



HAL
open science

Logic and Rational Languages of Scattered and Countable Series-Parallel Posets

Amazigh Amrane, Nicolas Bedon

► **To cite this version:**

Amazigh Amrane, Nicolas Bedon. Logic and Rational Languages of Scattered and Countable Series-Parallel Posets. Proceedings of LATA 2019, pp.275-287, 2019, 10.1007/978-3-030-13435-8_20. hal-02090421

HAL Id: hal-02090421

<https://hal.science/hal-02090421>

Submitted on 21 Jul 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Logic and Rational Languages of Scattered and Countable Series-Parallel Posets

Amazigh Amrane, Nicolas Bedon*

LITIS (EA 4108), Université de Rouen, Rouen, France

Abstract

Let A be an alphabet and $SP^\circ(A)$ denote the class of all countable N -free partially ordered sets labeled by A , in which chains are scattered linear orderings and antichains are finite. We characterize the rational languages of $SP^\circ(A)$ by means of logic. We define an extension of monadic second-order logic by Presburger arithmetic, named P-MSO, such that a language L of $SP^\circ(A)$ is rational if and only if L is the language of a sentence of P-MSO, with effective constructions from one formalism to the other. As a corollary, the P-MSO theory of $SP^\circ(A)$ is decidable.

Keywords: automata and logic, transfinite N -free posets, series-parallel posets, series-parallel rational languages, branching automata, monadic second-order logic, Presburger arithmetic

1. Introduction

Since they were independently established by Büchi [10], Elgot [16] and Trakhtenbrot [33], links between automata theory and formal logic have been widely developed. The fundamental result is the effective equivalence between Kleene automata [20] and sentences of Monadic Second-Order logic (MSO) for the description of languages of finite words. It provides, as examples, tools for languages classification, or decision algorithms on formal logic. This fundamental result have been later generalized to less restricted structures than finite words, relying on adapted notions of automata. Among them, let us cite words indexed by all the natural integers [11], ordinals [12], trees [29, 31, 14], linear orderings [5], etc. Among the first consequences of such generalizations are decision algorithms for second-order theories of countable ordinals [12], of two successors functions [29] and many others. Automata on a particular class C of structures were often used to obtain decision procedures for a theory of C . However, it is to notice that automata on C can also be used to obtain decision procedures for a theory of a seemingly unrelated class. As an example, the decidability of the second-order theory of countable linear orderings is reducible to the solvability of the decision problem of the second-order theory of two successors functions [29]. Apart from their interest in formal logic, generalization of Kleene automata are also used as models for processes. As an example, automata over ω -words provide a basement for the model-checking theory of sequential processes. Automata over finite N -free (or equivalently, series-parallel [34]) partially ordered sets (posets for short) can be used as models for concurrent programs, where concurrency relies on fork-join rules. Recall that a set partially ordered by $<$ is *N-free* if it has no subset $X = \{a, b, c, d\}$ such that $< \cap X \times X = \{(a, b), (c, b), (c, d)\}$.

In this paper we focus on the class $SP^\circ(A)$ of all countable N -free posets, labeled over an alphabet A , and whose chains are scattered linear orderings and antichains are finite. Recall that a linear ordering is *scattered* if it has no dense sub-ordering. A notion of automata, named *branching automata*, adapted to $SP^\circ(A)$ has been introduced in [6] as well as equivalent rational expressions. These automata and rational expressions are a generalisation of those on finite N -free posets of Lodaya and Weil [23, 24, 25, 26] and

*Corresponding author

those of Bruyère and Carton [9] on linear orderings. The logic P-MSO effectively equivalent to branching automata on finite N-free posets was introduced in [3]. Roughly speaking, P-MSO is a mix of MSO and Presburger arithmetic [28]. In [5] it was proved that the class of languages of countable linear orderings recognized by Bruyère and Carton automata is strictly included into MSO-definable languages, and that MSO and automata are effectively equivalent when linear orderings are restricted to be scattered. In [21, 22] Kuske proposed an extension of branching automata of Lodaya and Weil over finite N-free posets to N-free posets with finite antichains and ω -chains, together with a connection with MSO in the particular case of languages of N-free posets with bounded-size antichains. This is extended in [2] to N-free posets with bounded-size antichains and scattered and countable chains.

Our main result is that a language of $SP^\circ(A)$ is rational if and only if it is P-MSO definable (Theorem 36) with effective constructions from one formalism to the other. The decidability of the P-MSO theory of $SP^\circ(A)$ follows as a corollary. P-MSO is defined in Section 4.

It is known from [4] that the class of rational languages of $SP^\circ(A)$ is closed under boolean operations. Relying on this, the proof of the implication from right to left of Theorem 36 involves only well-known techniques: it is not developed in this paper. It is the purpose of the short Section 8. For the implication from left to right we had to develop new techniques. **The usual technique used in order to build directly a MSO-formula from an automaton in the case of finite or ω -words encodes by means of MSO the existence of a successful path labeled by some word. The formula attaches a transition to each letter of the word, consistently with the definition of the automaton. This can not be easily extended to the case of automata on linear orderings, since the MSO formula have to encode the use of limit transitions without the assumption that every factor has a least upper or a greater lower bound. This difficulty was avoided in [5] by constructing a formula from a rational expression instead of an automaton. The same difficulty occurs in the case of $SP^\circ(A)$, and we avoid it in this paper in the same way, but other problems arise.** This paper is essentially devoted to the construction of a sentence $\varphi_e(X)$ of P-MSO from a rational expression e such that P satisfies $\varphi_e(P)$ if and only if $P \in L(e)$. It is developed using the following scheme. The first step is to transform the rational expression into an equivalent form, named *>1-expression*. In *>1-expression*, sequential products guarantee the sequential composition of at least two non-empty posets. Thus, the sequential composition of two languages necessarily contains only non-trivial sequential posets. The *>1-expressions* are introduced together with rational expressions in Section 3. The next step consists in computing a graph by induction on a *>1-expression* (Section 5). The nodes of the graph are labeled by letters, sequential operations involved in *>1-expressions* and Presburger formulæ. The operations labeling the nodes are all P-MSO expressible; in particular the sequential operations can be expressed with MSO using techniques from [5]. During the induction, we enforce properties on the resulting graph in order to calculate the Presburger formulæ that will appear in the final P-MSO formula. The next step is the transformation of the graph into a P-MSO formula. To each node n we associate a P-MSO formula ϕ_n , and the idea is to make ϕ_n dependent of ϕ_m if there is an edge $n \rightarrow m$. Unfortunately the graph is not always acyclic. Cycles can be broken by avoiding particular edges, named *special*, that are identified in each inductive step of the construction of the graph. In order to avoid circular dependencies between the ϕ_n s we develop in Section 6 a technique named *s-coloring*, that permits to identify particular factors of posets by means of P-MSO. Using *s-coloring*, the effective transformation of the graph into a P-MSO formula is given in Section 7.

This paper is a long version, with proofs, of [1].

2. Notation, linear orderings and posets

We let $|E|$ denote the cardinality of a set E , 2^E its power-set, $[n]$ the set $\{1, \dots, n\}$ (for any non-negative integer $n \in \mathbb{N}$), and $\pi_i(c)$ the i^{th} component of a tuple c . We let 1_k denote the tuple of integers with 1 at the k^{th} position and 0 at all other positions. Every time we use this notation the size n of the tuple is clear from the context, and of course $k \in [n]$. Arithmetic operations on tuples of integers are componentwise. We also let $|s|$ denote the length k of a finite sequence s of k elements.

2.1. Unlabeled posets and linear orderings

Recall that an ordering $<$ over a set E is an irreflexive, transitive and asymmetric binary relation over the elements of E . It is *total*, or *linear*, when either $x < y$ or $y < x$ for all distinct $x, y \in E$. It is *dense* when for all x, y such that $x < y$ there is some z such that $x < z < y$. It is *scattered* if it has no dense non-trivial sub-ordering.

A partially ordered set, *poset* for short, $(P, <)$ consists of a set P and an ordering relation $<$ over the elements of P . For simplicity we often denote $(P, <)$ by P . We let $0 = (\emptyset, \emptyset)$ denote the empty poset. A *chain* is a totally ordered set. An *antichain* A is a set whose elements are pairwise incomparable: for all $x, y \in A$, neither $x < y$ nor $y < x$. An *interval* I of $(P, <)$ is a subset $I \subseteq P$ such that for all $i_1, i_2 \in I$ and $p \in P$, if $i_1 < p < i_2$ then $p \in I$.

A poset $(P, <)$ is *N-free* if there is no $X = \{x_1, x_2, x_3, x_4\} \subseteq P$ such that $< \cap X^2 = \{(x_1, x_2), (x_3, x_2), (x_3, x_4)\}$. N-free posets play a particular role in computer science, since they are related to scheduling of processes,

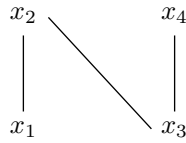


Figure 1: The poset N . Exceptionally, in this Figure, the (increasing) ordering is bottom-up

and model for example series and parallel electronic circuits, and concurrent processes relying on fork/join primitives. In this paper we particularly focus on N-free posets with finite antichains only.

