A taxonomy of finite automata construction algorithms

Citation for published version (APA):

Document status and date:
Published: 01/01/1993

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.
A taxonomy of finite automata construction algorithms

by

Bruce W. Watson

ISSN 0926-4515

All rights reserved
editors: prof.dr. J.C.M. Baeten
          prof.dr. M. Rem

Computing Science Report 93/43
Eindhoven, January 1995
A taxonomy of finite automata construction algorithms*

Bruce W. Watson
Faculty of Mathematics and Computing Science
Eindhoven University of Technology
P.O. Box 513, 5600 MB Eindhoven
The Netherlands
e-mail: watson@win.tue.nl
Tel: +31 40 474319
January 24, 1995

Abstract

This paper presents a taxonomy of finite automata construction algorithms. Each algorithm is classified into one of two families: those based upon the structure of regular expressions, and those based upon the automata-theoretic work of Myhill and Nerode.

Many of the algorithms appearing in the literature are based upon the structure of regular expressions. In this paper, we make this term precise by defining regular expressions as a $\Sigma$-term algebra, and automata constructions as various $\Sigma$-algebras of automata. Each construction algorithm is then presented as the unique natural homomorphism from the $\Sigma$-term algebra of regular expressions to the appropriate $\Sigma$-algebra of automata. The concept of duality is introduced and used to derive more practical construction algorithms. In this way, we successfully present (and relate) algorithms given by Thompson, Berry and Sethi, McNaughton and Yamada, Glushkov, and Aho, Sethi, and Ullman. Efficient implementations (including those due to Chang and Paige, and Brügge mann-Klein) are also treated. As a side-effect we derive several new algorithms.

A pair of impractical, but theoretically interesting, construction algorithms were presented by Myhill and Nerode. Some encoding techniques are used to make the algorithms practical — giving Brzozowski’s algorithm based upon derivatives. DeRemer’s algorithm is derived as an encoding of Brzozowski’s algorithm. Two new algorithms, related to DeRemer’s, are derived. Lastly, this family of algorithms is related to the first family.

In addition to classifying the algorithms, we identify (and abstract from) the coding tricks and implementation details present in many of the published algorithms. This paper also presents an introduction to finite automata, $\Sigma$-algebras, and their properties.

*Third printing.
## Contents

1 Introduction 3

2 Finite automata 6
   2.1 Properties of finite automata 7
   2.2 Transformations on finite automata 10
      2.2.1 Algorithms implementing the subset construction 12

3 $\Sigma$-algebras and regular expressions 14
   3.1 Some basic definitions 14
   3.2 Regular expressions as a $\Sigma$-term algebra 15

4 Constructions based on regular expression structure 18
   4.1 Thompson's construction 18
      4.1.1 A top-down version of Thompson's construction 22
      4.1.2 Constructing $\epsilon$-lookahead automata 23
   4.2 Towards the Berry-Sethi construction 27
      4.2.1 Reduced FA's 32
      4.2.2 The Berry-Sethi construction 35
      4.2.3 The McNaughton-Yamada-Glushkov construction 37
   4.3 The dual of the Berry-Sethi construction 39
      4.3.1 The Aho-Sethi-Ullman DFA construction 43
   4.4 Extending regular expressions 44
   4.5 Efficiently computing with RFA's 46
      4.5.1 A practical implementation of the RFA operators 46
      4.5.2 More efficient RFA operators 49

5 The Myhill-Nerode, Brzozowski and DeRemer constructions 53
   5.1 The Myhill-Nerode construction 54
   5.2 The minimal equivalence relation $R_L$ 56
      5.2.1 Encoding $R_L$ 56
   5.3 The Brzozowski construction 58
      5.3.1 Computing derivatives of an ERE 62
      5.3.2 Extending derivatives 63
   5.4 Relating the Brzozowski and Berry-Sethi constructions 63
   5.5 Towards DeRemer's construction 65
      5.5.1 Making the construction more efficient 69

6 Conclusions 73

A Some basic definitions 76

B Proofs of some $\Sigma$-algebra operators 79

References 81
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The family tree of finite automata constructions</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>A representative FA of the isomorphism class $Th((a \cup \epsilon)b^*)$</td>
<td>22</td>
</tr>
<tr>
<td>3</td>
<td>The LAFA $K((a \cup \epsilon)b^*)$</td>
<td>26</td>
</tr>
<tr>
<td>4</td>
<td>A representative LBFA of the isomorphism class $lbfa((a \cup \epsilon)b^*)$</td>
<td>32</td>
</tr>
<tr>
<td>5</td>
<td>A representative DFA of the isomorphism class $MYG((a \cup \epsilon)b^*)$</td>
<td>38</td>
</tr>
<tr>
<td>6</td>
<td>A representative FA of the isomorphism class $R \circ lbfa \circ R((a \cup \epsilon)b^*)$</td>
<td>40</td>
</tr>
<tr>
<td>7</td>
<td>A representative DFA of the isomorphism class $ASU((a \cup \epsilon)b^*)$</td>
<td>43</td>
</tr>
<tr>
<td>8</td>
<td>The DFA $MN_{min}([\epsilon, a] { b }^*)$.</td>
<td>58</td>
</tr>
<tr>
<td>9</td>
<td>The DFA $Brz((a \cup \epsilon)b^*)$.</td>
<td>62</td>
</tr>
<tr>
<td>10</td>
<td>The DFA $I_{const}((a \cup \epsilon)b^*)$.</td>
<td>69</td>
</tr>
<tr>
<td>11</td>
<td>The DFA $I_{const}(b^*)$.</td>
<td>69</td>
</tr>
<tr>
<td>12</td>
<td>The DFA $DeRemer(b^*)$.</td>
<td>70</td>
</tr>
<tr>
<td>13</td>
<td>The DFA $O_{const}((a \cup \epsilon)b^*)$.</td>
<td>71</td>
</tr>
<tr>
<td>14</td>
<td>The DFA's $Brz(ac \cup bc)$ and $O_{const}(ac \cup bc)$.</td>
<td>71</td>
</tr>
</tbody>
</table>
1 Introduction

The construction of finite automata (from regular expressions) is one of the oldest and most extensively developed areas of computing science. Just as the variety of applications has grown, so has the diversity of solutions. Some of the solutions were devised to deal with an extension of the problem, such as constructing a finite automaton from an extended regular expression\(^1\), while others were devised with efficiency in mind. Such a myriad of objectives in the algorithm design has lead to solutions that are difficult to compare. Frequently, people that study the algorithms (or constructions as they are called in this paper) marvel that two seemingly different algorithms construct isomorphic finite automata from the same regular expression. In order to differentiate these algorithms, a taxonomy of construction algorithms would be useful. This report presents such a taxonomy. A related taxonomy of finite automata minimization algorithms appears in [Wats93].

In developing a taxonomy, we have the luxury of rearranging the relationships between the algorithms, possibly introducing relationships that are not present in the history of an algorithm's development. In this paper, for example, we derive DeRemer's construction from Myhill and Nerode's construction. Historically, the theory of LR parsing had a much greater influence on DeRemer's construction.

Section 2 gives definitions of finite automata and some transformations on them. Section 3 introduces \(\Sigma\)-algebras, the foundations for the first family of finite automata constructions. Sections 4 and 5 include the two families of finite automata constructions. Appendix A gives the basic definitions required for reading this paper, while Appendix B presents some proofs related to Section 4. The construction relationships are summarized in the "family tree" shown in Figure 1. The main results of the taxonomy are summarized in the conclusions — Section 6.

In this taxonomy, the finite automata constructions are arranged into two families: those constructions that are based upon the structure of regular expressions, and those based upon the automata-theoretic results of Myhill and Nerode.

The first family of constructions is presented in Section 4:

- Thompson's construction as presented in [Thom68]. This algorithm constructs a (possibly nondeterministic) finite automaton (possibly with \(\epsilon\)-transitions). The description in this paper (Construction 4.3) is based upon those given in [AU92, HU79, Wood87, ASU86] as they are usually considered more readable than Thompson's original paper. Additionally, a more practical (top-down) version of Thompson's construction is presented (Construction 4.5).

- The \(\epsilon\)-lookahead finite automaton construction. This algorithm (Construction 4.11) constructs finite automata that are similar to those constructed by Thompson's construction. They may include so-called \(\epsilon\)-lookahead transitions.

- The guarded commands program construction. This algorithm (Construction 4.17) constructs a guarded commands program from a regular expression. The program is an acceptor for the regular language denoted by the regular expression. It is presented in this paper as a refinement (using hard-coded guarded commands) of the \(\epsilon\)-lookahead construction.

- The left-biased and right-biased constructions. These two constructions (Constructions 4.22 and 4.43 respectively) are related by being the mirror images (or duals) of one another. They both construct an \(\epsilon\)-free (possibly nondeterministic) finite automaton.

- Berry and Sethi's construction as presented in [BS86, Glus61, MY60]. This construction (Construction 4.32) uses some precomputation of sets to construct the same finite automaton as the left-biased construction. This construction is implicitly given by Glushkov [Glus61] and McNaughton and Yamada [MY60], where it is used as the nondeterministic finite automaton construction underlying a deterministic finite automaton construction. Berry and Sethi [BS86] explicitly present this algorithm, and they relate it to Brzozowski's construction.

\(^1\) An extended regular expression is one that includes either the intersection or complementation operator.
Figure 1: The family trees of finite automata constructions. The constructions fall into two broad categories: those based on the structure of regular expressions (descended from $\Sigma$-algebras), and those based on the Myhill-Nerode theorem. Each construction presented in this paper appears as a vertex in this tree, along with the name that it is given in this paper. If the construction is presented explicitly (in this paper), the construction number appears in parentheses (indicating where it appears in this paper). Solid edges denote refinements of the solution (and therefore explicit relationships between constructions). They are labeled with the name of the refinement. Dotted edges denote relationships (between algorithms) that are not elaborated upon in this paper. Some of the dotted edges are labeled with the name of the relationship or refinement. Vertices that are connected by a dashed edge are related by duality (they are the "mirror images" of one another).
We also present (in Construction 4.38) a variant of the Berry-Sethi construction that is more easily implemented in practice.

- McNaughton, Yamada and Glushkov's construction as presented in [MY60, Glus61]. This construction (Construction 4.39) produces a deterministic finite automaton.

- The dual of the Berry-Sethi construction. This construction (Construction 4.45) is the "mirror image" of Berry and Sethi's construction. A variant of this construction was also mentioned in passing by Aho, Sethi, and Ullman [ASU86, Example 3.22, pg. 140]; it appears in this paper as Construction 4.48. In our presentation of this construction, we correct an error appearing in Aho, Sethi, and Ullman's version (see Construction 4.48 of this paper).

- Aho, Sethi, and Ullman's construction as presented in [ASU86, Alg. 3.5, Fig. 3.44]. This construction (Construction 4.50) produces a deterministic finite automaton. It is the "mirror image" of the McNaughton-Yamada construction.

The second family of constructions (from regular expressions) are those based upon the automata-theoretic results of Myhill and Nerode [RS59]. They are presented in Section 5:

- Myhill and Nerode's construction as presented in [RS59]. This construction (Construction 5.11, which is given as part of the proof of the Myhill-Nerode theorem) uses some language theoretical results to construct a deterministic finite automaton. A version of this construction (Construction 5.19) gives the unique (up to isomorphism) minimal finite automaton. It is not a very practical construction (and usually is not even given as a construction), as it relies on the computation of possibly infinite sets. Certain encoding schemes can be used to represent these infinite sets, making the construction practical. Brzozowski's and DeRemer's constructions are two such encoding schemes.

- Brzozowski's construction as presented in [Brz64]. This construction (Construction 5.34) gives a deterministic finite automaton. We derive it as an encoding of the Myhill-Nerode construction, although Brzozowski's derivation was entirely independent.

- The item set construction. This construction (Construction 5.69, not appearing in the literature) produces a deterministic finite automaton, and is based upon the concept of "items" which is borrowed from LR parsing [Knut65]. In this paper, we present it as an encoding of the Myhill and Nerode construction.

- DeRemer's construction as presented in [DeRe74]. This construction (Construction 5.75) produces a deterministic finite automaton. In this paper, it is derived from the item set construction, although DeRemer made use of LR parsing in his derivation.

- An improvement of the item set construction. This construction (Construction 5.82, not appearing in the literature) produces a deterministic finite automaton, and is also based upon the item set construction. Furthermore, it is an improvement of DeRemer's construction. A variant (Construction 5.85) is also related to the Aho-Sethi-Ullman deterministic finite automaton construction.
2 Finite automata

In this section we define finite automata, some of their properties, and some transformations on finite automata.

Definition 2.1 (Finite automaton): A finite automaton (an FA) is a 6-tuple \((Q, V, T, E, S, F)\) where

- \(Q\) is a finite set of states,
- \(V\) is an alphabet,
- \(T \subseteq P(Q \times V \times Q)\) is a transition relation,
- \(E \subseteq P(Q \times Q)\) is an \(\epsilon\)-transition relation
- \(S \subseteq Q\) is a set of start states, and
- \(F \subseteq Q\) is a set of final states.

The definitions of an alphabet and function \(P\) are in Definition A.9 and Convention A.1 respectively. \(\Box\)

Remark 2.2: We will take some liberty in our interpretation of the signatures of the transition relations. For example, we also use the signatures \(T \subseteq V \rightarrow P(Q \times Q), T \subseteq Q \times Q \rightarrow P(V), T \subseteq Q \times V \rightarrow P(Q), T \subseteq Q \rightarrow P(V \times Q),\) and \(E \subseteq Q \rightarrow P(Q).\) In each case, the order of the \(Q\)'s from left to right will be preserved; for example, the function \(T \subseteq Q \rightarrow P(V \times Q)\) is defined as \(T(p) = \{(a,q) : (p,a,q) \in T\}.\) The signature that is used will be clear from the context. See Remark A.3. The definition of \(\rightarrow\) appears in Convention A.2. \(\Box\)

Remark 2.3: Our definition of finite automata differs from the traditional approach in three ways:

- multiple start states are permitted;
- the transition relations are presented in a symmetrical way (without any inherent left-to-right bias); and
- the \(\epsilon\)-transitions (relation \(E\)) are separate from transitions on alphabet symbols (relation \(T\)).

\(\Box\)

Since we only consider finite automata in this paper, we will frequently simply use the term automata.

Convention 2.4 (Finite automaton state graphs): When drawing the state graph corresponding to a finite automaton, we adopt the following conventions:

- All states are drawn as circles (vertices).
- Transitions are drawn as labeled (with \(\epsilon\) or alphabet symbol \(a \in V\)) directed edges between states.
- Start states have an in-transition with no source (the transition does not come from another state).
- Final states are drawn as two concentric circles.

For example, the FA below has two states, one is the start state, and other is the final state, with a transition on \(a:\)

\[\text{Diagram: Two connected states with an arrow labeled } a\]
2.1 Properties of finite automata

In this subsection we define some properties of finite automata. To make these definitions more concise, we introduce particular finite automata $M = (Q, V, T, E, S, F)$, $M_0 = (Q_0, V_0, T_0, E_0, S_0, F_0)$, and $M_1 = (Q_1, V_1, T_1, E_1, S_1, F_1)$.

**Definition 2.5 (Size of an FA):** Define the size of an FA as $|M| = |Q|$.

**Definition 2.6 (Isomorphism ($\cong$) of FA's):** We define isomorphism ($\cong$) as an equivalence relation on FA's. $M_0$ and $M_1$ are isomorphic (written $M_0 \cong M_1$) if and only if $V_0 = V_1$ and there exists a bijection $g \in Q_0 \rightarrow Q_1$ such that

- $T_1 = \{(g(p), a, g(q)) : (p, a, q) \in T_0\}$,
- $E_1 = \{(g(p), g(q)) : (p, q) \in E_0\}$,
- $S_1 = \{g(s) : s \in S_0\}$, and
- $F_1 = \{g(f) : f \in F_0\}$.

**Definition 2.7 (Extending the transition relation $T$):** We extend transition relation $T \in V \rightarrow P(Q \times Q)$ to $T^* \in V^* \rightarrow P(Q \times Q)$ as follows:

$$T^*(\varepsilon) = E^*$$

and (for $a \in V, w \in V^*$)

$$T^*(aw) = E^* \circ T(a) \circ T^*(w)$$

Operator $\circ$ (composition) is defined in Convention A.6. This definition could also have been presented symmetrically.

**Remark 2.8:** We also sometimes use the signature $T^* \in Q \times Q \rightarrow P(V^*)$.

**Remark 2.9:** If $E = \emptyset$ then $E^* = \emptyset^* = I_Q$ where $I_Q$ is the identity relation on the states of $M$.

**Definition 2.10 (The language between states):** The language between any two states $q_0, q_1 \in Q$ is $T^*(q_0, q_1)$.

**Definition 2.11 (Left and right languages):** The left language of a state (in $M$) is given by function $\overrightarrow{L}_M \in Q \rightarrow P(V^*)$, where

$$\overrightarrow{L}_M(q) = (\cup s : s \in S : T^*(s, q))$$

The right language of a state (in $M$) is given by function $\overrightarrow{L}_M \in Q \rightarrow P(V^*)$, where

$$\overrightarrow{L}_M(q) = (\cup f : f \in F : T^*(q, f))$$

The subscript $M$ is usually dropped when no ambiguity can arise.

**Definition 2.12 (Language of an FA):** The language of a finite automaton (with alphabet $V$) is given by the function $L_{FA} \in FA \rightarrow P(V^*)$ defined as:

$$L_{FA}(M) = (\cup s, f : s \in S \land f \in F : T^*(s, f))$$
Property 2.13 (Language of an FA): From the definitions of left and right languages (of a state), we can also write:

\[ \mathcal{L}_{FA}(M) = (\cup f : f \in F : \overrightarrow{\mathcal{L}}(f)) \]

and

\[ \mathcal{L}_{FA}(M) = (\cup s : s \in S : \overrightarrow{\mathcal{L}}(s)) \]

Definition 2.14 (Extension of \( \mathcal{L}_{FA} \)): Function \( \mathcal{L}_{FA} \) is extended to \([FA]_=\) as \( \mathcal{L}_{FA}([M]_=) = \mathcal{L}_{FA}(M) \). This use of brackets \([,]\) is defined in Convention A.7. The choice of representative is irrelevant, as isomorphic FA's accept the same language. 

Definition 2.15 (Complete): A Complete finite automaton is one satisfying the following:

\[
\text{Complete}(M) \equiv (\forall q,a : q \in Q \land a \in V : T(q,a) \neq \emptyset)
\]

Property 2.16 (Complete): For all Complete FA's \((Q,V,T,E,S,F)\):

\[
(\cup q : q \in Q : \overrightarrow{\mathcal{L}}(q)) = V^*
\]

Definition 2.17 (\( \epsilon \)-free): Automaton \( M \) is \( \epsilon \)-free if and only if \( E = \emptyset \). 

Remark 2.18: Even if \( M \) is \( \epsilon \)-free it is still possible that \( \epsilon \in \mathcal{L}_{FA}(M) \): in this case \( S \cap F \neq \emptyset \). 

Definition 2.19 (Reachable states): For \( M \) we can define a reachability relation \( \text{Reach}(M) \subseteq (Q \times Q) \) defined as

\[
\text{Reach}(M) = (\pi_2(T) \cup E)^*
\]

Functions \( \pi \) and \( \pi \) are defined in Convention A.4. Similarly the set of start-reachable states is defined to be:

\[
\text{SReachable}(M) = \text{Reach}(M)(S)
\]

and the set of final-reachable states is defined to be:

\[
\text{FReachable}(M) = (\text{Reach}(M))^R(F)
\]

Reversal of a relation is defined in Definition A.20. The set of useful states is:

\[
\text{Reachable}(M) = \text{SReachable}(M) \cap \text{FReachable}(M)
\]

Remark 2.20: For FA \( M = (Q,V,T,E,S,F) \), function \( \text{SReachable} \) satisfies the following interesting property:

\[
q \in \text{SReachable}(M) \equiv \overrightarrow{\mathcal{L}}_M(q) \neq \emptyset
\]

\( \text{FReachable} \) satisfies a similar property:

\[
q \in \text{FReachable}(M) \equiv \overrightarrow{\mathcal{L}}_M(q) \neq \emptyset
\]
2.1 Properties of finite automata

**Definition 2.21 (Useful automaton):** A *Useful* finite automaton is one with only reachable states:

\[ \text{Useful}(M) \equiv (Q = \text{Reachable}(M)) \]

\[ \square \]

**Definition 2.22 (Start-useful automaton):** A *Useful* finite automaton is one with only start-reachable states:

\[ \text{Useful}_{s}(M) \equiv (Q = \text{SReachable}(M)) \]

\[ \square \]

**Definition 2.23 (Final-useful automaton):** A *Useful* finite automaton is one with only final-reachable states

\[ \text{Useful}_{f}(M) \equiv (Q = \text{FReachable}(M)) \]

\[ \square \]

**Remark 2.24:** *Useful* and *Useful* are closely related by FA reversal (to be presented in Transformation 2.34). For all \( M \in \text{FA} \) we have \( \text{Useful}_{f}(M) \equiv \text{Useful}_{s}(M^{R}) \). \( \square \)

**Property 2.25 (Deterministic finite automaton):** A finite automaton \( M \) is deterministic if and only if

- it does not have multiple start states,
- it is \( \epsilon \)-free, and
- transition function \( T \in Q \times V \rightarrow P(Q) \) does not map pairs in \( Q \times V \) to multiple states.

Formally,

\[ \text{Det}(M) \equiv (|S| \leq 1 \land \text{\epsilon-free}(E) \land (\forall q, a : q \in Q \land a \in V : |T(q, a)| \leq 1)) \]

\[ \square \]

**Definition 2.26 (Deterministic FA’s):** \( \text{DFA} \) denotes the set of all deterministic finite automata. We call \( \text{FA} \setminus \text{DFA} \) the set of nondeterministic finite automata. \( \square \)

**Convention 2.27 (Transition function of a DFA):** For \( (Q, V, T, \emptyset, S, F) \in \text{DFA} \) we can consider the transition function to have signature \( T \in Q \times V \rightarrow P(Q) \). (A definition of \( \rightarrow \) appears in Convention A.2.) The transition function is total if and only if the DFA is Complete. \( \square \)

**Property 2.28 (Weakly deterministic automaton):** Some authors use a definition of a deterministic automaton that is weaker than \( \text{Det} \); it uses left languages and is defined as follows:

\[ \text{Det}'(M) \equiv (\forall q_{0}, q_{1} : q_{0} \in Q \land q_{1} \in Q \land q_{0} \neq q_{1} : \overrightarrow{L}(q_{0}) \cap \overrightarrow{L}(q_{1}) = \emptyset) \]

\[ \square \]

**Remark 2.29:** \( \text{Det}(M) \Rightarrow \text{Det}'(M) \) is easily proved. We can also demonstrate that there exists an \( M \in \text{FA} \) such that \( \text{Det}'(M) \land \neg \text{Det}(M) \):

\[ (\{q_{0}, q_{1}\}, \{b\}, \{(q_{0}, b, q_{0}), (q_{0}, b, q_{1})\}, \emptyset, \emptyset, \emptyset) \]

In this FA, \( \overrightarrow{L}(q_{0}) = \overrightarrow{L}(q_{1}) = \emptyset \), but state \( q_{0} \) has two out-transitions on symbol alphabet \( b \). \( \square \)
Definition 2.30 (Minimality of a DFA): An \( M \in DFA \) is minimal as follows:

\[
\text{Min}(M) \equiv (\forall M' : M' \in DFA \land L_{FA}(M) = L_{FA}(M') : |M| \leq |M'|)
\]

Predicate \( \text{Min} \) is defined only on DFA's. Some definitions are simpler if we define a minimal, but still complete, DFA as follows:

\[
\text{Min}_C(M) \equiv (\forall M' : M' \in DFA \land \text{Complete}(M') \land L_{FA}(M) = L_{FA}(M') : |M| \leq |M'|)
\]

Predicate \( \text{Min}_C \) is defined only on \( \text{Complete} \) DFA's. \( \square \)

Property 2.31 (Minimality of a DFA): An \( M \), such that \( \text{Min}(M) \), is the unique (modulo \( \equiv \)) minimal DFA, as will be shown in Section 5. There is no similar uniqueness property for nondeterministic finite automata. \( \square \)

Property 2.32 (An alternate definition of minimality of a DFA): For the purposes of minimizing a DFA, we use the definition (defined only on DFA's):

\[
\text{Minimal}(Q, V, T, \emptyset, S, F) \equiv (\forall q_0, q_1 : q_0 \in Q \land q_1 \in Q \land q_0 \neq q_1 : \overrightarrow{L}(q_0) \neq \overrightarrow{L}(q_1))
\]

\[
\land \text{Useful}(Q, V, T, \emptyset, S, F)
\]

We have the property that (for all \( M \in DFA \)) \( \text{Minimal}(M) \equiv \text{Min}(M) \). It is easy to prove that \( \text{Min}(M) \Rightarrow \text{Minimal}(M) \). The reverse direction follows from the Myhill-Nerode theorem (Theorem 5.7).

A similar definition that relates to \( \text{Min}_C \) is (also defined only on DFA's):

\[
\text{Minimal}_C(Q, V, T, \emptyset, S, F) \equiv (\forall q_0, q_1 : q_0 \in Q \land q_1 \in Q \land q_0 \neq q_1 : \overrightarrow{L}(q_0) \neq \overrightarrow{L}(q_1))
\]

\[
\land \text{Useful}_C(Q, V, T, \emptyset, S, F)
\]

We have the property that (for all \( M \in DFA \) such that \( \text{Complete}(M) \)) \( \text{Minimal}_C(M) \equiv \text{Min}_C(M) \). The contrapositive of \( \text{Min}_C(M) \Rightarrow \text{Minimal}_C(M) \) is easily proved, and the reverse direction also follows from Theorem 5.7. \( \square \)

Remark 2.33: In the literature the second conjunct in the definition of predicate \( \text{Minimal}_C \) is sometimes erroneously omitted. The necessity of the conjunct can be seen by considering the DFA

\[
\{(p, q), \{a\}, \{(p, a, p), (q, a, q)\}, \emptyset, \emptyset, \{p\}\}
\]

Here \( \overrightarrow{L}(p) = \overrightarrow{L}(q) = \emptyset \) (which is also the language of the DFA), \( \overrightarrow{L}(p) = \{a\}^* \), and \( \overrightarrow{L}(q) = \emptyset \). Without the second conjunct, this DFA would be considered \( \text{Minimal}_C \); clearly this is not the case, as the minimal \( \text{Complete} \) DFA accepting \( \emptyset \) is \( \emptyset, \{a\}, \emptyset, \emptyset, \emptyset, \emptyset \). \( \square \)

2.2 Transformations on finite automata

Transformation 2.34 (FA reversal): FA reversal is given by postfix (superscript) function \( R \in FA \longrightarrow FA \), defined as:

\[
\]

Function \( R \) satisfies

\[
(\forall M : M \in FA : (L_{FA}(M))^R = L_{FA}(M^R)).
\]

and preserves \( \epsilon \)-free and \( \text{Useful} \).

Reversal functions are defined in Definition A.19, and preservation is defined in Definition A.18. \( \square \)
2.2 Transformations on finite automata

Remark 2.35: The property \((\mathcal{L}_{FA}(M^R))^R = \mathcal{L}_{FA}(M)\) means that function \(\mathcal{L}_{FA}\) is its own dual, and is therefore symmetrical (see Definitions A.21 and A.22). 

Definition 2.36 (Extending reversal to \([FA]_\equiv\)): We extend reversal to \(R \in [FA]_\equiv \rightarrow [FA]_\equiv\) defined as \(((M)_{\equiv})^R = (M^R)_{\equiv}\). The definition is independent of the choice of representative (of an equivalence class of \(\equiv\)) since \(R\) and isomorphism commute. 

Transformation 2.37 (Useless state removal): There exists a function \(useful \in FA \rightarrow FA\) that removes states that are not reachable. A definition of this function is not given here, as it is not needed. Function \(useful\) satisfies

\[
(\forall M : M \in FA : Useful(useful(M)) \land \mathcal{L}_{FA}(useful(M)) = \mathcal{L}_{FA}(M))
\]

and can be defined so as to preserve \(\varepsilon\)-free, Useful, Det, and Min. 

Transformation 2.38 (Removing start state unreachable states): Transformation \(useful_s \in FA \rightarrow FA\) removes those states that are not start-reachable:

\[
useful_s(Q, V, T, E, S, F) = \begin{cases} 
U = SReachable(Q, V, T, E, S, F) 
\quad & \text{in} 
\quad (U, V, T \cap (U \times V \times U), E \cap (U \times U), S \cap U, F \cap U) 
\end{cases}
\]

Function \(useful_s\) satisfies

\[
(\forall M : M \in FA : Useful_s(useful_s(M)) \land \mathcal{L}_{FA}(useful_s(M)) = \mathcal{L}_{FA}(M))
\]

and preserves Complete, \(\varepsilon\)-free, Useful, Det, and (trivially) Min\(_C\) and Min. 

Remark 2.39: A function \(useful_f \in FA \rightarrow FA\) could also be defined, removing states that are not final-reachable. Such a function is not needed in this paper. 

Transformation 2.40 (Completing an FA): Function \(complete \in FA \rightarrow FA\) is defined as:

\[
complete(Q, V, T, E, S, F) = \begin{cases} 
s \text{ be a new (sink) state} 
\quad & \text{let} 
\quad T' = \{{(p, a, s) : \neg(\exists q : q \in Q : (p, a, q) \in T)}\} 
\quad & \text{in} 
\quad T'' = \begin{cases} 
\{s\} \times V \times \{s\} & \text{if } T' \neq \emptyset 
\emptyset & \text{else} 
\end{cases} 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\quad & \text{else} \emptyset \emptyset, V, 
\end{cases}
\]

It satisfies the requirement that:

\[
(\forall M : M \in FA : Complete(complete(M)) \land \mathcal{L}_{FA}(complete(M)) = \mathcal{L}_{FA}(M))
\]

In general, this transformation adds a sink state to the FA. This transformation preserves \(\varepsilon\)-free, (trivially) Complete, Det, and Min\(_C\). 

Transformation 2.41 (\(\varepsilon\) removal): An \(\varepsilon\) removal transformation \(remove_\varepsilon \in FA \rightarrow FA\) is one that satisfies

\[
(\forall M : M \in FA : \varepsilon\text{-free}(remove_\varepsilon(M)) \land \mathcal{L}_{FA}(remove_\varepsilon(M)) = \mathcal{L}_{FA}(M))
\]

There are several possible implementations of \(remove_\varepsilon\). One implementation is:

\[
remove_{\varepsilon,sym}(Q, V, T, E, S, F) = \begin{cases} 
T'(a) = E^* \circ T(a) \circ E^* 
\quad & \text{in} 
\quad (Q, V, T', \emptyset, E^*(S), (E^*)^R(F)) 
\end{cases}
\]

This implementation preserves Complete and Useful and is symmetrical. 

Transformation 2.42 (Subset construction): The function subset transforms an \( \epsilon \)-free FA into a DFA (in the let clause \( T' \in \mathcal{P}(Q) \times V \rightarrow \mathcal{P}(\mathcal{P}(Q)) \))

\[
\text{subset}(Q, V, T, \emptyset, S, F) = \text{let } T'(U, a) = \{(q : q \in U : T(q, a))\} \\
F' = \{U : U \in \mathcal{P}(Q) \land U \cap F \neq \emptyset\} \\
\text{in } (\mathcal{P}(Q), V, T', \emptyset, \{S\}, F')
\]

In addition to the obvious property that (for all \( M \in \text{FA} \)) \( \mathcal{L}_{\text{FA}}(\text{subset}(M)) = \mathcal{L}_{\text{FA}}(M) \), function subset satisfies

\[
(\forall M : M \in \text{FA} \land \epsilon\text{-free}(M) : \text{Det}(\text{subset}(M)) \land \text{Complete}(\text{subset}(M)))
\]
and preserves Complete, \( \epsilon \)-free, Det, and \( \text{Min}_{\epsilon} \). It is also known as the “powerset” construction. \( \square \)

Property 2.43 (Subset construction): Let \( M_0 = (Q_0, V, T_0, \emptyset, S_0, F_0) \) and \( M_1 = \text{subset}(M_0) \) be finite automata. By the subset construction, the state set of \( M_1 \) is \( \mathcal{P}(Q_0) \). We have the following property:

\[
(\forall p : p \in \mathcal{P}(Q_0) : \overrightarrow{L}_{M_1}(p) = (\cup q : q \in p : \overrightarrow{L}_{M_0}(q)))
\]

\( \square \)

Definition 2.44 (Optimized subset construction): The function subsetopt transforms an \( \epsilon \)-free FA into a DFA. This function is an optimized version of subset.

\[
\text{subsetopt}(Q, V, T, \emptyset, S, F) = \text{let } T'(U, a) = \{(q : q \in U : T(q, a))\} \\
Q' = \mathcal{P}(Q) \setminus \emptyset \\
F' = \{U : U \in \mathcal{P}(Q) \land U \cap F \neq \emptyset\} \\
\text{in } (Q', V, T' \cap (Q' \times V \times Q'), \emptyset, \{S\}, F')
\]

In addition to the property that (for all \( M \in \text{FA} \)) \( \mathcal{L}_{\text{FA}}(\text{subsetopt}(M)) = \mathcal{L}_{\text{FA}}(M) \), function subsetopt satisfies

\[
(\forall M : M \in \text{FA} \land \epsilon\text{-free}(M) : \text{Det}(\text{subsetopt}(M)))
\]
and preserves \( \epsilon \)-free, Det. \( \square \)

2.2.1 Algorithms implementing the subset construction

Since many of the states in a subset-constructed DFA may be unreachable, we consider an algorithm implementing the composition \( \text{useful} \circ \text{subset} \).

