Chart semantics for regular expressions: characterization and axiomatization

Your Name

November 19, 2025

Abstract

We introduce Milner charts, a class of transition structures naturally associated with regular expressions. Our main contributions are: (i) a rewriting-based characterization that identifies Milner charts among all charts, and (ii) a completeness result establishing that bisimilarity of expressions coincides with their provable equivalence in Milner's equational axiomatization.

1 Introduction

2 Preliminaries

This section introduces the main concepts used throughout the paper: regular expressions and their chart-based semantics, bisimulation equivalence between charts, and Milner's equational proof system for reasoning about expression equivalence.

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:*

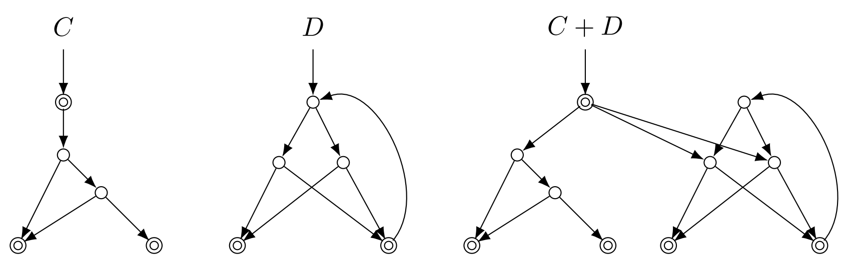
$$\chi(q) = \begin{cases} 1 & \text{if } q \text{ is final,} \\ 0 & \text{otherwise.} \end{cases}$$

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. As usual, we shall not distinguish between isomorphic charts.*

2.3 Milner Charts

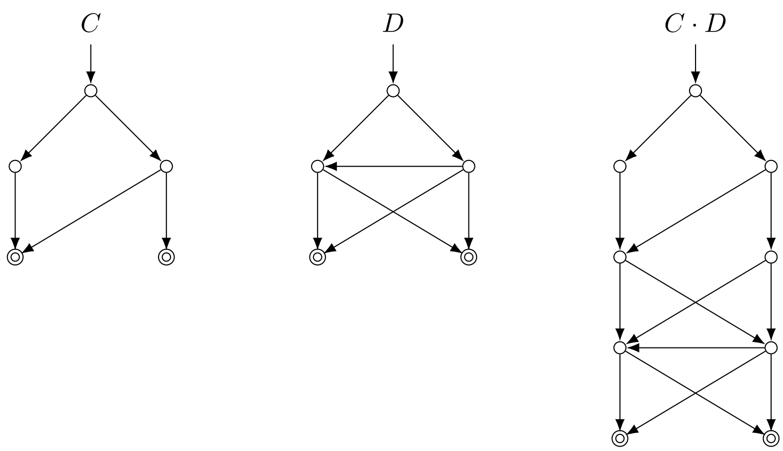
Definition 2.5. *The* union *of two charts* C *and* D*, denoted* $C + D$*, is obtained by taking their disjoint union, adding a fresh initial state* s *that inherits the derivatives of the initial states of* C *and* D*, and making* s *final whenever one of them is final. Finally, states not reachable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

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



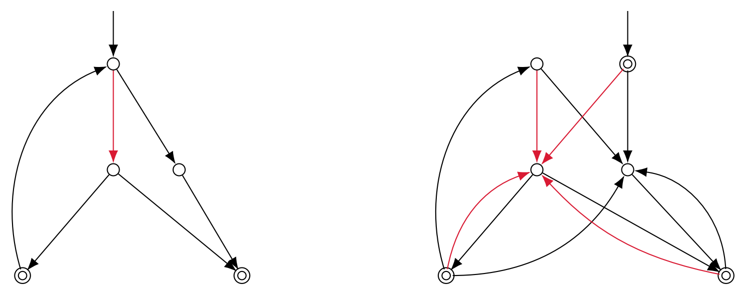
Definition 2.7. *The* sequential composition *of two charts* C *and* D*, denoted* $C \cdot D$*, is obtained from their disjoint union by adding to the derivatives of each final state of* C *the derivatives of the initial state of* D*, and marking it final whenever the initial state of* D *is final. The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



