Chart semantics for regular expressions: characterization and axiomatization

Your Name

November 12, 2025

Abstract

We define a Milner chart, a transition structure naturally associated with a regular expression. We provide a rewriting-based characterization of these charts and establish a precise correspondence between bisimilarity of Milner charts and equivalence in Milner's equational system for regular expressions.

1 Introduction

2 Preliminaries

2.1 Regular Expressions

Definition 2.1. *The set of* regular expressions *over an alphabet* Σ *is defined by the following grammar:*

$$e ::= 0 \mid 1 \mid a \mid e + e \mid e \cdot e \mid e^* \qquad (a \in \Sigma)$$

Notation 2.2. We sometimes write ef for $e \cdot f$. The precedence of operations is as usual: star, then concatenation, then union. For instance, $e + f^*g$ means $e + (f^* \cdot g)$.

While the semantics of regular expressions are traditionally defined as word languages, in this paper we will define them through charts, which we introduce in the following section.

2.2 Charts

Definition 2.3. *A* chart *over an alphabet* Σ *is a tuple* $(Q, \Sigma, \Delta, I, F)$ *where:*

- Q *is a finite set of* states*,*

- Σ *is a finite input alphabet,*
- $\Delta \subseteq Q \times \Sigma \times Q$ *is a set of* transitions*,*
- $I \in Q$ *is the* initial *state,*
- $F \subseteq Q$ *is the set of* final *states.*

The set of derivatives of a state q is defined by:

$$D(q) = \{(a, q') \mid (q, a, q') \in \Delta\},$$

and its finality is the function χ defined as follows:

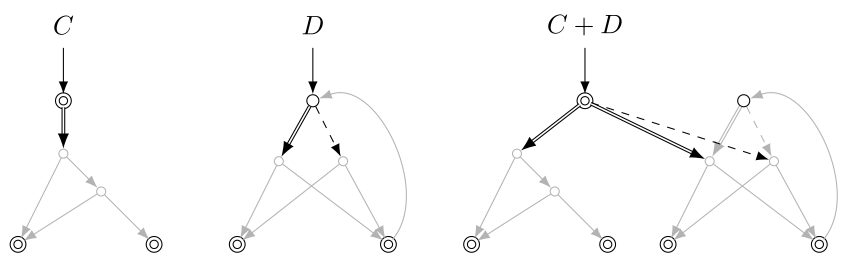
$$\chi(q) = \begin{cases} 1 & \text{if } q \text{ is final,} \\ 0 & \text{otherwise.} \end{cases}$$

Remark 2.4. *When clear from context or irrelevant, we omit the alphabet of a chart. We sometimes define a chart by specifying the derivatives of its states rather than its transitions set; this is equivalent.*

2.3 Milner Charts

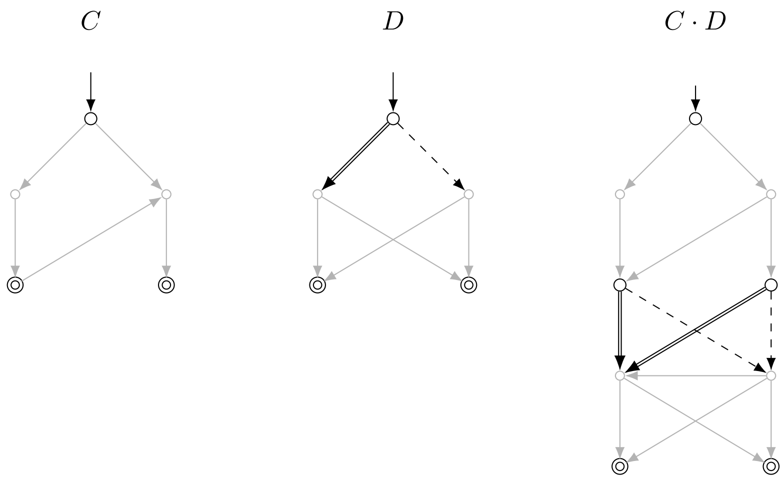
Definition 2.5. *The* union *of two charts C and D, denoted $C + D$, is obtained by taking their disjoint union, adding a fresh initial state s that inherits the derivatives of the initial states of C and D, and making s final whenever one of them is final. Finally, states not reachable from s are removed.*