In this algorithm, \( D \) \((\text{for done})\) is the set of states (of the DFA being constructed) already considered, and \( U \) \((\text{for un-done})\) is the set of states to be considered. The type of \( S', D, \) and \( U \) is \( \mathcal{P}(\mathcal{P}(Q)) \) (in particular, \( S' \) is a set of states in the constructed DFA). This algorithm will yield a Complete DFA. In the case that the language of the automaton (being subset constructed) is not \( V^* \), then there will be a state \( \emptyset \in D \) which is the sink state. The algorithm is implemented in Dijkstra’s guarded command language [Dijk76].
Algorithm 2.45:

\[
\{(Q, V, T, \emptyset, S, F) \in FA\} \\
S', T' := \{S\}, \emptyset; \\
D, U := \emptyset, S'; \\
do U \neq \emptyset \\
\quad \text{let } u : u \in U; \\
\quad D, U := D \cup \{u\}, U \setminus \{u\}; \\
\quad \text{for } a : a \in V \text{ do} \\
\quad \quad d := (\cup q : q \in u : T(q, a)); \\
\quad \quad \text{if } d \notin D \longrightarrow U := U \cup \{d\} \\
\quad \quad \text{skip} \\
\quad \text{u; \quad T' := T' \cup \{(u, a, d)\}} \\
\text{rof} \\
\text{od} \\
F' := \{d : d \in D \land d \cap F \neq \emptyset\} \\
\{(D, V, T', \emptyset, S', F') = \text{useful} \circ \text{subset}(Q, V, T, \emptyset, S, F)\} \\
\{(D, V, T', \emptyset, S', F') = \text{useful} \circ \text{subset opt}(Q, V, T, \emptyset, S, F)\} \\
\{\text{Complete}(D, V, T', \emptyset, S', F')\}
\]

An algorithm implementing \text{useful} \circ \text{subset opt}, yielding a (possibly non-\text{Complete}) DFA with no sink state is:

Algorithm 2.46:

\[
\{(Q, V, T, \emptyset, S, F) \in FA\} \\
S', T' := (\text{if } (S \neq \emptyset) \text{ then } \{S\} \text{ else } \emptyset, \emptyset); \\
D, U := \emptyset, S'; \\
do U \neq \emptyset \\
\quad \text{let } u : u \in U; \\
\quad D, U := D \cup \{u\}, U \setminus \{u\}; \\
\quad \text{for } a : a \in V \land (\exists q : q \in u : T(q, a) \neq \emptyset) \text{ do} \\
\quad \quad d := (\cup q : q \in u : T(q, a)); \\
\quad \quad \text{if } d \notin D \longrightarrow U := U \cup \{d\} \\
\quad \quad \text{skip} \\
\quad \text{u; \quad T' := T' \cup \{(u, a, d)\}} \\
\text{rof} \\
\text{od} \\
F' := \{d : d \in D \land d \cap F \neq \emptyset\} \\
\{(D, V, T', \emptyset, S', F') = \text{useful} \circ \text{subset opt}(Q, V, T, \emptyset, S, F)\}
\]

Remark 2.47: The algorithm given above can be made more efficient by removing the 3 quantification from the \text{for} guard, and implementing it in an \text{if-fi} structure within the \text{for} statement. The algorithm is left in this form since it is used in Construction 4.50 to present the Aho-Sethi-Ullman algorithm. \(\Box\)
3 Σ-algebras and regular expressions

Many of the known FA constructions have definitions that follow the syntactic structure of regular expressions. The best known (and perhaps the easiest to understand) is Thompson's construction [Thom68]. We would like to formalize the notion of "following the syntactic structure." This is done by introducing Σ-algebras in this section. Regular expressions are then defined as a Σ-algebra.

3.1 Some basic definitions

This subsection provides the basic definitions required for Σ-algebras. Most of these definitions are taken, with slight modification, from [EM85].

Definition 3.1 (Sorts): Given set S (the elements of which are called sorts), a set of sets X is called S-sorted if the elements of X correspond one-to-one with S. The element of X corresponding to $s \in S$ is written $X,s$.

Definition 3.2 (Signature Σ): A signature Σ is a pair $(S, \Gamma)$ where

- S is a finite set, and
- $\Gamma$ is an $(S^* \times S)$-sorted set called the operators.

We write elements of $S^* \times S$ as $<< s_1, \ldots, s_k, s >>$. We can make a couple of notational simplifications:

- Given $\gamma \in \Gamma_{<< s_1, \ldots, s_k, s >>}$ we write $\gamma : s_1 \times \ldots \times s_k \rightarrow s$. Constant $k$ is known as the arity of operator $\gamma$.
- For $\gamma \in \Gamma_{<< s, s >>}$ we write $\gamma : s$, and call $\gamma$ a constant; that is, constants are operators of arity zero.

Remark 3.3: Although the set $S^* \times S$ is infinite (for $S \neq \emptyset$), this does not imply that there are infinitely many operators. There may be $<< s_1, \ldots, s_m, s >> \in S^* \times S$ such that $\Gamma_{<< s_1, \ldots, s_m, s >>} = \emptyset$; in that case, there is no operator $\gamma : s_1 \times \ldots \times s_m \rightarrow s$.

Several of the following definitions are with respect to signature $\Sigma = (S, \Gamma)$.

Definition 3.4 (TermΣ): The S-sorted set $Term_{\Sigma}$ is the smallest S-sorted set such that

- if $\gamma : s_1 \times \ldots \times s_k \rightarrow s$ (for some $k \geq 0$) ($s, s_1, \ldots, s_k \in S$) and (for all $1 \leq i \leq k$) $t_i \in Term_{\Sigma_{s_i}}$ then $\gamma[t_1, \ldots, t_k] \in Term_{\Sigma}$. We adopt the convention that $\gamma[\] = \emptyset$.

Definition 3.5 (Σ-algebra): A Σ-algebra is a pair $(V, F)$ such that

- V is an S-sorted set, and
- F is a set of functions $f_{\gamma}$ (with $\gamma \in \Gamma$) such that
  - if $\gamma : s_1 \times \ldots \times s_k \rightarrow s$ then $f_{\gamma} \in V_{s_1} \times \ldots \times V_{s_k} \rightarrow V_s$.

Set V is called the carrier set of the Σ-algebra. Set F is called the operator set of the Σ-algebra.

---

2 Square brackets ([ and ]) are used syntactically here.
3.2 Regular expressions as a $\Sigma$-term algebra

**Definition 3.6** ($\Sigma$-term algebra): The $\Sigma$-term algebra is the $\Sigma$-algebra $(\text{Term}_\Sigma, F)$ such that

$$F = \{ f_\gamma : (\gamma : s_1 \times \ldots \times s_k \rightarrow s) \}$$

where (for all $f_\gamma \in F$) $f_\gamma \in \text{Term}_{\Sigma_1} \times \ldots \times \text{Term}_{\Sigma_k} \rightarrow \text{Term}_{\Sigma_\gamma}$ is defined as $f_\gamma(t_1, \ldots, t_k) = \gamma[t_1, \ldots, t_k]$. □

**Definition 3.7** ($\Sigma$-homomorphism): Given $\Sigma$-algebras $(V, F)$ and $(W, G)$, a $\Sigma$-homomorphism from $(V, F)$ to $(W, G)$ is an $S$-indexed set of functions $h$ such that

- for all $s \in S$ we have $h_s \in V_s \rightarrow W_s$, and
- for all $\gamma : s_1 \times \ldots \times s_k \rightarrow s$, $f_\gamma \in F$, $g_\gamma \in G$, and $e_1 \in V_{s_1}, \ldots, e_k \in V_{s_k}$
  $$h_s(f_\gamma(e_1, \ldots, e_k)) = g_\gamma(h_{s_1}(e_1), \ldots, h_{s_k}(e_k))$$

□

**Remark 3.8:** In the case that there is only one sort, a $\Sigma$-homomorphism is a singleton set and we speak of the homomorphic function. □

**Definition 3.9** (Initial $\Sigma$-algebra): A $\Sigma$-algebra is initial if there is a unique $\Sigma$-homomorphism from it to all other $\Sigma$-algebras. □

**Proposition 3.10** ($\Sigma$-term algebras): $\Sigma$-term algebras are initial. □

**Example 3.11** ($\Sigma$-algebras): Consider signature $\Sigma = (S, \Gamma)$ where $S$ consists only of sort $\text{expr}$, and $\Gamma$ consists of constant $a : \text{expr}$ and operator $\text{plus} : \text{expr} \times \text{expr} \rightarrow \text{expr}$. Some examples of terms in the $\Sigma$-term algebra are $\text{plus}[a, a]$ and $\text{plus}[\text{plus}[a, \text{plus}[a, a]], a]$.

We define another $\Sigma$-algebra $X$ with the natural numbers as the carrier set, 0 (the natural number) as the constant, and $f_{\text{plus}}(x, y) = (x \max y) + 1$ as the operator.

As an example of a $\Sigma$-homomorphism, we define the “expression tree height” function as a homomorphism from the $\Sigma$-term algebra to algebra $X$. With only one sort, we define function $h_{\text{expr}}$ as $h_{\text{expr}}(a) = 0$ and $h_{\text{expr}}(\text{plus}[e, f]) = f_{\text{plus}}(h_{\text{expr}}(e), h_{\text{expr}}(f)) = (h_{\text{expr}}(e) \max h_{\text{expr}}(f)) + 1$. □

### 3.2 Regular expressions as a $\Sigma$-term algebra

**Definition 3.12** (Regular expressions): We define regular expressions (over alphabet $V$) as the $\Sigma$-term algebra over signature $\Sigma = (S, \Omega)$ where

- $S$ consists of a single sort $\text{Reg}$ (for regular expression), and
- $\Omega$ is a set of several constants: $\epsilon, \emptyset, a_1, \ldots, a_n : \text{Reg}$ (where $V = \{a_1, \ldots, a_n\}$) and five operators: $\cdot : \text{Reg} \times \text{Reg} \rightarrow \text{Reg}$ (the dot operator), $\cup : \text{Reg} \times \text{Reg} \rightarrow \text{Reg}$, $*: \text{Reg} \rightarrow \text{Reg}$, $+: \text{Reg} \rightarrow \text{Reg}$, and $? : \text{Reg} \rightarrow \text{Reg}$.

Signature $\Sigma$ will be used throughout the remainder of this paper. We make the following notational simplification when writing terms in the $\Sigma$-term algebra:

- operators $\cdot$ (the dot) and $\cup$ are written as infix operators;
- operator $\cdot$ is usually not written, juxtaposition is used instead;
- operators $*, +$, and $?$ are written as postfix (superscript) operator.

The following will also be used for conciseness:

- a term in the $\Sigma$-term algebra is called a regular expression;
• the set $\text{Term}_\Sigma$ is denoted by $\text{RE}$;

• the operators have (ascending) precedence: $\cup$, $\cdot$, $\ast$ and $+$ and $?$; $\epsilon$, $\emptyset$, and $a_1, \ldots, a_n \in V$ are constants;

• regular expressions are usually fully parenthesized; parentheses can be omitted where the operator precedence allows.

Remark 3.13: The ? operator is non-standard. It will be used to denote union with the language containing the empty string $\epsilon$. See Definition 3.17.

Remark 3.14: Some authors write $\cup$ as (infix) $+$ or as $|$. 

Example 3.15 (A regular expression): Given alphabet $V = \{a, b\}$ the regular expression $\cdot [(a, c), \ast [b]]$ is usually written as $(a \cup c)b^\ast$. This particular regular expression will be used in running examples of FA construction.

Remark 3.16: Some authors leave $\emptyset$, $\ast$, or $+$ out of the definition of regular expressions. Strictly speaking, operators $\epsilon$, $\ast$, and $?$ are not needed in the signature, since they can be constructed from the other operators. There are some FA constructions (from REs) that have running time dependent on the size of the regular expression. In these cases, treating the extra operators fully (instead of as abbreviations) becomes advantageous.

Definition 3.17 (The $\Sigma$-algebra of regular languages): We define a $\Sigma$-algebra of regular languages (over alphabet $V$), with carrier $\mathcal{P}(V^*)$ and constants:

- $\{\epsilon\} \in \mathcal{P}(V^*)$ (the language containing only the empty string);
- $\emptyset \in \mathcal{P}(V^*)$ (the empty language);
- $\{a\} \in \mathcal{P}(V^*)$ (for all $a \in V$),

and operators:

- $\cup \in \mathcal{P}(V^*) \times \mathcal{P}(V^*) \rightarrow \mathcal{P}(V^*)$ (language union);
- $\cdot \in \mathcal{P}(V^*) \times \mathcal{P}(V^*) \rightarrow \mathcal{P}(V^*)$ (language concatenation);
- $\ast \in \mathcal{P}(V^*) \rightarrow \mathcal{P}(V^*)$ (Kleene closure);
- $+ \in \mathcal{P}(V^*) \rightarrow \mathcal{P}(V^*)$ (+ closure), and
- $? \in \mathcal{P}(V^*) \rightarrow \mathcal{P}(V^*)$ (union with $\{\epsilon\}$, see Definition A.14).

Each of these operators corresponds (in the obvious way) to the operators of signature $\Sigma$.

Definition 3.18 (Language denoted by an RE): The function $\mathcal{L}_{RE}$ is the (unique) homomorphism from the $\Sigma$-term algebra of REs to the $\Sigma$-algebra of regular languages. Function $\mathcal{L}_{RE}$ maps regular expressions to the languages they denote.

Definition 3.19 (Equivalence ($\equiv$) of REs): Two regular expressions, $E_0$ and $E_1$, are said to be equivalent (written $E_0 \equiv E_1$, note the dot above the $=$) if and only if they denote the same language.

Definition 3.20 (The nullable $\Sigma$-algebra): We define the nullable $\Sigma$-algebra as follows:

- The carrier set is $\{\text{true, false}\}$.
The constants are: \textit{true}, \textit{false}, and \textit{false} (corresponding respectively to $\epsilon$, $\emptyset$, and $a : a \in V$). Here the constant \textit{false} corresponds to $\emptyset$ and to all $a \in V$. The operators are: $\lor$ (disjunction), $\land$ (conjunction), the constant function \textit{true}, the identity function, and (again) the constant function \textit{true} (corresponding respectively to $\lor$, $\land$, $\ast$, $+$, and $?$). The operators corresponding to $\ast$ and to $?$ are interesting because they map their argument to the constant \textit{true}.

We denote the (unique) homomorphism from $RE$ to this $\Sigma$-algebra as $\text{Null}$. \hfill \Box

\textbf{Property 3.21 (The nullable $\Sigma$-algebra):} The homomorphism $\text{Null}$ has the property that for all $E \in RE$

$$\epsilon \in \mathcal{L}_{RE}(E) \equiv \text{Null}(E)$$

\hfill \Box

\textbf{Definition 3.22 (RE reversal):} Regular expression reversal is given by the postfix (superscript) isomorphism $R : RE \rightarrow RE$

$$\begin{align*}
\epsilon^R &= \epsilon \\
\emptyset^R &= \emptyset \\
a^R &= a \\
(E_0 \cup E_1)^R &= (E_0^R) \cup (E_1^R) \\
(E_0 \cdot E_1)^R &= (E_1^R) \cdot (E_0^R) \\
(E^*)^R &= (E^R)^* \\
(E^+)^R &= (E^R)^+ \\
(E^?)^R &= (E^R)^?
\end{align*}$$

Function $R$ satisfies the obvious property that

$$\forall E : E \in RE : (E^R)^R = E \land (\mathcal{L}_{RE}(E^R))^R = \mathcal{L}_{RE}(E)$$

\hfill \Box

\textbf{Remark 3.23:} The property satisfied by regular expression reversal implies that $\mathcal{L}_{RE}$ is an example of a symmetrical function (according to Definition A.22). \hfill \Box
4 Constructions based on regular expression structure

A finite automaton construction is any function \( f \), such that the following diagram commutes:

\[
\begin{align*}
RE & \xrightarrow{f} FA \\
\mathcal{L}_{RE} & \xrightarrow{L_{FA}} \mathcal{L}_{FA} \\
\mathcal{L}_{reg} &
\end{align*}
\]

In this section, we will be defining some \( \Sigma \)-algebras with \([FA]_{\Sigma}\) as the carrier set; the idea behind the above commuting diagram still holds in this case, as all isomorphic \( FA \)'s accept the same language. The isomorphism class of an \( FA \) corresponding to a given regular expression is the image of the regular expression under the (unique) homomorphism from \( RE \) to the other \( \Sigma \)-algebras. Such a homomorphism is a \( FA \) construction. Thompson's construction is considered first, followed by a derivation of Berry and Sethi's, McNaughton, Yamada and Glushkov's, and Aho, Sethi, and Ullman's constructions. We also consider methods of efficiently implementing some of the constructions, and methods of constructing \( FA \)'s from extended regular expressions (see Definition 4.53).

4.1 Thompson's construction

One \( \Sigma \)-algebra is based upon an \( RE \) to \( FA \) construction given by Thompson in [Thom68]. The explanations given in textbooks such as [HU79, Wood87, AU92, ASU86] are generally considered more readable than Thompson's original paper. None of those presentations made use of \( \Sigma \)-algebras.

**Definition 4.1 (Thompson's \( \Sigma \)-algebra of \( FA \)'s):** The carrier set is \([FA]_{\Sigma}\). The operator requirement\(^3\) is:

- For the binary operators, the representatives of the arguments must have disjoint state sets. For any two equivalence classes (under \( \cong \)) we can always choose a representative of each such that they satisfy this requirement.

The correctness of the operators\(^4\) is not included here, but is discussed in Theorem B.1. Along with each operator we present a graphic representation of the operator. The operators are separated by horizontal lines for clarity. The operators (with subscript \( Th \), for Thompson) are:

\[
C_{\epsilon, Th} = \begin{cases}
\text{let} & q_0, q_1 \text{ be new states} \\
\text{in} & [(\{q_0, q_1\}, V, \emptyset, \{(q_0, q_1)\}, \{q_0\}, \{q_1\})]_{\Sigma}
\end{cases}
\]

\[
C_{\emptyset, Th} = \begin{cases}
\text{let} & q_0, q_1 \text{ be new states} \\
\text{in} & [(\{q_0, q_1\}, V, \emptyset, \{q_0\}, \{q_1\})]_{\Sigma}
\end{cases}
\]

\(^3\)\( \Sigma \)-algebras presented in this section may have a list of items such as this, stating the requirements on the arguments for the correctness of the operators.

\(^4\)For example, the concatenation operator is correct when (for all \( M_0, M_1 \) in Thompson's \( \Sigma \)-algebra)

\[
\mathcal{L}_{FA}(C_{\cdot, Th} ([M_0]_{\Sigma}, [M_1]_{\Sigma})) = \mathcal{L}_{FA}([M_0]_{\Sigma}) \circ \mathcal{L}_{FA}([M_1]_{\Sigma}).
\]
4.1 Thompson’s construction

\[ C_{\omega,T_h} = \text{let } q_0, q_1 \text{ be new states} \]
\[ \quad \text{in} \]
\[ \quad \quad [(\{q_0, q_1\}, V, \{(q_0, a, q_1)\}, \emptyset, \{q_0\}, \{q_1\})]_{\equiv} \]
\[ \quad \text{end} \]
\[ \text{for all } a \in V. \]

\[ C_{\cdot,T_h}(\{M_0\}_{\equiv}, \{M_1\}_{\equiv}) = \text{let } (Q_0, V, T_0, E_0, S_0, F_0) = M_0 \]
\[ \quad (Q_1, V, T_1, E_1, S_1, F_1) = M_1 \]
\[ \quad \text{in} \]
\[ \quad \quad E' = E_0 \cup E_1 \cup (F_0 \times S_1) \]
\[ \quad \quad \text{in} \]
\[ \quad \quad \quad [(Q_0 \cup Q_1, V, T_0 \cup T_1, E', S_0, F_1)]_{\equiv} \]
\[ \quad \quad \text{end} \]

\[ C_{\cup,T_h}(\{M_0\}_{\equiv}, \{M_1\}_{\equiv}) = \text{let } (Q_0, V, T_0, E_0, S_0, F_0) = M_0 \]
\[ \quad (Q_1, V, T_1, E_1, S_1, F_1) = M_1 \]
\[ \quad q_0, q_1 \text{ be new states} \]
\[ \quad \text{in} \]
\[ \quad \quad Q' = Q_0 \cup Q_1 \cup \{q_0, q_1\} \]
\[ \quad \quad E' = E_0 \cup E_1 \cup (\{q_0\} \times (S_0 \cup S_1)) \]
\[ \quad \quad \cup ((F_0 \cup F_1) \times \{q_1\}) \]
\[ \quad \quad \text{in} \]
\[ \quad \quad \quad [(Q', V, T_0 \cup T_1, E', \{q_0\}, \{q_1\})]_{\equiv} \]
\[ \quad \quad \text{end} \]
\[ C_{\ast T_h}(\{M\}) = \text{let } (Q, V, T, E, S, F) = M \ \ \\
\text{q_0, q_1 \text{ be new states}} \ \\
\text{in} \ \\
\text{let } Q' = Q \cup \{q_0, q_1\} \ \\
E' = E \cup \{(q_0) \times S\} \cup (F \times S) \cup (F \times \{q_1\}) \cup \{(q_0, q_1)\} \ \\
\text{in} \ \\
[(Q', V, T, E', \{q_0\}, \{q_1\})]_{\ast} \ \\
\text{end} \ \\
\text{end} \]

\[ C_{+ T_h}(\{M\}) = \text{let } (Q, V, T, E, S, F) = M \ \ \\
\text{q_0, q_1 \text{ be new states}} \ \\
\text{in} \ \\
\text{let } Q' = Q \cup \{q_0, q_1\} \ \\
E' = E \cup \{(q_0) \times S\} \cup (F \times \{q_1\}) \ \\
\text{in} \ \\
[(Q', V, T, E', \{q_0\}, \{q_1\})]_{\ast} \ \\
\text{end} \ \\
\text{end} \]

\[ C_{\tau T_h}(\{M\}) = \text{let } (Q, V, T, E, S, F) = M \ \ \\
\text{q_0, q_1 \text{ be new states}} \ \\
\text{in} \ \\
\text{let } Q' = Q \cup \{q_0, q_1\} \ \\
E' = E \cup \{(q_0) \times S\} \cup (F \times \{q_1\}) \cup \{(q_0, q_1)\} \ \\
\text{in} \ \\
[(Q', V, T, E', \{q_0\}, \{q_1\})]_{\ast} \ \\
\text{end} \ \\
\text{end} \]
4.1 Thompson's construction

These operators are symmetrical (see Definition A.22 for a definition of symmetrical operators and functions). Furthermore, they do not depend upon the choice of representative of the equivalence classes (under $\cong$). An automaton in Thompson's $\Sigma$-algebra (here we speak of a representative $FA_i$ instead of the isomorphism class) has the following properties:

- It has a single start state with no in-transitions.
- It has a single final state with no out-transitions.
- Every state has either a single in-transition on a symbol (in $V$), or at most two $\varepsilon$-in-transitions.
- Every state has either a single out-transition on a symbol (in $V$), or at most two $\varepsilon$-out-transitions.

These properties are symmetrical because the operators are symmetrical. Hopcroft and Ullman have shown [HU79] that in practice these properties facilitate the quick simulation of $M$. For the remainder of this paper we will not duplicate properties such as these, but rather state whether the operator is symmetrical.

Remark 4.2: In the literature, these operators are usually presented as having arguments and results of type $FA$ instead of $[FA]_\Sigma$. Such a presentation is given in terms of particular representatives, and ignores the nondeterminism in choosing new states.

Construction 4.3 (Thompson): Thompson's construction is the (unique) homomorphism $Th$ from $RE$ to Thompson's $\Sigma$-algebra of $FA$'s.

Example 4.4 (Thompson's construction): We construct a particular representative$^5$ of

$$Th((a \cup \varepsilon)b^*) = Th(a \cup \varepsilon), Th(b^*)$$
$$= Th(\cup, Th(a), Th(\varepsilon), Th(b))$$
$$= Th(\cup, Th(a), Th(\varepsilon), Th(b))$$

(The regular expression is taken from Example 3.15.) The representative is shown in Figure 2.

In the next two subsections, we consider two algorithms that construct an $FA$ (from a regular expression) based on the top-down syntactic structure of the regular expression. In these two constructions, we use regular expressions as syntactic objects denoting regular languages.

The first construction is a top-down version of Thompson's construction. The second one is also top-down, but constructs a so-called $\varepsilon$-lookahead automaton. Such an automaton can be efficiently simulated or it can be converted to an efficient program, accepting the language of the automaton.

$^5$Obviously, constructing the entire equivalence class of isomorphic $FA$'s is not possible.
4.1.1 A top-down version of Thompson's construction

The top-down version of Thompson's construction is a practical implementation of homomorphism $Th$. It is a function of three parameters: a start state $s$, a regular expression $E$, and a final state $f$. It produces an FA, with start state $s$ and final state $f$, accepting the language $L_{RE}(E)$.

Construction 4.5 (Top-down Thompson's): We assume a universe of available states $U$, to define function

$$td : U \times RE \times U \rightarrow FA$$

The function is defined recursively on the structure of regular expressions:

$$td(s, \varepsilon, f) = (\{s, f\}, V, \emptyset, \{(s, f)\}, \{s\}, \{f\})$$
$$td(s, \emptyset, f) = (\{s, f\}, V, \emptyset, \emptyset, \{s\}, \{f\})$$
$$td(s, a, f) = (\{s, f\}, V, \{(s, a, f)\}, \emptyset, \{s\}, \{f\})$$

(for all $a \in V$)

$$td(s, E_0 \cdot E_1, f) = \begin{cases} \text{let } p, q \text{ be new states} \\ \text{let } (Q_0, V, T_0, E_0, \{p\}, \{q\}) = td(s, E_0, p) \\ (Q_1, V, T_1, E_1, \{f\}) = td(q, E_1, f) \\ \text{in } (Q_0 \cup Q_1, V, T_0 \cup T_1, E_0 \cup E_1 \cup \{(p, q)\}, \{s\}, \{f\}) \end{cases}$$

$$td(s, E_0 \cup E_1, f) = \begin{cases} \text{let } p, q, r, t \text{ be new states} \\ \text{let } (Q_0, V, T_0, E_0, \{p\}, \{q\}) = td(p, E_0, q) \\ (Q_1, V, T_1, E_1, \{r\}, \{t\}) = td(r, E_1, t) \\ \text{in } (Q_0 \cup Q_1 \cup \{s, f\}, V, T_0 \cup T_1, E_0 \cup E_1 \\ \cup \{(p, q)\} \cup (q, t) \times \{f\}), \{s\}, \{f\}) \end{cases}$$

$$td(s, E^*, f) = \begin{cases} \text{let } p, q \text{ be new states} \\ \text{let } (Q, V, T, E, \{p\}, \{q\}) = td(p, E, q) \\ \text{in } (Q \cup \{s, f\}, V, T, E \cup \{(s, p), (q, p), (q, f), (s, f)\}, \{s\}, \{f\}) \end{cases}$$
4.1 Thompson's construction

\[ \text{td}(s, E^+, f) = \text{let} \quad p, q \text{ be new states} \]
\[ \quad \text{in} \]
\[ \quad \text{let} \quad (Q, V, T, E, \{p\}, \{q\}) = \text{td}(p, E, q) \]
\[ \quad \text{in} \]
\[ \quad (Q \cup \{s, f\}, V, T, E \cup \{(s, p), (q, p), (q, f)\}, \{s\}, \{f\}) \]
\[ \text{end} \]
\[ \text{end} \]

\[ \text{td}(s, E^c, f) = \text{let} \quad p, q \text{ be new states} \]
\[ \quad \text{in} \]
\[ \quad \text{let} \quad (Q, V, T, E, \{p\}, \{q\}) = \text{td}(p, E, q) \]
\[ \quad \text{in} \]
\[ \quad (Q \cup \{s, f\}, V, T, E \cup \{(s, p), (q, f), (s, f)\}, \{s\}, \{f\}) \]
\[ \text{end} \]
\[ \text{end} \]

Function \( \text{td} \) satisfies the property that, for all \( E \in RE \):

\[ \text{Th}(E) = \text{let} \quad s, f \text{ be new states} \]
\[ \quad \text{in} \]
\[ \quad [\text{td}(s, E, f) \varepsilon] \]
\[ \text{end} \]

\( \square \)

The advantage of function \( \text{td} \) over homomorphism \( \text{Th} \) (Construction 4.3) is one of implementation. In Thompson’s construction, the subparts of the final \( FA \) are constructed in isolation; when two subparts are combined some states may have to be renamed to ensure that the subparts have disjoint state sets. In the top-down construction, more global knowledge is available concerning the final \( FA \) and this type of problem is avoided. (In practice, function \( \text{td} \) would make use of a global variable: the set of remaining available states.)

We do not prove the correctness of construction \( \text{td} \) in this paper.

4.1.2 Constructing \( \epsilon \)-lookahead automata

In this subsection, we extend the top-down Thompson construction (function \( \text{td} \)) to construct \( \epsilon \)-lookahead finite automata (LAFA). In an LAFA, every \( \epsilon \)-transition is qualified by a symbol of \( V \) (known as the lookahead symbol). When simulating an LAFA, an \( \epsilon \)-transition can be taken if the next symbol of the input string matches the lookahead symbol of the \( \epsilon \)-transition. Naturally, for any given state, it is desirable that there only be one \( \epsilon \)-transition from the state on any given symbol. The following definitions formalize this.

**Definition 4.6 (\( \epsilon \)-lookahead automata):** An \( \epsilon \)-lookahead finite automaton (LAFA) is a 6-tuple \( (Q, V, T, E, S, F) \) which is a normal \( FA \) with one exception:

- \( \epsilon \)-transition relation is now \( E \in P(Q \times V \times Q) \) instead of \( E \in P(Q \times Q) \).

\( \square \)

**Remark 4.7:** A more commonly presented definition of LAFA’s involves both \( \epsilon \)-lookahead and normal \( \epsilon \)-transitions (also called don’t-care transitions). Since we have combined the two, we implement normal \( \epsilon \)-transitions as lookahead transitions, where the lookahead set is \( V \) (the entire alphabet). \( \square \)

**Remark 4.8:** Naturally, we extend such functions as \( \mathcal{L}, \mathcal{L}^c, \) and \( \mathcal{L}_{FA} \) to use the definition of a LAFA. As a result, the language accepted by an LAFA is in accordance with the intuitive interpretation of an LAFA. \( \square \)
Remark 4.9: In order to present the following definition, we require the definition of function \( \text{First} \in \mathcal{R}E \rightarrow \mathcal{P}(V) \). Function \( \text{First} \) is defined in Definition 4.60. Informally, \( \text{First}(E) \) is the set of all alphabet symbols that can occur as the first symbol of a word in \( \mathcal{L}_{\mathcal{R}E}(E) \). □

Definition 4.10 (Lookahead function): In order to make the definition of the LAFA construction readable, we introduce function \( \text{look} \in \mathcal{R}E \times \mathcal{P}(V) \rightarrow \mathcal{P}(V) \), defined as:

\[
\text{look}(E, L) = \text{First}(E) \cup \text{if } (\text{Null}(E)) \text{ then } L \text{ else } \emptyset \text{ fi}
\]

Argument \( L \) is called the set of follow symbols. □

We now define an LAFA construction, based on the top-down version of Thompson's construction.

Construction 4.11 (\( \epsilon \)-lookahead finite automaton): We define function \( K \) which takes four parameters: a start state \( s \), a regular expression, final state \( f \), and a lookahead set \( L \in \mathcal{P}(V) \). As with the top-down version of Thompson's construction, we assume a universe of states \( U \).