Definition 2.11. *The constant chart* **0** *has a single state, which is initial and non-final, with no transitions. The constant chart* **1** *has a single state, both initial and final, with no transitions. For each letter a, the chart* **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 $a \in \Sigma$:*

$$\mathrm{M}(0) = \mathbf{0} \qquad \mathrm{M}(1) = \mathbf{1} \qquad \mathrm{M}(a) = \mathbf{a}$$
$$\mathrm{M}(e+f) = \mathrm{M}(e) + \mathrm{M}(f) \qquad \mathrm{M}(e \cdot f) = \mathrm{M}(e) \cdot \mathrm{M}(f) \qquad \mathrm{M}(e^*) = \mathrm{M}(e)^*$$

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,* $\mathrm{M}(e)$ *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:*

- $(p,q) \in R \quad \Rightarrow \quad \chi(p) = \chi(q)$,
- $\xleftarrow{a}_C \cdot R \;=\; R \cdot \xleftarrow{a}_D$ *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 expression e, we say that e is* final*, and write $e \Downarrow$, if the initial state of $\mathrm{M}(e)$ is final. We write $e \Uparrow$ otherwise.*

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

Algebraic laws:

(1) $e + f = f + e$	*(5)* $(ef)g = e(fg)$	*(9)* $e + 0 = e$
(2) $(e + f) + g = e + (f + g)$	*(6)* $1e = e = e1$	*(10)* $0e = 0 = e0$
(3) $e + e = e$	*(7)* $(e + f)g = eg + fg$	*(11)* $0^* = 1$
(4) $e^* = 1 + ee^*$	*(8)* $e^* = (e + 1)^*$	

Salooma's induction rule:

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

Congruence rules:

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

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

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. *A* network *has the same data as a chart, but instead of a single initial state, it has a set of initial states.*

Definition 3.2. *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.

Definition 3.3. *Let φ be a function from a finite set S to regular expressions. The* network $\mathrm{N}(\varphi)$ *is defined as follows. Its set of states is S plus a fresh state f. Its set of initial states is S and its unique final state is f. For every $s \in S$, there is a transition $(s, \varphi(s), f)$.*

Definition 3.4. *Let N be a network. If there is a function φ from the set of initial states of N to regular expressions, such that $N \twoheadrightarrow^* \mathrm{N}(\varphi)$, we say that N* reduces *to φ and write:*

$$N \twoheadrightarrow^* \varphi.$$

Remark 3.5. *When N is a chart, and thus φ can be identified with a single expression e, in that case, we write:*

$$N \twoheadrightarrow^* e.$$

Definition 3.6. *A regular expression is* star-safe *if every sub-expression f^* satisfies $f \Uparrow$.*

Proposition 3.7. *Every expression is provably equivalent to a star-safe one.*

Proposition 3.8. *For every star-safe expression e,* $\mathrm{M}(e)$ *reduces to e.*

Lemma 3.9. *Let M and N be two rooted charts. We have:*

$$M \twoheadrightarrow^* e \text{ and } N \twoheadrightarrow^* f \implies (M+N) \twoheadrightarrow^* (e+f) \text{ and } (M \cdot N) \twoheadrightarrow^* (e \cdot f)$$
$$M \Uparrow \text{ and } M \twoheadrightarrow^* e \implies M^* \twoheadrightarrow^* e^*.$$

4 Labeling Charts

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

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

The restriction of ℓ to the set of initial states of N is called the interface *of the labeling. If N admits a labeling with interface φ, we write*

$$N \models \varphi.$$

The rewriting rules interact well with the notion of labelability. Indeed, labelability is preserved by the inverse of the reduction relation, and as a corollary we deduce that reducibility implies labelability (Section 4.1). Labelability is also preserved under reduction (under a safety hypothesis), and a corollary of this is a form of confluence (Section 4.2).

4.1 Reducible implies labelable

Lemma 4.2. *Let M and N be two networks and φ a function. We have:*

$$M \twoheadrightarrow N \text{ and } N \models \varphi \implies M \models \varphi.$$

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

By noticing that for every finite domain function φ, we have $\mathrm{N}(\varphi) \models \varphi$, we obtain the following proposition.