We need the following operations on posets. The reversal $-(P, <)$ is $(P, <')$ defined by $x <' y$ if and only if $y < x$. Let $(P, <_P)$ and $(Q, <_Q)$ be two disjoint posets. The *union* (or *parallel composition*) $P \cup Q$ of $(P, <_P)$ and $(Q, <_Q)$ is the poset $(P \cup Q, <_P \cup <_Q)$. The *sum* (or *sequential composition*) $P + Q$ of P and Q is the poset $(P \cup Q, <_P \cup <_Q \cup P \times Q)$.

Example 1. Figure 2 represents two posets $P_1 = (\{x_1, x_2, x_3, x_4\}, x_1 <_1 x_2, x_3, x_4)$ and $P_2 = (\{y_1, y_2, y_3, y_4\}, y_1, y_2 <_2 y_3, y_4)$, the parallel composition $P_1 \cup P_2$ and the sequential composition $P_1 + P_2$.

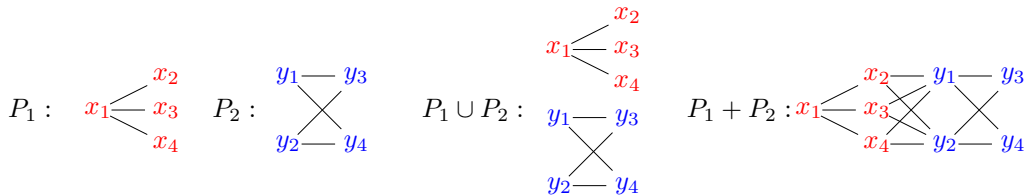


Figure 2: The Hasse diagrams of two N-free posets P_1 and P_2 , and their parallel and sequential compositions. In this Figure and all the others below, the ordering is from left to right

Before continuing let us focus on the particular cases of $(P, <)$ where the ordering $<$ is total. Linear orderings have a particular place in set theory. We refer the reader eg. to Rosenstein's book [30], entirely devoted to the subject. As it is conventional in the literature we let ω denote a representative of the class of linear orderings isomorphic to the linear ordering $(\mathbb{N}, <_{\mathbb{N}})$ of the natural integers. Ordinals are a particular case of scattered linear orderings. We let \mathcal{O} and \mathcal{S} denote respectively the class of countable ordinals and the class of countable scattered linear orderings (up to isomorphism). A *cut* (K, L) of a linear ordering $(J, <)$ consists of a pair of two disjoint intervals K and L of J such that $K \cup L = J$ and $k <_J l$ for all $(k, l) \in K \times L$. The set \hat{J} of all cuts of J is naturally equipped with the linear ordering $(K_1, L_1) <_{\hat{J}} (K_2, L_2)$ if and only if $K_1 \subsetneq K_2$. By extension, we equip the set $J \cup \hat{J}$ with the linear ordering $<$ containing $<_J$ and $<_{\hat{J}}$, and such

that $j < (K, L)$ (resp. $(K, L) < j$) whenever $j \in K$ (resp. $j \in L$), for any $j \in J$ and $(K, L) \in \hat{J}$. We let \hat{J}^* denote $\hat{J} \setminus \{(\emptyset, J), (J, \emptyset)\}$.

The sum of two posets can be generalized to any linearly ordered sequence $(P_j, <_j)_{j \in J}$ of pairwise disjoint posets by

$$\sum_{j \in J} P_j = \left(\bigcup_{j \in J} P_j, \left(\bigcup_{j \in J} <_j \right) \cup \left(\bigcup_{j, j' \in J, j < j'} P_j \times P_{j'} \right) \right)$$

In this paper we consider posets of the following class:

Definition 2. *The class SP^\diamond of series-parallel scattered and countable posets is the smallest class of posets containing 0, the singleton and being closed under disjoint finite parallel composition and disjoint sum indexed by countable scattered linear orderings. In this definition, posets are considered up to isomorphism.*

The following is an extension of a well-known result [34] on finite N-free posets.

Theorem 3 ([6]). *Let NF^\diamond be the class of countable N-free posets with scattered chains only and without infinite antichains (up to isomorphism). Then $NF^\diamond = SP^\diamond$.*

Example 4. *For any $i \in \mathbb{N}$, let A_i be an antichain of cardinality $|A_i| = i$. Set $P = \sum_{i \in \omega} A_i$. Then $P \in SP^\diamond$. Observe that there is no $n \in \mathbb{N}$ such that $|A| < n$ for all antichain A of P .*

We let $SP^{\diamond+}$ denote $SP^\diamond \setminus \{0\}$.

F. Hausdorff proposed in [19] an inductive definition of scattered linear orderings. In fact, each countable and scattered linear ordering is obtained using sums indexed by finite linear orderings, ω and $-\omega$. This has been adapted in [6] to SP^\diamond .

We denote by $\mathcal{C}_{\cup,+}(E)$ the closure of a set E under finite disjoint union and finite disjoint sum. Recall that 0 denotes the empty poset. We also let 1 denote the singleton poset.

Definition 5. *The classes of countable and scattered posets (equivalent up to isomorphism) V_α and W_α are defined inductively as follows:*

$$\begin{aligned} V_0 &= \{0, 1\} \\ W_\alpha &= \mathcal{C}_{\cup,+}(V_\alpha) \\ V_\alpha &= \left\{ \sum_{i \in J} P_i : J \in \{\omega, -\omega\} \text{ and } \forall i \in J, P_i \in \bigcup_{\beta < \alpha} W_\beta \right\} \cup \bigcup_{\beta < \alpha} W_\beta \end{aligned}$$

and the class S_{sp} of countable and scattered posets by $S_{sp} = \bigcup_{\alpha \in \mathcal{O}} W_\alpha$.

The following theorem extends a result of Hausdorff on linear orderings [19].

Theorem 6 ([6]). $S_{sp} = SP^\diamond$.

A unique ordinal can be associated to every poset of SP^\diamond , which can be used to prove properties inductively:

Definition 7. *The rank $r(P)$ of $P \in S_{sp}$ is the smallest ordinal α such that $P \in W_\alpha$.*

The notion of a rank for posets of SP^\diamond can be refined as follows. For every $\alpha \in \mathcal{O}$, the class W_α can be decomposed as the closure of V_α by finite disjoint union and finite disjoint sum.

Theorem 8 ([6]). *For all $\alpha \in \mathcal{O}$, $i \in \mathbb{N}$, let*

$$\begin{aligned} X_{\alpha,0} &= V_\alpha \\ Y_{\alpha,i} &= \left\{ P : \exists n \in \mathbb{N} P = \sum_{j \leq n} P_j \text{ such that } P_j \in X_{\alpha,i} \text{ for all } j \leq n \right\} \\ X_{\alpha,i+1} &= \left\{ P : \exists n \in \mathbb{N} P = \bigcup_{j \leq n} P_j \text{ such that } P_j \in Y_{\alpha,i} \text{ for all } j \leq n \right\} \end{aligned}$$

Then $W_\alpha = \bigcup_{i \in \mathbb{N}} X_{\alpha,i}$.

Define a well-ordering on $\mathcal{O} \times \mathbb{N}$ by $(\beta, j) < (\alpha, i)$ if and only if $\beta < \alpha$ or $\beta = \alpha$ and $j < i$. As a consequence of Theorems 6 and 8, for any $P \in SP^\circ$ there exists a pair $(\alpha, i) \in \mathcal{O} \times \mathbb{N}$ as small as possible such that $P \in X_{\alpha, i}$.

Definition 9. The X-rank $r_X(P)$ of $P \in SP^\circ$ is the smallest pair $(\alpha, i) \in \mathcal{O} \times \mathbb{N}$ such that $P \in X_{\alpha, i}$.

We finish this section on unlabeled posets by some definitions that generalize to posets the usual notions of factors of words.

Definition 10. An interval I of $P \in SP^\circ$ is good if it is non empty and, for all $p \in P$, if there are $x, y \in I$ such that $p < x$ or $x < p$ and neither $p < y$ nor $y < p$, then $p \in I$.

As emphasized by the following proposition, the notion of a good interval is deeply related with the usual notion of a factor. For that reason, in the sequel good intervals are called *factors*. Strictness of factors is relative to inclusion.

Proposition 11. Let $P \in SP^\circ$ and $I \subseteq P$. Then I is a good interval of P if and only if there exist a non-empty $J \in \mathcal{S}$ and a sequence of non-empty posets $(P_j)_{j \in J}$ such that $P = \sum_{j \in J} P_j$, or J is finite, $P = \bigcup_{j \in J} P_j$, and there exists $j \in J$ such that either $I = P_j$ or $I \subsetneq P_j$ and I is a good interval of P_j .