Function \( K \in U \times \mathcal{R}E \times U \times \mathcal{P}(V) \rightarrow \text{LAFA} \) is defined recursively on the structure of regular expressions:

\[
\begin{align*}
K(s, \epsilon, f, L) &= ((s, f), V, \emptyset, \{s\} \times L \times \{f\}, \{s\}, \{f\}) \\
K(s, \emptyset, f, L) &= ((s, f), V, \emptyset, \{s\}, \{f\}) \\
K(s, a, f, L) &= ((s, f), V, \{s, a, f\}, \emptyset, \{s\}, \{f\}) \\
\quad \text{(for all } a \in V) \\
K(s, E_1 \cdot E_2, f, L) &= \text{let } p, q \text{ be new states} \\
&= \text{let } (Q_0, V, T_0, E_0, \{p\}, \{q\}) = K(s, E_0, p, \text{look}(E_1, L)) \\
&= \text{let } (Q_1, V, T_1, E_1, \{q\}, \{f\}) = K(q, E_1, f, L) \\
&= \text{let } (Q_0 \cup Q_1, V, T_0 \cup T_1, E_0 \cup E_1 \\
&= \quad \text{U } \{p\} \times \text{look}(E_1, L) \times \{q\}, \{s\}, \{f\}) \\
\quad \text{end} \\
K(s, E_0 \cup E_1, f, L) &= \text{let } p, q, r, t \text{ be new states} \\
&= \text{let } (Q_0, V, T_0, E_0, \{p\}, \{q\}) = K(p, E_0, q, L) \\
&= \text{let } (Q_1, V, T_1, E_1, \{r\}, \{t\}) = K(r, E_1, t, L) \\
&= \text{let } (Q_0 \cup Q_1 \cup \{s, f\}, V, T_0 \cup T_1, E_0 \cup E_1 \\
&= \quad \text{U } \{s\} \times \text{look}(E_0, L) \times \{p\} \\
&= \quad \text{U } \{q, t\} \times L \times \{f\}, \{s\}, \{f\}) \\
\quad \text{end} \\
K(s, E^*, f, L) &= \text{let } p, q \text{ be new states} \\
&= \text{let } (Q, V, T, E, \{p\}, \{q\}) = K(p, E, q, L \cup \text{First}(E)) \\
&= \text{let } (Q \cup \{s, f\}, V, T, E \cup \{s, q\} \times \text{First}(E) \times \{p\}) \\
&= \quad \text{U } \{s, q\} \times L \times \{f\}, \{s\}, \{f\}) \\
\quad \text{end} \\
\end{align*}
\]
\[ K(s, E^+, f, L) = \begin{align*}
& \text{let } p, q \text{ be new states} \\
& \text{in} \\
& \quad (Q, V, T, E, \{p\}, \{q\}) = K(p, E, q, L \cup \text{First}(E)) \\
& \quad (Q \cup \{s\}, V, T \cup ((s, q) \times \text{First}(E) \times \{p\}) \\
& \quad \cup ([q] \times L \times \{f\}, \{s\}, \{f\}) \\
& \end{align*} \]

\[ K(s, E', f, L) = \begin{align*}
& \text{let } p, q \text{ be new states} \\
& \text{in} \\
& \quad (Q, V, T, E, \{p\}, \{q\}) = K(p, E, q, L) \\
& \quad (Q \cup \{s\}, V, T \cup ((s) \times \text{First}(E) \times \{p\}) \\
& \quad \cup ([s, q] \times L \times \{f\}, \{s\}, \{f\}) \\
& \end{align*} \]

Remark 4.12: Since we make use of a single symbol of lookahead, we assume that the input string always has an end-marker \$ concatenated on its right. We assume that \$ \in V and that \$ does not appear elsewhere in the regular expression. This means that, for \( E \in RE: \)

\[ \begin{align*}
& \text{let } s, f \text{ be new states} \\
& \text{in} \\
& \quad K(s, E, f, \{\$\}) \\
& \end{align*} \]

is a \( LAFA \) accepting \( L_{RE}(E). \) □

Definition 4.13 (Deterministic \( LAFA'\)'s): A \( LAFA \) is deterministic if and only if it has at most one start state and no state has more than one out-transition (either an ε-lookahead or a normal transition) on any given alphabet symbol. □

We present some determinism conditions that ensure that Construction 4.11 produces deterministic \( LAFA'\)'s.

Definition 4.14 (Determinism conditions): In order for function \( K \) to produce a deterministic \( LAFA \) we impose the following requirements for particular cases of \( K: \)

- For \( K(s, E_0 \cup E_1, f, L) \) we require that \( \text{look}(E_0, L) \cap \text{look}(E_1, L) = \emptyset. \)
- For \( K(s, E^*, f, L), K(s, E^+, f, L), \) and \( K(s, E', f, L) \) we require that \( \text{First}(E) \cap L = \emptyset. \)

□

Remark 4.15: The lookahead transitions in \( LAFA'\)'s make them are more efficient to simulate than an equivalent \( FA \) constructed with Thompson's construction. Simulation of a deterministic \( LAFA \) is as efficient as the simulation of a \( DFA. \) □

Example 4.16 (\( LAFA \)): Given new states \( s, f, \) we construct the deterministic \( LAFA \) \( K(s, (a \cup \epsilon)b^*, f, \{\$\}). \) The ε-lookahead transitions are labeled with both ε and the lookahead symbols. The state graph is given in Figure 3. □
Construction 4.17 (Creating a program from a LAFA): A deterministic LAFA can also be converted into a program which is a hard-coded simulation of the LAFA. We now describe a mapping \( N \in \text{RE} \times \mathcal{P}(V) \rightarrow \text{GCL} \), where \( \text{GCL} \) denotes the set of all guarded commands programs [Dijk76]. This construction is based upon the LAFA construction. The created programs are correct when the determinism conditions of Definition 4.14 hold. In the generated program, we assume that variable \( w \in V^* \) is the input string (with an end-marker $ concatenated on its right), and that \( \text{hd}(w) \) refers to the first symbol of \( w \) and \( \text{tl}(w) \) refers to the remainder of \( w \). We annotate the program fragments (in the definition of \( N \)) with the state names (in braces) in the corresponding definition of Construction 4.11. (The semantics of the guarded commands specify that if none of the guards in an if-fi statement are true, the statement is equivalent to abort.)

\[
\begin{align*}
N(&e, L) = \{s\} \\
&\quad \text{if } \text{hd}(w) \in L \rightarrow \text{skip} \\
&\quad \text{fi} \\
N(&\emptyset, L) = \{s\} \text{ abort}\{f\} \\
N(&a, L) = \{s\} &\quad \text{(for all } a \in V) \\
&\quad \text{if } \text{hd}(w) = a \rightarrow w := \text{tl}(w) \\
&\quad \text{fi} \\
N(&E_0 \cdot E_1, L) = \{s\} N(E_0, \text{look}(E_1, L))\{p\}; \\
&\quad \{q\} N(E_1, L)\{f\} \\
N(&E_0 \cup E_1, L) = \{s\} \\
&\quad \text{if } \text{hd}(w) \in \text{look}(E_0, L) \rightarrow \{p\} N(E_0, L)\{q\} \\
&\quad \| \text{hd}(w) \in \text{look}(E_1, L) \rightarrow \{r\} N(E_1, L)\{t\} \\
&\quad \text{fi} \\
N(&E^*, L) = \{s\} \\
&\quad \text{do } \text{hd}(w) \in \text{First}(E) \rightarrow \{p\} N(E, \text{First}(E) \cup L)\{q\} \\
&\quad \text{od} \\
N(&E^+, L) = \{s\} \\
&\quad \text{repeat } \{p\} N(E, \text{First}(E) \cup L)\{q\} \\
&\quad \text{until } \text{hd}(w) \notin \text{First}(E) \\
&\quad \{f\} \\
N(&E^?, L) = \{s\} \\
&\quad \text{if } \text{hd}(w) \in \text{First}(E) \rightarrow \{p\} N(E, L)\{q\} \\
&\quad \| \text{hd}(w) \in L \rightarrow \text{skip} \\
&\quad \text{fi} \\
&\quad \{f\}
\end{align*}
\]

As with Construction 4.11 we concatenate an end-marker $ on the right of the input string \( w \).

The entire program of the acceptor (for \( E \in \text{RE} \)) is:
4.2 Towards the Berry-Sethi construction

Termination of the program is equivalent to \( w \in \mathcal{L}_{RE}(E) \). □

Example 4.18 (Programs from LAFA's): We construct the program corresponding to \((a \cup \epsilon)b^*\).

\[
\{ w \in V^* \{ \$ \} \} \\
N(E, \{ \$ \}); \\
\text{if } w = \$ \rightarrow \text{skip} \\
\text{fi} \\
\{ w \in \mathcal{L}_{RE}(E) \}
\]

\[
\{ w \in V^* \{ \$ \} \} \\
\text{if } \text{hd}(w) \in \{ a \} \rightarrow \\
\quad \text{if } \text{hd}(w) = a \rightarrow w := \text{tl}(w) \\
\quad \text{fi} \\
\text{fi} \\
\text{if } \text{hd}(w) \in \{ b, \$ \} \rightarrow \\
\quad \text{if } \text{hd}(w) \in \{ b, \$ \} \rightarrow \text{skip} \\
\quad \text{fi} \\
\text{fi} \\
\text{do} \\
\quad \text{if } \text{hd}(w) \in \{ b \} \rightarrow \\
\quad \quad \text{if } \text{hd}(w) = b \rightarrow w := \text{tl}(w) \\
\quad \quad \text{fi} \\
\text{od} \\
\text{if } w = \$ \rightarrow \text{skip} \\
\text{fi} \\
\{ w \in \{ a, \epsilon \} \{ b \}^* \}
\]

□

4.2 Towards the Berry-Sethi construction

We now consider \( \Sigma \)-algebras of \( \epsilon \)-free FA's. One such \( \Sigma \)-algebra can be given with symmetrical operators.

Definition 4.19 (Symmetrical \( \epsilon \)-free \( \Sigma \)-algebra operators): The carrier set is \([FA]_{\epsilon}\). The operator requirement is (as with Thompson's \( \Sigma \)-algebra):

- For the binary operators, the representatives of the arguments must have disjoint state sets.

The symmetrical \( \epsilon \)-free preserving operators of the \( \Sigma \)-algebra are defined using Thompson's \( \Sigma \)-algebra operators and symmetrical function \( \text{remove}_{\epsilon, \text{sym}} \) (which is extended to \([FA]_{\epsilon} \longrightarrow [FA]_{\epsilon}\)):

\[
\begin{align*}
C_{e, \text{sym}} &= \text{remove}_{\epsilon, \text{sym}} \circ C_e \circ T_h \\
C_{\$\epsilon, \text{sym}} &= \text{remove}_{\epsilon, \text{sym}} \circ C_{\$ \epsilon} \circ T_h \\
C_{a, \text{sym}} &= \text{remove}_{\epsilon, \text{sym}} \circ C_{a, \epsilon} \circ T_h \\
C_{\epsilon, \text{sym}} &= \text{remove}_{\epsilon, \text{sym}} \circ C_{\epsilon, \epsilon} \circ T_h \\
C_{+, \text{sym}} &= \text{remove}_{\epsilon, \text{sym}} \circ C_{+, \epsilon} \circ T_h \\
C_{\text{?,sym}} &= \text{remove}_{\epsilon, \text{sym}} \circ C_{\text{?,} \epsilon} \circ T_h
\end{align*}
\]

(for all \( a \in V \))

These operators are symmetrical since they are compositions of symmetrical operators (see Proposition A.23). An FA in this \( \Sigma \)-algebra has the property that it is \( \epsilon \)-free. □
These operators are cumbersome to present fully. Furthermore, they are not particularly useful in practice. For this reason, we now consider asymmetrically defined \(e\)-free preserving operators.

The first asymmetrical \(e\)-free preserving \(\Sigma\)-algebra operators that we consider are the left-biased ones. The image of a \(RE\) in this \(\Sigma\)-algebra is easier to compute than its image in the \(\Sigma\)-algebra given in Definition 4.19).

**Definition 4.20 (\(\Sigma\)-algebra of left-biased \(e\)-free operators):** The carrier set is \([FA]\). The operator requirements are:

- For binary operators, the representatives of the arguments must have disjoint state sets.
- The following is required of the representatives of each argument:
  - it is \(e\)-free,
  - it has a single start state, and
  - the single start state has no in-transitions.

A proof of the correctness of these operators is outlined in Theorem B.2. As in Thompson’s \(\Sigma\)-algebra, each operator is presented here with a graphic representation of the operator\(^6\). Parts of the operator definitions are intentionally clumsy or verbose. This is done to facilitate the derivation of a \(\Sigma\)-algebra of reduced \(FA\)'s (in Definition 4.29). The operators are:

\[
C_{a, LBFA} = \text{let } q_0, q_1 \text{ be new states } \\
\text{in } \\
\begin{array}{ll}
\{\{q_0, q_1\}, V, \emptyset, \emptyset, \emptyset, \{q_0, q_1\}\} \\
\end{array}
\]

\[
C_{\emptyset, LBFA} = \text{let } q_0 \text{ be a new state } \\
\text{in } \\
\begin{array}{ll}
\{\{q_0\}, V, \emptyset, \emptyset, \emptyset, \{q_0\}\} \\
\end{array}
\]

\[
C_a, LBFA = \text{let } q_0 \text{ be a new state } \\
\text{in } \\
\begin{array}{ll}
\{\{q_0\}, V, \emptyset, \emptyset, \emptyset, \{q_0\}\} \\
\end{array}
\]

for all \(a \in V\).

\(^6\)The graphic representations of the operators depict only the simplest cases of each operator. Thick arrowed lines are intended to depict multiple transitions, while dotted arrowed lines are transitions that are removed from the constructed \(FA\). In the case of the non-constant operators, the start states (of the arguments) is struck out indicating that it is removed.
4.2 Towards the Berry-Sethi construction

\( C_{\text{LBFA}}([M_0]_\{\varepsilon\}, [M_1]_\{\varepsilon\}) = \) let

\( (Q_0, V, T_0, \emptyset, \{s_0\}, F_0) = M_0 \)
\( (Q_1, V, T_1, \emptyset, \{s_1\}, F_1) = M_1 \)
\( N_0 = \varepsilon \in \mathcal{L}_{\text{FA}}(M_0) \)
\( N_1 = \varepsilon \in \mathcal{L}_{\text{FA}}(M_1) \)
\( N = \varepsilon \in (\mathcal{L}_{\text{FA}}(M_0) \cup \mathcal{L}_{\text{FA}}(M_1)) \)
\( q_0 \) be a new state

in

let

\( Q' = Q_0 \setminus \{s_0\} \cup Q_1 \setminus \{s_1\} \cup \{q_0\} \)
\( T' = T_0 \cup T_1 \cup (F_0 \times T_1(s_1)) \)
\( \cup \{q_0\} \times (T_0(s_0)) \)
\( \cup \text{if} (N_0) \text{ then } T_1(s_1) \text{ else } \emptyset \fi \)
\( F' = F_1 \cup \text{if} (N_1) \text{ then } F_0 \text{ else } \emptyset \fi \)
\( \cup \text{if} (N) \text{ then } \{q_0\} \text{ else } \emptyset \fi \)
in

\[ [(Q', V, T' \cap (Q' \times V \times Q'), \emptyset, \{q_0\}, F' \cap Q')]_\{\varepsilon\} \quad \]
end


\( C_{\text{LBFA}}([M_0]_\{\varepsilon\}, [M_1]_\{\varepsilon\}) = \) let

\( (Q_0, V, T_0, \emptyset, \{s_0\}, F_0) = M_0 \)
\( (Q_1, V, T_1, \emptyset, \{s_1\}, F_1) = M_1 \)
\( N = \varepsilon \in (\mathcal{L}_{\text{FA}}(M_0) \cup \mathcal{L}_{\text{FA}}(M_1)) \)
\( q_0 \) be a new state

in

let

\( Q' = Q_0 \setminus \{s_0\} \cup Q_1 \setminus \{s_1\} \cup \{q_0\} \)
\( T'' = T_0 \cup T_1 \cup (\{q_0\} \times (T_0(s_0) \cup T_1(s_1))) \)
\( F'' = F_0 \cup F_1 \cup \text{if} (N) \text{ then } \{q_0\} \text{ else } \emptyset \fi \)
in

\[ [(Q', V, T'' \cap (Q' \times V \times Q'), \emptyset, \{q_0\}, F'' \cap Q')]_\{\varepsilon\} \quad \]
end

\[ \rightarrow \quad M_0 \quad M_1 \]
\[ C_{\star, LBFA}([M]) = \begin{array}{c}
(Q, V, T, \emptyset, \{s\}, F) = M \\
N = \epsilon \in \mathcal{L}_{FA}(M)^* \\
q_0 \text{ be a new state}
\end{array}
\]

\[ \text{(see Remark 4.21)} \]

\[ \begin{array}{c}
in \\
(\text{let } Q' = Q \setminus \{s\} \cup \{q_0\} \\
T' = T \cup (F \cup \{q_0\}) \times T(s) \\
F' = F \cup \{q_0\} \text{ if (N) then } \{q_0\} \text{ else } \emptyset \\
\text{in} \\
[(Q', V, T' \cap (Q' \times V \times Q'), \\
\emptyset, \{q_0\}, F' \cap Q')]_a \\
\text{end} \\
end
\]

\[ C_{+, LBFA}([M]) = \begin{array}{c}
(Q, V, T, \emptyset, \{s\}, F) = M \\
N = \epsilon \in \mathcal{L}_{FA}(M)^+ \\
q_0 \text{ be a new state}
\end{array}
\]

\[ \begin{array}{c}
in \\
(\text{let } Q' = Q \setminus \{s\} \cup \{q_0\} \\
T' = T \cup (F \cup \{q_0\}) \times T(s) \\
F' = F \cup \text{if (N) then } \{q_0\} \text{ else } \emptyset \\
\text{in} \\
[(Q', V, T' \cap (Q' \times V \times Q'), \\
\emptyset, \{q_0\}, F' \cap Q')]_a \\
\text{end} \\
end
\]
4.2 Towards the Berry-Sethi construction

\[ C_{\text{LBFA}}([M]) = \begin{array}{ll}
\text{let} & (Q, V, T, \emptyset, \{s\}, F) = M \\
& N = \epsilon \in \mathcal{L}_{\text{FA}}(M) \quad \text{(see Remark 4.21)} \\
& q_0 \text{ be a new state} \\
\text{in} & \begin{array}{ll}
Q' = Q \setminus \{s\} \cup \{q_0\} \\
T' = T \cup (\{q_0\} \times T(s)) \\
F' = F \cup \text{if } (N) \text{ then } \{q_0\} \text{ else } \emptyset \text{ fi}
\end{array} \\
\text{end} & \end{array} \]

\[ ((Q', V, T', \emptyset, \{q_0\}, F', \emptyset, \{q_0\}, F' \cap Q')) \]

The choice of representatives in these operators is irrelevant. For construction purposes, we note that \( \epsilon \in \mathcal{L}_{\text{FA}}(M) \equiv s \in F \).

Let \( \text{LBFA} \) (where \( \text{LBFA} \subset FA \)) denote the set of all finite automata that are images\(^7\) in this \( \Sigma \)-algebra of some \( E \in RE \). (That is, \( \text{LBFA} \) is the smallest set that contains the \( \text{LBFA} \) constants and is closed under the \( \text{LBFA} \) operators.) An \( \text{LBFA} \) has the following properties:

- It is \( \epsilon \)-free.
- It has a single start state.
- The single start state has no in-transitions.
- All in-transitions to a state are on the same symbol (in \( V \)). This can be seen by considering the constants \( C_{a,\text{LBFA}} \) (for all \( a \in V \)), which are the only operators introducing new transitions on an alphabet symbol.

Only the constants are symmetrical. □

Remark 4.21: Parts of the operator definitions of Definition 4.20 are intentionally clumsy; they are presented this way to facilitate the derivation of a \( \Sigma \)-algebra of reduced FA’s (Definition 4.29). □

Construction 4.22 (Left-biased finite automata): Define construction \( \text{lbfa} \in RE \rightarrow [\text{LBFA}]_\Sigma \) to be the unique homomorphism from \( RE \)'s to \([\text{LBFA}]_\Sigma \). □

Example 4.23 (\( \Sigma \)-algebra of \( \text{LBFA} \)'s): We construct a representative of the isomorphism class \( \text{lbfa}(a \cup \epsilon) \) (the regular expression is from Example 3.15). The representative is shown in Figure 4. □

Computing within the \( \Sigma \)-algebra of \( \text{LBFA} \)'s is inefficient. Each operator defined above does much redundant work. For example, the start states of the arguments to the operators are always removed, with only the out-transitions from the argument’s start state being used. Additionally, the if-fi structures within the final states definition are of the same structure in each operator. We wish to introduce an encoding of \( \text{LBFA} \)'s that will allow us to find cheap constructions that are equivalent to \( \text{lbfa} \). We now describe such an encoding.

A method of encoding an \( \text{LBFA} \) \((Q, V, T, \emptyset, \{s\}, F)\) is:

\(^7\)The images are really elements of \([\text{LBFA}]_\Sigma \). We consider a particular representative of the image.
The $\epsilon$-transitions are not needed (since LBFA's are $\epsilon$-free).

State $s$ has no in-transitions; only $T(s)$ (the out-transitions from the start state) and $s \in F$ are needed.

All in-transitions to a state are on the same symbol (in $V$). Therefore, a state-to-symbol map can be used, and the symbol components of $T$ and $T(s)$ can be removed.

In the following subsection, we introduce reduced FA's as an encoding of LBFA's.

### 4.2.1 Reduced FA's

**Definition 4.24 (RFA):** A reduced FA (RFA) is a 7-tuple $(Q, V, \text{follow}, \text{first}, \text{last}, \text{null}, Qmap)$ where

- $Q$ is a finite set of states,
- $V$ is an alphabet,
- $\text{follow} \subseteq \mathcal{P}(Q \times Q)$ is a follow relation (replacing the transition relation),
- $\text{first} \subseteq Q$ is a set of initial states (replacing $T(s)$ in an LBFA),
- $\text{last} \subseteq Q$ is a set of final states,
- $\text{null} \in \{\text{true, false}\}$ is a Boolean value (encoding $s \in F$ in an LBFA), and
- $Qmap \in \mathcal{P}(Q \times V)$ maps each state to exactly one symbol (it is also viewed as $Qmap \in Q \rightarrow V$, and its inverse as $Qmap^{-1} \in V \rightarrow \mathcal{P}(Q)$).

\[\square\]

**Definition 4.25 (Isomorphism of RFA's):** We extend isomorphism ($\cong$) to RFA's. \[\square\]

**Definition 4.26 (Reversal of RFA’s):** Reversal of RFA's is given by postfix (superscript) function $R \in \text{RFA} \rightarrow \text{RFA}$ defined as:

\[(Q, V, \text{follow}, \text{first}, \text{last}, \text{null}, Qmap)^R = (Q, V, \text{follow}^R, \text{last}, \text{first}, \text{null}, Qmap)\]

\[\square\]

**Definition 4.27 (Extending reversal to $[RFA]_\Sigma$):** We extend reversal to $[RFA]_\Sigma \rightarrow [RFA]_\Sigma$ as $([M]_\Sigma)^R = [M^R]_\Sigma$. \[\square\]

We can now give isomorphisms between $[LBFA]_\Sigma$ and $[RFA]_\Sigma$. These isomorphisms will be used to present a $\Sigma$-algebra of RFA's.
4.2 Towards the Berry-Sethi construction

Definition 4.28 (An isomorphism between \([\text{LBFA}]_{\Sigma}\) and \([\text{RFA}]_{\Sigma}\)): We define isomorphism

\[\text{encode} \in \left[\text{LBFA}\right]_{\Sigma} \rightarrow \left[\text{RFA}\right]_{\Sigma}\]

\[
\text{encode}([(Q, V, T, \emptyset, \{s\}, F)]_{\Sigma}) = \begin{cases} 
Q' = Q \setminus \{s\} \\
([Q', V, \pi_2(T) \cap (Q' \times Q'), \pi_2(T(s)), F \cap Q', s) \in F, (\pi_1(T))^R)]_{\Sigma}
\end{cases}
\]

and its inverse \(\text{decode} \in \left[\text{RFA}\right]_{\Sigma} \rightarrow \left[\text{LBFA}\right]_{\Sigma}\) as

\[
\text{decode}([M]_{\Sigma}) = \begin{cases} 
(Q, V, \text{follow}, \text{first}, \text{last}, \text{null}, Qmap) = M \\
s \text{be a new state}
\end{cases}
\]

\[
\begin{align*}
&\text{in} \quad T = \{(q_0, Qmap(q_1), q_1) : (q_0, q_1) \in \text{follow}\} \\
&\text{and } T' = \{(s, Qmap(q), q) : q \in \text{first}\} \\
&F = \text{last} \cup \text{if } (\text{null}) \text{ then } \{s\} \text{ else } \emptyset \text{ fi}
\end{align*}
\]

\[
\begin{cases} 
\text{in} \quad [(Q \cup \{s\}, V, T \cup T', \emptyset, \{s\}, F)]_{\Sigma}
\end{cases}
\]

It is easy to verify that both of these functions are isomorphisms, and that \(\text{decode}\) is the inverse of \(\text{encode}\). □

Given function \(\text{encode}\) and \(\text{decode}\), we would like to obtain a \(\Sigma\)-algebra with \([\text{RFA}]_{\Sigma}\) as carrier (and a corresponding unique homomorphism \(\text{rfa} \in \text{RE} \rightarrow \left[\text{RFA}\right]_{\Sigma}\)) such that the following diagram commutes:

\[
\begin{array}{ccc}
\text{RE} & \xrightarrow{\text{Bf}} & \text{[LBFA]}_{\Sigma} \\
\downarrow{\text{rfa}} & & \downarrow{\text{decode}} \\
\text{[RFA]}_{\Sigma} & & \\
\end{array}
\]

We can now define a \(\Sigma\)-algebra of \(\text{RFA}\)'s; it will be cheaper to compute the \(\text{RFA}\) image of a regular expression and map the \(\text{RFA}\) to an \(\text{LBFA}\), than to compute the \(\text{LBFA}\) directly. The operators of the \(\Sigma\)-algebra of \(\text{RFA}\)'s are defined using the \(\text{LBFA}\) operators and the isomorphisms \(\text{encode}\) and \(\text{decode}\).

Definition 4.29 (\(\Sigma\)-algebra of \(\text{RFA}\)'s): The carrier is \([\text{RFA}]_{\Sigma}\). Given the operator requirement in the \(\Sigma\)-algebra of \(\text{LBFA}\)'s, the operator requirement in this \(\Sigma\)-algebra is:

- For binary operators, the argument representatives must have disjoint state sets.

The operators of the \(\Sigma\)-algebra of \(\text{RFA}\)'s are defined in terms of the operators of \(\text{LBFA}\)'s:

\[
\begin{align*}
C_a,\text{RFA} &= \text{encode}(C_a,\text{LBFA}) \\
C_0,\text{RFA} &= \text{encode}(C_0,\text{LBFA}) \\
C_\times,\text{RFA} &= \text{encode}(C_\times,\text{LBFA}) \\
C_\times,\text{RFA}([M_0]_{\Sigma}, [M_1]_{\Sigma}) &= \text{encode} \circ C_\times,\text{LBFA}(\text{decode}([M_0]_{\Sigma}), \text{decode}([M_1]_{\Sigma})) \\
C_\cup,\text{RFA}([M_0]_{\Sigma}, [M_1]_{\Sigma}) &= \text{encode} \circ C_\cup,\text{LBFA}(\text{decode}([M_0]_{\Sigma}), \text{decode}([M_1]_{\Sigma})) \\
C_\times,\text{RFA}([M]_{\Sigma}) &= \text{encode} \circ C_\times,\text{LBFA}(\text{decode}([M]_{\Sigma})) \\
C_+,\text{RFA}([M]_{\Sigma}) &= \text{encode} \circ C_+,\text{LBFA}(\text{decode}([M]_{\Sigma})) \\
C_\times,\text{RFA}([M]_{\Sigma}) &= \text{encode} \circ C_\times,\text{LBFA}(\text{decode}([M]_{\Sigma}))
\end{align*}
\]
CONSTRUCTIONS BASED ON REGULAR EXPRESSION STRUCTURE

In full:

\[ C_{\ast,RFA} = [(\emptyset, V, \emptyset, \emptyset, 0, \text{true, 0})]_{\Sigma} \]

\[ C_{\emptyset,RFA} = [(\emptyset, V, \emptyset, \emptyset, 0, \text{false, 0})]_{\Sigma} \]

For all \( a \in V \):

\[ C_{a,RFA} = \text{let } q_0 \text{ be a new state} \]
\[ \text{in } [(\{q_0\}, V, \emptyset, \{q_0\}, \{q_0\}, \text{false, \{(q_0, a)\}})]_{\Sigma} \]

\[ C_{\ast,RFA}([M_0]_{\Sigma}, [M_1]_{\Sigma}) = \text{let } (Q_0, V, \text{follow}_0, \text{first}_0, \text{last}_0, \text{null}_0, \text{Qmap}_0) = M_0 \]
\[ (Q_1, V, \text{follow}_1, \text{first}_1, \text{last}_1, \text{null}_1, \text{Qmap}_1) = M_1 \]
\[ \text{in } \]
\[ \text{let } \begin{align*}
  \text{first}' & = \text{first}_0 \cup \text{if } (\text{null}_0) \text{ then } \text{first}_1 \text{ else } \emptyset \text{ fi} \\
  \text{last}' & = \text{last}_1 \cup \text{if } (\text{null}_1) \text{ then } \text{last}_0 \text{ else } \emptyset \text{ fi}
\end{align*} \]
\[ \text{in } [(Q_0 \cup Q_1, V, \text{follow}_0 \cup \text{follow}_1 \cup (\text{last}_0 \times \text{first}_1), \]
\[ \text{first}', \text{last}', \text{null}_0 \land \text{null}_1, \text{Qmap}_0 \cup \text{Qmap}_1)]_{\Sigma} \]

\[ C_{\cup,RFA}([M_0]_{\Sigma}, [M_1]_{\Sigma}) = \text{let } (Q_0, V, \text{follow}_0, \text{first}_0, \text{last}_0, \text{null}_0, \text{Qmap}_0) = M_0 \]
\[ (Q_1, V, \text{follow}_1, \text{first}_1, \text{last}_1, \text{null}_1, \text{Qmap}_1) = M_1 \]
\[ \text{in } [(Q_0 \cup Q_1, V, \text{follow}_0 \cup \text{follow}_1, \text{first}_0 \cup \text{first}_1, \]
\[ \text{last}_0 \cup \text{last}_1, \text{null}_0 \lor \text{null}_1, \text{Qmap}_0 \cup \text{Qmap}_1)]_{\Sigma} \]

\[ C_{\ast,RFA}([M]_{\Sigma}) = \text{let } (Q, V, \text{follow}, \text{first}, \text{last}, \text{null}, \text{Qmap}) = M \]
\[ \text{in } [(Q, V, \text{follow} \cup (\text{last} \times \text{first}), \text{first}, \text{last}, \text{true, Qmap})]_{\Sigma} \]

\[ C_{1,RFA}([M]_{\Sigma}) = \text{let } (Q, V, \text{follow}, \text{first}, \text{last}, \text{null}, \text{Qmap}) = M \]
\[ \text{in } [(Q, V, \text{follow} \cup (\text{last} \times \text{first}), \text{first}, \text{last}, \text{null, Qmap})]_{\Sigma} \]

\[ C_{7,RFA}([M]_{\Sigma}) = \text{let } (Q, V, \text{follow}, \text{first}, \text{last}, \text{null, Qmap}) = M \]
\[ \text{in } [(Q, V, \text{follow}, \text{first}, \text{last}, \text{true, Qmap})]_{\Sigma} \]

An \( M \in \text{RFA} \) (the image of some \( E \in \text{RE} \)) in this \( \Sigma \)-algebra has the following interesting property:

- The number of states in \( M \) equals the number of (not necessarily distinct) symbols (of \( V \)) occurring in \( E \). This follows from the fact that the operators \( C_{a,RFA} \) (for all \( a \in V \)) are the only \( \text{RFA} \) operators that introduce new states. This property will be used in Section 4.5 to derive a practical implementation of the \( \text{RFA} \) operators.

We can also note the following about the operators:

- The operators do not depend on the choice of representatives of the equivalence classes.
4.2 Towards the Berry-Sethi construction

- An important fact is that the operators of this $\Sigma$-algebra are symmetrical. That is, each operator is its own dual.

□

**Definition 4.30** (Homomorphism from $RE$ to $[RFA]_{\Sigma}$): We define $rfa \in RE \rightarrow [RFA]_{\Sigma}$ to be the unique homomorphism from $RE$'s to $[RFA]_{\Sigma}$. □

**Property 4.31** (Homomorphism $rfa$): Since the operators of the $\Sigma$-algebra of $RFA$'s are symmetrical (symmetrical functions are defined in Definition A.22), so is $rfa$. That is, $rfa \circ R(E) = R \circ rfa(E)$. □

In Section 4.5 practical implementations of the $\Sigma$-algebra of $RFA$'s (in particular, of homomorphism $rfa$) are discussed.

4.2.2 The Berry-Sethi construction

Given the $\Sigma$-algebra of $RFA$'s, we have the desired property that (for $E \in RE$):

$$ibfa(E) = decode \circ rfa(E)$$

We now present Berry and Sethi's $FA$ construction.

**Construction 4.32** (Berry-Sethi): Construction $BS \in RE \rightarrow [FA]_{\Sigma}$ is defined as:

$$BS(E) = decode \circ rfa(E)$$

An automaton constructed using this function has the same properties as one constructed with function $lbfa$, namely:

- It is $\epsilon$-free.
- It has a single start state.
- The single start state has no in-transitions.
- All in-transitions to a state are on the same symbol (of $V$).