Proposition 4.3. *For every network N and function φ, we have:*

$$N \twoheadrightarrow^* \varphi \implies N \models \varphi.$$

4.2 Confluence of the rewriting system

Definition 4.4. *Let N be a network over regular expressions. A cycle in N is* final *if all its transition labels are final. We say that N is* safe for reduction *if it has no final cycles.*

Remark 4.5. *Note that charts over the alphabet Σ are safe for reduction.*

Lemma 4.6. *Let M and N be networks safe for reduction and let φ be a function. We have:*

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

Proof. By case analysis on the rewriting rule. ☐

Proposition 4.7. *Let N be a network safe for reduction and let φ and ψ be two functions. We have:*

$$N \models \varphi \quad \text{and} \quad N \twoheadrightarrow^* \psi \quad \Longrightarrow \quad \varphi \equiv \psi.$$

Proof. By iterating lemma 4.6, and noticing that:

$$\mathrm{N}(\psi) \models \varphi \quad \Rightarrow \quad \varphi \equiv \psi.$$

☐

As a consequence, we show a sort of confluence for the rewriting system.

Corollary 4.8. *Let N be a network. We have:*

$$N \twoheadrightarrow^* \varphi \quad \text{and} \quad N \twoheadrightarrow^* \psi \quad \Longrightarrow \quad \varphi \equiv \psi.$$

for every functions φ and ψ.

5 Adding, removing and duplicating initial states

Promoting states to initial states, or conversely removing their initial-state status, does not change a network's reducibility. Likewise, duplicating an initial state does not change it either.

Proposition 5.1. *Let N be a network and φ be a function such that:*

$$N \twoheadrightarrow^* \varphi.$$

Let M be the network obtained from N by restricting its set of initial states, while all other data remain identical.. If ψ is the restriction of φ to the initial states of M, then:

$$M \twoheadrightarrow^* \psi.$$

Proposition 5.2. *Let N be a network and φ be a function such that:*

$$N \twoheadrightarrow^* \varphi.$$

Let M be the network obtained from N by extending its set of initial states, while all other data remain identical. There is an extension ψ of φ such that:

$$M \twoheadrightarrow^* \psi.$$

Proposition 5.3. *Let N be a network and φ be a function such that:*

$$N \twoheadrightarrow^* \varphi.$$

Let i be an initial state of N and let M be the network obtained from N by adding a fresh initial state j with no ingoing transitions and the same outgoing transitions as i. There is an extension ψ of φ such that:

$$M \twoheadrightarrow^* \psi \quad \text{and} \quad \psi(i) = \psi(j).$$

6 Chart quotients

6.1 Definition and properties of quotients

Definition 6.1. *The quotient of a chart M, denoted $[\![M]\!]$, is the chart obtained from M by identifying all bisimilar states.*

Proposition 6.2. *For every chart M and expression e, we have:*

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

In the remainder of this section, we show that the quotient of a Milner chart is reducible.

Proposition 6.3. *For every expression e, $[\![M(e)]\!]$ is reducible.*

6.2 Traps

Definition 6.4. *A* trap *in a chart is a set of states closed under transitions: if a state is in the trap, then all its successors are also in the trap.*

The frontier *of a trap is the set of states outside the trap which are either: i) final or ii) have an outgoing transition to a state in the trap.*

A summary *of a trap is a function whose domain is the frontier of the trap.*

Definition 6.5. *Let C be a chart, T a trap of C, and φ a summary of T.*

We denote by $C[T, \varphi]$ the chart obtained as follows. First, all states belonging to the trap T are removed. Then, for each state q on the frontier, we create a fresh copy of q, we declare this copy final, we make q non-final if it was, and we add a transition from q to its copy labelled by $\varphi(q)$. All other states and transitions remain unchanged.