Proof. Observe that when I is a good interval of P and X is such that $I \subseteq X \subseteq P$, then I is also a good interval of X . As a consequence, the implication from left to right holds. For the converse, it suffices to note that when I is a good interval of I' which is itself a good interval of P , then I is a good interval of P . \square

Definition 12. Let P be a poset and $J \in \mathcal{S}$. A J -sequential-factorization of P , also called J -factorisation or sequential factorization for short, is a sequence $(P_j)_{j \in J}$ of posets such that $P = \sum_{j \in J} P_j$.

A poset P is sequential if it admits a J -factorization where J contains at least two elements $j \neq j'$ with $P_j, P_{j'} \neq 0$, or P is a singleton. It is parallel when $P = P_1 \parallel P_2$ for some $P_1, P_2 \neq 0$. A sequential factorization is irreducible when all the P_j are either singletons or parallel posets. The notions of parallel factorization and irreducible parallel factorization are defined similarly.

We let Seq denote the class of all sequential posets of $SP^{\circ+}$. Note that every $P \in SP^{\circ+}$ is either sequential or parallel, but not both. The empty poset is the only poset of SP° which is neither sequential nor parallel. The poset N of Figure 1 is neither sequential nor parallel, but $N \notin SP^\circ$.

2.2. Labeled posets

An *alphabet* A is a non-empty finite set whose elements are called *letters*. Recall also that a *language* of a set S is a subset of S .

A poset $(P, <, l)$ labeled by A , also denoted by P for short, consists of a poset $(P, <)$ and a *labeling* total map $l: P \rightarrow A$. Considered up to an isomorphism, labeled posets, also named *pomsets* in the literature, are a generalization of the usual notion of a word, since a word can be seen as a finite linear ordering labeled by A . The finite case was first investigated in [35, 18] from a systematic point of view. In order to be consistent with the usual notation on words, we let ϵ denote the unique empty labeled poset, and a the singleton poset labeled by a . The class of posets of SP° labeled by A (or over A) is denoted by $SP^\circ(A)$, and $SP^{\circ+}(A) = SP^\circ(A) \setminus \{\epsilon\}$. We also denote by A° the restriction of $SP^\circ(A)$ to posets with antichains of cardinality at most 1. Again, this notation is consistent with the words indexed by countable and scattered linear orderings, see eg. [9]. In order to match the words case, we also adapt the notation and nomenclature previously introduced for posets for the labeled case. The sequential and parallel compositions $+$ and \cup of posets are named *products*, and respectively denoted by \cdot and \parallel , when labeled posets are considered:

Definition 13. Let $(P, <_P, l_P)$ and $(Q, <_Q, l_Q)$ be two disjoint posets labeled by an alphabet A . The *sequential product*, or *concatenation*, $(P, <_P, l_P) \cdot (Q, <_Q, l_Q)$, or $(P, <_P, l_P)(Q, <_Q, l_Q)$ for short, is the poset $(P, <_P) + (Q, <_Q)$ labeled by $l_P \cup l_Q$. Similarly, the *parallel product* $(P, <_P, l_P) \parallel (Q, <_Q, l_Q)$ is the poset $(P, <_P) \cup (Q, <_Q)$ labeled by $l_P \cup l_Q$. The *sequential product* of a linearly ordered sequence of labeled posets is denoted by \prod .

Let A and B be two alphabets, $P \in SP^\circ(A)$, $L \subseteq SP^\circ(B)$ and $\xi \in A$. The language of $SP^{\circ+}(A \setminus \{\xi\} \cup B)$ consisting of the labeled poset P in which each element labeled by the letter ξ is non-uniformly replaced by a labeled poset of L is denoted by $L \circ_\xi P$. By *non-uniformly* we mean that the elements labeled by ξ may be replaced by different elements of L . This substitution $L \circ_\xi$ is the homomorphism from $(SP^\circ(A), \parallel, \prod)$ into the power-set algebra $(2^{SP^\circ(A \cup B)}, \parallel, \prod)$ with $\xi \mapsto L$ and $a \mapsto a$ for all $a \in A \setminus \{\xi\}$. In other words:

Definition 14. Let A and B be two alphabets, $P \in SP^\circ(A)$, $L \subseteq SP^\circ(B)$ and $\xi \in A$. Let S_P , $<_P$ and l_P denote respectively the set of elements, the ordering relation, and the labeling map of P . Then $L \circ_\xi P$ consists of all $R \in SP^\circ(A \setminus \{\xi\} \cup B)$ such that there exists $\nu_R: l_P^{-1}(\xi) \rightarrow L$ and

$$\begin{aligned} S_R &= (S_P \setminus l_P^{-1}(\xi)) \cup_{x \in l_P^{-1}(\xi)} S_{\nu_R(x)} \\ <_R = <_P|_{S_R} \cup_{x \in l_P^{-1}(\xi)} <_{\nu_R(x)} \cup_{(x,y) \in <_P \cap P \times l_P^{-1}(\xi)} \{x\} \times S_{\nu_R(y)} \cup_{(y,x) \in <_P \cap l_P^{-1}(\xi) \times P} S_{\nu_R(y)} \times \{x\} \\ l_R &= l_P|_{S_R} \cup_{x \in l_P^{-1}(\xi)} l_{\nu_R(x)} \end{aligned}$$

where $f|_X$ is the restriction of the map or relation f to X .

Example 15. Let $B = \{a, b\}$, $A = B \cup \{\xi\}$, $P = b \parallel (\xi \cdot \xi) \in SP^\circ(A)$ and $L = \{a \parallel b, b \cdot a\} \subseteq SP^\circ(B)$. Then $L \circ_\xi P = \{b \parallel ((a \parallel b) \cdot (a \parallel b)), b \parallel ((b \cdot a) \cdot (b \cdot a)), b \parallel ((a \parallel b) \cdot (b \cdot a)), b \parallel ((b \cdot a) \cdot (a \parallel b))\}$.

Sequential and parallel products are extended from labeled posets to languages of labeled posets in the usual way: when L and L' are languages of labeled posets and op is either the sequential or the parallel product, then $L op L' = \{P op P' : P \in L, P' \in L'\}$.

3. Rational languages

In this section we define languages by means of expressions. Let op be operations over a class of languages of structures labeled by an alphabet A . Recall that an *expression* e is a term of the free algebra over $\{\emptyset\} \cup A$ using the operations of op as functions. The language $L(e)$ of e is defined inductively using the definitions of the operations of op . Rational expressions describe rational languages. In a rational expression e , the union is usually denoted by $+$ instead of \cup .

In the first sub-section, we recall the definition of the class of rational languages of $SP^\circ(A)$. By extension of a well-known result of Kleene [20] on languages of finite words, it is known from [6] that a language $L \subseteq SP^\circ(A)$ is rational if and only if L is the language of some automaton. In this paper, we need a slightly modified definition of rational languages of $SP^\circ(A)$. In Sub-section 3.2, we define the class of >1 -rational languages. A language $L \subseteq SP^\circ(A)$ is rational if and only if it is >1 -rational. Finally, in Sub-section 3.3, we recall the link between rational and semi-linear languages of finite commutative words.

3.1. Rational languages of $SP^\diamond(A)$

Let A be an alphabet and $\xi \in A$. Define the following (disjoint) operations on the languages L, L' of $SP^\diamond(A)$:

$$\begin{aligned}
L \circ_\xi L' &= \bigcup_{P \in L'} L \circ_\xi P & L^* &= \left\{ \prod_{j \in [n]} P_j : n \in \mathbb{N}, P_j \in L \right\} \\
L^{*\xi} &= \bigcup_{i \in \mathbb{N}} L^{i\xi} \text{ with } L^{0\xi} = \{\xi\} \text{ and } L^{(i+1)\xi} = \left(\bigcup_{j \leq i} L^{j\xi} \right) \circ_\xi L \\
L^\omega &= \left\{ \prod_{j \in \omega} P_j : P_j \in L \right\} & L^{-\omega} &= \left\{ \prod_{j \in -\omega} P_j : P_j \in L \right\} \\
L^\natural &= \left\{ \prod_{j \in \alpha} P_j : \alpha \in \mathcal{O}, P_j \in L \right\} & L^{-\natural} &= \left\{ \prod_{j \in -\alpha} P_j : \alpha \in \mathcal{O}, P_j \in L \right\} \\
L \diamond L' &= \left\{ \prod_{j \in J \cup \hat{J}^*} P_j : J \in \mathcal{S} \setminus \{0\} \text{ and } P_j \in L \text{ if } j \in J \text{ and } P_j \in L' \text{ if } j \in \hat{J}^* \right\}
\end{aligned}$$

Definition 16. Let A be an alphabet and $\xi \in A$. The class of rational languages [6] of $SP^\diamond(A)$ is the smallest class containing $\emptyset, \{\epsilon\}, \{a\}$ for all $a \in A$, and being closed under the operations of $op = \{\|, \circ_\xi, *^\xi, \cup, \cdot, *, \diamond, \omega, -\omega, \natural, -\natural\}$ under the following conditions:

- $\epsilon \notin L$ in $L \circ_\xi L'$ and $L^{*\xi}$;
- in $L^{*\xi}$, each element labeled by ξ in a poset of L must be incomparable with another element.