In practice, function $BS$ is cheaper to compute than $lbfa$. □

**Remark 4.33**: The history of this algorithm is somewhat complicated. The following account is given by Brüggemann-Klein [B-K93b]. Glushkov and McNaughton and Yamada simultaneously (and independently) discovered the same DFA construction [Glus61, MY60]. These papers use the same underlying $\epsilon$-free FA construction to which they apply the subset construction. Unfortunately, neither of them present the $\epsilon$-free FA construction (without the subset construction) explicitly. The underlying $\epsilon$-free FA construction was presented in some depth (with correctness arguments) by Berry and Sethi in [BS86, Alg. 4.4]. In their paper, Berry and Sethi also relate the construction to the Brzozowski construction (Brzozowski's construction appears as Construction 5.34 in this paper).

In this paper, we adopt the convention that the $\epsilon$-free FA construction (without subset construction) is named after Berry and Sethi, while the construction with the subset construction is named after McNaughton, Yamada, and Glushkov. □

**Example 4.34** (Berry-Sethi): A representative of the equivalence class $BS((a \cup \epsilon)b^*)$ is shown in Figure 4 appearing on page 32. This is the same FA as in Example 4.23. (This follows from the fact that the Berry-Sethi construction and the LBFA $\Sigma$-algebra are commuting ways of arriving at the same FA isomorphism class). □

---

*The underlying construction may actually produce a nondeterministic finite automata.*
It is possible to find a composition of functions that commutes with \( lbfa \) (and therefore \( decode \circ rfa \)) and is cheaper to compute in practice. We first give some required definitions.

**Definition 4.35** (Non-isomorphic mapping from \([RFA]_\omega\) to \([FA]_\omega\)): Function \( convert \in [RFA]_\omega \rightarrow [FA]_\omega \) is defined as:

\[
\text{convert}([M]_\omega) = \begin{array}{l}
\text{let } (Q, V, follow, first, last, null, Qmap) = M \\
\text{let } T = \{(q_0, Qmap(q_1), q_1) : (q_0, q_1) \in \text{follow}\} \\
\text{in } ([Q, V, T, \emptyset, \text{first}, \text{last}])_\omega
\end{array}
\]

An important property of this function is that:

\[
(\forall E : E \in RE : \mathcal{L}_{FA} \circ \text{convert} \circ rfa(E) = V^{-1}\mathcal{L}_{RE}(E))
\]

This follows from the fact that \( \text{convert} \) simply discards the transitions that would be out of the start state. Function \( \text{convert} \) does not add any new states, unlike function \( \text{decode} \) which adds a new start state. □

**Definition 4.36** (Adding a begin-marker): Define function \( \text{marker}_b \in RE \rightarrow RE \) as:

\[
\text{marker}_b(E) = \$ \cdot E
\]

Where \( \$ \) is an alphabet symbol, called a begin-marker. (In the literature, it is usually assumed — for no particular reason — that symbol \( \$ \) does not occur in regular expression \( E \).) This function satisfies the obvious property that:

\[
(\forall E : E \in RE : \mathcal{L}_{RE}(\text{marker}_b(E)) = \{\$\}\mathcal{L}_{RE}(E))
\]

□

Given functions \( \text{marker}_b, rfa, \) convert, and the following important property, we can construct an efficient alternative to homomorphism \( lbfa \).

**Property 4.37** (Functions \( \text{marker}_b, rfa, \) and \( \text{convert} \)): Because of the properties of \( \text{convert} \) and \( \text{marker}_b \), we can show that:

\[
\begin{align*}
\text{convert} \circ rfa \circ \text{marker}_b(E) \\
= & \quad \{ \text{Definition of } \text{marker}_b \} \\
= & \quad \{ \text{Definitions of } rfa, C_{,RFA} \} \\
= & \quad \{ \text{Definitions of } C_{,RFA}(rfa(\$), rfa(E)) \} \\
= & \quad \{ \text{Definitions of } rfa, C_{,RFA}(\$_{,RFA}, rfa(E)) \} \\
= & \quad \{ \text{Definitions of } \text{convert}, C_{,RFA}, C_{,RFA}, rfa, \text{and } \text{decode} \} \\
\text{decode} \circ rfa(E) \\
= & \quad \{ \text{Commutativity} \} \\
\text{lbfa}(E)
\end{align*}
\]

The composite \( \text{convert} \circ rfa \circ \text{marker}_b \) is an alternative (and in practice, cheaper) implementation of \( lbfa \). □
Towards the Berry-Sethi construction

The fact that $convert \circ rfa \circ marker_b$ is a construction is depicted in the following commuting diagram:

$$\begin{align*}
\text{RE} & \xrightarrow{lbfa} [LBFA]_\equiv \\
\text{marker}_b & \downarrow \quad \quad \quad \downarrow convert \\
\text{RE} & \xrightarrow{rfa} [RFA]_\equiv
\end{align*}$$

Construction 4.38 (A variation on the Berry-Sethi construction): Instead of constructing an FA using the functions $lbfa$ or $BS$, it is cheaper in practice to use the composite function

$$convert \circ rfa \circ marker_b(E)$$

4.2.3 The McNaughton-Yamada-Glushkov construction

Since the Berry-Sethi construction produces an $\epsilon$-free (possibly nondeterministic) FA, we now consider making the resulting FA deterministic.

Construction 4.39 (McNaughton-Yamada-Glushkov): (We assume that the composite function $useful_a \circ subset$ is extended to $[FA]_\equiv \rightarrow [DFA]_\equiv$. The McNaughton-Yamada-Glushkov DFA construction is $MYG \in \text{RE} \rightarrow [DFA]_\equiv$, defined as:

$$MYG(E) = useful_a \circ subset \circ decode \circ rfa(E)$$

A DFA produced by $MYG$ is Complete (by a property of $useful_a \circ subset$). A practical implementation is given in Algorithm 4.42 (given below), which implements $useful_a \circ subset \circ decode$. Homomorphism $rfa$ can be implemented using the techniques described in Section 4.5. This algorithm is the same as that given by McNaughton and Yamada [MY60, Construction method on pg. 44].

Example 4.40 (McNaughton-Yamada-Glushkov): In the case of $(a \cup \epsilon)b^* \in \text{RE}$, the Berry-Sethi construction produces a deterministic FA. Function $MYG$ produces a similar DFA, with a sink state added to make it Complete. The state graph of a representative DFA of isomorphism class $MYG((a \cup \epsilon)b^*)$ is shown in Figure 5.

Remark 4.41: The variation on the Berry-Sethi construction (Construction 4.38) can be used for a practical implementation of the McNaughton-Yamada-Glushkov construction. This would yield a construction not appearing in the literature.
Figure 5: A representative DFA of the isomorphism class $MYG((a \cup \epsilon)b^*)$.

Composite function $useful \circ subset \circ decode$ can be implemented using Algorithm 2.45 (which implements $useful \circ subset$). Here, the first iteration is unrolled to accommodate the definition of function $decode$, and some obvious improvements have not yet been made.

\[
\begin{align*}
(Q, V, follow, first, last, null, Qmap) &\in RFA \\
let S = \{\{s\} : s \text{ is a new state, } s \not\in Q; \quad T := \emptyset; \\
D, U := \emptyset, S; \\
let u : u \in U; \\
D, U := D \cup \{u\}, U \setminus \{u\}; \\
for a : a \in V do &
\begin{align*}
&d := (U \cup p : p \in u : \{q : q \in first \land Qmap(q) = a\}); \\
&\text{if } d \not\in D \quad U := U \cup \{d\} \\
&\text{if } d \in D \quad \text{skip}
\end{align*} \\
&fi; \\
&T := T \cup \{(u, a, d)\}
\end{align*}
\]

\begin{align*}
\text{do } U \neq \emptyset \rightarrow &
\begin{align*}
&\text{let } u : u \in U; \\
&D, U := D \cup \{u\}, U \setminus \{u\}; \\
&\text{for } a : a \in V do \\
&\begin{align*}
&d := (U \cup p : p \in u : \{q : (p, q) \in follow \land Qmap(q) = a\}); \\
&\text{if } d \not\in D \quad U := U \cup \{d\} \\
&\text{if } d \in D \quad \text{skip}
\end{align*} \\
&fi; \\
&T := T \cup \{(u, a, d)\}
\end{align*}
\]

\[
\text{od;}
\]

\begin{align*}
F := \{d : d \in D \land d \cap \text{last} \neq \emptyset\} \cup \text{if (null) then } S \text{ else } \emptyset \text{ fi} \\
\{(D, V, T, \emptyset, S, F) \mid useful \circ subset \circ decode([[Q, V, follow, first, last, null, Qmap]]) \}
\end{align*}

\{Complete(D, V, T, \emptyset, S, F)\}
Some simplification gives the algorithm:

**Algorithm 4.42 (McNaughton-Yamada-Glushkov):**

\[
\begin{align*}
&\{(Q,V,\text{follow}, \text{first}, \text{last}, \text{null}, \text{Qmap}) \in \text{RFA}\} \\
&\text{let } S = \{s\} : s \text{ is a new state, } s \notin Q; \\
&\quad T := \emptyset; \\
&\quad D, U := S, \emptyset; \\
&\quad \text{for } a : a \in V \text{ do} \\
&\quad \quad d := \{q : q \in \text{first} \land \text{Qmap(q)} = a\}; \\
&\quad \quad U := U \cup \{d\}; \\
&\quad \quad T := T \cup \{(s), (a, d)\} \\
&\quad \text{rof;} \\
&\quad \text{do } U \neq \emptyset \rightarrow \\
&\quad \quad \text{let } u : u \in U; \\
&\quad \quad D, U := D \cup \{u\}, U \setminus \{u\}; \\
&\quad \quad \text{for } a : a \in V \text{ do} \\
&\quad \quad \quad d := (\cup p : p \in u : \{q : (p, q) \in \text{follow} \land \text{Qmap(q)} = a\}); \\
&\quad \quad \quad \text{if } d \notin D \rightarrow U := U \cup \{d\} \\
&\quad \quad \quad \quad \text{skip} \\
&\quad \quad \quad T := T \cup \{(u, a, d)\} \\
&\quad \text{rof} \\
&\quad F := \{d : d \in D \land \text{last } \neq \emptyset\} \cup \text{if (null) then } S \text{ else } \emptyset \text{ fi} \\
&\quad \{(D,V,T,\emptyset,\text{null},\text{Qmap})\}_{=} = \text{useful} \circ \text{subset} \circ \text{decode}((Q,V,\text{follow}, \text{first}, \text{last}, \text{null}, \text{Qmap})|_{=} ) \\
&\quad \{(\text{Complete}(D,V,T,\emptyset,\text{null},\text{Qmap})|_{=} )\} 
\end{align*}
\]

This algorithm is used in the McNaughton-Yamada construction [MY60].

### 4.3 The dual of the Berry-Sethi construction

The following commuting diagram gives a property of regular expressions and regular languages that will prove to be useful:

\[
\begin{align*}
RE & \xrightarrow{L_{RE}} L_{\text{reg}} \\
\downarrow R & & \uparrow R \\
RE & \xrightarrow{L_{RE}} L_{\text{reg}}
\end{align*}
\]

In this diagram, the two reversal operators are different: one is reversal of \(RE\)'s, while the other is reversal of languages.

Given the definition of an FA construction \(f\) and the above diagram, we have the property that the dual of a construction is again a construction. That is, \(R \circ f \circ R\) is also a construction. Such a dual construction is less efficient than \(f\) (since it involves two reversal functions), and we explore ways to efficiently implement the dual constructions.

**Construction 4.43 (Right-biased):** We can use \(R \circ \text{lbfa} \circ R\) as a construction. For any given \(E \in RE\), a representative of \(R \circ \text{lbfa} \circ R(E)\) has the following properties (the properties are based upon those of the left-biased \(\Sigma\)-algebra):
CONSTRUCTIONS BASED ON REGULAR EXPRESSION STRUCTURE

Figure 6: A representative FA of the isomorphism class $R \circ \lambda bfa \circ R((a \cup c)b^*)$.

- It is $\epsilon$-free.
- It has a single final state.
- The single final state has no out-transitions.
- All out-transitions from a state are on the same symbol (in $V$).

\[ \square \]

Example 4.44 (Right-biased construction): We construct a representative FA of the isomorphism class $R \circ \lambda bfa \circ R((a \cup c)b^*)$ The representative is shown in Figure 6. \[ \square \]

To consider the dual of the Berry-Sethi construction, we combine the commuting diagrams of duals of a construction (above) and construction $\text{decode} \circ rfa$, giving:

\[
\begin{array}{c}
\text{RE} \\
\uparrow R \\
\downarrow R \\
\text{RE} \\
\end{array}
\quad \begin{array}{c}
[FA]_{\text{FA}} \\
\uparrow \mathcal{L}_{\text{FA}} \\
\downarrow \mathcal{L}_{\text{reg}} \\
[FA]_{\text{FA}} \\
\end{array}
\quad \begin{array}{c}
\text{RE} \\
\uparrow R \\
\downarrow \text{decode} \\
\text{RE} \\
\end{array}
\]

The source is the upper-left vertex, and the sink is the upper FA vertex.

The construction $R \circ \text{decode} \circ rfa \circ R$ (in this diagram) is still inefficient, requiring two redundant reversal operations. We can make it more efficient, by finding functions that form new paths in the commuting diagram.

From the definitions of the $\Sigma$-algebra of $RFA$'s (Definition 4.29) and homomorphism $rfa$ (Definition 4.30) we know that the $RFA$ operators are symmetrical, and so is $rfa$. In other words $rfa \circ R(E) = R \circ rfa(E)$. This allows us to add two new edges to the above commuting diagram; the resulting diagram is:
Construction 4.45 (The dual of Berry-Sethi): The construction is:

\[ R \circ \text{decode} \circ R \circ \text{rf}\](E) \]

This construction is the dual of the Berry-Sethi construction (Construction 4.32). □

We give \( R \circ \text{decode} \circ R \) in full:

\[
R \circ \text{decode} \circ R([M]_E) = \text{let } (Q, V, \text{follow}, \text{first}, \text{last}, \text{null}, Qmap) = M \text{ be a new state} \\
\quad \text{in} \\
\quad \quad \text{let } T = \{(q_0, Qmap(q_0), q_1) : (q_0, q_1) \in \text{follow} \} \\
\quad \quad \quad T' = \{(q, Qmap(q), f) : q \in \text{last} \} \\
\quad \quad \quad S = \text{first} \cup \text{if } (\text{null}) \text{ then } \{f\} \text{ else } \emptyset \text{ fi} \\
\quad \quad \text{in} \\
\quad \quad \quad [(Q \cup \{f\}, V, T \cup T', \emptyset, \{f\})]_E \\
\quad \text{end}
\]

The FA resulting from this construction is the same as given in Example 4.44.

We can also consider improving the dual of the variation on the Berry-Sethi construction (Construction 4.38). We combine the commuting diagram showing the dual of a construction, with construction \( \text{convert} \circ \text{rf} \circ \text{marker}_b \), giving:

\[
RE \quad \text{marker}_b \quad RE \quad \text{rf} \quad [RFA]_E \quad \text{convert} \quad [FA]_E \quad \mathcal{L}_E \quad \mathcal{L}_{\text{reg}}
\]

Again, the source is the upper-left \( RE \) vertex, while the sink is the upper-right \( FA \) vertex.

Consider the composite function \( R \circ \text{convert} \circ \text{rf} \circ \text{marker}_b \circ R \). We begin with the two rightmost functions:

\[
\text{marker}_b \circ R(E) \\
= \{ \text{Writing } R \text{ as postfix and superscript} \} \\
\text{marker}_b(E^R) \\
= \{ \text{Definition of } \text{marker}_b \text{ (Definition 4.36)} \} \\
\cdot (E^R) \\
= \{ \text{Function } R \circ R \text{ is the identity (see Definition A.19)} \} \\
R \circ R(\cdot (E^R)) \\
= \{ \text{Definition of } R \text{ on } \cdot \text{ regular expressions} \}
\]
To make this definition more concise, we define an end-marker function.

**Definition 4.46 (Adding an end-marker):** Define function $\text{marker}_e \in RE \rightarrow RE$ as:

$$\text{marker}_e(E) = E \cdot \$$$

where $\$ is assumed to be a symbol in the alphabet. \(\square\)

**Property 4.47 (marker):** Function $\text{marker}_e$ is the dual of function $\text{marker}_b$ (Definition 4.36):

$$\text{marker}_b \circ R(E) = R \circ \text{marker}_e(E)$$

\(\square\)

With the above property, we can transform the above commuting diagram, by adding two new edges:

The composite $R \circ \text{convert} \circ rfa \circ R \circ \text{marker}_e$ is no more efficient even with the use of $\text{marker}_e$. Fortunately, since $rfa$ is symmetrical, we can replace $rfa \circ R$ by $R \circ rfa$, giving:

The composite function $R \circ \text{convert} \circ R$ is particularly easy to present, using the definitions of $R$ and $\text{convert}$ (Definition 4.35):

$$R \circ \text{convert} \circ R([R]_\equiv) = \begin{array}{l}
\text{let} \quad (Q, V, \text{follow}, \text{first}, \text{last}, \text{null}, Qmap) = R \\
\text{in} \quad \begin{array}{l}
\text{let} \quad T = \{(q_0, Qmap(q_0), q_1) : (q_0, q_1) \in \text{follow}\} \\
\text{in} \quad [[Q, V, T, \emptyset, \text{first}, \text{last}]]_\equiv \\
\text{end}
\end{array}
\end{array}$$

This leads to the following construction.
4.3 The dual of the Berry-Sethi construction

Construction 4.48 (The dual of the variation on the Berry-Sethi construction): The construction is:

\[ R \circ \text{convert} \circ R \circ rfa \circ \text{marker}_e(E) \]

This construction is also presented very informally by Aho, Sethi, and Ullman [ASU86, Example 3.22, pg. 140]. There appears to be an error in item two of the three steps describing the construction in [ASU86]. Instead of

2. label each directed edge \((i,j)\) by the symbol at position \(j\), and

the step should read

2. label each directed edge \((i,j)\) by the symbol at position \(i\), and

Example 4.49 (The dual of the variation on the Berry-Sethi construction): A representative FA of the isomorphism class \(R_0 \circ \text{convert} \circ R \circ rfa \circ \text{marker}_e((a \cup \epsilon)b^*)\) is shown in Figure 6 on page 40. This is the same FA as in Example 4.44.

4.3.1 The Aho-Sethi-Ullman DFA construction

In order to obtain a (possibly non-Complete) DFA we use the composite function \(\text{useful}_\epsilon \circ \text{subsetopt}\) (given in Definition 2.44), extended to \(\left[DFA\right]_\epsilon \rightarrow [DFA]_\omega\).

We can immediately give the Aho-Sethi-Ullman DFA construction using this composite function.

Construction 4.50 (Aho-Sethi-Ullman): The construction is \(ASU \in RE \rightarrow [DFA]_\omega\) defined as:

\[ ASU(E) = \text{useful}_\epsilon \circ \text{subsetopt} \circ R \circ \text{convert} \circ R \circ rfa \circ \text{marker}_e(E) \]

Algorithm 4.52 (given below) is an imperative program implementing

\[ \text{useful}_\epsilon \circ \text{subsetopt} \circ R \circ \text{convert} \circ R \]

Homomorphism \(rfa\) can be implemented using the techniques described in Section 4.5, and function \(\text{marker}_\epsilon\) is trivial to implement. The Aho-Sethi-Ullman algorithm is given in [ASU86, Alg. 3.5, Fig. 3.44].

Example 4.51 (Aho-Sethi-Ullman): We give a representative DFA of the isomorphism class \(ASU((a \cup \epsilon)b^*)\). The state graph is shown in Figure 7.
Algorithm 4.52 (Aho-Sethi-Ullman):

\[
\{(Q, V, \text{follow}, \text{first}, \text{last}, \text{null}, \text{Qmap}) \in \text{RFA}\}
\]

\[S, T := \text{if} \ (\text{first} \neq \emptyset) \ \text{then} \ \{\text{first}\} \ \text{else} \ \emptyset, \emptyset;\]

\[D, U := \emptyset, S;\]

\[\text{do } U \neq \emptyset \rightarrow \]

\[
\begin{align*}
&\text{let } u : u \in U; \\
&D, U := D \cup \{u\}, U \setminus \{u\}; \\
&\text{for } a : a \in V \land (\exists q : q \in u : \text{Qmap}(q) = a \land \text{follow}(q) \neq \emptyset) \ \text{do} \\
&\quad d := (\cup q : q \in u \land \text{Qmap}(q) = a : \text{follow}(q)); \\
&\quad \text{if } d \notin D \rightarrow U := U \cup \{d\} \\
&\quad \cap d \in D \rightarrow \text{skip} \\
&\quad \text{fi;} \\
&\quad T := T \cup \{(u, a, d)\}
\end{align*}
\]

\[\text{rof}\]

\[F := \{d : d \in D \land d \cap \text{last} \neq \emptyset\}
\]

\[
\{(D, V, T, \emptyset, S, F)\} = \text{useful} \circ \text{subsetopt} \circ R \circ \text{convert} \circ R((\{(Q, V, \text{follow}, \text{first}, \text{last}, \text{null}, \text{Qmap})\})\}
\]

4.4 Extending regular expressions

For some regular languages, the regular expressions denoting the language can be considerably more succinct when operators such as intersection (\(\cap\)) and complement (\(\neg\)) are available in \(\text{RE}\)'s. Without formally adding them to the signature \(\Sigma\), we briefly consider how to implement operator \(\cap\) in the left-biased \(\Sigma\)-algebra of \(\text{FA}\)'s.

Definition 4.53 (Extended regular expressions and their languages): The set of extended regular expressions (over alphabet \(V\), \(\text{ERE}\), and the languages they denote, are exactly as \(\text{RE}\), with the addition of the operators \(\cap \in \text{ERE} \times \text{ERE} \rightarrow \text{ERE}\) (an infix operator) and \(\neg \in \text{ERE} \rightarrow \text{ERE}\) (a prefix operator). Operator \(\cap\) has the same precedence as \(\cup\), while \(\neg\) has higher precedence than \(\ast\). The language of an \(\text{ERE}\) is defined using the function \(L_{\text{ERE}} \in \text{ERE} \rightarrow L_{\text{reg}}\) which is as function \(L_{\text{RE}}\), with the extensions

\[
L_{\text{ERE}}(E_0 \cap E_1) = L_{\text{ERE}}(E_0) \cap L_{\text{ERE}}(E_1)
\]

\[
L_{\text{ERE}}(\neg E_0) = V^* \setminus L_{\text{ERE}}(E_0)
\]

\(\Box\)

Remark 4.54: The \(\Sigma\)-algebra definition of regular expressions are not used in this section as the algebraic structure is not needed. \(\Box\)

Definition 4.55 (Intersection of \(\text{LBFA}\)'s): In defining intersection, we assume that the two arguments have been constructed in the \(\Sigma\)-algebra of \(\text{LBFA}\). In particular, we require that for each state, all in-transitions are on the same symbol. Assuming the argument representatives
4.4 Extending regular expressions

have disjoint state sets, one possible implementation of the operator \( \cap \) is:

\[
C_{\cap, LBFA}([M_0]_{\equiv}, [M_1]_{\equiv}) = \begin{align*}
&\text{let } (Q_0, V, T_0, \emptyset, \{s_0\}, F_0) = M_0 \\
&\quad (Q_1, V, T_1, \emptyset, \{s_1\}, F_1) = M_1 \\
&\quad q_0 \text{ be a new state} \\
&\quad N = \epsilon \in (L_{FA}(M_0) \cap L_{FA}(M_1)) \\
&\quad \text{in} \\
&\quad \text{let } Q' = \{q_0\} \cup (\cup b : b \in V : \pi_2(T_0(b)) \times \pi_2(T_1(b))) \\
&\quad T'(a) = \{q_0\} \times (T_0(s_0, a) \times T_1(s_1, a)) \\
&\quad \cup \{((p, q), (p', q')) : (p, p') \in T_0(a) \land p \neq s_0 \\
&\quad \land (q, q') \in T_1(a) \land q \neq s_1 \\
&\quad \land (\exists b : b \in V : p \in \pi_2(T_0(b)) \land q \in \pi_2(T_1(b)))\}
&\quad \text{in} \\
&\quad [(Q', V, T', \emptyset, \{q_0\}, (F_0 \times F_1) \cap Q') \\
&\quad \cup \text{if } (N) \text{ then } \{q_0\} \text{ else } \emptyset]\}_\equiv \\
&\quad \text{end}
\end{align*}
\]

The expression

\[
Q' = \{q_0\} \cup (\cup b : b \in V : \pi_2(T_0(b)) \times \pi_2(T_1(b)))
\]

in the let clause deserves some explanation. A state in the constructed LBFA is either the new state \( q_0 \), or a pair of states \((p, q)\) where \( p \) and \( q \) \((p \neq s_0, q \neq s_1)\) are from \( M_0 \) and \( M_1 \) respectively. If

\( p \) and \( q \) do not have an in-transition on the same symbol, the state \((p, q)\) will be start-unreachable in the constructed LBFA. For this reason, it is omitted. The definition of the transition relation is similar. The constructed LBFA is sometimes called the cross-product LBFA. Although the operator removes most start-unreachable states, some may still remain.

We can now present an intersection operator for RFA's.

**Definition 4.56 (Intersection of RFA's):** We define intersection of RFA's as:

\[
C_{\cap, RFA}([M_0]_{\equiv}, [M_1]_{\equiv}) = \text{encode} \circ C_{\cap, LBFA}(\text{decode}([M_0]_{\equiv}), \text{decode}([M_1]_{\equiv}))
\]

In full:

\[
C_{\cap, RFA}([M_0]_{\equiv}, [M_1]_{\equiv}) = \begin{align*}
&\text{let } (Q_0, V, \text{follow}_0, \text{first}_0, \text{last}_0, \text{null}_0, \text{Qmap}_0) = M_0 \\
&\quad (Q_1, V, \text{follow}_1, \text{first}_1, \text{last}_1, \text{null}_1, \text{Qmap}_1) = M_1 \\
&\quad \text{in} \\
&\quad \text{let } Q' = (\cup b : b \in V : \text{Qmap}_0^{-1}(b) \times \text{Qmap}_1^{-1}(b)) \\
&\quad \text{follow}' = \{((p, q), (p', q')) : (p, p') \in \text{follow}_0 \\
&\quad \land (q, q') \in \text{follow}_1 \\
&\quad \land \text{Qmap}_0(p) = \text{Qmap}_1(q) \\
&\quad \land \text{Qmap}_0(p') = \text{Qmap}_1(q')\}
&\quad \text{first}' = \{(p, q) : p \in \text{first}_0 \land q \in \text{first}_1 \\
&\quad \land \text{Qmap}_0(p) = \text{Qmap}_1(q)\}
&\quad \text{Qmap}'^{-1}(a) = \text{Qmap}_0^{-1}(a) \times \text{Qmap}_1^{-1}(a)
&\quad \text{in} \\
&\quad [(Q', V, \text{follow}', \text{first}', (\text{null}_0 \times \text{null}_1) \cap Q', \\
&\quad \text{null}_0 \land \text{null}_1, \text{Qmap}', \text{Qmap}')]_{\equiv}
\end{align*}
\]

Note that this operator is symmetrical. \( \Box \)

---

\(^{10}\)The definition presented here is intentionally clumsy, making it easier to present intersection of RFA's.
In their original paper [MY60, Section IV], McNaughton and Yamada attempt to define intersection. Unfortunately, their informal presentation is difficult to understand. Subsequent presentations of the RFA operators have all omitted intersection. For example, Berry and Sethi note (in [BS86, Remark 3.1]):

"The approach ... does not extend to regular expressions with intersection and complementation operators."

We can construct an RFA operator for any regular operator (an operator on languages that preserves the regularity property) for which we can construct an LBFA operator. Examples of such operators are intersection, symmetrical difference, complement, asymmetrical difference, and prefix closure.

4.5 Efficiently computing with RFA's

In this section, we consider some practical methods for constructing RFA's. The first subsection considers a practical implementation of the (Σ-algebra of) RFA operators, while the second subsection introduces some improvements (due to Chang, Paige, and Brüggemann-Klein) to the RFA operators.

4.5.1 A practical implementation of the RFA operators

In an RFA, the states are mapped to their corresponding symbol (of V) by the seventh component (usually called Qmap) of the RFA. This seventh component would be redundant if the states and symbols were in a one-to-one correspondence. Furthermore, the symbols could then be used as the states. In this subsection, we explore this encoding method, and the requirements on the RE's for this method to work. We will also be defining a new, restricted, mapping rfa' ∈ RE → RFA. We will be able to use this mapping for regular expressions in which each alphabet symbol occurs no more than once.

We first define an important auxiliary function.

**Definition 4.57 (Occurrences of symbols in RE's):** We define function \( \text{Occ} : \text{RE} \rightarrow \mathcal{P}(V) \) such that \( \text{Occ}(E) \) is the set of symbols (of V) occurring in \( E \). We can also define \( \text{Occ} \) recursively as follows:

\[
\begin{align*}
\text{Occ}(\epsilon) &= \emptyset \\
\text{Occ}(\emptyset) &= \emptyset \\
\text{Occ}(a) &= \{a\} \quad \text{(for } a \in V) \\
\text{Occ}(E \cdot F) &= \text{Occ}(E) \cup \text{Occ}(F) \\
\text{Occ}(E \cup F) &= \text{Occ}(E) \cup \text{Occ}(F) \\
\text{Occ}(E^*) &= \text{Occ}(E) \\
\text{Occ}(E^+) &= \text{Occ}(E) \\
\text{Occ}(E^+) &= \text{Occ}(E)
\end{align*}
\]

\( \square \)

**Definition 4.58 (RRE):** We define \( \text{RRE} \subset \text{RE} \) as the smallest set satisfying:

- \( \epsilon \in \text{RRE} \),
- \( \emptyset \in \text{RRE} \),
- \( a \in \text{RRE} \) \quad (for \( a \in V \)),
- if \( E, F \in \text{RRE} \), and \( \text{Occ}(E) \cap \text{Occ}(F) = \emptyset \) then \( E \cdot F \in \text{RRE} \) and \( E \cup F \in \text{RRE} \), and
- if \( E \in \text{RRE} \) then \( E^* \in \text{RRE} \), \( E^+ \in \text{RRE} \), and \( E^+ \in \text{RRE} \).
Intuitively, \( RRE \) (for restricted regular expressions) denotes the set of all \( E \in RE \) such that each symbol (of \( V \)) occurs no more than once in \( E \). □

**Example 4.59 (RRE):** A \( RRE \) is \( (a \cup \epsilon)b^* \). □

In order to give our alternative \( RE \) to \( RFA \) mapping, \( \text{rfa}' \in RRE \rightarrow RFA \), we first define some more auxiliary functions. The definitions of these functions also follow directly from the \( RFA \) operators.

**Definition 4.60 (First):** We define \( \text{First} \in RE \rightarrow \mathcal{P}(V) \) recursively (recall from Example 3.20 that \( \text{Null}(E) \equiv (\epsilon \in \mathcal{L}_{RE}(E)) \)):

\[
\text{First}(\epsilon) = \emptyset \\
\text{First}(\emptyset) = \emptyset \\
\text{First}(a) = \{a\} \quad \text{(for } a \in V\text{)} \\
\text{First}(E \cdot F) = \text{First}(E) \cup \{a\} \quad \text{if } \text{Null}(E) \text{ then } \text{First}(F) \text{ else } \emptyset \fi \\
\text{First}(E \cup F) = \text{First}(E) \cup \text{First}(F) \\
\text{First}(E^*) = \text{First}(E) \\
\text{First}(E^+) = \text{First}(E) \\
\text{First}(E^?) = \text{First}(E)
\]

This definition follows directly from the first tuple element of the \( RFA \) operator definitions. □

**Remark 4.61:** It is useful to have an intuitive understanding of function \( \text{First} \). \( \text{First}(E) \) is the set of all symbols that can occur as the first symbol of a string in \( \mathcal{L}_{RE}(E) \). □

**Definition 4.62 (Last):** Function \( \text{Last} \) is defined to be the dual of \( \text{First} \). □

**Remark 4.63:** \( \text{Last}(E) \) is the set of all symbols that can occur as the last symbol of a string in \( \mathcal{L}_{RE}(E) \). □

**Definition 4.64 (Follow):** We define \( \text{Follow} \in RE \rightarrow \mathcal{P}(V \times V) \) recursively:

\[
\text{Follow}(\epsilon) = \emptyset \\
\text{Follow}(\emptyset) = \emptyset \\
\text{Follow}(a) = \emptyset \quad \text{(for } a \in V\text{)} \\
\text{Follow}(E \cdot F) = \text{Follow}(E) \cup \text{Follow}(F) \cup (\text{Last}(E) \times \text{First}(F)) \\
\text{Follow}(E \cup F) = \text{Follow}(E) \cup \text{Follow}(F) \\
\text{Follow}(E^*) = \text{Follow}(E) \cup (\text{Last}(E) \times \text{First}(E)) \\
\text{Follow}(E^+) = \text{Follow}(E) \cup (\text{Last}(E) \times \text{First}(E)) \\
\text{Follow}(E^?) = \text{Follow}(E)
\]

This definition follows directly from the follow tuple element of the \( RFA \) operator definitions. □

**Remark 4.65:** For \( a,b \in V \), \( (a,b) \in \text{Follow}(E) \) is equivalent to \( ab \) being a substring of some string in \( \mathcal{L}_{RE}(E) \). □

**Example 4.66 (First, Last, Null, Follow):** We use the regular expression \( (a \cup \epsilon)b^* \) (from Example 3.15):

\[
\text{First}((a \cup \epsilon)b^*) = \{a,b\} \\
\text{Last}((a \cup \epsilon)b^*) = \{a,b\} \\
\text{Null}((a \cup \epsilon)b^*) = \text{true} \\
\text{Follow}((a \cup \epsilon)b^*) = \{(a,b),(b,b)\}
\]

□

We now have the auxiliary functions required for the definition of \( \text{rfa}' \).
Definition 4.67 (Function \( rfa' \in RRE \rightarrow RFA \)): The definition of \( rfa' \) is straightforward:

\[
\text{rfa}'(E) = (\text{Occ}(E), V, \text{Follow}(E), \text{First}(E), \text{Last}(E), \text{Null}(E), I_v)
\]

where \( I_v \) is the identity function on alphabet symbols.

