Generalization of the double-reversal method of finding a canonical residual finite state automaton

Hellis Tamm
Institute of Cybernetics
Tallinn University of Technology

DCFS 2015
Waterloo, ON, June 25-27, 2015

This work was supported by the ERDF funded CoE project EXCS and the Estonian Ministry of Education and Research institutional research grant IUT33-13.

Residual finite state automata

- Residual finite state automata (RFSA), introduced by Denis, Lemay, and Terlutte in 2001, are a subclass of NFAs, where every state defines a residual language (a left quotient) of the language accepted by the automaton.
- Every regular language has a unique canonical RFSA that is a state-minimal RFSA.
- The size of the canonical RFSA is at least the size of a minimal NFA and at most the size of the minimal DFA.

Canonical RFSA

Let L be a regular language over Σ. The left quotient of a language L by a word w is the language $w^{-1}L = \{x \in \Sigma^* \mid wx \in L\}$.

Let $K = \{K_0, \ldots, K_{n-1}\}$ be the set of quotients of L.

A quotient K_i of L is prime if it is non-empty and if it cannot be obtained as a union of other quotients of L.

The canonical RFSA of L is the NFA $\mathcal{C} = (K', \Sigma, \delta, I, F)$, where $K' \subseteq K$ is the set of prime quotients of L, Σ is an input alphabet,
$I = \{K_i \in K' \mid K_i \subseteq L\}$, $F = \{K_i \in K' \mid \varepsilon \in K_i\}$, and
$\delta(K_i, a) = \{K_j \in K' \mid K_j \subseteq a^{-1}K_i\}$ for every $K_i \in K'$ and $a \in \Sigma$.

Among all RFSAs of L, the canonical RFSA is minimal regarding to the number of states, with a maximal number of transitions.

Obtaining the canonical RFSA: saturation

The saturation operation S, applied to an NFA $\mathcal{N} = (Q, \Sigma, \delta, I, F)$:

$\mathcal{N}^S = (Q, \Sigma, \delta_S, I_S, F)$, where $I_S = \{q \in Q \mid L_{q,F}(\mathcal{N}) \subseteq L(\mathcal{N})\}$ and $\delta_S(q, a) = \{q' \in Q \mid aL_{q',F}(\mathcal{N}) \subseteq L_{q,F}(\mathcal{N})\}$ for all $q \in Q$ and $a \in \Sigma$.

- 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.
- If $\mathcal{D}$ is a DFA, then $\mathcal{D}^S$ is an RFSA.

Obtaining the canonical RFSA: reduction

For any state q of $\mathcal{N}$, let $R(q) = \{q' \in Q \setminus \{q\} \mid L_{q',F}(\mathcal{N}) \subseteq L_{q,F}(\mathcal{N})\}$.

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 $Q' = Q \setminus \{q\}$, $I' = I$ if $q \notin I$, and $I' = (I \setminus \{q\}) \cup R(q)$ otherwise, $F' = F \cap Q'$, $\delta'(q', a) = \delta(q', a)$ if $q \notin \delta(q', a)$, and $\delta'(q', a) = (\delta(q', a) \setminus \{q\}) \cup R(q)$ 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 $\mathcal{N}$.

If $\mathcal{N}$ is an RFSA, then $\phi(\mathcal{N}, q)$ is an RFSA.

Obtaining the canonical RFSA

An NFA $\mathcal{N}$ is reduced if there is no erasable state in $\mathcal{N}$.

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

The canonical RFSA can be obtained from a DFA by using saturation and reduction operations.

Atoms of a regular language

Let $K_0, \ldots, K_{n-1}$ be the quotients of L.

An atom of L is any non-empty language of the form

$$A = \widetilde{K_0} \cap \widetilde{K_1} \cap \cdots \cap \widetilde{K_{n-1}},$$

where $\widetilde{K_i}$ is either K_i or $\overline{K_i}$.

Atoms of L are regular languages uniquely determined by L, they define a partition of Σ^*.

Atoms are pairwise disjoint, and every quotient of L (including L itself) is a union of atoms.

Átomaton

Let $K_0 = L$ be the initial quotient of L.

- An atom is initial if it has K_0 (rather than $\overline{K_0}$) as a term.
- An atom is final if and only 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 $\varepsilon \in K_i$, $\widehat{K_i} = \overline{K_i}$ otherwise.

Átomaton

Let $K_0 = L$ be the initial quotient of L.

- An atom is initial if it has K_0 (rather than $\overline{K_0}$) as a term.
- An atom is final if and only 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 $\varepsilon \in K_i$, $\widehat{K_i} = \overline{K_i}$ otherwise.

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

The átomaton of L is the NFA $\mathcal{A} = (\mathbf{A}, \Sigma, \alpha, \mathbf{I}_A, \{\mathbf{A}_{m-1}\})$, where $\mathbf{A} = \{\mathbf{A}_i \mid A_i \in A\}$, $\mathbf{I}_A = \{\mathbf{A}_i \mid A_i \in I_A\}$, and $\mathbf{A}_j \in \alpha(\mathbf{A}_i, a)$ if and only if $A_j \subseteq a^{-1}A_i$, for all $\mathbf{A}_i, \mathbf{A}_j \in \mathbf{A}$ and $a \in \Sigma$.