The last condition excludes from the rational languages those of the form $(a\xi b)^{*\xi} = \{a^n \xi b^n : n \in \mathbb{N}\}$, for example, which is not rational in the usual language theory of finite words.

The definition above generalizes the notions of rational languages of several classes of structures. Actually:

- with $op = \{\cup, \cdot, *\}$ we get the rational languages of finite words of Kleene [20];
- with $op = \{\cup, \cdot, *, \|, \circ_\xi, *^\xi\}$ we get the rational languages of finite N-free posets of Lodaya and Weil [23, 24, 25, 26];
- with $op = \{\cup, \cdot, *, \diamond, \omega, -\omega, \natural, -\natural\}$ we get the rational languages of scattered and countable words of Bruyère and Carton [9].

For convenience we use the shortcut L^\diamond for $L \diamond \epsilon + \epsilon$.

Example 17. Let $A = \{a\}$ and $L = a \circ_\xi (a(\xi \| \xi))^{*\xi}$. Then L is the smallest language containing a and such that if $x, y \in L$ then $a(x \| y) \in L$. Thus $L = \{a, a(a \| a), a(a(a \| a) \| a), a(a(a \| a) \| a(a \| a)), \dots\}$. Furthermore, L is a rational language of $SP^\diamond(A)$, since $L_1 = a(\xi \| \xi)$ and $L_2 = a$ are rational, $\epsilon \notin L_1, L_2$ and each ξ is in parallel with the other ξ in L_1 .

The following result is fundamental in the sequel:

Theorem 18 ([4]). Let A be an alphabet. The class of rational languages of $SP^\diamond(A)$ is effectively closed under boolean operations.

3.2. >1 -rational languages of $SP^\diamond(A)$

In the remainder of the paper, we need sequentially operations to compose at least two non-empty posets. This is not the case for operations of the previous section, since for example $\{\epsilon\} \cdot \{a\} = \{a\}$. When

$L, L' \subseteq SP^\circ(A)$ define

$$\begin{aligned}
L^{\cdot > 1} L' &= (L \setminus \{\epsilon\}) \cdot (L' \setminus \{\epsilon\}) & L^{* > 1} &= \left\{ \prod_{i \in [n]} P_i : n > 1, P_i \in L \setminus \{\epsilon\} \right\} \\
L^{\omega > 1} &= \left\{ \prod_{i \in \omega} P_i : P_i \in L \text{ for all } i \in \omega \text{ and } P_i, P_j \neq \epsilon \text{ for some } i, j \text{ with } i \neq j \right\} \\
L^{\diamond > 1} L' &= \left\{ \prod_{j \in J \cup \hat{J}^*} P_j : J \in \mathcal{S} \setminus \{0\}, P_j \in L \text{ if } j \in J, P_j \in L' \text{ if } j \in \hat{J}^* \right. \\
&\quad \left. \text{and } P_i, P_j \neq \epsilon \text{ for some } i, j \in J \cup \hat{J}^*, i \neq j \right\}
\end{aligned}$$

Define similarly $L^{-\omega > 1}$, $L^{\natural > 1}$ and $L^{-\natural > 1}$.

Definition 19. Let A be an alphabet and $\xi \in A$. The class of >1 -rational languages of $SP^\circ(A)$ is the smallest class containing \emptyset , $\{\epsilon\}$, $\{a\}$ for all $a \in A$, and being closed under the operations of $op^{>1} = \{\|, \circ_\xi, *^\xi, \cup, \cdot > 1, * > 1, \diamond > 1, \omega > 1, -\omega > 1, \natural > 1, -\natural > 1\}$ under the following conditions:

- $\epsilon \notin L$ in $L \circ_\xi L'$ and $L^{*\xi}$;
- in $L^{*\xi}$, each element labeled by ξ in a poset of L must be incomparable with another element.

We let $L +_c L'$ denote $L + L'$ when condition c is verified, L otherwise. Then

$$\begin{aligned}
L \cdot L' &= L^{\cdot > 1} L' +_{\epsilon \in L} L' +_{\epsilon \in L'} L & L^* &= L^{* > 1} + L + \epsilon \\
L^\omega &= L^{\omega > 1} +_{\epsilon \in L} L^* & L \diamond L' &= L^{\diamond > 1} L' + L +_{\epsilon \in L} L'
\end{aligned}$$

Similar equalities hold for $-\omega$, \natural and $-\natural$. Every rational expression can be transformed into a >1 -expression. Considering the equalities above as rewriting rules this transformation is unique (up to associativity). As a consequence of Theorem 18, a language of $SP^\circ(A)$ is rational if and only if it is >1 -rational.

Example 20. Consider the rational expressions $e_1 = a \circ_\xi (a(\xi \| \xi))^{*\xi}$ and $e_2 = e_1 \diamond b^\diamond$. Then $e'_1 = a \circ_\xi (a^{\cdot > 1}(\xi \| \xi))^{*\xi}$ is the >1 -expression of e_1 . The >1 -expression of e_2 is $e'_1 \diamond^{>1} (b \diamond^{>1} \epsilon + b + \epsilon) + e'_1$.

3.3. Rational languages of finite commutative words

Recall that in a monoid (S, \cdot) , the class of rational languages is the smallest containing the empty set, $\{s\}$ for all $s \in S$ and closed under the operations of $\{\cup, \cdot, *\}$. When the monoid is commutative we usually denote its product by $\|$ instead of \cdot and its Kleene closure $^\circledast$ instead of $*$. When A is an alphabet, A^\circledast is the class of all finite antichains over A , or equivalently, the class of all finite commutative words over A .

Definition 21. Let $(S, \|)$ be a commutative monoid. A subset L of S is linear if it has the form $L = a_1 \| \cdots \| a_k \| \left(\bigcup_{i \in I} (a_{i,1} \| \cdots \| a_{i,k_i}) \right)^\circledast$ where I is a finite set, $k, k_i \in \mathbb{N}$, $a_i, a_{i,j} \in S$ for all $i \in I$ and $j \in [k_i]$. It is semi-linear if it is a finite union of linear sets.

Theorem 22 (see e.g. [15]). In a commutative monoid, a language is rational if and only if it is semi-linear.

Example 23. Consider the language L over the alphabet $A = \{a, b\}$ consisting of all commutative finite words with strictly more a s than b s. Then L is the language of the rational expression $a \| a^\circledast \| (a \| a^\circledast \| b)^\circledast$ over the commutative monoid A^\circledast . It is linear, hence semi-linear, since it also has the form $a \| (a \| b \cup a)^\circledast$.

The following lemma is a direct consequence of a Parikh's Theorem [27, Theorem 2]:

Lemma 24. Let A be an alphabet, $\xi \in A$, L and L' semi-linear languages of A^\circledast . Then $L \circ_\xi L'$ is a semi-linear language of A^\circledast , and if $\epsilon \notin L$ then so is $L^{*\xi}$.

In the remainder of the paper, when an ordering is required over an alphabet $A = \{a_1, \dots, a_n\}$ and is not specified explicitly, it is implicitly given by the indexes of the a_i s. Let $A = \{a_1, \dots, a_n\}$ and $B = \{b_1, \dots, b_m\}$ be two disjoint and totally ordered alphabets. For any $i \in [m]$ let us denote by $A \circ_{b_i} B$ the totally ordered alphabet $\{b_1, \dots, b_{i-1}, a_1, \dots, a_n, b_{i+1}, \dots, b_m\}$. When L and L' are rational sets of respectively A^\otimes and B^\otimes , A and B disjoint, then $L \circ_{b_i} L'$ is a rational set of $(A \circ_{b_i} B)^\otimes$. If $A = B$ then $L \circ_{b_i} L'$ is a rational set of A^\otimes .

Recall that every $u \in A^\otimes$ can be thought of as its Parikh's commutative image, that is a n -tuple $(|u|_{a_1}, \dots, |u|_{a_n}) \in \mathbb{N}^n$, where $|u|_{a_i}$ denotes the number of occurrences of letter a_i in u . Thus, A^\otimes is isomorphic to $(\mathbb{N}^n, +)$.

4. Logic

Presburger arithmetic and Monadic Second-Order logic (MSO) are two classical logics in computer science. In this section we briefly recall those two logics and introduce P-MSO, which extends MSO with Presburger arithmetic.

4.1. Presburger arithmetic

Recall that Presburger arithmetic is the first-order logic of $(\mathbb{N}, +)$. The Presburger set $L(\rho)$ of a Presburger formula $\rho(x_1, \dots, x_n)$ whose free variables are x_1, \dots, x_n consists of all interpretations of (x_1, \dots, x_n) which satisfy ρ . A language $L \subseteq \mathbb{N}^n$ is a *Presburger set* of \mathbb{N}^n if it is the Presburger set of some Presburger formula. We let \mathcal{P}_n denote the class of all Presburger formulæ with n free variables and we set $\mathcal{P} = \bigcup_{i \in \mathbb{N}} \mathcal{P}_i$.