Example 4.68 (\( rfa' \)): Using the results of the above example, we have:

\[
\text{rfa}'((a \cup e)b^*) = (\{a, b\}, \{a, b\}, \{(a, b), (b, b)\}, \{a, b\}, \{a, b\}, \text{true}, \{(a, a), (b, b)\})
\]

Property 4.69 (\( rfa' \)): Given \( E \in RRE \) then \( rfa(E) = [rfa'(E)]_\approx \).

Function \( rfa' \) is convenient, as all of the auxiliary functions can easily be computed bottom-up on the structure of \( E \).

Construction 4.70 (An encoding of \( BS \)): The method of constructing an \( RFA \) (using \( rfa' \)) leads to a particularly concise definition of \( BS \). For example, we define \( BSenc \in RRE \rightarrow FA \):

\[
BSenc(E) = \begin{cases}
\text{let } s \text{ be a new state} & \\
\text{let } T = \{(a, b, b) : (a, b) \in \text{Follow}(E)\} & \\
T' = \{(s, a, a) : a \in \text{First}(E)\} & \\
F = \text{Last}(E) \cup \text{if (Null(E)) then } \{s\} \text{ else } \emptyset \text{ fi} & \\
\text{in } & \\
(\text{Occ}(E) \cup \{s\}, V \cup T \cup T', \emptyset, \{s\}, F) & \\
\end{cases}
\]

Remark 4.71: Compare the definition of \( BSenc \) to the definition of \( decode \) (Definition 4.28).

Property 4.72 (Construction \( BSenc \)): For \( E \in RRE \):

\[
[BSenc(E)]_\approx = BS(E)
\]

Remark 4.73: By inspection, we see that (for \( E \in RRE \)) the FA \( BSenc(E) \) (equivalently \( BS(E) \)) is deterministic. This implies that:

\[
MYG(E) = \text{complete } \circ BS(E)
\]

Remark 4.74: In Section 5.4 we will show that Brzozowski's construction (with an appropriate encoding) produces a \( DFA \) (from an \( E \in RRE \)) that is isomorphic to the one produced by \( BSenc \) (and therefore \( BS \)).
4.5 Efficiently computing with RFA's

Similarly, the Aho-Sethi-Ullman algorithm becomes quite concise (from Algorithm 4.52):

\[
\{E \in RRE\}
\]

\[
E' := \text{marker}_r(E);
\]

\[
S, T := \begin{cases} \text{if} \ (\text{First}(E') \neq \emptyset) \ \text{then} \ \{\text{First}(E')\} \ \text{else} \ \emptyset, \emptyset; \\
D, U := \emptyset, S;
\end{cases}
\]

\[
\text{do } U \neq \emptyset \rightarrow \\
\quad \text{let } u : u \in U; \\
\quad D, U := D \cup \{u\}, U \setminus \{u\}; \\
\quad \text{for } a : a \in u \land \text{Follow}(E')(a) \neq \emptyset \ \text{do} \\
\quad \quad d := (\text{Follow}(E'))(a); \\
\quad \quad \text{if } d \notin D \rightarrow U := U \cup \{d\} \\
\quad \quad \text{if } d \in D \rightarrow \text{skip} \\
\quad \quad T := T \cup \{(u, a, d)\}
\]

\[
\text{rof}
\]

\[
D, T := 0, S; \\
\text{do} \\
\quad D, U := D \cup \{u\}, U \setminus \{u\}; \\
\quad \text{for } a : a \in u \land \text{Follow}(E')(a) \neq \emptyset \ \text{do} \\
\quad \quad d := (\text{Follow}(E'))(a); \\
\quad \quad \text{if } d \notin D \rightarrow U := U \cup \{d\} \\
\quad \quad \text{if } d \in D \rightarrow \text{skip} \\
\quad \quad T := T \cup \{(u, a, d)\}
\]

\[
\text{rof}
\]

\[
F := \{d : d \in D \land S \in d\}
\]

\[
\{L_{FA}(D, V, T, \emptyset, S, F) = L_{RE}(E)\}
\]

This algorithm is very similar to the one given in [ASU86].

The only problem remaining is how to deal with an \( E \in RE \) when \( E \notin RRE \). The method usually used is to "mark" the symbols of \( E \) (perhaps with an integer subscript), making each symbol unique. For example, \((a^+ \cup ab) \notin RRE\) but after marking we get \((a_1^+ \cup a_2b_3) \in RRE\). Once the corresponding \( FA \) is constructed from the marked regular expression, the marks are removed (the \( FA \) is "unmarked") and the \( FA \) accepts \( L_{RE}(E) \). There are a few different styles of marking. For example, consider \( a^+ \cup ab \): McNaughton-Yamada mark this as \( a_1^+ \cup a_2b_1 \), Berry-Sethi use \( a_2^+ \cup a_3b_3 \), and Aho-Sethi-Ullman use \( a_1^+ \cup a_2b_3 \).

The only disadvantage to the use of marking to encode \( RFA \) computation is that marking is unable to deal with some of the other regular operators, such as intersection, and complementation. For all \( E, F \in RRE \) we have the property that \( L_{RE}(E \cap F) = L_{RE}(E) \cap L_{RE}(F) = \emptyset \). For example, given\(^{11} \) \( aa \cap a^* \) (with language \( L_{RE}(aa \cap a^*) = L_{RE}(aa) \cap L_{RE}(a^*) = \{aa\} \)). After marking we get \( a_1a_2 \cap a_3^* \) after marking (with \( L_{RE}(a_1a_2 \cap a_3^*) = \emptyset \)). In Section 4.4 we saw how these operators can be readily implemented with \( RFA \)'s (without the encoding scheme of this section).

The approach presented in this subsection is essentially due to McNaughton and Yamada [MY60], Glushkov [Glus61], and Berry and Sethi [BS86]. The presentations in [B-K93a, Section 2], [BS86], [tEvG93], and [ASU86, Fig. 3.40, pp. 134-141] are particularly clear. Those interested in a rigorous treatment of this approach to \( RFA \)'s can refer to the paper of ten Eikelder and van Geldrop [tEvG93].

4.5.2 More efficient \( RFA \) operators

The definition of the \( RFA \) operators may still result in inefficient implementation. In particular, Brüggemann-Klein and Chang and Paige found that the implementation of the \((U)\) in the \( RFA \) operators may require more than constant time [B-K93a, Chan92, CP92]. In most cases the arguments (of \( U \)) are disjoint; the only possible exception is the union \( \text{follow} \cup (\text{last} \times \text{first}) \), appearing in the \( C_{+,RFA} \) and \( C_{+,RFA} \) operators. Two solutions to this problem will be presented here.

Convention 4.75 (Constant time union): We use the symbol \( \cup \) to denote union where the arguments to \( \cup \) are assumed to be disjoint.

\(^{11}\) Here we assume, for the moment, that \( L_{RE} \) can deal with the intersection operator.
The first solution was proposed by Chang and Paige [Chang92, CP92].

**Definition 4.76 (Chang-Paige RFA):** We add an eighth component $W$ to each RFA $(Q, V, \text{follow}, \text{first}, \text{last}, \text{null}, Qmap, W)$ such that

$$W = (\text{last} \times \text{first}) \setminus \text{follow}$$

These modified RFA’s will be called Chang-Paige RFA’s, and are denoted by $RFA’$. □

We only give the new operators, instead of the $\Sigma$-algebra. The operators follow directly from the above definition.

**Construction 4.77 (Operators of the $\Sigma$-algebra of Chang-Paige RFA’s):** As usual, the operator requirement is:

- For binary operators, the representatives have disjoint state sets.

\[
C_{\cup, RFA'} = [([\emptyset, V, \emptyset, \emptyset, \text{null}, \text{true}, 0, 0]) \setminus \\
C_{\cap, RFA'} = [([\emptyset, V, \emptyset, \emptyset, \text{false}, 0, 0]) \setminus \\
C_{\cap V, RFA'} = \text{let } q_0 \text{ be a new state} \text{ in} \\
\quad [([q_0], V, \emptyset, \{q_0\}, \{q_0\}, \text{false}, \{(q_0, a)\}, \{(q_0, q_0)\})] \setminus \\
C_{\cap \cap, RFA'}([M_0]_{\Sigma}, [M_1]_{\Sigma}) = \text{let } (Q_0, V, \text{follow}_0, \text{first}_0, \text{last}_0, \text{null}_0, Qmap_0, W_0) = M_0 \text{ in} \\
\quad (Q_1, V, \text{follow}_1, \text{first}_1, \text{last}_1, \text{null}_1, Qmap_1, W_1) = M_1 \\
\quad \text{let } \text{first}' = \text{first}_0 \cup \text{null}_0 \text{ if (null)} \text{ then first}_1 \text{ else } \emptyset \text{ if } \\
\quad \text{last}' = \text{last}_0 \cup \text{null}_0 \text{ if (null)} \text{ then last}_1 \text{ else } \emptyset \text{ if } \\
\quad W' = \text{if (null)} \text{ then } W_0 \text{ else } \emptyset \text{ if } \\
\quad W'' = \text{if (null)} \text{ then } W_1 \text{ else } \emptyset \text{ if } \\
\quad \text{in} \\
\quad [(Q_0 \cup Q_1, V, \text{follow}_0 \cup \text{follow}_1 \cup (\text{last}_0 \times \text{first}_1), \\
\quad \text{first}', \text{last}', \text{null}_0 \land \text{null}_1, Qmap_0 \cup Qmap_1, \\
\quad (\text{last}_1 \times \text{first}_0) \cup W' \cup W'')]_{\Sigma} \\
\quad \text{end} \\
\quad \text{end} \\
C_{\cap \cup, RFA'}([M_0]_{\Sigma}, [M_1]_{\Sigma}) = \text{let } (Q_0, V, \text{follow}_0, \text{first}_0, \text{last}_0, \text{null}_0, Qmap_0, W_0) = M_0 \text{ in} \\
\quad (Q_1, V, \text{follow}_1, \text{first}_1, \text{last}_1, \text{null}_1, Qmap_1, W_1) = M_1 \\
\quad [(Q_0 \cap Q_1, V, \text{follow}_0 \cap \text{follow}_1, \text{first}_0 \cap \text{first}_1, \\
\quad \text{last}_0 \cap \text{last}_1, \text{null}_0 \lor \text{null}_1, Qmap_0 \cup Qmap_1, \\
\quad (\text{last}_1 \times \text{first}_0) \cup (\text{last}_1 \times \text{first}_0) \cup W_0 \cup W_1)]_{\Sigma} \\
\quad \text{end} \\
\quad \text{end} \\
C_{\cap, RFA'}([M]_{\Sigma}) = \text{let } (Q, V, \text{follow}, \text{first}, \text{last}, \text{null}, Qmap, W) = M \text{ in} \\
\quad [(Q, V, \text{follow} \cup W, \text{first}, \text{last}, \text{true}, Qmap \cup \emptyset)]_{\Sigma} \\
\quad \text{end}
4.5 Efficiently computing with RFA's

\[
C_{+, \text{RFA}}([M]_{\infty}) = \begin{array}{l}
\text{let} (Q, V, \text{follow}, \text{first}, \text{last}, \text{null}, \text{Qmap}, W) = M \\
\text{in} \\
[(Q, V, \text{follow} \cup W, \text{first}, \text{last}, \text{null}, \text{Qmap}, \emptyset)]_{\infty} \\
\end{array}
\]

\[
C_{-, \text{RFA}}([M]_{\infty}) = \begin{array}{l}
\text{let} (Q, V, \text{follow}, \text{first}, \text{last}, \text{null}, \text{Qmap}, W) = M \\
\text{in} \\
[(Q, V, \text{follow}, \text{first}, \text{last}, \text{null}, \text{Qmap}, W)]_{\infty} \\
\end{array}
\]

These operators are symmetrical. The correctness of these operators is shown in Theorem B.3. □

Chang and Paige make additional running-time savings by computing the components of a \text{RFA}' only as needed in the operators \text{C}_{+, \text{RFA}} and \text{C}_{-, \text{RFA}}. The running-time and space savings, along with implementation details are given in [Chang92, CP92].

The second solution also involves adding an eighth tuple element to \text{RFA}'s, giving \text{RFA}''.

**Definition 4.78 (RFA''):** We add an eighth component \( W \) to each \text{RFA} 

\[
(Q, V, \text{follow}, \text{first}, \text{last}, \text{null}, \text{Qmap}, W)
\]

such that

\[
W = \text{follow} \setminus (\text{last} \times \text{first})
\]

These modified \text{RFA}'s are denoted by \text{RFA}''. □

As before, the new operators follow directly from the above definition.

**Construction 4.79 (Operators of the \( \Sigma \)-algebra of RFA''):** As usual, the requirement for binary operators is that the representatives of the arguments are chosen such that they have disjoint state sets.

\[
C_{\epsilon, \text{RFA}''} = [(\emptyset, V, \emptyset, \emptyset, \emptyset, \text{true}, \emptyset, \emptyset)]_{\infty}
\]

\[
C_{0, \text{RFA}''} = [(\emptyset, V, \emptyset, \emptyset, \emptyset, \text{false}, \emptyset, \emptyset)]_{\infty}
\]

\[
C_{\epsilon \in \nu, \text{RFA}''} = \begin{array}{l}
\text{let} q_0 \text{ be a new state} \\
\text{in} \\
[(\{q_0\}, V, \emptyset, \{q_0\}, \{q_0\}, \text{false}, \{(q_0, a)\}, \emptyset)]_{\infty} \\
\end{array}
\]

\[
C_{-, \text{RFA}''}([M_0]_{\omega}, [M_1]_{\infty}) = \begin{array}{l}
\text{let} (Q_0, V, \text{follow}_0, \text{first}_0, \text{last}_0, \text{null}_0, \text{Qmap}_0, W_0) = M_0 \\
(Q_1, V, \text{follow}_1, \text{first}_1, \text{last}_1, \text{null}_1, \text{Qmap}_1, W_1) = M_1 \\
\text{in} \\
\text{let} \\
\text{first}' = \text{first}_0 \cup \text{if} (\text{null}_0) \text{ then } \text{first}_1 \text{ else } \emptyset \text{ fi} \\
\text{last}' = \text{last}_0 \cup \text{if} (\text{null}_1) \text{ then } \text{last}_1 \text{ else } \emptyset \text{ fi} \\
W' = \text{if} (\text{null}_1) \text{ then } W_0 \text{ else follow}_0 \text{ fi} \\
W'' = \text{if} (\text{null}_0) \text{ then } W_1 \text{ else follow}_1 \text{ fi} \\
W''' = \text{if} (\text{null}_0 \land \text{null}_1) \text{ then } \emptyset \text{ else } (\text{last}_0 \times \text{first}_1) \text{ fi} \\
\text{in} \\
[(Q_0 \cup Q_1, V, \text{follow}_0 \cup \text{follow}_1 \cup (\text{last}_0 \times \text{first}_1), \\
\text{first}', \text{last}', \text{null}_0 \land \text{null}_1, \text{Qmap}_0 \cup \text{Qmap}_1, \\
W' \cup W'' \cup W''')]_{\infty} \\
\end{array}
\]
CONSTRUCTIONS BASED ON REGULAR EXPRESSION STRUCTURE

\[ C_{U,RFA'}([M_0]_\Xi,[M_1]_\Xi) = \text{let} \begin{align*}
(Q_0,V,\text{follow}_0,\text{first}_0,\text{last}_0,\text{null}_0,\text{Qmap}_0,W_0) &= M_0 \\
(Q_1,V,\text{follow}_1,\text{first}_1,\text{last}_1,\text{null}_1,\text{Qmap}_1,W_1) &= M_1
\end{align*} \text{in} \]
\[ [(Q_0 \uplus Q_1, V, \text{follow}_0 \uplus \text{follow}_1, \text{first}_0 \uplus \text{first}_1, \\
\text{last}_0 \uplus \text{last}_1, \text{null}_0 \lor \text{null}_1, \text{Qmap}_0 \uplus \text{Qmap}_1, \\
W_0 \uplus W_1)]_\Xi \]
\text{end} \]

\[ C_{+,RFA'}([M]_\Xi) = \text{let} \begin{align*}
(Q, V,\text{follow},\text{first},\text{last},\text{null},\text{Qmap},W) &= M \\
\end{align*} \text{in} \]
\[ [(Q, V, W \uplus (\text{last} \times \text{first}), \text{first}, \text{last}, \text{true}, \text{Qmap}, W)]_\Xi \]
\text{end} \]

\[ C_{\gamma,RFA'}([M]_\Xi) = \text{let} \begin{align*}
(Q, V,\text{follow},\text{first},\text{last},\text{null},\text{Qmap},W) &= M \\
\end{align*} \text{in} \]
\[ [(Q, V, \text{follow}, \text{first}, \text{false}, \text{Qmap}, W)]_\Xi \]
\text{end} \]

These operators are symmetrical. Their correctness is shown in Theorem B.4. □

Definition 4.80 (Mapping \([RFA']_\Xi\) and \([RFA''']_\Xi\) to \([RFA]_\Xi\)): The mapping is \(\pi_8\). □

Remark 4.81: Although this construction does not appear in the literature, a related one does: Brüggemann-Klein describes a transformation on regular expressions which closely parallels the \(RFA''\) operators. A regular expression \(E\) is first transformed into a star normal-form expression, denoted by \(E^*\); the \(RFA\) image of \(E^*\) has similar properties to the \(RFA''\) image of \(E\). The details of the star normal-form transformations (and the running time improvements resulting from them) are described in [B-K93a]. □
In this section, we explore a DFA construction method — due to Myhill and Nerode — from which we derive Brzozowski’s construction. First, we make some observations about determinism in finite automata.

Recall a property of weakly deterministic automata — Property 2.28. Given that the set of left languages (of the states) in a DFA are disjoint, we will be exploring methods of computing a set of left languages to construct an automaton.

**Definition 5.1 (Left languages of a DFA):** Define the left languages of a DFA as:

\[ \mathcal{L}(Q, V, T, E, S, F) = \{ \mathcal{L}(q) : q \in Q \} \]

Since the elements of this set are pairwise disjoint (by Property 2.28), we can also view it as the (finite) set of equivalence classes of some equivalence relation on \( V^* \). There are two potential problems with this:

- In the case that an \( M \in DFA \) is not Complete then \( \mathcal{L}(M) \) is a partial partition of \( V^* \). (This follows from Property 2.16.) To make the definitions in this section easier to present, we restrict ourselves to Complete DFA’s.
- It may be that \( \emptyset \) is a left language of some state — corresponding to a start-unreachable state. In this section, we will not be interested in DFA’s with start-unreachable states.

Since \( \mathcal{L}_{PA}(Q, V, T, E, S, F) = (\cup f : f \in F : \mathcal{L}(f)) \) (see Definition 2.12) we also note that the language of an automaton \( M \) is the union of some of the equivalence classes in \( \mathcal{L}(M) \).

**Definition 5.2 (Right invariance of an equivalence relation):** An equivalence relation \( \mathcal{E} \) on \( V^* \) is right-invariant if and only if

\[ (\forall u, a : u \in V^* \land a \in V : (\exists v : v \in V^* : [u]_{\phi} \cdot \{a\} \subseteq [v]_{\phi}) \]

**Property 5.3 (Right invariance of an equivalence relation):** Sometimes right invariance of equivalence relation \( \mathcal{E} \) on \( V^* \) is given as

\[ (\forall u, z : u \in V^* \land z \in V^* : (\exists v : v \in V^* : [u]_{\phi} \cdot \{z\} \subseteq [v]_{\phi}) \]

This is equivalent to the definition given above (by induction on the length of \( z \in V^* \)).

We can now formulate an important property of \( \mathcal{L}(M) \), the partition of \( V^* \).

**Property 5.4 (Right invariance of a partition of \( V^* \)):** Partition \( \mathcal{L}(Q, V, T, E, S, F) \) is right-invariant if and only if

\[ (\forall p, a : p \in Q \land a \in V : (\exists q : q \in Q : \mathcal{L}(p) \cdot \{a\} \subseteq \mathcal{L}(q))) \]

**Remark 5.5:** It should be clear that for all \( M \in DFA \) such that Complete(\( M \)), \( \mathcal{L}(M) \) is right-invariant; this follows since for all states \( p, \) (and transition relation \( T \in Q \times V \rightarrow Q, \) since \( M \in DFA \)) (and \( a \in V \)):

\[ \mathcal{L}(p) \cdot \{a\} \subseteq \mathcal{L}(T(p, a)) \]
Remark 5.6: Had we been considering non-Complete DFA's, we would not have a partition on $V^*$; right invariance could still be defined (for partial partition $\overrightarrow{\mathcal{L}}(Q, V, T, E, S, F)$) as:

$$(\forall p, a : p \in Q \land a \in V \land T^*(p, a) \neq \emptyset : (\exists q : q \in Q : \overrightarrow{\mathcal{L}}(p) \cdot \{a\} \subseteq \overrightarrow{\mathcal{L}}(q))$$

We will not be using this definition. We give it to point out that the techniques of this section are also usable in constructing non-Complete DFA's (as Brzozowski demonstrated in [Brz064]). □

5.1 The Myhill-Nerode construction

Before considering how to construct finite automata, we first present the Myhill-Nerode theorem. A good textbook introduction to the theorem is [HU79].

Theorem 5.7 (Myhill-Nerode): The Myhill-Nerode theorem states that the following statements are equivalent [Myhi57, Nero58, RS59, HU79]:

1. $L$ is a regular language.

2. $L$ is the union of some of the equivalence classes of a right-invariant equivalence relation (on $V^*$) of finite index.

3. Let $R_L$ be the right-invariant equivalence relation defined by

$$(x, y) \in R_L \equiv (\forall z : z \in V^*: (xz \in L) \equiv (yz \in L))$$

Relation $R_L$ is of finite index.

Proof:

A following proof is given in [HU79, Theorem 3.9].

(1) ⇒ (2): Assume $L$ is accepted by $M \in DFA$ such that $Complete(M)$. Let $E$ be the equivalence relation corresponding to $\overrightarrow{\mathcal{L}}(M)$. $\forall E$ is finite, and $L = (\cup f : f \in F : \overrightarrow{\mathcal{L}}(f))$. (See Definition A.8 for the definition of $\forall$.)

(2) ⇒ (3): We show that for an equivalence relation $E$ satisfying (2) that $E \subseteq R_L$. (Here $\subseteq$ denotes equivalence relation refinement, see Definition A.10.) We start the derivation using the right invariance property of $E$ (Property 5.4, written slightly differently):

$$(\forall u : u \in V^* : (\forall w : w \in V^* : ([u]_E \cdot \{w\}) \subseteq [v]_E))$$

$\Rightarrow$

{ Assumption that $E$ satisfies (2), for all $v \in V^*$: $([v]_E \subseteq L) \lor ([v]_E \cap L = \emptyset)$ }

$$(\forall u : u \in V^* : (\forall w : w \in V^* : ([u]_E \cdot \{w\}) \subseteq L) \lor (([u]_E \cdot \{w\}) \cap L = \emptyset))$$

$\Rightarrow$

{ Definition of $R_L$ }

$$(\forall u : u \in V^* : (\exists v : v \in V^* : [u]_E \subseteq [v]_{R_L}))$$

$\equiv$

{ Definition of refinement ($\subseteq$) — Definition A.11 }

$[V^*]_E \subseteq [V^*]_{R_L}$

$\Rightarrow$

{ Definition of refinement ($\subseteq$) — Definition A.10 }

$E \subseteq R_L$

$\Rightarrow$

{ Property of refinement — Property A.12 }

$\forall E \geq \forall R_L$

It follows that since $\forall E$ is finite, so is $\forall R_L$. 
5.1 The Myhill-Nerode construction

(3) ⇒ (1): We can construct the following Complete DFA (from \( R_L \)) accepting \( L \):

\[
\begin{align*}
\text{let } & T([w]_{R_L}, a) = \{[wa]_{R_L}\} \\
\text{in } & F = \{[w]_{R_L} : w \in L\} \\
\text{end}
\end{align*}
\]

\( ([V^*]_{R_L}, V, T, \emptyset, \{[e]_{R_L}\}, F) \)

It follows that \( L \) is a regular language.

\[ \square \]

Property 5.8 (Index of equivalence relation \( E \)): Given \( L \in \mathcal{L}_{\text{Reg}} \), \( M \in \text{DFA} \) accepting \( L \), and \( E \) (satisfying statement 2 and constructed as in the proof of (1) ⇒ (2) of the Myhill-Nerode theorem) then \( \#E \leq |M| \) (since \( \#E = \#\mathcal{L}(M) \) by definition of \( E \) and \( \text{Det}' \)). \( \square \)

Property 5.9 (Uniqueness and minimality of equivalence relation \( R_L \)): Of all equivalence relations satisfying statement 2 of the Myhill-Nerode theorem, \( R_L \) is the unique minimal one. This follows from the fact that all others are refinements of \( R_L \) (see the proof of (2) ⇒ (3)). \( \square \)

The theorem does not say much about how to find equivalence relations satisfying statement 2, other than providing a definition of the unique minimal one, \( R_L \).

We can formalize statement 2 of the Myhill-Nerode theorem:

Definition 5.10 (Predicate \( MN \)): For regular language \( L \) and equivalence relation \( E \) (on \( V^* \)) \( MN(L, E) \) is equivalent to

- \( \#E \) is finite,
- \( L = (\cup v : v \in L : [v]_E) \), and
- \( E \) is right-invariant.

\[ \square \]

Note that \( MN(L, R_L) \).

The DFA construction given in the (3) ⇒ (1) proof can be used with other right-invariant equivalence relations.

Construction 5.11 (Myhill-Nerode): Given a language \( L \) and right-invariant equivalence relation \( E \) such that \( MN(L, E) \) we can construct an automaton accepting \( L \) using the function

\[
\begin{align*}
\text{MNC} & \text{onstr}(L, E) = \text{let } & T([w]_E, a) = \{[wa]_E\} \\
\text{in } & F = \{[w]_E : w \in L\} \\
\text{end}
\end{align*}
\]

\( ([V^*]_E, V, T, \emptyset, \{[e]_E\}, F) \)

This construction has the following properties:

- The definition is independent of the choice of representatives of the equivalence classes of \( E \).
- By inspection we can see that the FA constructed by \( \text{MNC} \) is a Complete DFA.
- For any state \( U \in [V^*]_E \) we have \( \mathcal{L}(U) = U \).
- All states in the constructed automaton are start-reachable.
- The number of states is \( \#E \).
- The construction satisfies the property

\( (\forall L, E : MN(L, E) : \mathcal{L}_{\text{FA}}(\text{MNC}onstr(L, E)) = L) \)

\[ \square \]
5.2 The minimal equivalence relation $R_L$

The only relation (corresponding to $L \in \mathcal{L}_{\text{reg}}$) that Myhill and Nerode actually defined was $R_L$. This relation is particularly important, being the unique one of minimal index.

**Theorem 5.12 (Unique minimal DFA):** Given $L \in \mathcal{L}_{\text{reg}}$, $M = MN_{\text{constr}}(L, R_L)$ is the unique minimal Complete DFA accepting $L$; that is, $\text{Minc}(M)$.

**Proof:**
Assume there exists $M' \in \text{DFA}$ such that $\text{Complete}(M')$, $\mathcal{L}_{\text{FA}}(M') = L$ and $|M'| \leq |M|$. From the proof of $(2) \Rightarrow (3)$ (and Property 5.8), $\sharp R_L \leq \sharp \mathcal{L}(M')$. In summary, $\sharp R_L \leq \sharp \mathcal{L}(M') \leq |M'| \leq |M| = \sharp R_L$, and so (by Property A.12) $|M'| = |M|$, $E = R_L$ and $M' \equiv M$. □

**Property 5.13 (Reformulating $R_L$):** We can rewrite the definition of $R_L$ using derivatives (see Definition A.15) as follows:

$$(x, y) \in R_L \quad \equiv \quad \{ \text{Property of derivatives and definition of } R_L \}$$

$$\quad \equiv \quad \{ \forall z : z \in V^* : (z \in x^{-1}L) \equiv (z \in y^{-1}L) \}$$

$$\quad \equiv \quad \{ \text{Definition of } = \text{ on languages} \}$$

$$x^{-1}L = y^{-1}L$$

□

We could combine this definition of $R_L$ with $MN_{\text{constr}}$ to get a minimal DFA construction. Such a function would have a clumsy definition, and therefore we explore some encoding tricks.

5.2.1 Encoding $R_L$

An encoding trick is hinted at by Property 5.13: every equivalence class $[w]_{R_L}$ of $R_L$ can be characterized by the language $w^{-1}L$.

**Definition 5.14 (Derivative set of a language):** We define the set of derivatives of language $L$ as

$$\text{deriv}(L) = \{ v^{-1}L : v \in V^* \}$$

□

We have the following theorem relating to $\text{deriv}$

**Theorem 5.15 (Finiteness of derivatives):** If $L \in \mathcal{L}_{\text{reg}}$ then $|\text{deriv}(L)|$ is finite.

**Proof:**
$\sharp R_L$ is finite (from the Myhill-Nerode theorem), and since $|\text{deriv}(L)| = \sharp R_L$, $|\text{deriv}(L)|$ is also finite. □

This theorem has also been given by Brzozowski [Brzo64]. His proof is, however, somewhat more complicated, and is by induction on the structure of language $L$.

**Definition 5.16 (Encoding an equivalence class):** We define a derivative encoding function (for a given $L \in \mathcal{L}_{\text{reg}}$) $\text{encderiv}_L : [V^*]_{R_L} \rightarrow \text{deriv}(L)$ as

$$\text{encderiv}_L([w]_{R_L}) = w^{-1}L$$

This function has inverse $\text{encderiv}_L^{-1}(v^{-1}L) = [v]_{R_L}$. Both of these functions are independent of the choice of representative of equivalence class of $R_L$. □
5.2 The minimal equivalence relation $R_L$

Remark 5.17: In Construction 5.11 (with $R_L$ as the equivalence class parameter) the equivalence classes of $R_L$ are the left languages of the states of a DFA constructed from $R_L$. The function $\text{encderiv}_L$ maps these left languages (equivalence classes) to their corresponding right languages (the derivatives). The right language $(\text{encderiv}_L(U))$ of a particular equivalence class $U \in [V^*]_{R_L}$ is called the continuation of $U$ (in language $L$) by Berry and Sethi [BS86].

Property 5.18 (Derivatives and function $\text{encderiv}_L$): Note that (for $a \in V$, $w \in V^*$):

- $\text{encderiv}_L([e]_{R_L}) = \epsilon^{-1}L = L$,
- $\text{encderiv}_L([wa]_{R_L}) = (wa)^{-1}L = a^{-1}(w^{-1}L)$.

These properties follow from Property A.16, and the definition of $\text{encderiv}_L$. □

Noting the form of function $\text{MNconstr}$, we use the encoding (function $\text{encderiv}_L$) to obtain construction $\text{MNmin} \in \mathcal{L}_{\text{reg}} \rightarrow \text{DFA}$.

Construction 5.19 ($\text{MNmin}$): Combining $\text{MNconstr}$ (Construction 5.11) with $\text{encderiv}_L$ (and its inverse) gives construction $\text{MNmin} \in \mathcal{L}_{\text{reg}} \rightarrow \text{DFA}$:

$$\text{MNmin}(L) = \text{let } T(w^{-1}L,a) = \{a^{-1}(w^{-1}L)\}$$
$$F = \{w^{-1}L : \epsilon \in w^{-1}L\}$$
$$\text{in}$$
$$\text{deriv}(L), V, T, \emptyset, \{L\}, F$$
$$\text{end}$$

Since $\text{MNmin}$ is defined using $\text{MNconstr}$, the properties are similar:

- By inspection we can see that the FA constructed by $\text{MNmin}$ is a Complete DFA.
- For any state $U \in \text{deriv}(L)$ we have $\overline{L}(U) = U$.
- All states in the constructed automaton are start-reachable.
- The only state that is not final-reachable is $\emptyset$. The state $\emptyset$ exists in automaton $\text{MNmin}(L)$ if and only if $L \neq V^*$. It follows that we can remove the sink state $\emptyset$, to obtain a (possibly) non-Complete DFA with only useful states.
- The constructed DFA is the unique (up to isomorphism) minimal Complete DFA accepting $L$ (since $R_L$ is implicit in the definition).
- The construction satisfies the property $(\forall L : L \in \mathcal{L}_{\text{reg}} : L_{FA}(\text{MNmin}(L)) = L)$

□

Example 5.20 ($\text{MNmin}$ construction): We construct the minimal Complete DFA corresponding to the regular language $\{e,a\}b^*$, denoted by regular expression $(a \cup e)b^*$ (the regular expression from Example 3.15).