Maximized atoms

Let $K_0, \dots, K_{n-1}$ be the quotients of L, and let $A_0, \dots, A_{m-1}$ be the atoms of L.

For every atom A_i, we define the maximized atom M_i to be a union of atoms: $M_i = \bigcup\{A_h \mid A_h \subseteq \bigcap_{A_i \subseteq K_k} K_k\}$.

Since atoms are pairwise disjoint, and every quotient is a union of atoms, $M_i = \bigcap_{A_i \subseteq K_k} K_k$ holds.

Maximized átomaton

Let $A = \{A_0, \ldots, A_{m-1}\}$ be the set of atoms of L, with the set of initial atoms $I_A \subseteq A$, and the final atom A_{m-1}.

Let $M = \{M_0, \ldots, M_{m-1}\}$ be the set of the maximized atoms of L, let $I_M = \{M_i \mid A_i \in I_A\}$, and $F_M = \{M_i \mid A_{m-1} \subseteq M_i\}$.

The maximized átomaton of L is the NFA defined by $\mathcal{M} = (\mathbf{M}, \Sigma, \mu, \mathbf{I}_M, \mathbf{F}_M)$, where $\mathbf{M} = \{\mathbf{M}_i \mid M_i \in M\}$, $\mathbf{I}_M = \{\mathbf{M}_i \mid M_i \in I_M\}$, $\mathbf{F}_M = \{\mathbf{M}_i \mid M_i \in F_M\}$, and $\mathbf{M}_j \in \mu(\mathbf{M}_i, a)$ if and only if $M_j \subseteq a^{-1} M_i$ for all $\mathbf{M}_i, \mathbf{M}_j \in \mathbf{M}$ and $a \in \Sigma$.

Maximized átomaton

Let $A = \{A_0, \ldots, A_{m-1}\}$ be the set of atoms of L, with the set of initial atoms $I_A \subseteq A$, and the final atom A_{m-1}.

Let $M = \{M_0, \ldots, M_{m-1}\}$ be the set of the maximized atoms of L, let $I_M = \{M_i \mid A_i \in I_A\}$, and $F_M = \{M_i \mid A_{m-1} \subseteq M_i\}$.

The maximized átomaton of L is the NFA defined by $\mathcal{M} = (\mathbf{M}, \Sigma, \mu, \mathbf{I}_M, \mathbf{F}_M)$, where $\mathbf{M} = \{\mathbf{M}_i \mid M_i \in M\}$, $\mathbf{I}_M = \{\mathbf{M}_i \mid M_i \in I_M\}$, $\mathbf{F}_M = \{\mathbf{M}_i \mid M_i \in F_M\}$, and $\mathbf{M}_j \in \mu(\mathbf{M}_i, a)$ if and only if $M_j \subseteq a^{-1}M_i$ for all $\mathbf{M}_i, \mathbf{M}_j \in \mathbf{M}$ and $a \in \Sigma$.

Recently, Myers, Adámek, Milius, and Urbat (2014) introduced a new canonical NFA, the distromaton, which appears to be the same NFA as the maximized átomaton.

Maximized átomaton

Let $\mathcal{A}$ be the átomaton of L, and let $\mathcal{D}$ be the minimal DFA of L^R.

Proposition. The maximized átomaton $\mathcal{M}$ of L is isomorphic to $\mathcal{D}^{SR}$.

Proposition (Brzozowski and Tamm, 2011, 2014). The reverse $\mathcal{A}^R$ of $\mathcal{A}$ is the minimal DFA of L^R.

Corollary. The maximized átomaton $\mathcal{M}$ is isomorphic to $\mathcal{A}^{RSR}$.

Operation C

Denis, Lemay, and Terlutte (2001) introduced a modified subset construction operation C to be applied to an NFA:

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 $Q_s \subseteq Q_D \setminus \{s\}$ such that $s = \bigcup_{s' \in Q_s} s'$.

The NFA $\mathcal{N}^C = (Q_C, \Sigma, \delta_C, I_C, F_C)$ is defined as follows:
$Q_C = \{s \in Q_D \mid s \text{ is not coverable }\}$, $I_C = \{s \in Q_C \mid s \subseteq I\}$,
$F_C = \{s \in Q_C \mid s \cap F \neq \emptyset\}$, 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 $\mathcal{N}^C$.

This RFSA is not necessarily a canonical RFSA.

Another method to obtain the canonical RFSA

Theorem (Brzozowski, 1962). For an NFA $\mathcal{N}$ without empty states, if $\mathcal{N}^R$ is a DFA, then $\mathcal{N}^D$ is a minimal DFA.

Brzozowski's double-reversal DFA minimization:

For any NFA $\mathcal{N}$ of L, $\mathcal{N}^{RDRD}$ is the minimal DFA of L.

