_i, b) = \{q_{i-1}\}$ for $1 < i < n$. The NFA $\mathcal{B}_4$ is shown in Fig. 1 and its reverse $\mathcal{B}_4^R$ is in Fig. 2.

We claim that the NFA $\mathcal{B}_n^R$ is a canonical RFSA of $L(\mathcal{B}_n)^R$. Indeed, $\mathcal{B}_n^R$ is an RFSA, because the right languages of $\mathcal{B}_n^R$ are quotients of $L(\mathcal{B}_n^R)$: $L_{q_0,F}(\mathcal{B}_n^R) = \varepsilon^{-1}L(\mathcal{B}_n^R)$ and $L_{q_i,F}(\mathcal{B}_n^R) = (a^{n-i})^{-1}L(\mathcal{B}_n^R)$, for $i = 1, \dots, n-1$. Denoting $K_i = (a^{(n-i) \bmod n})^{-1}L(\mathcal{B}_n^R)$ and noticing that $a^{(i-\lceil n/2 \rceil+1) \bmod n}, \dots, a^{i \bmod n} \in K_i$, and $a^{(i+1) \bmod n}, \dots, a^{(i+\lfloor n/2 \rfloor) \bmod n} \notin K_i$, for $i = 0, \dots, n-1$, it is easy to see that K_i's are pairwise incomparable. Therefore, $\mathcal{B}_n^{RC}$ is isomorphic to $\mathcal{B}_n^R$, and it is clear that $\mathcal{B}_n^R$ is a canonical RFSA of $L(\mathcal{B}_n)^R$.

Hence, by Proposition 5, $\mathcal{B}_n$ is the maximized prime átomaton of $L(\mathcal{B}_n)$. Also, by Theorem 3, $\mathcal{B}_n^C$ is the canonical RFSA of $L(\mathcal{B}_n)$. The automaton $\mathcal{B}_n^C$ has $\binom{n}{\lceil n/2 \rceil}$ states, because any candidate state of $\mathcal{B}_n^C$ with more than $\lceil n/2 \rceil$ elements can be covered by those with exactly $\lceil n/2 \rceil$ elements. Thus, for $n \geqslant 4$, $\mathcal{B}_n$ is smaller than the canonical RFSA for $L(\mathcal{B}_n)$, and the difference between the sizes of these two NFAs grows with n. Moreover, $\mathcal{B}_n$ is a minimal NFA for $L(\mathcal{B}_n)$, as can be seen by the *fooling set method* [2] using the fooling set $\{(\varepsilon, a^n), (a, a^{n-1}), \ldots, (a^{n-1}, a)\}$ of size n.

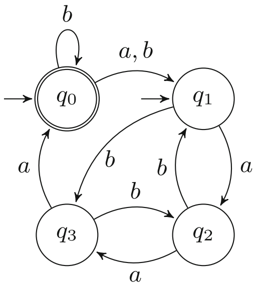


Fig. 1. The automaton $\mathcal{B}_4$.

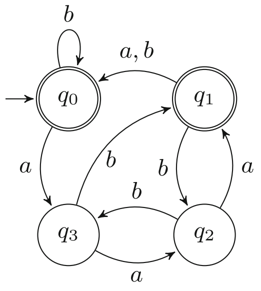


Fig. 2. The automaton $\mathcal{B}_4^R$.

6 Observation Tables

We now turn to observation tables known from the L^* learning algorithm [1] and how to read out various canonical automata from suitable observation tables. These tables can be seen as submatrices of the quotient-atom matrix [13] of a language, which is used, for example, in finding a minimal NFA of the language [8,13].

The L^* algorithm works by performing *membership* (whether a word belongs to the unknown language) and *equivalence* (whether a hypothesis is equivalent to the unknown language) queries. Informally, an observation table is used in the L^* algorithm to collect the observations that have been made so far and also to organize the observations in such a manner that it can be determined which observations need to be performed next. The membership queries are always performed for words composed from a prefix s and a suffix e. If the result of the membership query for the word se is positive, then the entry in the table at row s and column e is set to 1, otherwise it is set to 0.