After some calculation (using Property A.17):

- $a^{-1}(\{e,a\}\{b\}^*) = (a^{-1}\{e,a\})\{b\}^* \cup a^{-1}\{b\}^* = \{e\}\{b\}^* \cup (a^{-1}\{b\})\{b\}^* = \{b\}^*$
- $b^{-1}(\{e,a\}\{b\}^*) = (b^{-1}\{e,a\})\{b\}^* \cup b^{-1}\{b\}^* = \emptyset\{b\}^* \cup (b^{-1}\{b\})\{b\}^* = \{b\}^*$
Figure 8: The DFA \( \text{MNmin}(\{e,a\}\{b\}^*) \).

\[
\begin{align*}
    a^{-1}(\{b\}^*) &= (a^{-1}\{b\})\{b\}^* \\
                    &= \emptyset \{b\}^* \\
                    &= \emptyset \\
    b^{-1}(\{b\}^*) &= (b^{-1}\{b\})\{b\}^* \\
                    &= \{b\}^* \\
    a^{-1} \emptyset &= \emptyset \\
    b^{-1} \emptyset &= \emptyset
\end{align*}
\]

we determine that the three derivatives are: \( L = \{e,a\}\{b\}^* \), \( L_1 = \{b\}^* \), and \( \emptyset \). The state graph is shown in Figure 8. □

5.3 The Brzozowski construction

We now concentrate on constructing DFA’s from extended regular expressions, as opposed to constructing them from regular languages. In Property A.17, a method is given for computing a derivative of a regular language (based upon the structure of the language). Being able to compute derivatives in this way also provides us with a definition of derivatives of extended regular expressions (EREs). Extended regular expressions were defined in Definition 4.53.

Remark 5.21: The \( \Sigma \)-algebra definition of regular expressions is not used in this section as the algebraic structure is not needed. Regular expressions are used only as syntactic objects, denoting regular languages. □

Remark 5.22: The remaining constructions in this section do not necessarily depend on extended regular expressions (normal regular expressions can also be used). They are introduced because some regular languages have more succinct descriptions as ERE’s than as RE’s. □

Definition 5.23 (Derivatives of ERE’s): Assuming \( a \in V \) and \( E, E_0, E_1 \in ERE \)

\[
\begin{align*}
    a^{-1} \emptyset &= \emptyset \\
    a^{-1} e &= \emptyset \\
    a^{-1} b &= \begin{cases} 
        \emptyset & \text{if } (a = b) \text{ then } e \text{ else } \emptyset \text{ fi} 
    \end{cases} \quad \text{(for all } b \in V) \\
    a^{-1}(E_0 E_1) &= (a^{-1} E_0) E_1 \cup \begin{cases} 
        \emptyset & \text{if } (e \in \mathcal{L}_{ERE}(E_0)) \text{ then } a^{-1} E_1 \text{ else } \emptyset \text{ fi} 
    \end{cases} \\
    a^{-1}(E_0 \cup E_1) &= (a^{-1} E_0) \cup (a^{-1} E_1)
\end{align*}
\]
5.3 The Brzozowski construction

\[
\begin{align*}
 a^{-1}(E^*) &= (a^{-1}E)E^* \\
 a^{-1}(E^+) &= (a^{-1}E)E^* \\
 a^{-1}(E^0) &= a^{-1}E \\
 a^{-1}(E_0 \cap E_1) &= (a^{-1}E_0) \cap (a^{-1}E_1) \\
 a^{-1}(\neg E) &= \neg(a^{-1}E)
\end{align*}
\]

\[\blacksquare\]

**Property 5.24 (Derivatives of \(ER^E\)'s):** By inspecting the definition of derivatives of \(ER^E\)'s, we can verify that (for \(a \in V, E \in ER^E\) \(a^{-1} L_{ER^E}(E) = L_{ER^E}(a^{-1}E)\). \[\blacksquare\]

**Remark 5.25:** Given equivalence relation \(\equiv\) (equivalence of regular expressions, extended to \(ER^E\)), an \(E \in ER^E\) will have a finite number of derivatives. More formally, \(|\{w^{-1}E : w \in V^*\}|\) is finite. Brzozowski gave the same result by induction on the structure of a \(ER^E\) in [Brz64, Theorem 4.3a]. \[\blacksquare\]

**Definition 5.26 (Similarity \((\sim)\) of regular expressions):** Similarity (written \(\sim\)) is an equivalence relation on \(ER^E\)'s. Two \(ER^E\)'s are similar if and only if they are identical or one can be transformed into the other using the following rules:

1. \(E_0 \cup E_1 = E_1 \cup E_0\) (commutativity of \(\cup\)),
2. \(E_0 \cup (E_1 \cup E_2) = (E_0 \cup E_1) \cup E_2\) (associativity of \(\cup\)), and
3. \(E \cup E = E\) (idempotence of \(\cup\)).

\[\blacksquare\]

**Property 5.27 (Similarity):** \(\sim \subseteq \equiv\); that is, \(\sim\) is a refinement of \(\equiv\). \[\blacksquare\]

**Definition 5.28 (The derivatives of an \(ER^E\)):** Function \(deriv_{ER^E} \in ER^E \rightarrow \mathcal{P}([ER^E]_\sim)\) is defined as

\[deriv_{ER^E}(E) = \{v^{-1}E : v \in V^*\}_\sim = \{v^{-1}E : v \in V^*\}\]

\[\blacksquare\]

Before proving that \(deriv_{ER^E}(E)\) is finite (for all \(E \in ER^E\)), we need the following proposition.

**Proposition 5.29 (Similarity equivalence class of a union \(ER^E\)):** Assume a finite set \(H \subseteq ER^E\) and a fully parenthesized regular expression

\[J = (\cdots (((h_1 \cup h_2) \cup h_3) \cup h_4) \cdots) \cup h_k)\]

where (for \(1 \leq i \leq k\) \(h_i \in H\); each \(h_i\) is called a term of \(J\). Using \(\sim\) we can always find a similar (and, of course, equivalent) regular expression \(K\), where \(K\) is the union of at most \(|H|\) terms of \(J\). This is because the rules defining \(\sim\) can be used to reassociate and commute the terms of \(J\) (to place identical terms adjacent to one another), while the idempotence rule of \(\sim\) can be used to remove identical terms. \[\blacksquare\]

**Proposition 5.30 (Similarity):** Given a finite set \(H \subseteq ER^E\), the set of all non-similar \(ER^E\)'s that are unions of terms \(h_i \in H\) is finite. \[\blacksquare\]

**Theorem 5.31 (Finiteness of derivatives under similarity):** For all \(E \in ER^E\), \(|deriv_{ER^E}(E)|\) is finite.

**Proof:** This proof is similar to the one given in Brzozowski's original paper [Brz64, Theorem 5.2]. The proof is by induction on the number of operators in \(E \in ER^E\).
Basis: The theorem is true for each of the constants: \( \text{deriv}_{ER}(\epsilon) = \{[\epsilon], [\emptyset]\} \), \( \text{deriv}_{ER}(\emptyset) = \{[\emptyset]\} \), and \( \text{deriv}_{ER}(a) = \{[a], [\epsilon], [\emptyset]\} \) (for \( a \in V \)).

Induction hypothesis: Assume that \( |\text{deriv}_{ER}(E)| \) is finite for all \( E \in ERE \) where \( E \) has fewer than \( k \) operators.

Induction step: Assume \( E \in ERE \) has \( k \) operators. We use case analysis to deal with the possible forms of \( E \):

- \( E = E_0 \cup E_1 \): It is possible to show that
  \[
  \{ [w^{-1}(E_0 \cup E_1)] : w \in V^* \} = \{ [w^{-1}E_0 \cup w^{-1}E_1] : w \in V^* \}
  \]
  By the induction hypothesis, \( |\text{deriv}_{ER}(E_0)| \) and \( |\text{deriv}_{ER}(E_1)| \) are finite, and so is \( |\text{deriv}_{ER}(E)| \).

- \( E = E_0 \cap E_1 \) or \( E = E_0' \) or \( E = \neg E_0 \): An argument similar to that for \( \cup \) applies to these cases.

- \( E = E_0 \cdot E_1 \): In order to analyze \( |\text{deriv}_{ER}(E_0 \cdot E_1)| \), we consider a particular \( w^{-1}(E_0 \cdot E_1) \) (for \( w \in V^* \)). Let \( w = a_1 \cdots a_n \) where each \( a_i \in V \). Writing out \( w^{-1}(E_0 \cdot E_1) \) we get
  \[
  w^{-1}(E_0 \cdot E_1) = (a_2 \cdots a_n)^{-1}(a_1^{-1}E_0) \cup \text{if } (\epsilon \in L_{ER}(E_0)) \text{ then } a_1^{-1}E_1 \text{ else } \emptyset \text{ fi}
  \]
  Had we been able to continue this rewriting, we would see that \( [w^{-1}(E_0 \cdot E_1)] \) is equal to
  \[
  [(w^{-1}E_0) \cdot E_1 \cup (w \cup u \cdot uv : w = w : \text{if } (\epsilon \in L_{ER}(w^{-1}E_0)) \text{ then } v^{-1}E_1 \text{ else } \emptyset \text{ fi})].
  \]
  That is, \( w^{-1}(E_0 \cdot E_1) \) is the union of a set of terms, one of which is \( (w^{-1}E_0) \cdot E_1 \), and the remaining ones are either a derivative of \( E_1 \), or \( \emptyset \in ERE \). By the induction hypothesis (the set of derivatives of \( E_1 \) is finite) the set of possible terms is finite. It follows from Propositions 5.29 and 5.30 that \( |\text{deriv}_{ER}(E)| \) is finite.

- \( E = E_0^* \) or \( E = E_0^+ \): As in the \( E_0 \cdot E_1 \) case, we could write out \( w^{-1}(E_0^*) \) for a particular \( w \in V^* \). If we do this, we see that it is the union of terms, each of which is a derivative of \( E_0 \) concatenated with \( E_0^* \). The set of possible terms is finite — by the induction hypothesis. Again, it follows from Propositions 5.29 and 5.30 that \( |\text{deriv}_{ER}(E)| \) is finite.

\( \square \)

**Remark 5.32:** Unfortunately, using similarity (in computing derivatives) may yield more derivatives than recognizing equivalence (=) of derivatives (as shown in Example 5.37). The rules defining similarity can be augmented with others to decrease the redundancy of the derivatives (and therefore the size of the constructed DFA). Any equivalence relation \( G \) such that \( \sim \subseteq G \subseteq = \) is usable for this. Examples of additional rules are (for \( E_0, E_1, E_2 \in ERE \)):

1. \( E_0 \cdot \emptyset = \emptyset \) (\( \emptyset \) is the zero of concatenation),
2. \( E_0 \cup \emptyset = E_0 \) (\( \emptyset \) is the unit of \( \cup \)),
3. \( E_0 \cdot \epsilon = E_0 \) (\( \epsilon \) is the unit of concatenation),
4. \( \emptyset^* = \epsilon \) (a property of *),
5. \( E_0 \cdot (E_1 \cup E_2) = E_0 \cdot E_1 \cup E_0 \cdot E_2 \) (\( \cdot \) distributes over \( \cup \)),
6. \( E_0 \cap \emptyset = \emptyset \) (\( \emptyset \) is the zero of \( \cap \)),
7. \( E_0 \cap E_1 = E_1 \cap E_0 \) (commutativity of \( \cap \)),
8. \( E_0 \cdot E_1 = E_1 \cdot E_0 \) (associativity of \( \cdot \)),
5.3 The Brzozowski construction

8. \( E_0 \cap (E_1 \cap E_2) = (E_0 \cap E_1) \cap E_2 \) (associativity of \( \cap \)).

9. \( E \cap E = E \) (idempotence of \( \cap \)).

There is a property of similarity that will be needed to present the Brzozowski construction.

Definition 5.33 (Derivative of a similarity equivalence class): For \( E \in ERE \) and \( a \in V \) we have \( a^{-1}[E]_\sim = [a^{-1}E]_\sim \). This definition does not depend on the choice of representative of the equivalence class (under \( \sim \)). □

The Brzozowski construction is an encoding of \( MNmin \) to use \( ERE \) and equivalence classes of \( \sim \).

Construction 5.34 (Brzozowski): Function \( Brz \in ERE \rightarrow DFA \) is defined as:

\[
Brz(E) = \text{let } T([w^{-1}E]_\sim, a) = \{a^{-1}[w^{-1}E]_\sim\} \\
F = \{[w^{-1}E]_\sim : \epsilon \in \mathcal{L}_{ERE}(w^{-1}E)\} \\
in \{\text{deriv}_{ERE}(E), \mathcal{L}, \emptyset, ([E]_\sim), F\} \\
end
\]

The properties of \( Brz \) correspond to those of \( MNmin \):

- The construction is independent of the representatives of equivalence classes.
- By inspection we can see that \( Brz \) constructs Complete DFA's.
- For any state \( E' \in \text{deriv}_{ERE}(E) \) we have \( \mathcal{L}(E') = \mathcal{L}_{ERE}(E') \).
- All states in the constructed automaton are start-reachable.
- There may be a state that is not final-reachable; this sink state will exists if and only if \( \mathcal{L}_{ERE}(E) \neq V^* \). The sink state corresponds to the derivative \( \emptyset \in ERE \).
- The construction satisfies the property

\[
(\forall E : E \in ERE : \mathcal{L}_{FA}(Brz(E)) = \mathcal{L}_{ERE}(E))
\]

□

Remark 5.35: Any equivalence relation \( G \) (on \( ERE \)'s) such that \( \sim \subseteq G \subseteq \sim \) can be used in place of \( \sim \) in Brzozowski's construction. □

Remark 5.36: In Brzozowski's original paper [Brzo64], the sink state (corresponding to derivative \( \emptyset \in ERE \)) was always omitted from the constructed DFA, producing a possibly non-Complete DFA. □

Example 5.37 (Brzozowski's construction): We construct a Complete DFA corresponding to regular expression \((a \cup e)b^*\) (the regular expression from Example 3.15). The derivatives are:

\[
\begin{align*}
a^{-1}((a \cup e)b^*) & = (a^{-1}(a \cup e))b^* \cup a^{-1}(b^*) \\
& = (a^{-1}a \cup a^{-1}e)b^* \cup (a^{-1}b)b^* \\
& = (e \cup \emptyset)b^* \cup \emptyset b^*
\end{align*}
\]

\[
\begin{align*}
b^{-1}((a \cup e)b^*) & = (b^{-1}(a \cup e))b^* \cup b^{-1}(b^*) \\
& = (b^{-1}a \cup b^{-1}e)b^* \cup (b^{-1}b)b^* \\
& = (\emptyset \cup \emptyset)b^* \cup \emptyset b^*
\end{align*}
\]
The four derivatives (under $\sim$) are: $d_0 = (a \cup e)b^*$, $d_1 = (e \cup \emptyset)b^* \cup \emptyset b^*$, $d_2 = \emptyset b^* \cup eb^*$, and $d_3 = \emptyset b^*$. The state graph is shown in Figure 9. Had we been able to recognize equivalence of ERE's, we would have had a smaller DFA since $(e \cup \emptyset)b^* \cup \emptyset b^* = \emptyset b^* \cup eb^*$, and we could have identified states $d_1$ and $d_2$. 

5.3.1 Computing derivatives of an ERE

Brzozowski also shows [Brz064] if $E \in ERE$ has $n$ derivatives (including $E$, under any equivalence relation $G$ such that $\sim \subseteq G \subseteq \equiv$) then they are all of the form $v^{-1}E$ where $|v| < n$ [Brz064, Theorem 4.3b]. Also part of this theorem is if all derivatives (of $E$) with respect to strings of length not greater than $n$ have been found, and no new ones are found with respect to strings of length $n + 1$, then no new ones will be found with respect to strings of length greater than $n$. This useful property of derivatives (in fact a slightly stronger property) can be stated as follows:

![Figure 9: The DFA $Brz((a \cup e)b^*)$.](image-url)
5.4 Relating the Brzozowski and Berry-Sethi constructions

Theorem 5.38 (Finding derivatives): For all $r \geq 0$

\[
\{[w^{-1}E]_\sim : w \in V^* \land |w| = r\} \subseteq \{[w^{-1}E]_\sim : w \in V^* \land |w| < r\}
\]

\[
\Rightarrow \quad \{[w^{-1}E]_\sim : w \in V^* \land |w| \geq r\} \subseteq \{[w^{-1}E]_\sim : w \in V^* \land |w| < r\}
\]

This gives the following algorithm (in the guarded commands of [Dijk76]) which computes the derivatives of a regular expression $E$ ($D \subseteq [ERE]_\sim$ and $next \subseteq [ERE]_\sim$):

Algorithm 5.39:

\[
\{E \in ERE\}
\]

\[
D, next, k := \emptyset, \{[E]_\sim\}, 0;
\]

\[
\{\text{invariant: } D = \{[w^{-1}E]_\sim : w \in V^* \land |w| < k\} \land next = \{[w^{-1}E]_\sim : w \in V^* \land |w| = k\}\}
\]

\[
\text{do next} \notin D \rightarrow
\]

\[
D, next, k := D \cup next, \{a^{-1}F : a \in V \land F \in next\}, k + 1
\]

\[
\text{od}\{D = \text{deriv}_{ERE}(E)\}
\]

5.3.2 Extending derivatives

It is sometimes useful to extend derivatives to deal with additional operators: prefix closure and certain functions on languages. We now briefly give the definition of derivatives of regular languages (and thus regular expressions) with these operators.

The prefix closure of a language is defined as:

\[
\text{pref}(L) = \{u : u^{-1}L \neq \emptyset\}
\]

and the derivative of a prefix closed language is:

\[
a^{-1}(\text{pref}(L)) = \text{pref}(a^{-1}L)
\]

For certain functions $f \in L_{\text{reg}} \times L_{\text{reg}} \rightarrow L_{\text{reg}}$, derivatives are defined as:

\[
a^{-1}(f(L_0, L_1)) = f(a^{-1}L_0, a^{-1}L_1)
\]

Some examples of such functions are $\cap$, $\cup$, asymmetrical difference, and symmetrical difference. For more on this see [Brz064].

5.4 Relating the Brzozowski and Berry-Sethi constructions

It turns out that for $RRE$s (recall from Definition 4.58 that an $E \in RRE$ is an $RE$ such that each symbol of $V$ occurs at most once in $E$), the Brzozowski construction (with sink state removal — as in Brzozowski’s original paper — and a suitable encoding) and the Berry-Sethi construction produce isomorphic $DFA$’s. In this section, we consider only $RRE$s. Berry and Sethi first presented this result in [BS86].

We will be using the following version of Brzozowski’s construction (for $E \in RRE$), which does not introduce a sink state (the sink state is equivalent $(\sim)$ to $\emptyset \in RRE$ — its language under $L_{RE}$ is $\emptyset$).

Construction 5.40 (Brzozowski — without sink state): Given $E \in RE$, the following constructs a DFA accepting $L_{RE}(E)$:

\[
\text{let } Q = \{[w^{-1}E]_\sim : w \in V^* \land L_{RE}(w^{-1}E) \neq \emptyset\}
\]

\[
T([v^{-1}E]_\sim, a) = \begin{cases} a^{-1}[v^{-1}E]_\sim & \text{if } (L_{RE}(a^{-1}(v^{-1}E)) \neq \emptyset) \text{ then } \{a^{-1}[v^{-1}E]_\sim\} \text{ else } \emptyset \end{cases}
\]

\[
F = \{[w^{-1}E]_\sim : \text{Null}(w^{-1}E)\}
\]

\[
in (Q, V, T, \emptyset, ([E]_\sim), F)
\]

end
In the let clause, the transition function has signature $T : Q \times V \rightarrow \mathcal{P}(Q)$. Recall from Definition 3.20 that $\text{Null}(E) \equiv \varepsilon \in \mathcal{L}(E)$. □

For any $E \in RRE$, the only way that $\mathcal{L}(b^{-1}((wa)^{-1}E)) \neq \emptyset$ is if $\mathcal{L}(wa)^{-1}E \neq \emptyset$ and a $b$ can follow a $wa$ in some string in $\mathcal{L}(E)$.

In order to make the above construction practical, we explore the possibility of characterizing all of the derivatives of an $E \in RE$ (except for the derivative $E$ itself) by the symbols occurring in $E$.

**Definition 5.41 (Unambiguous regular expressions):** An $E \in RE$ is said to be unambiguous if and only if, for all $a \in \text{Occ}(E)$ (function $\text{Occ}$ is defined in Definition 4.57), the following set is a singleton set:

$$\{((wa)^{-1}E) : w \in V^* \land \mathcal{L}(wa)^{-1}E \neq \emptyset\}$$

In other words, all derivatives of $E$ by $wa$ (for $w \in V^*$ and $a \in V$) are either equivalent to $\emptyset \in RRE$, or are similar to one another. □

**Remark 5.42:** If an $E \in RE$ is unambiguous, its derivatives are either $E$ or $\emptyset$, or can be characterized by an element of $\text{Occ}(E)$. □

**Remark 5.43:** The regular expression $(a \cup a)$ is unambiguous, but is not an $RRE$. □

**Remark 5.44:** Unambiguous regular expressions are also defined by Champarnaud [Cham93], although he characterizes them quite differently, and he does not make use of derivatives. Champarnaud calls such regular expressions local. □

**Theorem 5.45 (Characterizing derivatives of $RRE$'s):** For any $E \in RRE$, $E$ is unambiguous. This theorem is also given by Berry and Sethi [BS86, Theorem 3.4].

**Proof:**
We proceed by induction on the number of operators in $E \in RRE$.

**Basis:** The theorem is trivially true for the $RRE$ base cases $\varepsilon$ and $\emptyset$ since $\text{Occ}(\varepsilon) = \text{Occ}(\emptyset) = \emptyset$.

It is also trivially true for the $RRE$ $a \in V$.

**Induction hypothesis:** Assume that the theorem is true for any $E \in RRE$ with fewer than $k$ operators.

**Induction step:** We now consider $E \in RRE$ with $k$ operators. We now examine the possible structure of $E$ (assuming $a \in \text{Occ}(E)$):

$E = E_0 \cup E_1$: Given $w \in V^*$ such that $\mathcal{L}(wa)^{-1}E \neq \emptyset$

$$(wa)^{-1}(E_0 \cup E_1) = ((wa)^{-1}E_0) \cup ((wa)^{-1}E_1)$$

Since $E \in RRE$, then either $a \in \text{Occ}(E_0)$ or $a \in \text{Occ}(E_1)$ (but not both). It follows that $(wa)^{-1}(E_0 \cup E_1)$ is similar to $(wa)^{-1}E_0 \cup \emptyset$ or similar to $(wa)^{-1}E_1 \cup \emptyset$ (but not both). The theorem then follows from the induction hypothesis.

$E = E_0 \cdot E_1$: From Theorem 5.31 we know that (for $w \in V^*$, $a \in V$) $[(wa)^{-1}(E_0 \cdot E_1)]$ is equal to:

$$[((wa)^{-1}E_0)E_1 \cup (\cup u, v : uwa = wa : \text{if } \varepsilon \in \mathcal{L}(u^{-1}E_0) \text{ then } (va)^{-1}E_1 \text{ else } \emptyset)]$$

Since $E \in RRE$, then either $a \in \text{Occ}(E_0)$ or $a \in \text{Occ}(E_1)$ (but not both). It follows, by an argument similar to the $E = E_0 \cup E_1$ case (above), that the theorem holds from the induction hypothesis.

$E = E_0^+$, $E = E_0^*$ or $E = E_0^0$: The argument for these cases proceeds similarly to the $E = E_0 \cdot E_1$ case. For more on this type of argument see Theorem 5.31.
5.5 Towards DeRemer's construction

Remark 5.46: The above theorem implies that each derivative of $E \in RRE$ (under similarity) is either $\emptyset$ (which we ignore since it corresponds to the sink state) or $E$, or it can be characterized by an $a \in Occ(E)$. □

Definition 5.47 (Encoding the derivatives of an $RRE$): Given the theorem above, $E \in RRE$, we can now give a partial encoding function $ene_E$ (corresponding to the $E \in RRE$) from the derivatives of $E$ to $Occ(E)$ such that (for $w \in V^*$, $a \in V$) $ene_E((wa)^{-1}E) = a$ when $L_{RE}((wa)^{-1}E) \neq \emptyset$, and $ene_E$ is undefined otherwise. Note that this function is not defined on $E$. □

The following property makes use of the definitions of $Null$, $Occ$, $First$, $Last$, and $Follow$ (Definitions 3.20, 4.57, 4.60-4.64 respectively).

Property 5.48 (Functions $Follow$, $First$, and $Last$): The following properties will be used:

- $(a, b) \in Follow(E) \equiv (\exists w \in V^* : L_{RE}(b^{-1}((wa)^{-1}E)) \neq \emptyset)$.
- $a \in First(E) \equiv L_{RE}(a^{-1}E) \neq \emptyset$.
- $a \in Last(E) \equiv L_{RE}(Ea^{-1}) \neq \emptyset \equiv (\exists w \in V^* : \epsilon \in L_{RE}((wa)^{-1}E))$.

Derivatives on the right are mentioned in Definition A.15. □

We can rewrite our sink stateless version of Brzozowski's construction, using the above properties, to obtain the construction now following.

Construction 5.49 (Encoding Brzozowski for $RRE$'s): We can now give our encoded version (using $Occ$, $First$, $Last$, $Null$, and $Follow$) of Construction 5.40, as $Brzenc E \rightarrow DFA$:

$$Brzenc(E) = \begin{array}{l}
\text{let } s \text{ be a new state (characterizing } E \in RRE) \\
\text{in} \\
\text{let } T = \{(a, b, b) : (a, b) \in Follow(E)\} \\
T' = \{(s, a, a) : a \in First(E)\} \\
F = Last(E) \cup \{s\} \text{ if } (Null(E)) \text{ then } \{s\} \text{ else } \emptyset \text{ fi} \\
\text{in} \\
(Occ(E) \cup \{s\}, V, T \cup T', \emptyset, \{s\}, F) \\
\end{array}$$

Remark 5.50: Using the set $Occ(E)$ as the set of states can yield a DFA with start-unreachable states. For example, in the DFA $Brzenc(\emptyset \cdot a)$, we have start-unreachable state $a$. □

Remark 5.51: By inspection we see that, for all $E \in RRE$, $Brzenc(E) \equiv BSenc(E)$ (Construction 4.70). It follows from Remark 4.73, that for $E \in RRE$, $[Brz(E)]_{\omega} = MYG(E)$. □

Remark 5.52: Finally, we note that the construction $Brzenc$ produces a correct DFA for any $E \in RE$ such that $E$ is unambiguous. That is, $E \in RRE$ is not required. This property is not noted in the literature. This follows from Definition 5.41 and the definition of $Brzenc$. □

5.5 Towards DeRemer's construction

In this subsection, we consider several more constructions based upon the $MNconstr$ and $MNmin$ constructions (Constructions 5.11 and 5.19). The idea is to characterize the derivatives of a regular expression by so-called dotted regular expressions. We only consider constructing a DFA from an $RE$, as opposed to an $ERE$. Since some of the proofs are tedious to present, we give this construction in an informal manner.

We begin by introducing dotted regular expressions, which are essentially regular expressions with a dot (\*) appearing in each of them.
Remark 5.53: We will be characterizing derivatives, not equivalence classes (as is required for \( MN\text{constr} \)). Using dotted \( RE\)'s, it is considerably easier to characterize the derivatives than the equivalence classes. □

The dot should not be confused with the concatenation dot, the star normal-form dot of Brüggemann-Klein (presented in Section 4.5), or the bullet used in typesetting lists.

Definition 5.54 (Dotted regular expressions, their languages, and \( \text{undot} \)): We recursively define dotted regular expressions (\( DRE\)s), function \( \mathcal{R} \in DRE \rightarrow \mathcal{P}(V^*) \), and function \( \text{undot} \in DRE \rightarrow RE \). Function \( \mathcal{R} \) maps \( DRE\)s to the (regular) language to the right of the dot, and function \( \text{undot} \) removes the dot in a \( DRE \).

1. If \( E \in RE \) then
   
   (a) \( \bullet E \in DRE, \mathcal{R}(\bullet E) = \mathcal{L}_RE(E), \) and \( \text{undot}(\bullet E) = E; \)
   
   (b) \( E\bullet \in DRE, \mathcal{R}(E\bullet) = \{ e \}, \) and \( \text{undot}(E\bullet) = E. \)

2. If \( E \in RE \) and \( D \in DRE \) then
   
   (a) \( E \cup D \in DRE, D \cup E \in DRE, \mathcal{R}(E \cup D) = \mathcal{R}(D \cup E) = \mathcal{R}(D), \) \( \text{undot}(E \cup D) = E \cup \text{undot}(D), \) and \( \text{undot}(D \cup E) = \text{undot}(D) \cup E; \)
   
   (b) \( E \cdot D \in DRE, D \cdot E \in DRE, \mathcal{R}(E \cdot D) = \mathcal{R}(D), \mathcal{R}(D \cdot E) = \mathcal{R}(D) \cdot \mathcal{L}_RE(E), \) \( \text{undot}(E \cdot D) = E \cdot \text{undot}(D), \) and \( \text{undot}(D \cdot E) = \text{undot}(D) \cdot E; \)
   
   (c) \( D^* \in DRE, \mathcal{R}(D^*) = \mathcal{R}(D) \cdot \mathcal{L}_RE(\text{undot}(D))^*, \) and \( \text{undot}(D^*) = \text{undot}(D)^*; \)
   
   (d) \( D^+ \in DRE, \mathcal{R}(D^+) = \mathcal{R}(D) \cdot \mathcal{L}_RE(\text{undot}(D))^+, \) and \( \text{undot}(D^+) = \text{undot}(D)^+; \)
   
   (e) \( D' \in DRE, \mathcal{R}(D') = \mathcal{R}(D), \) and \( \text{undot}(D') = \text{undot}(D)' \).

3. Nothing else is a \( DRE \).

A dotted regular expression is also known as an \textit{item}, from LR parsing [Knut65]; we will frequently use this name. □

We also require a function mapping a regular expression to all of its dottings.

Definition 5.55 (Function \( \text{dots} \)): We define function \( \text{dots} \in RE \rightarrow \mathcal{P}(DRE) \) as follows:

\[
\text{dots}(E) = \{ D : D \in DRE \land \text{undot}(D) = E \}
\]

Remark 5.56: For a given \( E \in RE \), we will be using sets of items (elements of \( \mathcal{P}(\text{dots}(E)) \)) to characterize the derivatives of \( \mathcal{L}_RE(E) \) when constructing a DFA accepting \( \mathcal{L}_RE(E) \). □

Property 5.57 (\( \text{dots} \)): For all \( E \in RE, |\text{dots}(E)| \) is finite, and so is \( |\mathcal{P}(\text{dots}(E))| \). □

We define an item set, and its language as follows:

Definition 5.58 (Item sets and their languages): An item set \( J \) is a subset of \( DRE \) such that:

\[
(\exists E : E \in RE : J \subset \text{dots}(E))
\]

\( IS \) denotes the set of all item sets. Essentially, an \( IS \) is a set of items, all of which are dottings of the same regular expression. We also extend \( \text{undot} \) to \( IS \). □

Definition 5.59 (Language of an \( IS \)): The language of a \( J \in IS \) is given by function \( \mathcal{L}_IS \in IS \rightarrow \mathcal{P}(V^*) \) defined as:

\[
\mathcal{L}_IS(J) = (\cup I : I \in J : \mathcal{R}(I))
\]

□
Remark 5.60: We will use item sets to characterize the derivatives of an RE. □

We now define derivatives of item sets.

**Definition 5.61 (Derivative of an item set):** Given \( J \in IS \) and \( a \in V \) we define \( a^{-1}J \) to be the following set:

1. If \( I \in J \) has a subexpression \( \cdot a \) then \( I' \) is in \( a^{-1}J \), where \( I' \) is the same as \( I \) with the subexpression \( \cdot a \) replaced by \( a\cdot \).

2. Nothing else is in \( a^{-1}J \).

□

We can now define a special type of function, which we call a **closure function**.