Presburger logic provides tools to manipulate semi-linear sets of commutative monoids with formulæ.

Theorem 25 ([17]). *Let $A = \{a_1, \dots, a_n\}$ be a totally ordered alphabet. A language L of A^\otimes is semi-linear if and only if it is the Presburger set $L(\rho)$ of some Presburger formula $\rho(x_1, \dots, x_n)$, i.e. $(|u|_{a_1}, \dots, |u|_{a_n}) \in L(\rho)$ if and only if $u \in L$. Furthermore, the constructions from one formalism to the other are effective.*

Observe that the ordering of the free variables x_1, \dots, x_n of ρ is related to the ordering of A . By convention $\{()\}$ is the Presburger set of any closed tautology ρ . When ρ is used to define some language, we identify the empty tuple $()$ with ϵ .

Example 26. *The Presburger set of the formula $\rho(x, y) \equiv \exists k, k' \ k = k' + 1 \wedge x = y + k$ of \mathcal{P}_2 is $\{(x, y) : y < x\}$. Let A and L be as in Example 23. Then $\rho(x, y)$ defines L in A^\otimes .*

With the help of Theorem 25 and Lemma 24 we get the following definition.

Definition 27. *Let $\rho(x_1, \dots, x_k)$ and $\rho'(x'_1, \dots, x'_{k'})$ be Presburger formulæ and let $A = \{a_1, \dots, a_k\}$ and $B = \{b_1, \dots, b_{k'}\}$ be two totally ordered alphabets, either disjoint or equal. Consider the Presburger sets of ρ and ρ' as semi-linear languages L and L' of respectively A^\otimes and B^\otimes . When A and B are disjoint, for all $i \in [k']$, $L \circ_{b_i} L'$ is a semi-linear language of $(A \circ_{b_i} B)^\otimes$ and also the Presburger set of some formula that we denote by $\rho \circ_{x'_i} \rho'$. When $A = B$, for all $i \in [k']$, $L \circ_{b_i} L'$ is a semi-linear language of A^\otimes and the Presburger set of some formula that we denote by $\rho \bullet_{x'_i} \rho'$. If $(0, \dots, 0) \notin L(\rho)$, for all $j \in \mathbb{N}$ and $i \in [k]$, L^{*a_i} and L^{ja_i} are semi-linear languages of A^\otimes and also the Presburger sets of some formulæ that we denote respectively by ρ^{*x_i} and ρ^{jx_i} .*

Example 28. *Let $\rho(x_1, x_2) = x_1 + x_2 = 1$ and $\rho'(x'_1, x'_2, x'_3) = x'_1 = 1 \wedge x'_2 + x'_3 \leq 1$. Then $L(\rho) = \{(1, 0), (0, 1)\}$ and $L(\rho') = \{(1, 0, 1), (1, 1, 0), (1, 0, 0)\}$. Let $A = \{a_1, a_2\}$ and $B = \{b_1, b_2, b_3\}$ disjoint. Consider the Presburger sets of ρ and ρ' as semi-linear languages L and L' of respectively A^\otimes and B^\otimes . Then L , L' and $L \circ_{b_3} L'$ are respectively the languages of the rational expressions $a_1 + a_2$, $b_1 \parallel (b_2 + b_3 + \epsilon)$ and $b_1 \parallel (b_2 + a_1 + a_2 + \epsilon)$ of A^\otimes , B^\otimes , and $(A \circ_{b_3} B)^\otimes$. The Parikh commutative image $\{(1, 0, 0, 0), (1, 1, 0, 0), (1, 0, 1, 0), (1, 0, 0, 1)\}$ of $L \circ_{b_3} L'$ is also the Presburger set of $\rho \circ_{x'_3} \rho' \equiv x_1 = 1 \wedge x_2 + x_3 + x_4 \leq 1$.*

Lemma 29. *Let $\rho(x_1, \dots, x_n)$ be a Presburger formula and $i \in [n]$. For all $j \in \mathbb{N}$, $\rho^{(j+1)x_i} \equiv (\bigvee_{k \leq j} \rho^{kx_i}) \bullet_{x_i} \rho$. Furthermore, $L(\rho^{*x_i}) = \bigcup_{j \in \mathbb{N}} L(\rho^{jx_i})$.*

Proof. Consider the Presburger set of ρ as a semi-linear language over the totally ordered alphabet $\{a_1, \dots, a_n\}$. Let us start with the first part of the lemma. Following definitions,

$$\begin{aligned} L(\rho^{(j+1)x_i}) &= L(\rho)^{(j+1)a_i} = (\cup_{k \leq j} L(\rho)^{ka_i}) \circ_{a_i} L(\rho) \\ &= (\cup_{k \leq j} L(\rho^{kx_i}) \circ_{a_i} L(\rho)) = L(\vee_{k \leq j} \rho^{kx_i}) \circ_{a_i} L(\rho) = L((\vee_{k \leq j} \rho^{kx_i}) \bullet_{x_i} \rho) \end{aligned}$$

For the second part of the lemma, $L(\rho^{*x_i}) = L(\rho)^{*a_i} = \cup_{j \in \mathbb{N}} L(\rho)^{ja_i} = \cup_{j \in \mathbb{N}} L(\rho^{jx_i})$. \square

Example 30. Let $\rho(x_1, x_2) = (x_1 = 1 \wedge x_2 = 0) \vee (x_1 = 0 \wedge x_2 = 2)$. Let $A = \{a_1, a_2\}$ and consider the Presburger set of ρ as a language L of A^\otimes . Then $L = \{a_1, a_2 \parallel a_2\}$. Consider $L^{*a_2} = \bigcup_{i \in \mathbb{N}} L^{ia_2}$. By induction on i it can be checked that

$$\bigcup_{j \leq i} L^{ja_2} = \{u \in A^\otimes : 1 \leq |u| \leq 2^i \text{ and } |u|_{a_1} \leq 2^{i-1}\}$$

It follows that $L^{*a_2} = A^\otimes \setminus \{\epsilon\}$, and as a consequence $\rho^{*x_2}(x_1, x_2) \equiv x_1 + x_2 \geq 1$. Now for any $i \in \mathbb{N}$ let $\rho_i(x_1, x_2) \equiv 1 \leq x_1 + x_2 \leq 2^i \wedge x_1 \leq 2^{i-1}$. Then $L(\rho_i)$ is the Parikh commutative image of $\bigcup_{j \leq i} L^{ja_2}$. Thus $\rho^{(i+1)x_2} \equiv \rho_i \bullet_{x_2} \rho$ and $\rho^{0x_2}(x_1, x_2) \equiv x_1 = 0 \wedge x_2 = 1$.

The following lemmas are direct consequences of the definitions above.

Lemma 31. Let $\rho(x_1, \dots, x_k)$ and $\rho'(x'_1, \dots, x'_{k'})$ be two Presburger formulæ where $k \in \mathbb{N}$, $k' \in \mathbb{N} \setminus \{0\}$ and $L(\rho), L(\rho') \neq \emptyset$, and let $i \in [k']$. Then $(v''_1, \dots, v''_{k+k'-1}) \in L(\rho \circ_{x'_i} \rho')$ if and only if $(v''_1, \dots, v''_{i-1}, v'_i, v''_{i+k}, \dots, v''_{k+k'-1}) \in L(\rho')$, for some $v'_i \in \mathbb{N}$, and there exist $(v_{1,1}, \dots, v_{1,k}), \dots, (v_{v'_i,1}, \dots, v_{v'_i,k}) \in L(\rho)$ such that $(v''_i, \dots, v''_{i+k-1}) = (v_{1,1}, \dots, v_{1,k}) + \dots + (v_{v'_i,1}, \dots, v_{v'_i,k})$.

Lemma 32. Let $\rho(x_1, \dots, x_k)$ and $\rho'(x'_1, \dots, x'_k)$ be two Presburger formulæ where $k \in \mathbb{N} \setminus \{0\}$ and $L(\rho), L(\rho') \neq \emptyset$, and let $i \in [k]$. Then $(v''_1, \dots, v''_k) \in L(\rho \bullet_{x'_i} \rho')$ if and only if there exist $(v'_1, \dots, v'_k) \in L(\rho')$ and $(v_{j,1}, \dots, v_{j,k}) \in L(\rho)$ for all $j \in [v'_i]$ such that for all $r \in [k]$

$$v''_r = \begin{cases} \sum_{j \in [v'_i]} v_{j,r} & \text{when } r = i; \\ v'_r + \sum_{j \in [v'_i]} v_{j,r} & \text{otherwise.} \end{cases}$$

Lemma 33. Let $\rho(x_1, \dots, x_k)$ be a Presburger formula for some $k \in \mathbb{N} \setminus \{0\}$ and such that $(0, \dots, 0) \notin L(\rho)$. Let $i \in [k]$ and $t \in [k] \setminus \{i\}$. Then $1_t \in L(\rho)$ if and only if $1_t \in L(\rho^{*x_i})$.