Definition 6. *An* observation table *is a triple* $\mathcal{T} = (S, E, T)$ *where* $S \subseteq \Sigma^*$ *is a prefix-closed set of words,* $E \subseteq \Sigma^*$ *is a suffix-closed set of words and* $T : \Sigma^* \to 2$ *is a finite function. The* proper *part of the table consists of* S *rows and*

E columns. The row extensions *of the table consist of the rows* $S \cdot \Sigma \backslash S$*. The* column extensions *of the table consist of the columns* $\Sigma \cdot E \backslash E$*. The entry in the table at row* s *and column* e *is* $T(se)$*.*

A *row* of $\mathcal{T} = (S, E, T)$ is an E-indexed vector consisting of the corresponding entries of the table. That is, for $s \in S$ and $e \in E$, $row(s)(e) = T(se)$. A *column* of $\mathcal{T}$ is an S-indexed vector. That is, for $e \in E$ and $s \in S$, $col(e)(s) = T(se)$. Note that $row(sa)(e) = row(s)(ae) = col(ae)(s) = col(e)(sa)$.

Definition 7. *An observation table* $\mathcal{T} = (S, E, T)$ *is called*

- row-closed *when, for every* $s \in (S \cdot \Sigma)\backslash S$*, there exists* $s' \in S$ *such that* $row(s) = row(s')$*;*
- column-closed *when, for every* $e \in (\Sigma \cdot E)\backslash E$*, there exists* $e' \in E$ *such that* $col(e) = col(e')$*;*
- row-consistent *when, for every* $s, s' \in S$*, if* $row(s) = row(s')$*, then, for every* $a \in \Sigma$*,* $row(sa) = row(s'a)$*.*
- column-consistent *when, for every* $e, e' \in E$*, if* $col(e) = col(e')$*, then, for every* $a \in \Sigma$*,* $col(ae) = col(ae')$*.*

Note that what are called *closed* and *consistent* in [1] are respectively called *row-closed* and *row-consistent* in our setting.

We also use $row(S)$ to denote the set $\{row(s) \mid s \in S\}$ and $col(E)$ for $\{col(e) \mid e \in E\}$. Two indices s_1 and s_2 are equivalent when $row(s_1) = row(s_2)$. This partitions S and we write $[s]$ for the equivalence class of s as well as its representative. Similarly, we have an equivalence relation on E and we write $[e]$ for the equivalence class of e and its representative. We can use the lexicographically minimal element as the representative.

6.1 Row Automaton

Let $\mathcal{T} = (S, E, T)$ be a row-closed and row-consistent observation table. Define a function $suc : row(S) \times \Sigma \to row(S)$ as $suc(r, a) = row([r]a)$. The co-domain is $row(S)$ as for any $r \in row(S)$, we have $[r] \in S$ and by being row-closed, there is an $s \in S$ such that $row([r]a) = row(s)$. Since the table is consistent, this function respects the equivalence classes.

Definition 8. *The* row automaton *of* $\mathcal{T}$*, denoted by* $A_{row}(\mathcal{T})$*, is the automaton* $(Q, \Sigma, \delta, q_0, F)$ *where* $Q = row(S)$*,* $\delta(q, a) = suc(q, a)$*,* $q_0 = row(\varepsilon)$ *and* $F = \{q \in Q \mid q(\varepsilon) = 1\}$*. The transition function* δ *extends to words by* $\delta(q, \varepsilon) = q$ *and* $\delta(q, ua) = \delta(\delta(q, u), a)$*. The language of the automaton is* $L(A_{row}(\mathcal{T})) = \{u \in \Sigma^* \mid \delta(q_0, u) \in F\}$*.*

Proposition 7. *If* $\mathcal{T}$ *is row-closed and row-consistent, then* $A_{row}(\mathcal{T})$ *is the minimal DFA accepting* $L(A_{row}(\mathcal{T}))$*.*

Since $A_{row}(\mathcal{T})$ is minimal, the left language of a state $row(s)$ is a right congruence class of $L(A_{row}(\mathcal{T}))$ and we denote it by $[s]_{row}$. Furthermore, this congruence class contains the equivalence class $[s]$ of S. This is the automaton constructed by the L^* algorithm.