**Definition 5.62 (Closure functions):** Any function \( \varepsilon \in IS \rightarrow IS \) can be used as a closure function, provided that

\[
(\forall J : J \in IS : L_{IS}(\varepsilon(J)) = L_{IS}(J) \land \varepsilon \circ \varepsilon(J) = \varepsilon(J))
\]

and

\[
(\forall E, J : E \in RE \land J \subseteq dots(E) : \varepsilon \in L_{IS}(J) \equiv (E\cdot) \in \varepsilon(J))
\]

and

\[
(\forall J, a : J \in IS \land a \in V : L_{IS}(a^{-1}J) = a^{-1}L_{IS}(J))
\]

□

We are now in a position to define our first closure function, and an auxiliary relation.

**Definition 5.63 (Dot closure relation \( D \)):** We define a binary relation \( D \) on \( DRE \). \( D \) is the smallest relation such that:

1. If \( E, F \in RE \), then (here we use infix notation for relation \( D \)):

\[
\begin{align*}
   \cdot\varepsilon &\quad D &\quad \varepsilon\cdot \\
   (E \cdot F) &\quad D &\quad (E)\cdot F \\
   (E\cdot) \cdot F &\quad D &\quad (E\cdot)\cdot F \\
   E \cdot (F\cdot) &\quad D &\quad (E \cdot F)\cdot \\
   (E\cup F) &\quad D &\quad (E\cup F)\cdot \\
   (E\cup) F &\quad D &\quad (E\cup) F\cdot \\
   (E\cup(F\cdot)) &\quad D &\quad (E\cup F)\cdot \\
   (E\cdot) &\quad D &\quad (E\cdot) \\
   (E\cdot) &\quad D &\quad (E\cdot) \\
   (E\cdot) &\quad D &\quad (E\cdot) \\
   (E\cdot) &\quad D &\quad (E\cdot) \\
   (E\cdot) &\quad D &\quad (E\cdot) \\
   (E\cdot) &\quad D &\quad (E\cdot) \\
   (E\cdot) &\quad D &\quad (E\cdot) \\
   (E\cdot) &\quad D &\quad (E\cdot) \\
   (E\cdot) &\quad D &\quad (E\cdot) \\
   (E\cdot) &\quad D &\quad (E\cdot) \\
\end{align*}
\]

2. If \( E \in RE \) and \( D_0, D_1 \in DRE \) such that \((D_0, D_1) \in D \), then:

\[
(a) \ (E \cup D_0, E \cup D_1) \in D, \ (D_0 \cup E, D_1 \cup E) \in D, \ (E \cdot D_0, E \cdot D_1) \in D, \text{ and } (D_0 \cdot E, D_1 \cdot E) \in D.
\]
(b) \((D_0^*, D_1^*) \in D, (D_0^+, D_1^+) \in D, \) and \((D_0^0, D_1^1) \in D\).

\[\square\]

**Definition 5.64** (Closure function \(C\)): We define function \(C \in IS \rightarrow IS\) as:

\[C(J) = D^*(J)\]

Function \(C\) satisfies the Definition 5.62, making it a closure function. \(\square\)

**Remark 5.65**: The closure function \(C\) presented here is an extension (to deal with our definition of regular expressions) of the one usually given for LR parsing. \(\square\)

**Example 5.66** (Function \(C\)): \(C(\{.«aU€)b')\})\) is computed to be \(\{.«aU€)b'), (.(aU€»b', (a U \€)b', (a U \€)(b'), (a U \€)(.b'), ((a U \€)(b')\}). \square\)

In order to construct a \(DFA\) for an \(RE\), we require a set of item sets which characterize the set of derivatives of a the regular expression. The following definition gives the necessary conditions.

**Definition 5.67** (Derivative item set): Given \(E \in RE\), the set \(D \in \mathcal{P}(dots(E))\) (that is, \(D \subseteq IS\) and for each \(J \in D, E = undot(J)\)) characterizes (under some closure function \(E\)) the derivatives of \(E\) if and only if:

\[
\{w^{-1}L_{RE}(E) : w \in V^*\} = \{L_{IS}(J) : J \in D\}
\]

and

\[
(\forall J : J \in D : J = E(J))
\]

and

\[
(\forall J, a : J \in D \land a \in V : E(a^{-1}J) \in D)
\]

We write this property \(DIS(E, D, E)\). The set \(D\) is called a derivative item set for \(E\). \(\square\)

We are now in a position to modify Algorithm 5.39 to compute such a derivative item set (under some closure function \(E\)), instead of a set of derivatives. In the following algorithm, \(D, next \subseteq IS\).

**Algorithm 5.68**:

\[
\begin{align*}
\{E \in RE\} \\
D, next := \emptyset, \{E(\{E\})\} \\
do \text{next} \notin D \rightarrow \\
D, next := D \cup \text{next}, \{E(a^{-1}I) : a \in V \land I \in \text{next}\} \\
\text{od} \\
\{DIS(E, D, E)\}
\end{align*}
\]

This algorithm terminates since \(|\mathcal{P}(dots(E))|\) is finite (Property 5.57). With the set \(D\) computed above, we can now construct a \(DFA\) accepting \(L_{RE}(E)\).

**Construction 5.69** (Item set construction): Function \(Iconstr \in RE \times \mathcal{P}(IS) \rightarrow DFA\) takes a regular expression \((E)\) and a derivative item set \((D)\) for the \(RE\) (such that \(DIS(E, D, C)\)), and constructs a \(DFA\):

\[
Iconstr(E, D) = \text{let} \\
T(J, a) = \{C(a^{-1}J)\} \\
S = \{C(\{E\})\} \\
F = \{J : J \in D \land E \in J\} \\
in \\
(D, V, T, \emptyset, S, F) \\
\text{end}
\]

A \(DFA\) constructed with \(Iconstr\) has the following property:

\[
(\forall E, D : DIS(E, D, C) : L_{FA}(Iconstr(E, D)) = L_{RE}(E))
\]

The \(DFA\) is also Complete. \(\square\)
5.5 Towards DeRemer's construction

Figure 10: The DFA $\text{Iconstr}((a \cup e)b^*)$.

Figure 11: The DFA $\text{Iconstr}(b^*)$.

Example 5.70 ($\text{Iconstr}$): We construct the DFA corresponding to $(a \cup e)b^*$. The derivative item set is (the individual item sets have been compressed, as a notational convenience, and each item set is given a label):

$$
\{I_0 = \{e(a \cup e)e\}, I_1 = \{(a \cup e)e, (e)b^*e, (e)b^*e\}, I_2 = \{(a \cup e)e, (e)b^*e\}, I_3 = \emptyset\}
$$

The DFA is shown in Figure 10.

5.5.1 Making the construction more efficient

Because of the definition of $C$, function $\text{Iconstr}$ sometimes constructs a DFA which is larger than necessary, as shown in the following example.

Example 5.71 (A DFA that is not minimal): We use $\text{Iconstr}$ to construct a DFA for $b^* \in RE$. The two item sets are $D = \{(e)(b^*), (e)b^*e, (b^*)e\} = \{(b^*), (e)b^*e, (b^*)e\}$. The DFA is shown in Figure 11. The problem is that the two item sets should have been recognized as denoting equivalent derivatives since:

$$
L_{IS}(\{(e)(b^*), (e)b^*e, (b^*)e\}) = L_{IS}(\{(b^*), (e)b^*e, (b^*)e\})
$$

They only differ in the items $(e)b^*$ and $(e)b^*e$.

The problem is that for some $J \in IS$, there is much redundant information in $C(J)$. In particular, there may be a $J' \subset J$ such that $L_{IS}(J') = L_{IS}(J)$. We can introduce a function $\mathcal{X}$ such that $\mathcal{X} \circ C$ is a closure function. That is, $\mathcal{X}$ is used as a filter.

Definition 5.72 (Item set optimization function $\mathcal{X}$): Given $J \in IS$ such that $J = C(J)$, $\mathcal{X}(J)$ is the same as $J$, with the following removed: any item containing a subexpression of the form $e(E \cup F)$, $e(E^*)$, or $(Ee)^*$. 

Property 5.73 (Function $\mathcal{X} \circ C$): Function $\mathcal{X} \circ C$ satisfies Definition 5.62, and is a closure function.
Remark 5.74: The reason that such items are removed is that the definition of $C$ ensures that they are redundant; other items will have been added to the item set to ensure that these ones are not needed. For example, in the case of $\bullet(E \cup F)$ function $C$ will ensure that $\bullet E \cup F$ and $E \cup \bullet F$ are added to the item set. In this case $\mathcal{R}(\bullet(E \cup F)) = \mathcal{R}(\bullet E \cup F) \cup \mathcal{R}(E \cup \bullet F)$.

Using composite function $\mathcal{X} \circ \mathcal{C}$, we can now present a revised version of the above construction.

Construction 5.75 (DeRemer's construction): DeRemer's construction is function $\text{DeRemer}$, which is exactly as $\text{Iconstr}$, except that the composite function $\mathcal{X} \circ \mathcal{C}$ is used as the closure function wherever $\mathcal{C}$ was used in $\text{Iconstr}$. This construction is due to DeRemer [DeRe74], where he attributes the idea behind the definition of $\mathcal{X}$ to Earley [Earl70].

Remark 5.76: DeRemer presented this construction in a slightly different context: he extended an LR parser to deal with grammar production rules with regular expression as right hand sides [DeRe74]. Remark 5.78 points out a slight problem with the original presentation by DeRemer.

Example 5.77 (DeRemer): We use $\text{DeRemer}$ to construct a DFA for $b^* \in \mathcal{R}E$. The only item set is $\{(\bullet b)^*, (b^*)\bullet\}$ and the DFA is shown in Figure 12. With alphabet $V = \{b\}$, this is the minimal Complete DFA accepting $L_{RE}(b^*)$.

Remark 5.78: DeRemer and Earley specify that both the closure (function $\mathcal{C}$) and the optimization (function $\mathcal{X}$) operations are to be performed simultaneously. Unfortunately, when $\epsilon$ is permitted as an $\mathcal{R}E$ (as we have done) it is possible that the process never terminates. For example, consider the closure (with optimization) of $\{(\bullet(a \cup \epsilon))^*, (\bullet(a \cup \epsilon))\epsilon\}$. After the first step we have $\{(\bullet(a \cup \epsilon))^*, (\bullet(a \cup \epsilon))\epsilon\}$. After an optimization step, and a few more steps we have $\{(a \cup \epsilon)^*, (a \cup \epsilon)\epsilon\}$, after which we add $(\bullet(a \cup \epsilon))\epsilon$ which we had originally removed. The rewriting process begins again. In this paper, we avoid this problem by defining the closure and optimization steps separately.

We can devise an even more effective optimization function than $\mathcal{X}$.

Definition 5.79 (Function $\mathcal{Y}$): Given $J \in IS$ corresponding to $E \in \mathcal{R}E$ such that $J = \mathcal{C}(J)$, $\mathcal{Y}(J)$ is a subset of $J$, keeping only the following items:

1. Any item containing a subexpression of the form $\bullet a$ (for some regular expression $a \in V$).
2. The item $E \bullet$ (if present in $J$).

Property 5.80 (Function $\mathcal{Y} \circ \mathcal{C}$): Composite function $\mathcal{Y} \circ \mathcal{C}$ satisfies Definition 5.62, and is a closure function.

Remark 5.81: The function $\mathcal{Y}$ makes the computation of derivatives (of a $J \in IS$ such that $J = \mathcal{C}(J)$) particularly easy, as the items in $\mathcal{Y}(J)$ are precisely those required in the computation of derivatives and for determining if $\epsilon \in L_{IS}(J)$.

Construction 5.82 (Improved item sets): Our optimized construction, called $\text{Oconstr}$, is as $\text{Iconstr}$, except that the composite function $\mathcal{Y} \circ \mathcal{C}$ is used wherever $\mathcal{C}$ was used in $\text{Iconstr}$. This construction does not appear in the literature.
5.5 Towards DeRemer’s construction

Example 5.83 (Onestr): Using Onestr, we construct a DFA for \((a \cup \epsilon)b^*\). The derivative item sets are (each item set is given a label for use in the state graph): 

\[ I_0 = \{(\bullet \cup \epsilon)b^*, (a \cup \epsilon)(\bullet)b^*, (a \cup \epsilon)b^*\}, I_1 = \{(a \cup \epsilon)(\bullet)b^*, (a \cup \epsilon)b^*\}, I_2 = \emptyset \].

The DFA is shown in Figure 13. This is the minimal Complete DFA for the given regular expression. □

As seen in the above example, function Onestr constructs a smaller DFA than Brz did in Example 5.37. The two constructions seem difficult to compare, as the following example shows:

Example 5.84 (Comparing Brz to Onestr): We use Brz and Onestr to construct DFA’s for \(ac \cup bc\). The derivatives (under extended similarity — see Remark 5.32, each given a label) are: 

\[ d_0 = ac \cup bc, d_1 = c, d_2 = \epsilon, d_3 = \emptyset \].

The derivative item sets (using \(Y \circ C\), each given a label) are: 

\[ I_0 = \{ac \cup bc, ac \cup bc\}, I_1 = \{a \bullet c \cup bc\}, I_2 = \{ac \cup b \bullet c\}, I_3 = \{(ac \cup bc)\bullet\}, I_4 = \emptyset \].

The results are shown in Figure 14. Construction Onestr is unable to recognize that states \(I_1\) and \(I_2\) are equivalent. (An equivalence relation on \(IS\) — much like \(~\) on ERE’s — could be defined in order to identify such equivalent states.) □

We can make a more practical implementation by concatenating an end-marker $ onto \(E \in RE\) (using function \(\text{marker}_e\) — see Definition 4.46). The second rule defining \(Y\) (Definition 5.79) is then no longer required.
Construction 5.85 (Oconstr with end-marker): Using the end-marker, the body of the Oconstr construction becomes (assuming that \(DIS(E\$, D, Y \circ C)\)):

\[
\begin{align*}
&\text{let } \quad T(J, a) = \{ Y \circ C(a^{-1}J) \} \\
&\quad S = \{ Y \circ C(\{ E \}) \} \\
&\quad F = \{ J : J \in D \land S^{-1}J \neq \emptyset \} \\
&\quad \text{in} \\
&\quad (D, V, T, \emptyset, S, F) \\
&\text{end}
\end{align*}
\]

Remark 5.86: The Aho-Sethi-Ullman construction (Construction 4.50) can be viewed as a heavily encoded variation on the Oconstr construction. Each item in an item set of \( D \) is of the form \( \ldots \bullet a \ldots \) (for \( a \in V \)) and corresponds to the basis RFA's that are used in the construction of the RFA for \( E \). The subset construction (with start-unreachable state removal) of the Aho-Sethi-Ullman algorithm (Algorithm 4.52) is folded into the algorithm computing the derivative item set (using composite function \( Y \circ C \)) and the definition of a derivative of an item set. Compare the DFA produced in Example 4.51 to that produced in Example 5.83; the only difference is the sink state in the latter example. \( \Box \)
6 Conclusions

The conclusions of this paper fall into two groups, depending on the section to which they relate: constructions based upon the structure of regular expressions (Section 4), or constructions based upon the Myhill-Nerode theorem (Section 5).

The conclusions about constructions based on regular expression structure are:

- Finite automaton constructions are frequently said to be "based upon the structure of regular expressions." The $\Sigma$-algebra framework (given in Sections 3 and 4) was useful in formalizing this notion. The $\Sigma$-algebras were particularly useful in the following ways:
  - They placed Thompson's, the left-biased, the right-biased, and the reduced finite automata (RFA) constructions in a common framework.
  - They highlighted the fact that the type of object produced by the Thompson's, the left-biased and the right-biased constructions is actually the isomorphism class of a finite automaton, as opposed to a finite automaton.

- The concept of duality (that one construction can be the mirror image of another) played a central part in finding common parts in constructions. Duality was made more obvious through the use of $\Sigma$-algebras. The following constructions were found to be related by duality:
  - The Berry-Sethi nondeterministic finite automaton construction (also known in the literature as the McNaughton-Yamada or the Glushkov nondeterministic finite automata construction) and the dual of the Berry-Sethi construction (a variant of which is also known as the Aho-Sethi-Ullman nondeterministic finite automata construction [ASU86, Example 3.22, pg. 140]).
  - The McNaughton-Yamada-Glushkov deterministic finite automaton (DFA) construction and the Aho-Sethi-Ullman DFA construction\(^ {12}\).

- The use of end-markers (concatenated to either the left or the right of a regular expression) was found to be a simple coding trick, which may be useful in practice. End markers do not play a central role in any of the constructions, although they have previously been portrayed as important.

- The concept of marking a regular expression (each alphabet symbol occurring in the regular expression is given a unique mark, making all of the symbols unique — see Section 4.5) is an encoding trick. Marking is not central to the correctness of any of the constructions, although it is a useful technique in the practical implementation of some of the constructions.

- Marking was found to cause problems in some of the constructions. In particular, intersection, complementation, and language difference cannot be dealt with using marking\(^ {13}\). In the $\Sigma$-algebra framework, intersection, complementation, and language difference can easily be implemented for the Berry-Sethi, McNaughton-Yamada-Glushkov, and Aho-Sethi-Ullman constructions — constructions that are all traditionally defined using marking.

- Two interpretations of marking appear in the literature. In the first one, being "at a mark" (Aho, Sethi, and Ullman use the phrase "at a position" [ASU86]) means to be in the state resulting from making a transition on the alphabet symbol associated with the particular mark\(^ {14}\). The second interpretation equates being "at a mark" with being in the state which

\(^{12}\)Here we assume that the sink state (if it exists) is removed from a DFA produced by the McNaughton-Yamada-Glushkov construction.

\(^{13}\)Actually, McNaughton and Yamada [MY60] attempted to define intersection and complementation. Their informal descriptions are difficult to understand, and more recent papers use marking and have abandoned trying to define intersection or complementation. See Section 4.4.

\(^{14}\)This is the interpretation taken by Glushkov, McNaughton and Yamada, and Berry and Sethi [Glus61, MY60, BS86].
has an out-transition on the alphabet symbol associated with the particular mark (that is, the marked symbol is valid as the next input symbol). The two interpretations are duals of one another, and arise naturally from the duality of the left-biased and right-biased constructions. For example, the interpretations give rise to the duality between McNaughton, Yamada and Glushkov's DFA and Aho, Sethi, and Ullman's DFA constructions.

- The improvements to the Berry-Sethi construction due to Brüggemann-Klein [B-K93a] and Chang and Paige [Chan92, CP92] have been difficult to compare. This has been largely due to the fact that Chang and Paige's improvements are to the finite automaton construction itself, while Brüggemann-Klein's improvements involve transforming the regular expression. In Section 4.5, we presented an improvement to the construction (not found in the literature) that mirrors Brüggemann-Klein's improvements (not on regular expressions, but on finite automata), and is easy to compare to Chang and Paige's construction.

- Some relationships between the constructions were found that were not made obvious by the Σ-algebra derivations:
  - For restricted regular expressions (where each alphabet symbol occurs at most once — as in marked regular expressions) the Berry-Sethi construction produces a deterministic FA. As a consequence, the Berry-Sethi construction and the McNaughton-Yamada-Glushkov DFA construction produce isomorphic finite automata (with the exception of the sink state present in a McNaughton-Yamada-Glushkov DFA).
  - The Berry-Sethi construction (and therefore the McNaughton-Yamada-Glushkov DFA construction) and the Brzozowski construction (under an appropriate encoding) produce isomorphic finite automata for restricted regular expressions. This result was originally presented by Berry and Sethi [BS86].

The conclusions about the Myhill-Nerode, Brzozowski, and DeRemer constructions (Section 5) are:

- Deriving the second major family of constructions from the Myhill-Nerode theorem proved useful in a number of ways:
  - The use of equivalence classes makes the correctness argument for the Myhill-Nerode construction particularly clear.
  - The unique minimal DFA (for a particular language) can be easily constructed using a particular equivalence class as the parameter to the Myhill-Nerode construction.
  - Derivatives (of a language) are a useful encoding of the equivalence classes of Myhill and Nerode's unique minimal-index equivalence relation $R_E$.
  - The definition of derivatives provides an efficient method to compute finite sets which encode the infinite sets that are used in the Myhill-Nerode construction.

- The Brzozowski construction can be viewed as an ingenious encoding of the Myhill-Nerode minimal DFA construction.

- Brzozowski's original paper provided a proof (a similar one is given in this paper) that his construction also works when only similarity of regular expressions is recognized. Similarity is defined in his paper using four rules, and is defined in this paper using only three rules. The missing fourth rule (that $\epsilon$ is the unit of concatenation) is not required in the definition of similarity for the correctness of our presentation of Brzozowski's construction.

---

This is the interpretation taken by Aho, Sethi, and Ullman [ASU86].

16 Which is therefore an improvement to the McNaughton-Yamada-Glushkov and the Aho-Sethi-Ullman nondeterministic finite automaton constructions.
We defined the equivalence and the similarity of regular expressions as equivalence relations on regular expressions. We also demonstrated that any such equivalence relation $E$ can be used in Brzozowski's construction, provided that $E$ is a refinement of equivalence and similarity is a refinement of $E$.

For restricted regular expressions, Brzozowski's construction (with an encoding), the Berry-Sethi construction, and the McNaughton-Yamada-Glushkov construction produce isomorphic DFA's.

The use of dotted regular expressions (also known as items, from LR parsing) is a useful, if obscure, encoding of the derivatives of a regular expression. We obtained the following results on the use of dotted regular expressions:

- Computing the set of all dotted regular expressions (from a given regular expression) can be defined very simply. The derivatives of dotted regular expressions, and the construction of a DFA can be defined simply. This construction does not appear in the literature.
- The straightforward definition of dotted regular expressions is unable to deal with intersection and complementation. This is for the same reason that marking constructions are unable to deal with intersection and complementation.
- DeRemer specified a DFA construction that appears to be very easy to implement. It is an optimization over the straightforward dotted regular expression construction, and the constructed DFA is always smaller.
- The original specification of item closure, due to DeRemer and Earley, is incomplete. They attempted to define closure and optimization as a single step. This can lead to non-termination, as we have demonstrated. The problem can be easily solved by defining closure and optimization steps separately.
- We show that additional optimizations, added to DeRemer's construction, can reduce the size of the produced DFA. This construction is not given in the literature. Furthermore, the optimizations are arguably easier to understand than those of DeRemer, and likely easier to implement.
- It is possible to show that this improved construction is related to the Aho-Sethi-Ullman DFA construction.
A Some basic definitions

**Convention A.1** (Powerset): For any set $A$ we use $\mathcal{P}(A)$ to denote the set of all subsets of $A$. $\mathcal{P}(A)$ is called the powerset of $A$; it is sometimes written $2^A$.

**Convention A.2** (Sets of functions): For sets $A$ and $B$, $A \rightarrow B$ denotes the set of all total functions from $A$ to $B$, while $A \mapsto B$ denotes the set of all partial functions from $A$ to $B$.

**Remark A.3:** For sets $A$, $B$ and relation $C \subseteq A \times B$ we can interpret $C$ as a function $C \in A \rightarrow \mathcal{P}(B)$.

**Convention A.4** (Tuple projection): For an $n$-tuple $t = (x_1, x_2, \ldots, x_n)$ we use the notation $\pi_i(t)$ ($1 \leq i \leq n$) to denote tuple element $x_i$; we use the notation $\pi_i(t)$ ($1 \leq i \leq n$) to denote the $(n - 1)$-tuple $(x_1, x_{i-1}, x_{i+1}, \ldots, x_n)$. Both $\pi$ and $\pi$ extend naturally to sets of tuples.

**Convention A.5** (Tuple arguments to functions): For functions (or predicates) taking a single tuple as an argument, we usually drop one set of parentheses in a function application.

**Convention A.6** (Relation composition): Given sets $A, B, C$ (not necessarily different) and two relations, $E \subseteq A \times B$ and $F \subseteq B \times C$, we define relation composition (infix operator $\circ$) as:

$$E \circ F = \{(a, c) : (\exists b : b \in B : (a, b) \in E \land (b, c) \in F)\}$$

**Convention A.7** (Equivalence classes of an equivalence relation): For any equivalence relation $E$ on set $A$ we denote the set of equivalence classes of $E$ by $[A]_E$; that is

$$[A]_E = \{[a]_E : a \in A\}$$

Set $[A]_E$ is also called the partition of $A$ induced by $E$.

**Definition A.8** (Index of an equivalence class): For equivalence relation $E$ on set $A$, define $\#E = |[A]_E|$. $\#E$ is called the index of $E$.

**Definition A.9** (Alphabet): An alphabet is a non-empty set of finite size.

**Definition A.10** (Refinement of an equivalence relation): For equivalence relations $E$ and $E'$ (on set $A$), $E$ is a *refinement* of $E'$ if and only if $E \subseteq E'$. 

**Definition A.11** (Refinement (⊆) relation on partitions): For equivalence relations $E$ and $E'$ (on set $A$), $[A]_E$ is said to be a refinement of $[A]_{E'}$ (written $[A]_E \subseteq [A]_{E'}$) if and only if $E \subseteq E'$. An equivalent statement is that $[A]_E \subseteq [A]_{E'}$ if and only if every equivalence class (of $A$) under $E$ is entirely contained in some equivalence class (of $A$) under $E'$. 

**Property A.12** (Equivalence relations): Given two equivalence relations $E, F$, we have the following property:

$$(E \subseteq F) \land (\#E = \#F) \Rightarrow (E = F)$$

**Definition A.13** (Regular languages): $L_{reg V}$ denotes the set of all regular languages over alphabet $V$. That is, $L_{reg V} \subseteq \mathcal{P}(V^*)$ is the smallest set containing $V$ that is closed under $\cup$ (language union), $\cdot$ (a dot, language concatenation), and $*$ (Kleene closure). The subscript $V$ is dropped when no ambiguity arises.

**Definition A.14** (Operator ? on languages): We define $?$ as a postfix (superscript) operator on languages as $L^? = L \cup \{\epsilon\}$. 


Definition A.15 (Left derivatives): Given language \( A \subseteq V^* \) and \( w \in V^* \) we define the left derivative of \( A \) with respect to \( w \) as:

\[
w^{-1}A = \{ x \in V^* : wx \in A \}
\]

Sometimes derivatives are written as \( D_wA \) or as \( \frac{dA}{dw} \). Right derivatives are analogously defined. Derivatives can also be extended to \( B^{-1}A \) where \( B \) is also a language. □

Property A.16 (Left derivatives): The following two properties follow from Definition A.15 (assuming \( L \) is a language):

- \( \forall w \in L \quad \varepsilon \in w^{-1}L \), and
- \( (wa)^{-1}L = a^{-1}(w^{-1}L) \).

□

Property A.17 (Derivatives of regular languages): Assuming \( a \in V \) and \( L, L_0, L_1 \in \mathcal{L}_{\text{reg}} \), derivatives have the following properties (given with respect to the structure of regular languages):

\[
a^{-1}\emptyset = \emptyset
\]

\[
a^{-1}\{\varepsilon\} = \emptyset
\]

\[
a^{-1}\{b\} = \begin{cases} \{\varepsilon\} & \text{if } (a = b) \text{ then } \{\varepsilon\} \text{ else } \emptyset \\ \end{cases}
\]

\[
a^{-1}(L_0L_1) = \begin{cases} (a^{-1}L_0)L_1 & \text{if } (\varepsilon \in L_0) \text{ then } a^{-1}L_1 \text{ else } \emptyset \\ \end{cases}
\]

\[
a^{-1}(L_0 \cup L_1) = (a^{-1}L_0) \cup (a^{-1}L_1)
\]

\[
a^{-1}(L^*) = (a^{-1}L)L^*
\]

\[
a^{-1}(L^+) = a^{-1}L
\]

\[
a^{-1}(L^?) = a^{-1}L
\]

\[
a^{-1}(L \cdot L) = (a^{-1}L) \cap (a^{-1}L_1)
\]

\[
a^{-1}(L) = \neg(a^{-1}L)
\]

The definition related to Kleene closure is shown as follows:

\[
a^{-1}(L^*) = \begin{cases} \{ \text{Definition of } \ast \} \\ a^{-1}((L \setminus \{\varepsilon\})L^* \cup \{\varepsilon\}) \\ \{ \text{Definition of } a^{-1}(L_0 \cup L_1) \} \\ a^{-1}((L \setminus \{\varepsilon\})L^*) \cup a^{-1}\{\varepsilon\} \\ \{ \text{Definition of derivative of concatenation and } \{\varepsilon\} \} \\ (a^{-1}(L \setminus \{\varepsilon\}))L^* \\ \{ \text{Definition of derivative of } \{\varepsilon\} \} \\ (a^{-1}L)L^* \end{cases}
\]

The definition related to complementation \( \neg \) is as follows:

\[
a^{-1}(\neg L) = \begin{cases} \{ \text{Definition of derivative} \} \\ \{ x : ax \in \neg L \} \\ \{ \text{Definition of } \neg \text{ operator} \} \\ \neg\{ x : ax \in L \} \\ \{ \text{Definition of } a^{-1}L \} \\ \neg(a^{-1}L) \end{cases}
\]
Definition A.18 (Preserving a predicate): A (partial) function \( f \in B^n \rightarrow B \) (for fixed \( n \geq 0 \)) is said to preserve predicate (or property) \( P \) (on \( B \)) if and only if
\[
(\forall B' : B' \in B^n \cap \text{domain}(f)) \land (\forall k : 1 \leq k \leq n : P(\pi_k(B')) : P(f(B')))
\]
The set \( \text{domain}(f) \) refers to those elements of \( B^n \) on which \( f \) is defined. \( \square \)

Intuitively, a function \( f \) preserves a property \( P \) if, when every argument of \( f \) satisfies \( P \), the result of \( f \) applied to the arguments also satisfies \( P \).

Definition A.19 (Reversal operator): A reversal operator \( R \) (usually written postfix and superscript) for a set \( A \) is a function \( R : A \rightarrow A \) such that \( R \circ R \) (equivalently \( R^2 \)) is the identity function on \( A \). We sometimes write the reversal operator as a standard (prefix notation) function. \( \square \)

Definition A.20 (Tuple and relation reversal): For an \( n \)-tuple \( (x_1, x_2, \ldots, x_n) \) define reversal as (postfix and superscript) function \( R \):
\[
(x_1, x_2, \ldots, x_n)^R = (x_n, \ldots, x_2, x_1)
\]
Given a set \( A \) of tuples, we define \( A^R = \{x^R : x \in A}\). \( \square \)

Definition A.21 (Dual of a function): We assume two sets \( A \) and \( B \) whose reversal operators are \( R \) and \( R' \) respectively. Two functions, \( f \in A \rightarrow B \) and \( f_d \in A \rightarrow B \) are one another's dual if and only if
\[
f(a) = (f_d(a^R))^R'
\]
In some cases we relax the equality to isomorphism (when isomorphism is defined on \( B \)). \( \square \)

Definition A.22 (Symmetrical function): A symmetrical function is one that is its own dual. \( \square \)

Proposition A.23 (Symmetrical functions): The composition of two symmetrical functions is again symmetrical. \( \square \)
B  Proofs of some $\Sigma$-algebra operators

In this section, we sketch proofs of the correctness of the operators of Thompson’s $\Sigma$-algebra (Definition 4.1), the left-biased $\Sigma$-algebra operators (Definition 4.20), the Chang-Paige $\textit{RFA}'$ operators (Construction 4.77), and the $\textit{RFA}''$ operators (Construction 4.79).

Theorem B.1 (Correctness of Thompson’s $\Sigma$-algebra of FA’s): Recall the operator definitions of Definition 4.1. We only present a correctness proof for the operator $\textit{CU,Th}$.