Example 2.6. *Here is an example of the sum of two charts C and D:*



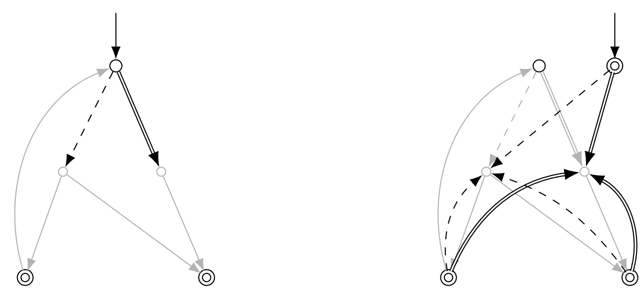
Definition 2.7. *The* sequential composition *of two charts C and D, denoted $C \cdot D$, is obtained from their disjoint union by adding to the derivatives of each final state of C the derivatives of the initial state of D, and marking it final whenever the initial state of D is final. The initial state of C becomes that of the composition, and unreachable states are removed.*

Example 2.8. *An example of sequential composition of two charts C and D:*



Definition 2.9. *The* Kleene star *of a chart C, denoted C^*, is obtained by adding to each final state of C the derivatives of its initial state, which remains initial and becomes also final. Unreachable states are then removed.*

Example 2.10. *Example of ietration of a chart.*



Definition 2.11. *The constant chart* $\mathbf{0}$ *has a single state, which is initial and non-final, with no transitions. The constant chart* $\mathbf{1}$ *has a single state, both initial and final, with no transitions. For each letter a, the chart* $\mathbf{a}$ *has two states, one initial and one final, connected by a single a-labeled transition.*

Definition 2.12. *The* Milner chart *of an expression e over the alphabet Σ, denoted $M(e)$, is defined by induction as follows, where $a \in \Sigma$:*

$$\begin{array}{rclrclrcl} M(0) & = & \mathbf{0} & M(1) & = & \mathbf{1} & M(a) & = & \mathbf{a} \\ M(e+f) & = & M(e)+M(f) & M(e \cdot f) & = & M(e) \cdot M(f) & M(e^*) & = & M(e)^* \end{array}$$

Example 2.13. *The milner chart of $a(b+c), (a(b+c))^*$ et $ab^* + c^*$.*

Definition 2.14. *A chart is* rooted *if its initial state has no ingoing transitions.*

Proposition 2.15. *For every regular expression e, $M(e)$ is rooted.*

2.4 Bisimulation

Notation 2.16. If C is a chart, we denote by $\xrightarrow{a}_C$ the relation on its set of states defined as:

$$\xrightarrow{a}_C \; = \; \{(p,q) \mid (p,a,q) \text{ is a transition of } C\}.$$

We denote by $\xleftarrow{a}_C$ the converse of $\xrightarrow{a}_C$.

Definition 2.17. *Let C and D be two charts with state sets Q and Q' and respectively. A relation $R \subseteq Q \times Q'$ is a* bisimulation *between C and D if:*

- $(p,q) \in R \Rightarrow \chi(p) = \chi(q)$,
- $\xleftarrow{a}_C \cdot R \; = \; R \cdot \xleftarrow{a}_D$ *and*
- $\xrightarrow{a}_C \cdot R \; = \; R \cdot \xrightarrow{a}_D$.

Let $(p,q) \in C \times D$. We say that p and q are bisimilar*, and write $p \sim q$, if $(p,q) \in R$ for some bisimilation R between C and D.*

We say that C and D are bisimilar*, and write $C \sim D$, if there initial states are bisimilar.*

2.5 Milner's Axiomatization

Definition 2.18. *For every regular expression e, we write $e \Downarrow$ if the initial state of $M(e)$ is final. We write $e \Uparrow$ otherwise.*

Definition 2.19. Milner's proof system *contains the following axioms and deduction rules:*

Algebraic laws:

(1) $e + f = f + e$

(2) $(e + f) + g = e + (f + g)$

(3) $e + e = e$

(4) $e^* = 1 + ee^*$

(5) $(ef)g = e(fg)$

(6) $1e = e = e1$

(7) $(e + f)g = eg + fg$

(8) $e^* = (e + 1)^*$

(9) $e + 0 = e$

(10) $0e = 0 = e0$

(11) $0^* = 1$

Salooma's induction rule:

$$\frac{f = ef + g \qquad e \Uparrow}{f = e^* g}$$

Congruence rules:

$$\frac{}{e = e} \qquad \frac{e = f}{f = e} \qquad \frac{e = f \;\; f = h}{e = h}$$

$$\frac{e_1 = f_1 \;\; e_2 = f_2}{e_1 + e_2 = f_1 + f_2} \qquad \frac{e_1 = f_1 \;\; e_2 = f_2}{e_1 e_2 = f_1 e_2} \qquad \frac{e = f}{e^* = f^*}$$