4.2. P-MSO

We refer the reader to [32] for a survey on MSO. Presburger Monadic Second-Order logic, P-MSO for short, is an extension of Monadic Second-Order logic (MSO) with Presburger arithmetic.

Let A be an alphabet, $V_1 = \{x, y, \dots\}$, $V_2 = \{X, Y, \dots\}$ and $V_{\mathbb{N}}$ be three disjoint sets of variables, respectively named *first-order variables of P-MSO*, *second-order variables of P-MSO*, and *(first-order) variables of Presburger arithmetic*. Formulæ of P-MSO are inductively built according to the following grammar, where $a \in A$, $x, y \in V_1$, $X, Y, Z \in V_2$, $x_1, \dots, x_n \in V_{\mathbb{N}}$, $\psi, \psi_1, \dots, \psi_n$ are P-MSO formulæ, and ρ a Presburger formula.

$$\begin{aligned} \psi ::= & a(x) \mid x \in X \mid x < y \mid \psi_1 \vee \psi_2 \mid \psi_1 \wedge \psi_2 \mid \neg \psi \\ & \mid \exists x \psi \mid \exists X \psi \mid \forall x \psi \mid \forall X \psi \mid \mathcal{Q}(Z, \psi_1, \dots, \psi_n, \rho(x_1, \dots, x_n)) \end{aligned}$$

This grammar restricted to the 10 first items produces exactly the formulæ of MSO. The last item extends MSO to P-MSO.

Let us turn to the semantics. Formulæ of P-MSO are interpreted over posets of $SP^\diamond(A)$. Note that we allow posets to be empty. The variables of V_1 , V_2 , $V_{\mathbb{N}}$ are respectively interpreted over the elements of the

posets, sets of elements of the posets, and \mathbb{N} . For convenience we often do not make the distinction between a variable and its interpretation when the context is sufficiently explicit.

When x and y are interpreted over the elements of a poset $(P, <_P, l_P) \in SP^\diamond(A)$ and X is interpreted as a subset of P , then the atomic formula $x < y$ is interpreted as $x <_P y$, and $a(x)$ as $l_P(x) = a$. The semantics of $x \in X$ is self-explanatory, and those of all the other forms of formulæ of the grammar, except the last one, are as usual. So let us turn to the semantics of $\psi(Z) \equiv \mathcal{Q}(Z, \psi_1, \dots, \psi_n, \rho(x_1, \dots, x_n))$. Let $P \in SP^{\diamond+}(A)$ and $Z \subseteq P$. Then Z satisfies ψ if it is a non-empty factor of P , and there exist $(v_1, \dots, v_n) \in L(\rho)$ and sequential posets $Z_{1,1}, \dots, Z_{1,v_1}, \dots, Z_{n,1}, \dots, Z_{n,v_n} \in SP^{\diamond+}(A)$ such that $Z = \parallel_{i \in [n]} \parallel_{j \in [v_i]} Z_{i,j}$ and $Z_{i,j}$ satisfies ψ_i for all $i \in [n]$ and $j \in [v_i]$.

Example 34. Let $A = \{a_1, a_2\}$, and let $\psi_i \equiv \exists x a_i(x)$, $i \in [2]$. Let $\rho(x_1, x_2) \equiv \exists k_1, k_2 x_1 = 2k_1 \wedge x_2 = 2k_2 + 1 \in \mathcal{P}$. Let $P_1 = a_1(a_1 \parallel a_1)a_1$, $P_2 = a_2a_2$ and $P_3 = (a_1 \parallel a_2)a_1$. Then $P = P_1 \parallel P_2 \parallel P_3$ satisfies $\mathcal{Q}(P, \psi_1, \psi_2, \rho(x_1, x_2))$ since there is $(K_1, K_2) = (\{1, 3\}, \{2\})$ with $(|K_1|, |K_2|) \in L(\rho)$, and for all $i \in K_j$, $j \in [2]$, P_i satisfying ψ_j .

In this paper we use a lot of common shortcuts in formulæ, such as for example $\psi_1 \rightarrow \psi_2$ for $\neg\psi_1 \vee \psi_2$. As usual the logical equivalence of formulæ is denoted by \equiv .

A relation $R \subseteq P^m \times (2^P)^n$ is P-MSO (resp. MSO) *definable* in $P \in SP^\diamond(A)$ if there is a P-MSO (resp. MSO) formula $\phi_R(x_1, \dots, x_m, X_1, \dots, X_n)$ which is true if and only if $(x_1, \dots, x_m, X_1, \dots, X_n) \in R$.

Example 35. Let $P \in SP^\diamond(A)$, $X, Y, Z \subseteq P$. The following relations are MSO definable:

- $Y \subseteq Z, Y \cup Z = X, Y \cap Z = \emptyset$ and $Y, Z \neq \emptyset$. For example, $Y \subseteq Z \equiv \forall y y \in Y \rightarrow y \in Z$;
- $X < Y \equiv \forall x \forall y (x \in X \wedge y \in Y) \rightarrow x < y$
- “ (Y, Z) is a partition of X ”:
 $\text{Partition}(X, Y, Z) \equiv Y \cup Z = X \wedge Y \cap Z = \emptyset \wedge Y, Z \neq \emptyset$
- $|X| = n$, for any $n \in \mathbb{N}$:
 - $|X| = 0 \equiv \forall x x \notin X$
 - $|X| = n + 1 \equiv \exists Y, Z \text{ Partition}(X, Y, Z) \wedge |Y| = 1 \wedge |Z| = n$
- $X = Y + Z \equiv |Y| \neq 0 \wedge |Z| \neq 0 \wedge \text{Partition}(X, Y, Z) \wedge Y < Z$
- $F(R, P)$ “ R is a factor of P ”, as a direct consequence of Proposition 11;
- $F_s(R, P)$ “ R is a sequential factor of P ”, since
 $F_s(R, P) \equiv F(R, P) \wedge (|R| = 1 \vee \exists R_1, R_2 \text{ Partition}(R, R_1, R_2) \wedge R_1 < R_2)$

The main result of this paper is the following:

Theorem 36. Let A be an alphabet. A language L of $SP^\diamond(A)$ is rational if and only if it is P-MSO definable. Furthermore the constructions from one formalism to the other are effective.

5. D-graphs

A *D-graph* is a particular case of rooted, directed and finite graph whose edges are partially ordered and nodes are labeled. More formally:

Definition 37. A D-graph $D = (V, E_S, E_N, r, \text{out}, A, \gamma)$ is a rooted, directed, finite and labeled graph whose edges are partially ordered:

- V is a finite set whose elements are the nodes of D ;

- the edges of D consist of two disjoint sets $E_S \subseteq V \times V$ (special edges) and $E_N \subseteq V \times V$ (normal edges). An edge (n, n') is also denoted $n \rightarrow n'$. Its source is n and its destination is n' ;
- $r \in V$ is the root of D ;
- out is a total map that associates to every $n \in V$ a totally ordered and finite sequence $e_1 \dots e_k$ containing all edges of D of source n (note that an edge may occur more than once in the sequence). We say that n is edged by $out(n) = e_1 \dots e_k$. A leaf is a node n such that $out(n)$ has no normal edge;
- A is the alphabet of D ;
- $\gamma: V \rightarrow A \cup \mathcal{P} \cup \{ \cdot >^1, * >^1, \diamond >^1, \omega >^1, -\omega >^1, \natural >^1, -\natural >^1 \}$ is a total map labeling the nodes of D verifying:
 - if $\gamma(n) \in A$ then $|out(n)| = 0$;
 - if $\gamma(n) \in \mathcal{P}_k$ then $|out(n)| = k$;
 - if $\gamma(n) \in \{ * >^1, \omega >^1, -\omega >^1, \natural >^1, -\natural >^1 \}$ then $|out(n)| = 1$;
 - if $\gamma(n) \in \{ \cdot >^1, \diamond >^1 \}$ then $|out(n)| = 2$.

The sets of special edges, normal edges, and the root of D are respectively denoted by $E_S(D)$, $E_N(D)$ and $r(D)$.

In a D-graph D , a sequence of nodes n_1, \dots, n_k is *consecutive* if there is an edge $n_i \rightarrow n_{i+1}$ for all $i \in [k-1]$. Similarly, a sequence of edges e_1, \dots, e_k is *consecutive* if the destination of e_i is the source of e_{i+1} , for all $i \in [k-1]$.