In the following derivation we assume the context of the innermost $\text{let}$ clause of the operator definition.

$$
\mathcal{L}_{\text{FA}}(\textit{CU,Th}([M_0]_\Sigma, [M_1]_\Sigma)) = \\
= \{ \text{Definition of } \mathcal{L}_{\text{FA}} \} \\
= \{ \text{Definition of } F' \} \\
= (\cup s, f : s \in S_0 \land f \in F_0 : T_0^*(s, f)) \cup (\cup s, f : s \in S_1 \land f \in F_1 : T_1^*(s, f)) \\
= \{ \text{Definitions of } \mathcal{L}_{\text{FA}}([M_0]_\Sigma), \mathcal{L}_{\text{FA}}([M_1]_\Sigma) \} \\
\mathcal{L}_{\text{FA}}([M_0]_\Sigma) \cup \mathcal{L}_{\text{FA}}([M_1]_\Sigma)
$$

\hfill ∎

Theorem B.2 (Correctness of the LBFA operators): Recall the operator definitions of Definition 4.20. We present a correctness proof of the operator $\textit{CU,LBPA}$. In the following derivation we assume the context of the innermost $\text{let}$ clause of the operator definition.

$$
\mathcal{L}_{\text{FA}}(\textit{CU,LBPA}([M_0]_\Sigma, [M_1]_\Sigma)) = \\
= \{ \text{Definitions of } \mathcal{L}_{\text{FA}} \text{ and } F' \} \\
= (\cup f : f \in F_0 \cap Q' : T_0^*(q_0, f)) \cup (\cup f : f \in F_1 \cap Q' : T_1^*(q_0, f)) \\
\cup \text{ if } (N) \text{ then } T''(q_0, q_0) \text{ else } 0 \\
= \{ \text{Definitions of } \mathcal{L}_{\text{FA}}([M_0]_\Sigma), \mathcal{L}_{\text{FA}}([M_1]_\Sigma), N, T'' \} \\
\mathcal{L}_{\text{FA}}([M_0]_\Sigma) \cup \mathcal{L}_{\text{FA}}([M_1]_\Sigma) \cup \text{ if } (\epsilon \in (\mathcal{L}_{\text{FA}}(M_0) \cup \mathcal{L}_{\text{FA}}(M_1))) \text{ then } \{ \epsilon \} \text{ else } 0 \\
= \{ \text{Definition of } \mathcal{L}_{\text{FA}}([M_0]_\Sigma) \cup \mathcal{L}_{\text{FA}}([M_1]_\Sigma) \} \\
\mathcal{L}_{\text{FA}}([M_0]_\Sigma) \cup \mathcal{L}_{\text{FA}}([M_1]_\Sigma)
$$

\hfill ∎

Theorem B.3 (Correctness of the Chang-Paige $\textit{RFA}'$ operators): Recall Definition 4.76 and Construction 4.77. We only present the derivation of the eighth component (usually called $W$) for operators $\textit{C.,RFA}'$ and $\textit{CU,RFA}'$ (the others are easy to prove). We assume the context of the innermost $\text{let}$ clause for both operators.

$$
\textit{C.,RFA'}: \quad \text{last'} \times \text{first'} \setminus (\text{follow}_0 \cup \text{follow}_1 \cup \text{last}_0 \times \text{first}_1) \\
= \{ \text{Definitions of first' and last' } \} \\
(\text{last}_1 \text{ if } (\text{null}_1) \text{ then } \text{last}_0 \text{ else } 0 \text{ if } (\text{null}_0) \text{ then } \text{first}_1 \text{ else } 0 \text{ if }) \\
\setminus (\text{follow}_0 \cup \text{follow}_1 \cup \text{last}_0 \times \text{first}_1) \\
= \{ \text{Rewriting } \} \\
(\text{last}_1 \times \text{first}_0 \text{ if } (\text{null}_0) \text{ then } \text{last}_1 \times \text{first}_1 \text{ else } 0 \text{ if } \text{null}_1) \text{ then } \text{last}_0 \times \text{first}_0 \text{ else } 0 \text{ if } \text{null}_0 \land \text{null}_1) \text{ then } \text{last}_0 \times \text{first}_1 \text{ else } 0 \text{ if } \text{null}_0 \text{ then } \text{follow}_0 \cup \text{follow}_1 \cup \text{last}_0 \times \text{first}_1 \\
\setminus (\text{follow}_0 \cup \text{follow}_1 \cup \text{last}_0 \times \text{first}_1) \\
= \{ \text{Assumption that } Q_0 \cap Q_1 = \emptyset \} \\
\text{last}_1 \times \text{first}_0 \text{ if } (\text{null}_0) \text{ then } \text{last}_1 \times \text{first}_1 \setminus \text{follow}_1 \text{ else } 0 \text{ if }$

\hfill ∎
\( \mathsf{if} (\mathsf{null}_1) \mathsf{then} \mathsf{last}_0 \times \mathsf{first}_0 \setminus \mathsf{follow}_0 \mathsf{else} \emptyset \mathsf{fi} \)

= \{ \text{Definitions of } \mathsf{W}_0 \text{ and } \mathsf{W}_1 \}

\( \mathsf{last}_1 \times \mathsf{first}_0 \mathsf{if} (\mathsf{null}_0) \mathsf{then} \mathsf{W}_1 \mathsf{else} \emptyset \mathsf{fi} \mathsf{if} (\mathsf{null}_1) \mathsf{then} \mathsf{W}_0 \mathsf{else} \emptyset \mathsf{fi} \)

= \{ \text{Definitions of } \mathsf{W}' \text{ and } \mathsf{W}'' \}

\( \mathsf{last}_1 \times \mathsf{first}_0 \mathsf{if} \mathsf{W}' \mathsf{or} \mathsf{W}'' \)

\( \mathsf{C}_{\mathsf{U}, \mathsf{RFA}}: \) \( (\mathsf{last}_0 \mathsf{or} \mathsf{last}_1) \times (\mathsf{first}_0 \mathsf{or} \mathsf{first}_1) \setminus (\mathsf{follow}_0 \mathsf{or} \mathsf{follow}_1) \)

= \{ \text{Rewriting} \}

\( (\mathsf{last}_0 \times \mathsf{first}_0 \mathsf{or} \mathsf{last}_0 \times \mathsf{first}_1 \mathsf{or} \mathsf{last}_1 \times \mathsf{first}_0 \mathsf{or} \mathsf{last}_1 \times \mathsf{first}_1) \setminus (\mathsf{follow}_0 \mathsf{or} \mathsf{follow}_1) \)

= \{ \text{Assumption that } \mathsf{Q}_0 \cap \mathsf{Q}_1 = \emptyset \}

\( (\mathsf{last}_0 \times \mathsf{first}_0 \setminus \mathsf{follow}_0) \mathsf{or} \mathsf{last}_1 \times \mathsf{first}_1 \mathsf{or} \mathsf{last}_0 \times \mathsf{first}_0 \mathsf{or} (\mathsf{last}_1 \times \mathsf{first}_1 \setminus \mathsf{follow}_1) \)

= \{ \text{Definitions of } \mathsf{W}_0 \text{ and } \mathsf{W}_1 \}

\( \mathsf{last}_0 \times \mathsf{first}_1 \mathsf{or} \mathsf{last}_1 \times \mathsf{first}_0 \mathsf{W}_0 \mathsf{or} \mathsf{W}_1 \)

\( \square \)

**Theorem B.4 (Correctness of the RFA" operators):** Recall Definition 4.78 and Construction 4.79. We only present the derivation of the eighth component (usually called \( \mathsf{W} \)) for operators \( \mathsf{C}_{\mathsf{RFA}''} \) and \( \mathsf{C}_{\mathsf{U}, \mathsf{RFA}''} \) (the others are easy to prove). We assume the context of the innermost let clause for both operators.

\( \mathsf{C}_{\mathsf{RFA}''}: \) \( (\mathsf{follow}_0 \mathsf{or} \mathsf{follow}_1 \mathsf{or} (\mathsf{last}_0 \times \mathsf{first}_1)) \setminus (\mathsf{last}' \times \mathsf{first}') \)

= \{ \text{Definitions of } \mathsf{first}' \text{ and } \mathsf{last}' \}

\( (\mathsf{follow}_0 \mathsf{or} \mathsf{follow}_1 \mathsf{or} (\mathsf{last}_0 \times \mathsf{first}_1)) \setminus (\mathsf{last}_1 \times \mathsf{first}_0 \mathsf{or} (\mathsf{null}_1 \mathsf{then} \mathsf{last}_0 \mathsf{else} \emptyset \mathsf{fi}) \)

\( \times (\mathsf{first}_0 \mathsf{or} (\mathsf{null}_0 \mathsf{then} \mathsf{first}_1 \mathsf{else} \emptyset \mathsf{fi}) \)

= \{ \text{Rewriting} \}

\( (\mathsf{follow}_0 \mathsf{or} \mathsf{follow}_1 \mathsf{or} (\mathsf{last}_0 \times \mathsf{first}_1)) \setminus (\mathsf{last}_1 \times \mathsf{first}_0 \mathsf{or} (\mathsf{null}_0 \mathsf{then} \mathsf{last}_1 \mathsf{else} \emptyset \mathsf{fi}) \)

\( \mathsf{or} (\mathsf{null}_1 \mathsf{then} \mathsf{last}_0 \mathsf{else} \emptyset \mathsf{fi}) \)

\( \mathsf{or} (\mathsf{null}_0 \mathsf{and} \mathsf{null}_1 \mathsf{then} \mathsf{last}_0 \times \mathsf{first}_1 \mathsf{else} \emptyset \mathsf{fi}) \)

= \{ \text{Assumption that } \mathsf{Q}_0 \cap \mathsf{Q}_1 = \emptyset \}

\( \mathsf{if} (\mathsf{null}_0) \mathsf{then} \mathsf{follow}_1 \times \mathsf{last}_1 \mathsf{else} \mathsf{follow}_1 \mathsf{fi} \)

\( \mathsf{if} (\mathsf{null}_1) \mathsf{then} \mathsf{follow}_0 \times \mathsf{last}_0 \mathsf{else} \mathsf{follow}_0 \mathsf{fi} \)

\( \mathsf{if} (\mathsf{null}_0 \mathsf{and} \mathsf{null}_1) \mathsf{then} \mathsf{last}_0 \times \mathsf{first}_1 \mathsf{else} \mathsf{last}_0 \times \mathsf{first}_1 \mathsf{fi} \)

= \{ \text{Definitions of } \mathsf{W}_0 \text{ and } \mathsf{W}_1; \text{ rewriting} \}

\( \mathsf{if} (\mathsf{null}_0) \mathsf{then} \mathsf{W}_1 \mathsf{else} \mathsf{follow}_1 \mathsf{fi} \mathsf{or} (\mathsf{null}_1) \mathsf{then} \mathsf{W}_0 \mathsf{else} \mathsf{follow}_0 \mathsf{fi} \)

\( \mathsf{or} (\mathsf{null}_0 \mathsf{and} \mathsf{null}_1) \mathsf{then} \mathsf{W}_0 \mathsf{else} \mathsf{first}_1 \mathsf{fi} \)

= \{ \text{Definitions of } \mathsf{W}', \mathsf{W}'', \text{ and } \mathsf{W}'' \}

\( \mathsf{W}' \mathsf{or} \mathsf{W}'' \mathsf{or} \mathsf{W}''' \)

\( \mathsf{C}_{\mathsf{U}, \mathsf{RFA}''}: \) \( (\mathsf{follow}_0 \mathsf{or} \mathsf{follow}_1) \setminus ((\mathsf{last}_0 \mathsf{or} \mathsf{last}_1) \times (\mathsf{first}_0 \mathsf{or} \mathsf{first}_1)) \)

= \{ \text{Rewriting} \}

\( (\mathsf{follow}_0 \mathsf{or} \mathsf{follow}_1) \setminus ((\mathsf{last}_0 \times \mathsf{first}_0 \mathsf{or} \mathsf{last}_0 \times \mathsf{first}_1 \mathsf{or} \mathsf{last}_1 \times \mathsf{first}_0 \mathsf{or} \mathsf{last}_1 \times \mathsf{first}_1) \)

= \{ \text{Assumption that } \mathsf{Q}_0 \cap \mathsf{Q}_1 = \emptyset \}

\( \mathsf{follow}_0 \setminus (\mathsf{last}_0 \times \mathsf{first}_0) \mathsf{or} \mathsf{follow}_1 \setminus (\mathsf{last}_1 \times \mathsf{first}_1) \)

= \{ \text{Definitions of } \mathsf{W}_0 \text{ and } \mathsf{W}_1 \}

\( \mathsf{W}_0 \mathsf{or} \mathsf{W}_1 \)

\( \square \)
References


REFERENCES


Index

\(\circ\), see composition
\(\equiv\), see regular expressions, equivalence of
\(\rightarrow\), see function, total
\(\not\rightarrow\), see function, partial
\(\cong\), see isomorphism
\(\pi\), see tuple, projection
\(\pi'\), see tuple, projection
\(\sim\), see regular expressions, similarity of

Aho, Sethi, and Ullman's construction, see finite automata, construction, Aho, Sethi, and Ullman's algebras, see \(\Sigma\)-algebras
algorithms
subset construction, 12
alphabet, 6, 15, 16
ASU, see finite automata, construction, Aho, Sethi, and Ullman's automata, see finite automata

Berry and Sethi's construction, see finite automata, construction, Berry and Sethi's bias
left-to-right, 6
Brüggemann-Klein's construction, see finite automata, construction, Brüggemann-Klein
Brz, see finite automata, construction, Brzozowski
Brzenc, see finite automata, construction, Brzozowski
Brzozowski's construction, see finite automata, construction, Brzozowski
BS, see finite automata, construction, Berry and Sethi's
BSenc, see finite automata, construction, Berry and Sethi's (encoded)

C, see item set, closure of
carrier set, 14, 16
Chang and Paige's construction, see finite automata, construction, Chang-Paige closure
Kleene, 16
of an item set, 67-72
Complete, see finite automata, complete
complete, see finite automata, transformations on, complete
composition, 7, 11-13, 27, 33, 35-45, 48, 67, 69-72, 76, 78
constants, 14, 17
construction, see finite automata, construction
Aho-Sethi-Ullman, 13
powerset, see finite automata, subset construction
subset, see finite automata, subset construction
convert, 36, 37, 41-44

D, see item set, relation on
decode, 33, 35-41, 45, 48
DeRemer, see finite automata, construction, DeRemer
DeRemer's construction, see finite automata, construction, DeRemer
deriv, 56, 57
derivative
left, 36, 56-65, 67, 68, 72, 77
derivative item set, 68, 72
deriv_{ERE}, 59-61
Det, see finite automata, deterministic
deterministic finite automata, see finite automata, deterministic
Det', see finite automata, deterministic, weak
DFA, see finite automata, deterministic
Dijkstra, E.W., 12
DIS, see derivative item set
dots, 66-68
dotted regular expressions, 66, 67
DRE, see dotted regular expressions
dual, 11

E, see item set, closure of
direction of edge, labeled directed, 6
Ehrig, E., 14
encoderiv, 56, 57
encoder, 65
encode, 33, 45
c-free, see finite automata, c-free
c-lookahead finite automata, 23-27
c-transition
relation, 6
removal of, see finite automata (transf. on), c removal
equivalence
derivation of regular expressions, see regular expressions, equivalence of
equivalence relation, see relation, equivalence
ERE, see regular expressions, extended expression tree height, 15
extended regular expressions, see regular expressions, extended
language of, 44, 58–61, 64

FA, see finite automata
false, 16, 17, 32, 34, 50, 51
final states, see states, final
final-reachable states, see states, final-reachable
finite automata, 6–14, 16, 18, 21–23, 25, 27,
28, 31, 32, 35–37, 39–44, 45, 49, 55,
57, 74, 79
complete, 8–13, 37–39, 43, 53–57, 61, 68,
70, 71
collection
Aho, Sethi, and Ullman’s, 43
Berry and Sethi’s, 35, 37, 48
Berry and Sethi’s (encoded), 48, 65
Brüggemann-Klein, 52
Brzozowski, 61, 62, 65, 71
Brzozowski (encoded), 65
Chang-Paige, 50
DeRemer, 70
guarded commands, 26, 27
item set, 68–70
lookahead automata, 24–26
McNaughton, Yamada, and Glushkov’s,
37, 38, 48, 65
Myhill and Nerode’s, 55–57, 65, 66
Myhill and Nerode’s minimal, 57, 58,
61, 65
optimized item set, 70–72
Thompson, 18, 21–23
Thompson top-down, 22, 23
definition of, 6
deterministic, 9–13, 25, 35, 37, 38, 43,
48, 53–58, 60–63, 65, 66, 68–75
definition of, 9
minimality of, 10–12, 56
weak, 9, 55
č-free, 3, 8–12, 27, 28, 31, 32, 35, 37, 40
isomorphism of, 7, 8, 11
language of, 7, 8, 10–12, 18, 23, 29–31,
36, 40–42, 45, 49, 53, 55–57, 61, 68,
79
extension, 8
property of, 8
left-biased, 28–33, 35, 37, 44–46, 79
powerset construction, see finite automata,
subset construction
properties of, 7–10
reduced, see reduced finite automata
reverse of, 9, 10
extension of, 11
size of, 7

subset construction, 12, 13, 37–39
implementation of, 12
optimized, 12, 13, 43, 44
transformations on, 10–13
complete, 11, 48
c-transition removal, 11, 27
final-unreach. removal, 11–13, 37–39,
43, 44
start-unreach. removal, 11
useless state removal, 11
useful, 9–11
final, 9
start, 9–11
First, 24–26, 47–49, 65
first, 32–34, 36, 38, 39, 41, 42, 43, 44, 45, 47,
49–52, 79, 80
Follow, 47–49, 65
follow, 32–34, 36, 38, 39, 41, 42, 44, 45, 47,
49–52, 79, 80
FReachable, see states, final-reachable
function
partial, 9, 32, 46, 65, 76, 78
total, 6, 7, 9–12, 14–17, 22, 24, 26, 27,
31–33, 35–37, 42–44, 46–48, 53, 56,
57, 59, 61, 63–68, 76, 78
GCL, see guarded command language
guarded command language, 12, 26
Iconstr, see finite automata, construction, item
set
identity relation, see relation, identity
initial Σ-algebra, see Sigma-algebra, initial
IS, see item sets
isomorphism, 7, 8, 10, 11, 17–21, 23, 27–45,
48, 50–52, 56, 65, 79
classes, 8, 11, 18–21, 23, 27–45, 48, 50–
52, 65, 79
of finite automata, see finite automata,
isomorphism of
item, see dotted regular expression
language of, 66, 70
item set
closure of, 67–72
optimization of, 4, 69–72
construction, see finite automata, con-
struction, item set
DeRemer’s, see finite automata, con-
struction, DeRemer’s
improved, see finite automata, construc-
tion, item set (high quality)
derivative, see derivative item set
language of, 66–70
relation on, 67, 68
INDEX

item sets, 66–71

K, see finite automata, construction, look-ahead automata

LAFA, see ε-lookahead finite automata

language

of an item, 66, 70
of extended regular expressions, 44, 58–61
of finite automata, 7, 8, 10–12, 18, 23, 29–31, 36, 40–42, 45, 49, 53, 55–57, 61, 68, 79
extension of, 8
of item sets, 66–70
of regular expressions, 16–18, 22, 24, 25, 27, 36, 39, 44, 47, 49, 63–66, 68, 70
regular, 18, 39–42, 44, 55–57, 63, 76, 77
Σ-algebra of, 16

Last, 47, 48, 65

last, 32–34, 36, 38, 39, 41, 42, 44, 45, 49–52, 79, 80

LBFA, see finite automata, left-biased

left, 31–33, 35–37, 39, 40
left derivative, 36, 56–65, 67, 68, 72, 77
left-biased finite automata, see finite automata, left-biased

LFE, see extended regular expressions, language denoted by

LF, see finite automata, language of

LIS, see item set, language of

look, 24–26
lookahead finite automata, see ε-lookahead finite automata

LRE, see regular expressions, language denoted by

LR, see language, regular

LRFE, see reduced finite automata, language of

Mahr, B., 14
convertb, 36, 37, 41, 42
converta, 42, 43, 49, 71

max, 15

McNaughton, Yamada and Glushkov's construction, see finite automata, construction, McNaughton, Yamada and Glushkov's
merge, see finite automata, transformations on, merge

Min, see finite automata, deterministic, minimality of

MinC, see finite automata, deterministic, minimality of

Minimal, see finite automata, deterministic, minimality of

MinimalC, see finite automata, deterministic, minimality of

MN, 55

MNconstr, see finite automata, construction, Myhill and Nerode's

MNmin, see finite automata, construction, Myhill and Nerode's minimal

MYG, see finite automata, construction, McNaughton, Yamada, and Glushkov's
Myhill and Nerode's construction, see finite automata, construction, Myhill and Nerode's
minimal, see finite automata, construction, Myhill and Nerode's minimal

Myhill-Nerode theorem, 10

N, see finite automata, construction, guarded commands

nondeterministic finite automata, see finite automata, nondeterministic

Null, 17, 24, 47, 48, 63–65

null, 32–34, 36, 38, 39, 41, 42, 44, 45, 50–52, 79, 80

Occ, 46, 48, 64, 65

Oconstr, see finite automata, construction, optimized item set

operator set, 14

operators, 14, 15

arity of, 14

constant, see constants

optimization

of closure of an item set, 4, 69–72

P, see powerset

powerset, 6, 7, 9, 12, 16, 23, 24, 26, 32, 46, 47, 59, 64, 66, 68, 76

powerset construction, see finite automata, subset construction

preservation, 10

projection, 8, 33, 45, 52, 76, 78

properties

of finite automata, 7–10

Qmap, 32–34, 36, 38, 39, 41, 42, 44–46, 50–52

R, see item, language of

RBFA, see finite automata, right-biased

RE, see regular expressions

Reach, see states, reachability relation on

reachability relation

on states, see states, reachability relation on
Reachable, see states, reachable
reachable states, see state, reachability relation on
reduced finite automata, 4, 32–42, 44–52, 72, 73, 79, 80
Reg, 15
regular expressions, 15–18, 21–26, 28, 31, 33–37, 39–44, 46, 47, 49, 58, 63–71
constants, 15
definition of, 15
dotted, see dotted regular expressions
equivalence of, 16, 59–63
extended, 44, 58–63, 65, 71
language denoted by, 44, 58–61, 64
language denoted by, 16–18, 22, 24, 25, 27, 36, 39, 44, 47, 49, 63–66, 68, 70
operators, 15
abbreviations, 16
restricted, 46–49, 63–65
reverse of, 17
similarity, 59–61, 63, 64
similarity of, 59–64, 71
regular languages, see language, regular relation
e-transition, see e-transition relation
equivalence, 7
identity, 7
reachability, of states, see states, reachability relation on
transition, see transition, relation
remove, see finite automata (transf. on), e removal
remove_e_sym, see finite automata (transf. on), e removal
restricted regular expressions, see regular expressions, restricted
reversal
of finite automata, see finite automata, reverse of
of regular expressions, see regular expressions, reverse of
RFA, see reduced finite automata
rfa, 33, 35–37, 40–43, 46–48
right-biased finite automata, see finite automata, right-biased
RRE, see regular expressions, restricted

INDEX

Null, 17
\(\Sigma\)-term algebra, 15
signature, 14, 15
sink state, see states, sink
size of finite automata, see finite automata, size of
sort, 14
SReachable, see states, start-reachable
start states, see states, start
start-reachable states, see states, start-reachable
state graphs, 6
states, 6
final, 6
final-reachable, 8, 9
final-unreachable, removal, see finite automata (transf. on), final-unreach. removal
language between, 7
left language of, 7, 8
right language of, 7, 8
sink, 11–13
start, 6, 9
start-reachable, 8, 9, 11
start-unreachable, removal, see finite automata (transf. on), start-unreach. removal
useless, removal, see finite automata, transformations on, useless state removal
subset, see finite automata, subset construction
subset construction, see finite automata, subset construction
implementation of, 12
subsetopt, see finite automata, subset construction, optimized
symmetrical, see dual
function, 11, 17
transformation, 11
td, see finite automata, construction, Thompson top-down
Term_\(\Sigma\), 14
Th, see finite automata, construction, Thompson
Thompson’s construction, see finite automata, construction, Thompson top-down
top-down, see finite automata, construction, Thompson
transformations
on finite automata, 10–13
transition
function, 9
relation, 6
extension of, 7
true, 16, 17, 32, 34, 47, 48, 50–52
tuple, 6
    projection, 8, 33, 45, 52, 76, 78

undot, 66, 68
Useful, see finite automata, useful
useful, see finite automata, transformations
    on, useless state removal
useful states, see states, useful
Useful_f, see finite automata, useful, final
useful_f, see finite automata (transf. on), final-
    unreach. removal
Useful_s, see finite automata, useful, start
useful_s, see finite automata (transf. on), start-
    unreach. removal

vertices, 6

X, see item set, optimization of closure of

Y, see item set, optimization of closure of
<table>
<thead>
<tr>
<th>Computing Science Reports</th>
<th>Department of Mathematics and Computing Science</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Eindhoven University of Technology</td>
</tr>
</tbody>
</table>

**In this series appeared:**

<table>
<thead>
<tr>
<th>Code</th>
<th>Authors</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>91/02</td>
<td>R.P. Nederpelt H.C.M. de Swart</td>
<td>Implication. A survey of the different logical analyses &quot;if...,then...&quot;, p. 26.</td>
</tr>
<tr>
<td>91/03</td>
<td>J.P. Katoen L.A.M. Schoenmakers</td>
<td>Parallel Programs for the Recognition of $P$-invariant Segments, p. 16.</td>
</tr>
<tr>
<td>91/05</td>
<td>D. de Reus</td>
<td>An Implementation Model for GOOD, p. 18.</td>
</tr>
<tr>
<td>91/06</td>
<td>K.M. van Hee</td>
<td>SPECIFICATIEMETHODEN, een overzicht, p. 20.</td>
</tr>
<tr>
<td>91/07</td>
<td>E. Poll</td>
<td>CPO-models for second order lambda calculus with recursive types and subtyping, p. 49.</td>
</tr>
<tr>
<td>91/12</td>
<td>E. van der Sluis</td>
<td>A parallel local search algorithm for the travelling salesman problem, p. 12.</td>
</tr>
<tr>
<td>91/14</td>
<td>P. Lemmens</td>
<td>The PDB Hypermedia Package. Why and how it was built, p. 63.</td>
</tr>
<tr>
<td>91/16</td>
<td>A.J.J.M. Marcelis</td>
<td>An example of proving attribute grammars correct: the representation of arithmetical expressions by DAGs, p. 25.</td>
</tr>
<tr>
<td>Session</td>
<td>Authors</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>91/18</td>
<td>Rik van Geldrop</td>
<td>Transformational Query Solving, p. 35.</td>
</tr>
<tr>
<td>91/19</td>
<td>Erik Poll</td>
<td>Some categorical properties for a model for second order lambda calculus with subtyping, p. 21.</td>
</tr>
<tr>
<td>91/23</td>
<td>K.M. van Hee, L.J. Somers, M. Voorhoeve</td>
<td>Z and high level Petri nets, p. 16.</td>
</tr>
<tr>
<td>91/24</td>
<td>A.T.M. Aerts, D. de Reus</td>
<td>Formal semantics for BRM with examples, p. 25.</td>
</tr>
<tr>
<td>91/25</td>
<td>P. Zhou, J. Hooman, R. Kuiper</td>
<td>A compositional proof system for real-time systems based on explicit clock temporal logic: soundness and completeness, p. 52.</td>
</tr>
<tr>
<td>91/27</td>
<td>F. de Boer, C. Palamidessi</td>
<td>Embedding as a tool for language comparison: On the CSP hierarchy, p. 17.</td>
</tr>
<tr>
<td>91/28</td>
<td>F. de Boer</td>
<td>A compositional proof system for dynamic process creation, p. 24.</td>
</tr>
<tr>
<td>91/30</td>
<td>J.C.M. Baeten, F.W. Vaandrager</td>
<td>An Algebra for Process Creation, p. 29.</td>
</tr>
<tr>
<td>91/31</td>
<td>H. ten Eikelder</td>
<td>Some algorithms to decide the equivalence of recursive types, p. 26.</td>
</tr>
<tr>
<td>91/33</td>
<td>W. v.d. Aalst</td>
<td>The modelling and analysis of queueing systems with QNM-ExSpect, p. 23.</td>
</tr>
<tr>
<td>91/34</td>
<td>J. Coenen</td>
<td>Specifying fault tolerant programs in deontic logic, p. 15.</td>
</tr>
<tr>
<td>Year</td>
<td>Authors</td>
<td>Title</td>
</tr>
<tr>
<td>------</td>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>92/01</td>
<td>J. Coenen, J. Zwiers, W.-P. de Roever</td>
<td>A note on compositional refinement, p. 27.</td>
</tr>
<tr>
<td>92/02</td>
<td>J. Coenen, J. Hooman</td>
<td>A compositional semantics for fault tolerant real-time systems, p. 18.</td>
</tr>
<tr>
<td>92/03</td>
<td>J.C.M. Baeten, J.A. Bergstra</td>
<td>Real space process algebra, p. 42.</td>
</tr>
<tr>
<td>92/05</td>
<td>J.P.H.W.v.d.Eijnde</td>
<td>Conservative fixpoint functions on a graph, p. 25.</td>
</tr>
<tr>
<td>92/06</td>
<td>J.C.M. Baeten, J.A. Bergstra</td>
<td>Discrete time process algebra, p. 45.</td>
</tr>
<tr>
<td>92/07</td>
<td>R.P. Nederpelt</td>
<td>The fine-structure of lambda calculus, p. 110.</td>
</tr>
<tr>
<td>92/10</td>
<td>P.M.P. Rambags</td>
<td>Composition and decomposition in a CPN model, p. 55.</td>
</tr>
<tr>
<td>92/13</td>
<td>F. Kamareddine</td>
<td>Set theory and nominalisation, Part II, p. 22.</td>
</tr>
<tr>
<td>92/14</td>
<td>J.C.M. Baeten</td>
<td>The total order assumption, p. 10.</td>
</tr>
<tr>
<td>92/15</td>
<td>F. Kamareddine</td>
<td>A system at the cross-roads of functional and logic programming, p. 36.</td>
</tr>
<tr>
<td>92/16</td>
<td>R.R. Seljée</td>
<td>Integrity checking in deductive databases; an exposition, p. 32.</td>
</tr>
<tr>
<td>92/17</td>
<td>W.M.P. van der Aalst</td>
<td>Interval timed coloured Petri nets and their analysis, p. 20.</td>
</tr>
<tr>
<td>92/18</td>
<td>R.Nederpelt, F. Kamareddine</td>
<td>A unified approach to Type Theory through a refined lambda-calculus, p. 30.</td>
</tr>
<tr>
<td>92/20</td>
<td>F. Kamareddine</td>
<td>Are Types for Natural Language? P. 32.</td>
</tr>
</tbody>
</table>
Non well-foundedness and type freeness can unify the interpretation of functional application, p. 16.

A useful lambda notation, p. 17.

Nominalization, Predication and Type Containment, p. 40.

Bottum-up Abstract Interpretation of Logic Programs, p. 33.

A Programming Logic for Fσ, p. 15.

A modelling method using MOVIE and SimCon/ExSpect, p. 15.

A taxonomy of keyword pattern matching algorithms, p. 50.

Deriving the Aho-Corasick algorithms: a case study into the synergy of programming methods, p. 36.

A continuous version of the Prisoner’s Dilemma, p. 17

Quicksort for linked lists, p. 8.

Deterministic and randomized local search, p. 78.

A congruence theorem for structured operational semantics with predicates, p. 18.

On the unavoidability of metastable behaviour, p. 29

Exercises in Multiprogramming, p. 97

A Formal Deterministic Scheduling Model for Hard Real-Time Executions in DEDOS, p. 32.


Systems Engineering: a Formal Approach Part II: Frameworks, p. 44.


93/14 J.C.M. Baeten
J.A. Bergstra

93/15 J.C.M. Baeten
J.A. Bergstra
R.N. Bol

93/16 H. Schepers
J. Hooman
A Trace-Based Compositional Proof Theory for Fault Tolerant Distributed Systems, p. 27

93/17 D. Alstein
P. van der Stok
Hard Real-Time Reliable Multicast in the DEDOS system, p. 19.

93/18 C. Verhoef
A congruence theorem for structured operational semantics with predicates and negative premises, p. 22.

93/19 G.J. Houben
The Design of an Online Help Facility for ExSpect, p. 21.

93/20 F.S. de Boer

93/21 M. Codish
D. Dams
G. Filé
M. Bruynooghe

93/22 E. Poll
A Typechecker for Bijective Pure Type Systems, p. 28.

93/23 E. de Kogel
Relational Algebra and Equational Proofs, p. 23.

93/24 E. Poll and Paula Severi
Pure Type Systems with Definitions, p. 38.

93/25 H. Schepers and R. Gerth

93/26 W.M.P. van der Aalst
Multi-dimensional Petri nets, p. 25.

93/27 T. Kloks and D. Kratsch
Finding all minimal separators of a graph, p. 11.

93/28 F. Kamareddine and R. Nederpelt
A Semantics for a fine λ-calculus with de Bruijn indices, p. 49.

93/29 R. Post and P. De Bra
GOLD, a Graph Oriented Language for Databases, p. 42.

93/30 J. Deogun
T. Kloks
D. Kratsch
H. Müller
On Vertex Ranking for Permutation and Other Graphs, p. 11.

93/31 W. Körver
Derivation of delay insensitive and speed independent CMOS circuits, using directed commands and production rule sets, p. 40.

93/32 H. ten Eikelder and H. van Geldrop

93/33 L. Loyens and J. Moonen
ILIAS, a sequential language for parallel matrix computations, p. 20.
<table>
<thead>
<tr>
<th>No.</th>
<th>Authors</th>
<th>Paper Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>93/34</td>
<td>J.C.M. Baeten and J.A. Bergstra</td>
<td>Real Time Process Algebra with Infinitesimals, p.39.</td>
</tr>
<tr>
<td>93/36</td>
<td>J.C.M. Baeten and J.A. Bergstra</td>
<td>Non Interleaving Process Algebra, p. 17.</td>
</tr>
<tr>
<td>93/38</td>
<td>C. Verhoef</td>
<td>A general conservative extension theorem in process algebra, p. 17.</td>
</tr>
<tr>
<td>93/41</td>
<td>A. Bijlsma</td>
<td>Temporal operators viewed as predicate transformers, p. 11.</td>
</tr>
<tr>
<td>93/42</td>
<td>P.M.P. Rambags</td>
<td>Automatic Verification of Regular Protocols in P/T Nets, p. 23.</td>
</tr>
</tbody>
</table>