We say that two regular expressions e and f are provably equivalent *in the Milner system, denoted $e \equiv f$, if the equation $e = f$ is derivable from the axioms (1-11), Salomaa's rule and congruence rules.*

Our goal in the remainder of this paper is to address two questions: whether Milner charts can be recognized among all charts, and whether Milner's axiomatization is complete with respect to bisimulation equivalence. We will answer these questions by means of a rewriting system introduced in the next section.

3 Rewriting system for charts

We work with a slightly broader class of structures, called *networks*, which extend ordinary charts by allowing multiple inputs. The rewriting system will be defined on these networks.

Definition 3.1. *We consider the rewriting rules of Figure 1, which transform networks over Σ^* into networks over Σ^*. In this picture,*

- *squares denote states for which only some transitions are shown,*
- *circles denote states whose transitions are all explicitly displayed,*
- *circles are distinct from surrounding nodes and are neither initial nor final.*

We write $\twoheadrightarrow$ for the union of all these rewriting rules.

As in other rewriting systems, such as state elimination for automata, we require that initial and final states have no incoming and no outgoing transitions, respectively. We introduce a *normalization* operation that enforces that.

Definition 3.2. *The* normalization *of a network C, denoted $|C|$, is the network obtained as follows:*

- *for each index i, if the i-th initial state has incoming transitions, add a fresh state which becomes the new i-th initial state and add a* 1*-transition from this new state to the former one,*
- *add a fresh state, called the* sink*, which becomes the unique final state,*
- *for each former final state, add a* 1*-transition to the sink.*

Definition 3.3. *Let φ be a function of domain $\{1, \ldots, n\}$. We define the network $N(\varphi)$ at follows: it has n distinct initial states denoted $1, \ldots, n$ and a single final state. The state named i is the i-th initial state. For each $i \in \{1, \ldots, n\}$ there is exactly one transition, labelled $\varphi(i)$, from state i to the final state.*

Definition 3.4. *For every regular expressions e, we let $N(e)$ be the chart:*

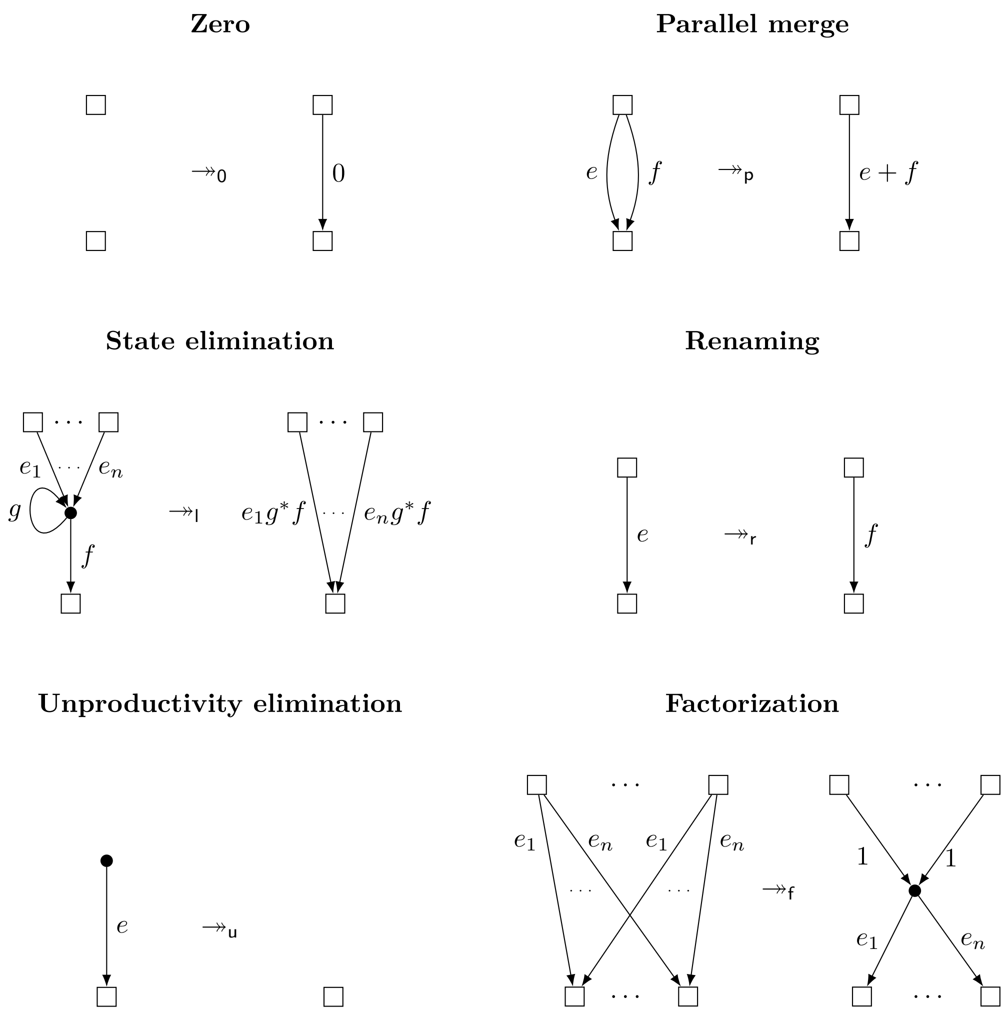
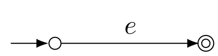