A trap abstraction *of C is a chart of the form $C[T, \varphi]$ for some trap T of C and some summary φ of T.*

Example 6.6. *Example of a trap abstraction.*

Proposition 6.7. *A trap abstraction of a Milner chart is a Milner chart.*

7 Proof of completeness

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

Theorem 7.1. *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 6.3, there is an expression g such that:

$$Q \models g.$$

Using Proposition 6.2, we have:

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

By Proposition **??**, we also have:

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

By Corollary 4.8, we have then:

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

Hence $e \equiv f$. □

8 Erasing transitions from Milner charts

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

Definition 8.1. *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 8.2. *A* letter erasing *of an expression e is an expression obtained from e by replacing some letter occurrences by 0.*

Proposition 8.3. *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:*

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

References

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

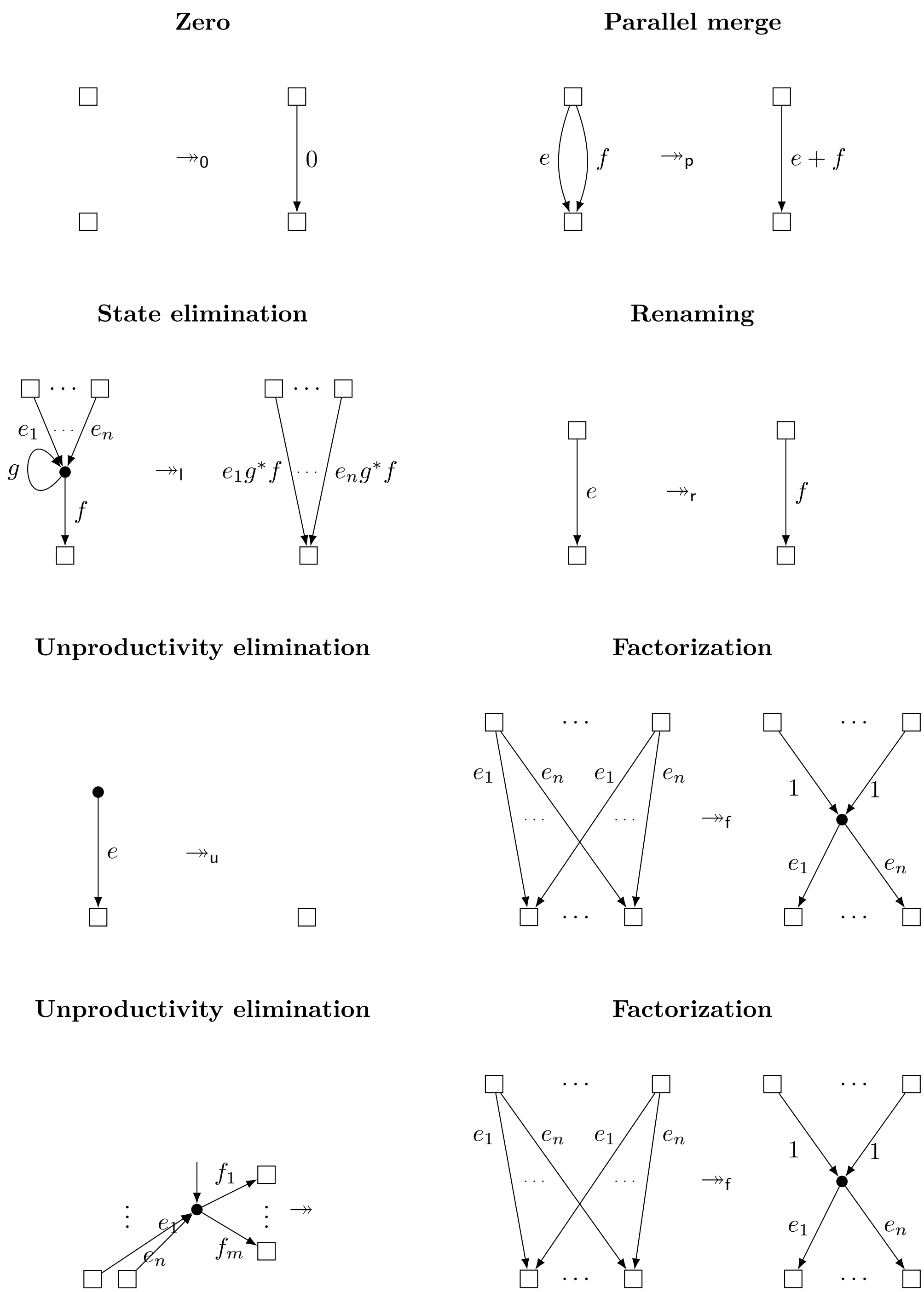


Figure 1: Rewriting rules for networks