6.2 Column Automaton

Let $\mathcal{T} = (S, E, T)$ be a column-closed and column-consistent observation table. Define a function $pre : \Sigma \times col(E) \to col(E)$ as $pre(a, c) = col(a[c])$. The co-domain is $col(E)$ as for any $c \in col(E)$, we have $[c] \in E$ and by being column-closed, there is an $e \in E$ such that $col(a[c]) = col(e)$. Since the table is consistent, this function respects the equivalence classes.

Definition 9. *The* column automaton *of $\mathcal{T}$, denoted by $A_{col}(\mathcal{T})$, is the automaton $(Q, \Sigma, \delta, I, f)$ where $Q = col(E)$, $\delta(q, a) = \{q' \in Q \mid q = pre(a, q')\}$, $I = \{q \in Q \mid q(\varepsilon) = 1\}$ and $f = col(\varepsilon)$. The transition function extends to sets of states and words in the usual way: $\delta(K, a) = \bigcup\{\delta(k, a) \mid k \in K\}$ and $\delta(K, \varepsilon) = K$ and $\delta(K, ua) = \delta(\delta(K, u), a)$. The language of the automaton is $L(A_{col}(\mathcal{T})) = \{u \in \Sigma^* \mid f \in \delta(I, u)\}$.*

Proposition 8. *If $\mathcal{T}$ is column-closed and column-consistent, then $A_{col}(\mathcal{T})$ is the átomaton of $L(A_{col}(\mathcal{T}))$.*

Since $A_{col}(\mathcal{T})$ is the átomaton, the right language of a state $col(e)$ is an atom and thus a left congruence class of $L(A_{col}(\mathcal{T}))$ which we denote by $[e]_{col}$. Furthermore, this congruence class contains the equivalence class $[e]$ of E. This automaton can be learned by a column-oriented variant of L^*. Recall that the reverse of the átomaton is the minimal DFA of the reverse language.

6.3 Rows and Columns

Let $\mathcal{T} = (S, E, T)$ be an observation table that is closed and consistent both for rows and columns. We have $A_{row}(\mathcal{T})$ and $A_{col}(\mathcal{T})$ associated with $\mathcal{T}$.

Proposition 9. *For any $u, v \in \Sigma^*$, we have $uv \in L(A_{row}(\mathcal{T}))$ if and only if $row([u]_{row})([v]_{col}) = 1$.*

We thus see that the right language of the state of the row automaton corresponding to u (that is $row([u]_{row})$) consists of those words v for which the entry at row $[u]_{row}$ and column $[v]_{col}$ is 1.

Proposition 10. *For any $u, v \in \Sigma^*$, we have $uv \in L(A_{col}(\mathcal{T}))$ if and only if $col([v]_{col})([u]_{row}) = 1$.*

Similarly, we see that the left language of the state of the column automaton corresponding to v (that is $col([v]_{col})$) consists of those words u for which the entry at column $[v]_{col}$ and row $[u]_{row}$ is 1. Since $row(s)(e) = col(e)(s)$, we can state the following:

Proposition 11. *For any observation table $\mathcal{T}$ that is closed and consistent both for rows and columns, the equality $L(A_{row}(\mathcal{T})) = L(A_{col}(\mathcal{T}))$ holds.*

6.4 Primes

Rows and columns are vectors of Booleans. We partially order such vectors by extending the order $0 \leqslant 1$ to vectors as the product order. For any $s, s' \in S$, we say $row(s) \leqslant row(s')$ when, for every $e \in E$, $row(s)(e) \leqslant row(s')(e)$. The *join* of two rows is given pointwise: $(row(s) \vee row(s'))(e) = row(s)(e) \vee row(s')(e)$. Column vectors are treated similarly.