Another method to obtain the canonical RFSA

Theorem (Brzozowski, 1962). For an NFA $\mathcal{N}$ without empty states, if $\mathcal{N}^R$ is a DFA, then $\mathcal{N}^D$ is a minimal DFA.

Brzozowski's double-reversal DFA minimization:

For any NFA $\mathcal{N}$ of L, $\mathcal{N}^{RDRD}$ is the minimal DFA of L.

Theorem (Denis, Lemay, and Terlutte, 2001). For an NFA $\mathcal{N}$ without empty states, if $\mathcal{N}^R$ is an RFSA, then $\mathcal{N}^C$ is a canonical RFSA.

For any NFA $\mathcal{N}$ of L, $\mathcal{N}^{RCRC}$ is the canonical RFSA of L.

So it seems that the operation C has a similar role for RFSAs as determinization D has for DFAs.

Brzozowski's theorem and its generalization

Theorem (Brzozowski, 1962). For an NFA $\mathcal{N}$ without empty states, if $\mathcal{N}^R$ is a DFA, then $\mathcal{N}^D$ is a minimal DFA.

An NFA $\mathcal{N}$ is atomic if for every state q of $\mathcal{N}$, the right language of q is a union of some atoms of $L(\mathcal{N})$.

Theorem (Brzozowski and Tamm, 2011, 2014). For any NFA $\mathcal{N}$, $\mathcal{N}^D$ is a minimal DFA if and only if $\mathcal{N}^R$ is atomic.

Brzozowski's theorem and its generalization

Theorem (Brzozowski, 1962). For an NFA $\mathcal{N}$ without empty states, if $\mathcal{N}^R$ is a DFA, then $\mathcal{N}^D$ is a minimal DFA.

An NFA $\mathcal{N}$ is atomic if for every state q of $\mathcal{N}$, the right language of q is a union of some atoms of $L(\mathcal{N})$.

Theorem (Brzozowski and Tamm, 2011, 2014). For any NFA $\mathcal{N}$, $\mathcal{N}^D$ is a minimal DFA if and only if $\mathcal{N}^R$ is atomic.

We want to generalize this:

Theorem (Denis, Lemay, and Terlutte, 2001). For an NFA $\mathcal{N}$ without empty states, if $\mathcal{N}^R$ is an RFSA, then $\mathcal{N}^C$ is a canonical RFSA.

Theorem ? For any NFA $\mathcal{N}$, $\mathcal{N}^C$ is a canonical RFSA if and only if ???.

Canonical left languages

Let $\mathcal{D}$ be the minimal DFA of L, with the state set $Q = \{q_0, \ldots, q_{m-1}\}$. Let the quotients of L be $Q_0, \ldots, Q_{m-1}$.

The NFA $\mathcal{D}^S$ is the saturated version of $\mathcal{D}$, and $\mathcal{D}^{SE}$ is the NFA, obtained after removing all erasable states and their transitions from $\mathcal{D}^S$. Since $\mathcal{D}^{SE}$ is saturated and reduced, it is the canonical RFSA for L.

We denote the canonical RFSA by $\mathcal{C}$. For every state q_i of $\mathcal{C}$, we denote the left language of q_i by L_i; the right language of q_i is Q_i.

Since the canonical RFSA $\mathcal{C}$ is uniquely determined by L, so are L_i's, and we call these L_i's canonical left languages of L.

Since the canonical RFSA is a state-minimal RFSA with maximal number of transitions, left languages of the states of any state-minimal RFSA are subsets of corresponding canonical left languages.

Properties of canonical left languages

Let $A_0, \ldots, A_{m-1}$ be the atoms of L^R, let $M_0, \ldots, M_{m-1}$ be their maximized atoms, and let $\mathcal{M}$ be the maximized átomaton of L^R.

Clearly, $\mathcal{D}^{SE}$ is a subautomaton of $\mathcal{D}^S$. Since $\mathcal{D}^S$ and $\mathcal{M}^R$ are isomorphic, $\mathcal{C}$ is isomorphic to a subautomaton of $\mathcal{M}^R$.

We can set a correspondence between L_i's and those M_i's which correspond to non-erasable states of $\mathcal{D}^S$.

The following properties hold for any canonical left languages L_i and L_j:

- $A_i^R \subseteq L_i \subseteq M_i^R$.
- If $A_i^R \cap L_j \neq \emptyset$, then $L_i \subseteq L_j$.

Main results

Let $\mathcal{N}$ be an NFA, and let $\mathcal{N}^C$ be the NFA obtained by applying the operation C to $\mathcal{N}$.

Proposition. The RFSA $\mathcal{N}^C$ is a canonical RFSA if and only if for every state s_i of $\mathcal{N}^C$, the left language of s_i is L_i for some L_i.

Theorem. For any NFA $\mathcal{N}$ of L, $\mathcal{N}^C$ is a canonical RFSA if and only if the left language of every state of $\mathcal{N}$ is a union of some canonical left languages L_i of L.