Let n and p be two nodes of D . In D , we say that p is an *ancestor* of n if D contains a sequence of consecutive nodes m_1, \dots, m_k for some $k \in \mathbb{N}$, such that p, m_1, \dots, m_k, n are consecutive too. When $k = 0$ we say that p is a *parent* of n . The definitions of the notions of *descendant* and *child* are symmetrical. When n is the destination of the i^{th} edge of $out(p)$ then n is the i^{th} *child* of p . Note that n may be simultaneously the i^{th} and the j^{th} *child* of p with $i \neq j$ (see Example 52). It may be also the *child* of more than one node.

We often see sequences as words. For example, we let $e \circ_{e'} s$ denote the sequence of edges obtained by replacing in the sequence of edges s each occurrence of the edge e' by the edge e . When i is the position of an element of s , then $s' \circ_i s$ is the sequence obtained by replacing the i^{th} element of s by s' . We also let $n' \circ_n^{\text{src}} s$ denote the sequence of edges obtained from s by replacing every occurrence of n in sources of edges by n' . When nothing is specified, replacing the source of a normal (resp. special) edge provides a normal (resp. special) edge.

In Section 5.1 a D-graph D_e is constructed from a >1 -expression e by induction on e . The inductive steps for substitutions, ie. $e \circ_\xi e'$ and $e^{*\xi}$ are key cases. Informally the main idea is as follows. For simple substitution $e \circ_\xi e'$, $D_{e \circ_\xi e'}$ is built from the union of D_e and $D_{e'}$ by replacing each node labeled by ξ in $D_{e'}$ by a copy of $r(D_e)$. The case of $e^{*\xi}$ is more tricky since the substitution is recursive and the base case must be ensured. The idea here is: first to guaranty that the root of D_e is labeled by a Presburger formula ρ with a child n labeled by ξ , and to substitute every node labeled by ξ , except n , by a copy of the root. Second, transform ρ consistently with the definition of $e^{*\xi}$. This requires n to be a child of the root only. In order to make the transformation of Presburger formulæ easily we need nodes labeled by Presburger formulæ to have at most one child labeled by ξ . Actually we need that D_e is in a normal form defined as follows, and has properties defined below.

Definition 38. Let A be an alphabet and $\xi \in A$. A D-graph $D = (V, E_S, E_N, r, out, A, \gamma)$ is ξ -normalized when the following conditions are verified:

1. if a node labeled by ξ is a *child* of the root then it cannot be *child* of another node;
2. for every $n \in V$ with $\gamma(n) \in \mathcal{P}$, if $out(n)$ is some $out(n) = s_1 \rightarrow d_1, \dots, s_k \rightarrow d_k$ there is at most one $i \in [k]$ such that $\gamma(d_i) = \xi$.

The ξ -normalization of D consists of the transformation given by Algorithm 1.

Algorithm 1 ξ -normalization of a D-graph

```
// First step
for each edge  $f: n \rightarrow x$  for some  $n \neq r$ ,  $x$  child of  $r$  and with  $\gamma(x) = \xi$  do
  Add a new node  $y$  labeled by  $\xi$  in  $V$  and  $n \rightarrow y$  in  $E_N(D)$ 
   $\text{out}(n) \leftarrow (n \rightarrow y) \circ_f \text{out}(n)$ 
// Second step
for all  $n \in V$  such that  $\gamma(n)$  is some  $\rho(x_1, \dots, x_k)$ 
  and  $\text{out}(n)$  is some  $e_1: n \rightarrow n_1, \dots, e_k: n \rightarrow n_k$  do
  if  $|\{i_1, \dots, i_m \in [k] : \gamma(n_{i_j}) = \xi\}| > 1$  then
    Add a new node  $n_0$  labeled by  $\xi$  in  $V$  and  $n \rightarrow n_0$  in  $E_N(D)$ 
     $\gamma(n) \leftarrow \exists x_{i_1}, \dots, x_{i_m} (x_0 = \sum_{j \in [m]} x_{i_j} \wedge \rho(x_1, \dots, x_k))$ 
     $\text{out}(n) \leftarrow (n \rightarrow n_0)(\epsilon \circ_{e_{i_1}} \dots \epsilon \circ_{e_{i_m}} \text{out}(n))$ 
    Remove all  $n_{i_j}$ ,  $j \in [m]$ , that are not the destination of some remaining edge
```

In Algorithm 1 by $\epsilon \circ_{e_j} \text{out}(n)$ (resp. $\text{out}(n) \leftarrow (n \rightarrow y) \circ_f \text{out}(n)$) we remove all the occurrences of the edge e_j (resp. f) from $\text{out}(n)$ and thus implicitly from the D-graph.

Definition 39. Let $D = (V, E_S, E_N, r, \text{out}, A, \gamma)$ be a D-graph. We say that D has Property PP if it has no edge $n \rightarrow m$ such that $\gamma(n), \gamma(m) \in \mathcal{P}$. It has Property SS if it has no special edge $n \rightarrow m$ such that m is labeled in \mathcal{P} . It has Property DAG if $(V, \emptyset, E_N, r, \text{out}, A, \gamma)$ is acyclic.

Algorithm 2 transforms a D-graph with no consecutive edges $n \rightarrow n'$, $n' \rightarrow n''$ such that $\gamma(n), \gamma(n'), \gamma(n'') \in \mathcal{P}$ into a D-graph with Property PP.

Algorithm 2 PP-suppression

```
for all node  $n$  such that  $\gamma(n)$  is some  $\rho(x_1, \dots, x_k)$  do
  while there is some  $(p, i)$  such that  $n$  is the  $i^{\text{th}}$  descendant of node  $p$ 
    and  $\gamma(p)$  is some  $\rho'(x'_1, \dots, x'_{k'})$  do
     $\gamma(p) \leftarrow \rho \circ_{x'_i} \rho'$ 
     $\text{out}(p) \leftarrow (p \circ_n^{\text{src}} \text{out}(n)) \circ_i \text{out}(p)$ 
  Remove  $n$  if it has no incoming edge
```

Obviously, ξ -normalization and PP-suppression preserve Properties PP and SS. The ξ -normalization also preserves Property DAG, but not PP-suppression.

5.1. From >1 -expressions to D-graphs

Let e be a >1 -expression. In this section we build a D-graph D_e by induction on e . During the construction, we assume the D-graphs obtained by the induction to be ξ -normalized. Thus, at each inductive step we construct a D-graph, and if it is not ξ -normalized we implicitly transform it with Algorithm 1. The construction ensures that the D-graph constructed at each inductive step has Properties PP, SS and DAG.

Let us start the construction of $D_e = (V, E_S, E_N, r, \text{out}, A, \gamma)$. Except when the contrary is specified, new edges added during the constructions of this section are normal.

- Case $e = \epsilon$ (resp. $e = a \in A$)
 D_e is just a node labeled by any closed Presburger tautology (resp. labeled by a), without edges.
- Case $e = e_1 \text{ op } e_2$ (resp. $e = e' \text{ op}$) with $\text{op} \in \{.>^1, \diamond>^1\}$ (resp. $\text{op} \in \{*>^1, \omega>^1, -\omega>^1, \natural>^1, -\natural>^1\}$)
Then D_e is built from the union of D_{e_1} and D_{e_2} (resp. from $D_{e'}$), with one more node n as a root, labeled by op , and edged by $n \rightarrow r(D_{e_1}), n \rightarrow r(D_{e_2})$ (resp. $n \rightarrow r(D_{e'})$).

Algorithm 3 Construction of $D_{e_1 \circ_\xi e_2}$

for all node n of D_{e_2} with $\gamma(n) = \xi$ **do**
 $\gamma(n) \leftarrow \gamma(r(D_{e_1}))$
 $\text{out}(n) \leftarrow n \circ_r^{\text{src}} \text{out}(r(D_{e_1}))$
 Apply PP-suppression
Remove $r(D_{e_1})$ from the D-graph

- Case $e = e_1 \circ_\xi e_2$
When $e_2 = \xi$ then D_e is identical to D_{e_1} . Otherwise, D_e is built from the union of D_{e_1} and D_{e_2} transformed by Algorithm 3. The root of the new D-graph is $r(D_{e_2})$.
- Case $e = e'^{*}\xi$
This is the only case where special edges are added. This is also the only case likely to add to the edging $\text{out}(n)$ of a node n an edge which is already in $\text{out}(n)$. The construction of $D_{e'^{*}\xi}$ from $D_{e'}$ relies on the same principle as the previous case: each node of $D_{e'}$ labeled by ξ should be replaced by a copy of the root. Again, this may be not so simple in some cases because of the properties we want to ensure on the resulting D-graph. The construction follows Algorithm 4, starting from $D_{e'}$. We proceed in two steps. The first step ensures in particular that the root is labeled by some Presburger formula ρ with a i^{th} **child** labeled by ξ for exactly one i , and that 1_i is in the Presburger set of ρ . In the second step we proceed with the replacement of nodes labeled by ξ by a copy of the root and we ensure Property PP. Note that after the first step, the D-graph fulfills Property PP. In particular,