We say that a vector v is covered by $\{v_1, \ldots, v_n\}$ when $v = v_1 \vee \ldots \vee v_n$. We say that a vector v is *prime* wrt. a set of vectors $V = \{v_1, \ldots, v_n\}$ if v is not zero and no subset $V' \subseteq V$ covers v. The set of prime vectors of a set V, denoted by $primes(V)$, consists of those $v \in V$ that are prime wrt. $V \backslash v$. Every $v \in V$ is covered by the vectors below it in $primes(V)$. The primes are also referred to as the *join-irreducible* elements [10].

6.5 Prime Row Automaton

From the prime rows of an observation table we can construct an NFA that accepts the same language as the row automaton.

Definition 10. *Let $\mathcal{T} = (S, E, T)$ be closed and consistent for rows and columns. The* prime row automaton *of $\mathcal{T}$, denoted by $A_{row'}(\mathcal{T})$, is the automaton given by $(Q, \Sigma, \Delta, I, F)$ where $Q = primes(row(S))$, $I = \{q \in Q \mid q \leqslant row(\varepsilon)\}$, $F = \{q \in Q \mid q(\varepsilon) = 1\}$, $\Delta(q, a) = \{q' \in Q \mid q' \leqslant \delta(q, a)\}$ and δ is the transition function of $A_{row}(\mathcal{T})$.*

Recall that the right language of a state $row(s)$ in the $A_{row}(\mathcal{T})$ consists of those left congruence classes (atoms) for which the corresponding entry in the vector $row(s)$ is 1. Thus a prime row corresponds to a state whose right language is prime, i.e., it is not a union of right languages of other states. Furthermore, the right language of a state $row(s)$ in $A_{row'}(\mathcal{T})$ is the same as in $A_{row}(\mathcal{T})$.

Proposition 12. *If $\mathcal{T}$ is closed and consistent for rows and columns, then $A_{row'}(\mathcal{T})$ is the canonical RFSA of $L(A_{row}(\mathcal{T}))$.*

The canonical RFSA can be learned with the NL^* algorithm [3] which, however, has different conditions on consistency and closedness of the table than the construction given here.

6.6 Prime Column Automaton

From the prime columns of an observation table we can construct an NFA that accepts the same language as the column automaton.

Definition 11. *Let $\mathcal{T} = (S, E, T)$ be closed and consistent for rows and columns. The* prime column automaton *of $\mathcal{T}$, denoted by $A_{col'}(\mathcal{T})$, is the automaton given by $(Q, \Sigma, \Delta, I, F)$ where $Q = primes(col(E))$, $I = \{q \in Q \mid q(\varepsilon) = 1\}$, $F = \{q \in Q \mid q \leqslant col(\varepsilon)\}$, $\Delta(q, a) = \{q' \mid \exists q''. q' \in \delta(q'', a) \wedge q \leqslant q''\}$ and δ is the transition function of $A_{col}(\mathcal{T})$.*

Recall that the left language of a state $col(e)$ in the column automaton consists of those right congruence classes for which the corresponding entry in the vector $col(e)$ is 1. Thus, a prime column corresponds to a state whose left language is prime, i.e., it is not a union of left languages of other states. Furthermore, the left language of a state $col(e)$ in $A_{col'}(\mathcal{T})$ is the same as in $A_{col}(\mathcal{T})$.

Proposition 13. *If $\mathcal{T}$ is closed and consistent for rows and columns, then $A_{col'}(\mathcal{T})$ is the maximized prime átomaton of $L(A_{col}(\mathcal{T}))$.*

The maximized prime átomaton can be learned with a column-oriented variant of NL^*, but, again, the conditions on consistency and closedness of the table would be different than the construction given here.