Figure 1: Rewriting rules for networks



A chart M is reducible *to an expression e, written $M \twoheadrightarrow^* e$, if:*

$$[\![M]\!] \twoheadrightarrow^* N(e).$$

Definition 3.5. *A regular expression is* star-safe *if in every sub-expression of the form f^*, we have $f \Uparrow$.*

Proposition 3.6. *Every expression is equivalent to a star-safe one.*

Proposition 3.7. *For every star-safe expression e, $M(e)$ is reducible to e.*

Lemma 3.8. *Let M and N be two rooted charts. We have:*

$$M \twoheadrightarrow^* e \text{ and } N \twoheadrightarrow^* f \quad \Longrightarrow \quad (M+N) \twoheadrightarrow^* (e+f) \quad \text{and} \quad (M \cdot N) \twoheadrightarrow^* (e \cdot f)$$

$$M \Uparrow \text{ and } M \twoheadrightarrow^* e \quad \Longrightarrow \quad M^* \twoheadrightarrow^* e^*.$$

4 Labeling Charts

Definition 4.1. *Let M be a chart whith a transition relation Δ. A* labeling *of M is a function $\ell : Q \mapsto Re(\Sigma^*)$ satisfying, for every state p:*

- $\ell(p) \equiv \sum_{(p,a,q)\in\Delta} a\ell(q) + \chi(\ell(p))$,
- $\chi(p) = \chi(\ell(p))$.

The image of the initial state by ℓ is called its root. If M has a labeling with root e, we write:

$$M \models e.$$

4.1 Reducible implies labelable

If a chart is reducible to an expression e, then it is labelable with root e.

Proposition 4.2. *For every chart M and expression e, we have:*

$$M \twoheadrightarrow^* e \quad \Longrightarrow \quad M \models e.$$

Proof. We show first that labelability is preserved by the converse of $\twoheadrightarrow$:

Claim 4.3. *Let M and N be two charts. We have:*

$$M \twoheadrightarrow N \quad \text{and} \quad N \models e \quad \Longrightarrow \quad M \models e.$$

Proof. By case analysis on the applied rewriting rule. ☐

We conclude by noticing that the chart $N(e) \models e$. ☐

Corollary 4.4. *For every expression e, $M(e) \models e$.*

4.2 Charts with bisimilar roots are bisimilar

Proposition 4.5. *Let M and N be charts and e and f expressions. We have:*

$$M \models e, \quad N \models f \quad \textit{and} \quad e \sim f \quad \Longrightarrow \quad M \sim N.$$

Proof. let ℓ and ℓ' be labelings of M and N with roots e and f respectively. We show that the relation R between the states of M and N defined as follows:

$$(p, q) \in R \quad \Leftrightarrow \quad \ell(p) \sim \ell'(q),$$

is a bisimulation between M and N. □

If a chart has a labeling with root e, then it is bisimilar to $M(e)$.

Corollary 4.6. *For every chart M and expression e, we have:*

$$M \models e \quad \Longrightarrow \quad M \sim M(e).$$

4.3 Confluence of the rewriting system

Definition 4.7. *We say that a cycle is* final *if all its labels are final. A chart is* safe for reduction *if it has no final cycle.*

Remark 4.8. *Note that charts over the alphabet Σ are safe for reduction.*

Proposition 4.9. *Let M be a chart safe for reduction. We have:*

$$M \models e \quad \textit{and} \quad M \twoheadrightarrow^* f \quad \Longrightarrow \quad e \equiv f,$$

for every expressions e and f.