6.7 Learning NFAs

An observation table that is closed and consistent for rows and columns can be obtained from a table that is closed and consistent only for rows or only for columns. For example, when L^* terminates, then we have a minimal DFA and an observation table that is row-closed and -consistent. We can then use the learned automaton to fill in the missing parts of the table to make it closed and consistent also for columns. From such a table we can construct the átomaton and also calculate the prime elements to construct the canonical RFSA and the maximized prime átomaton.

7 Conclusions

We introduced a new canonical NFA for regular languages, the maximized prime átomaton, and studied its properties. Being the dual automaton of the canonical RFSA, the maximized prime átomaton can be considered as a candidate for a small NFA representation of a language.

We described how four canonical automata – the minimal DFA, the canonical RFSA, the átomaton, and the maximized prime átomaton – can be obtained from suitable observation tables used in automata learning algorithms. We also believe that interpreting these observation tables in terms of quotients and atoms of a language can provide new insights on automata learning problems.

References

1. Angluin, D.: Learning regular sets from queries and counterexamples. Inf. Comput. **75**(2), 87–106 (1987). https://doi.org/10.1016/0890-5401(87)90052-6
2. Birget, J.: Intersection and union of regular languages and state complexity. Inf. Process. Lett. **43**(4), 185–190 (1992). https://doi.org/10.1016/0020-0190(92)90198-5
3. Bollig, B., Habermehl, P., Kern, C., Leucker, M.: Angluin-style learning of NFA. In: Boutilier, C. (ed.) IJCAI 2009, Proceedings of the 21st International Joint Conference on Artificial Intelligence, Pasadena, California, USA, 11–17 July 2009, pp. 1004–1009 (2009). http://ijcai.org/Proceedings/09/Papers/170.pdf

4. Brzozowski, J.A.: Canonical regular expressions and minimal state graphs for definite events. In: Proceedings of Symposium on Mathematical Theory of Automata. MRI Symposia Series, vol. 12, pp. 529–561. Polytechnic Press, Polytechnic Institute of Brooklyn, N.Y. (1963)
5. Brzozowski, J.A., Tamm, H.: Theory of átomata. Theor. Comput. Sci. **539**, 13–27 (2014)
6. Denis, F., Lemay, A., Terlutte, A.: Residual finite state automata. Fund. Inform. **51**, 339–368 (2002)
7. Iván, S.: Complexity of atoms, combinatorially. Inf. Process. Lett. **116**(5), 356–360 (2016)
8. Kameda, T., Weiner, P.: On the state minimization of nondeterministic finite automata. IEEE Trans. Comput. **19**(7), 617–627 (1970)
9. Lombardy, S., Sakarovitch, J.: The universal automaton. In: Flum, J., Grädel, E., Wilke, T. (eds.) Logic and Automata: History and Perspectives [in Honor of Wolfgang Thomas]. Texts in Logic and Games, vol. 2, pp. 457–504. Amsterdam University Press (2008)
10. Myers, R.S.R., Adámek, J., Milius, S., Urbat, H.: Coalgebraic constructions of canonical nondeterministic automata. Theor. Comput. Sci. **604**, 81–101 (2015)
11. Nerode, A.: Linear automaton transformations. Proc. Amer. Math. Soc. **9**, 541–544 (1958)
12. Tamm, H.: Generalization of the double-reversal method of finding a canonical residual finite state automaton. In: Shallit, J., Okhotin, A. (eds.) DCFS 2015. LNCS, vol. 9118, pp. 268–279. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-19225-3_23
13. Tamm, H.: New interpretation and generalization of the Kameda-Weiner method. In: 43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016). Leibniz International Proceedings in Informatics (LIPIcs), vol. 55, pp. 116:1–116:12. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, Dagstuhl (2016)
14. Tamm, H., van der Merwe, B.: Lower bound methods for the size of nondeterministic finite automata revisited. In: Drewes, F., Martín-Vide, C., Truthe, B. (eds.) LATA 2017. LNCS, vol. 10168, pp. 261–272. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-53733-7_19