Lemma 4.10. *Reduction safety is preserved by reduction.*

We show that reduction preserves the existence of a labeling with the same root.

Claim 4.11. *Let M and N be two charts safe for reduction. We have:*

$$M \twoheadrightarrow N \quad \textit{and} \quad M \models e \quad \Longrightarrow \quad N \models e,$$

for every expression e.

Proof. By case analysis on the rewriting rule. □

Proof. □

As a consequence, we show that a chart is reducible to two expressions, then these expressions are equivalent.

Corollary 4.12. *Let M be a chart, and e and f be two expressions. We have:*

$$M \twoheadrightarrow^* e \quad \textit{and} \quad M \twoheadrightarrow^* f \quad \Longrightarrow \quad e \equiv f.$$

5 Adding new initial states

We use the same rewriting system as for charts.

Definition 5.1. *Let M be a chart over Σ whose state set is Q and initial state set I, and let E be a set of fresh names.*

A bridge *of* external shore *E on M is a subset of $E \times \Sigma \times (Q \setminus I)$.*

The bridge application *of B to M, denoted $B \triangleright M$, is the generalized chart obtained from M by adding E to the initial states set and adding B to the transitions set, keeping the other components unchanged.*

Lemma 5.2. *Let M and M' be a charts and B a bridge for M. We have:*

$$M \twoheadrightarrow M' \quad \Longrightarrow \quad B \triangleright M \twoheadrightarrow B' \triangleright M'$$

where B' is a bridge for M' having the same external shore as B.

6 Reducing traps

7 Erasing transitions from Milner charts

In general, erasing transitions from a Milner chart does not yields a Milner Chart. However, we show that if the erased transitions are *coherent* and *border*, then it does.

Definition 7.1. *Let M be a chart whose transition set is Δ.*

A transition of M is border *if its source is either initial or final.*

A set $\Gamma \subseteq \Delta$ is coherent*, if for every transitions $t, u \in \Delta$, whose sources are final and having the same label and target, we have:*

$$t \in \Gamma \quad \Leftrightarrow \quad u \in \Gamma.$$

Definition 7.2. *A* letter erasing *of an expression e is an expression obtained from e by replacing some letter occurrences by 0.*

Proposition 7.3. *Let e be an expression and Γ be a coherent set of border transitions of $M(e)$. There is a letter erasing f of e such that:*

$$M(e) \setminus \Gamma = M(f).$$

8 Chart quotients

8.1 Definition and properties of quotients

Definition 8.1. *The quotient of a chart M, denoted $[\![M]\!]$, is the chart obtained from M by identifying all bisimilar states.*

Proposition 8.2. *For every chart M and expression e, we have:*

$$[\![M]\!] \models e \quad \Longrightarrow \quad M \models e.$$

Proposition 8.3. *Let M be a chart. We have:*

$$M \twoheadrightarrow e \;\; \text{and} \;\; [\![M]\!] \twoheadrightarrow f \quad \Longrightarrow \quad e \equiv f,$$

for every expressions e and f.

8.2 Quotients of Milner charts are reducible

Proposition 8.4. *For every expression e, $[\![M(e)]\!]$ is reducible.*

9 Proof of completeness

We can now prove the completeness of Milner's axiomatization.

Theorem 9.1. *For every expressions e and f, we have:*

$$e \sim f \quad \Longrightarrow \quad e \equiv f.$$

Proof. Let e and f be two expressions such that $e \sim f$. The charts $M(e)$ and $M(f)$ being trim and bisimilar, they have the same quotient Q. By Proposition 8.4, there is an expression g such that:

$$Q \models g.$$

Using Proposition 8.2, we have:

$$M(e) \models g \quad \text{and} \quad M(f) \models g.$$

By Proposition 4.4, we also have:

$$M(e) \models e \quad \text{and} \quad M(f) \models f$$

By Corollary 4.12, we have then:

$$e \equiv g \quad \text{and} \quad f \equiv g.$$

Hence $e \equiv f$. □

References

[1] V. Antimirov, "Partial derivatives of regular expressions and finite automaton constructions," *Theoretical Computer Science*, 155(2): 291–319, 1996.