Algorithm 4 Construction of $D_{e'^{*}\xi}$

// First step: root transformation
if $\gamma(r(D_{e'}))$ is some $\rho(x_1, \dots, x_k)$ and $\text{out}(r(D_{e'}))$ is some $e_1 \dots e_k$ **then**
 if there exists $e_i: r(D_{e'}) \rightarrow n_i, i \in [k]$, such that $\gamma(n_i) = \xi$ **then**
 $\gamma(r(D_{e'})) \leftarrow \rho^{*x_i}$ ▷ Since $D_{e'}$ is ξ -normalized i is unique
 else
 Add a new node x labeled by ξ and $r(D_{e'}) \rightarrow x$ in $E_N(D_{e'})$
 $\gamma(r(D_{e'})) \leftarrow (\rho(x_1, \dots, x_k) \wedge x_{k+1} = 0) \vee (\wedge_{i \in [k]} x_i = 0 \wedge x_{k+1} = 1)$
 $\text{out}(r(D_{e'})) \leftarrow \text{out}(r(D_{e'}))(r(D_{e'}) \rightarrow x)$
 else
 Consider $D_{e'+\xi}$ instead of $D_{e'}$ for the remainder of the construction
// Second step
for each node n labeled by ξ which is not a **child** of the root r **do**
 $\gamma(n) \leftarrow \gamma(r)$
 $\text{out}(n) \leftarrow n \circ_r^{\text{src}} \text{out}(r)$ ▷ All those new edges are special
 Apply PP-suppression

none of the **childs** n_1, \dots, n_k of the root r has its label in \mathcal{P} . Since the new special edges have their destinations in n_1, \dots, n_k then the construction preserves Property SS.

- Case $e = e_1 + e_2$ (resp. $e = e_1 \parallel e_2$)
 D_e is built from the union of D_{e_1} and D_{e_2} , with a new node n labeled by $\rho(x_1, x_2) \equiv \sum_{i \in [2]} x_i = 1$ (resp. $x_1 = x_2 = 1$), edged by $n \rightarrow r(D_{e_1}), n \rightarrow r(D_{e_2})$, which is the root of D_e . Apply PP-suppression if necessary.

Example 40. Let $e = (b \cdot (a \parallel \xi^*))^{*\xi}$. The corresponding >1 -expression is $f = (b \cdot >^1(a \parallel (\xi^{*>^1} + \xi + \epsilon)))^{*\xi}$. The different steps of the construction of D_f are detailed in Figure 3. The transformation of the Presburger formula during the PP-suppression step is detailed in Example 28.

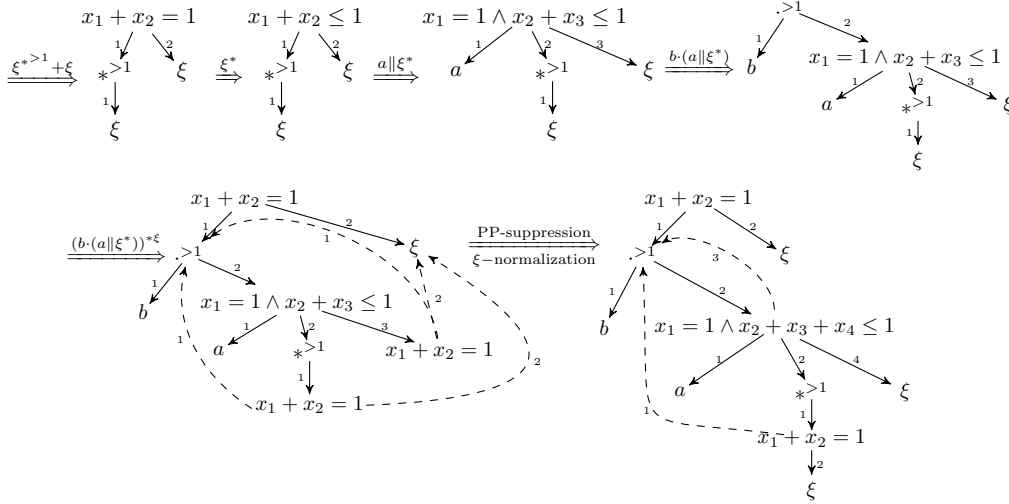


Figure 3: The step-by-step construction of the D-graph of $(b.>^1(a \parallel (\xi^{*>^1} + \xi + \epsilon))) * \xi$. PP-suppression is detailed only for the last step

In the remainder of the paper we let D_e denote the D-graph of e when e is a >1 -expression, or of the >1 -expression of e when e is a rational expression.

5.2. Properties

This section is devoted to some structural properties of D-graphs of rational expressions. The proofs of these properties are essentially verification by induction on the rational expression from which the D-graphs are built.

The following lemma characterizes the D-graph D_e of some >1 -expression e such that $\epsilon \in L(e)$.

Lemma 41. *Let e be a >1 -expression. Then $\epsilon \in L(e)$ if and only if $r(D_e)$ is labeled by some Presburger formula $\rho(x_1, \dots, x_k)$ and $(0, \dots, 0) \in L(\rho)$, or D_e has a unique node labeled by a closed Presburger tautology.*

Recall that by definition (Section 3.1), $L^{*\xi}$ is not rational when L contains a poset of the form $P\xi P'$. The following lemma characterizes the D-graph D_e of some rational expression e such that $P\xi P' \in L(e)$ for some P, P' .

Lemma 42. *Let f be a rational expression and e the corresponding >1 -expression. Then either D_e consists in a single node labeled by ξ , or there exists in D_e a sequence of consecutive edges $e_1: s_1 \rightarrow d_1, \dots, e_n: s_n \rightarrow d_n$ such that*

- $s_1 = r(D_e)$;
- d_n is labeled by ξ ;
- for all $i \in [n]$, if s_i is labeled by some Presburger formula $\rho(x_1, \dots, x_n)$, there is some $j \in [n]$ such that e_i is the j^{th} edge of $\text{out}(s_i)$ and $1_j \in L(\rho)$

if and only if $P_1\xi P_2 \in L(f)$ for some $P_1, P_2 \in SP^\circ(A)$.

Note that in the statement of Lemma 42, the D-graph is obtained from the >1 -expression of a rational expression, and not from an arbitrary >1 -expression. This is a necessary condition (consider for example the >1 -expression $\xi.>^1\epsilon$, whose language is empty, and its D-graph).

According to the inductive construction of D-graphs from >1 -expressions (Section 5.1), special edges are added only during the replacement of nodes labeled by ξ in the construction of $D_{e^*\xi}$ from D_e , for some >1 -expression e (Algorithm 4). Relying on Lemma 42, we have:

Proposition 43. *Let D_f be the D-graph of some rational expression f . For any sequence $\alpha = e_1 : s_1 \rightarrow d_1, \dots, e_l : s_l \rightarrow d_l$ of consecutive edges of D_f such that $s_1 = r(D_f)$ and $e_i \in E_S(D_f)$, there exists $i \in [l]$ such that s_i is labeled by some Presburger formula $\rho(x_1, \dots, x_m)$, and for all $r \in [m]$ such that e_i has index r in $\text{out}(s_i)$ and for all $(y_1, \dots, y_m) \in L(\rho)$, if $y_r > 0$ then $\sum_{i \in [m]} y_i > 1$.*

Starting the sequence of edges from any special edge we have:

Proposition 44. *Let D_f be the D-graph of some rational expression f . For any sequence $\alpha = e_1 \dots e_l$ of consecutive edges of D_f with $e_1, e_l \in E_S(D_f)$, $l > 1$, there exists a node n source of some e_i , $i \in [l]$, such that n is labeled by some Presburger formula $\rho(x_1, \dots, x_m)$, and for all $r \in [m]$ such that e_i has index r in $\text{out}(n)$ and for all $(y_1, \dots, y_m) \in L(\rho)$, if $y_r > 0$ then $\sum_{i \in [m]} y_i > 1$.*

5.3. Languages

In this section we define the language $L(D_e)$ of a D-graph D_e of a rational expression e , such that $L(D_e) = L(e)$.

Definition 45. *Let e be the >1 -expression of some rational expression. Let n be a node of D_e and $P \in SP^\circ(A)$. A path from n in D_e and labeled by P is an ordered labeled tree T_P such that:*

1. when P is a singleton labeled by $a \in A$ and n is a leaf labeled by a then T_P is a singleton labeled by (n, a) ;
2. when P is a sequential poset and n is labeled by $\cdot^{>1}$ and edged by $\text{out}(n) = n \rightarrow n_1, n \rightarrow n_2$ then T_P has the form $T_P = (m, T_{P_1}, T_{P_2})$ with m labeled by $(n, 2)$ and T_{P_i} is a path in D_e from n_i labeled by some non-empty P_i , $i \in [2]$, such that $P = P_1 P_2$;
3. when P is a sequential poset and n is labeled by $*^{>1}$ and edged by $\text{out}(n) = n \rightarrow n_1$ then T_P has the form $T_P = (m, T_{P_1}, \dots, T_{P_k})$ with m labeled by (n, k) and T_{P_i} is a path in D_e from n_1 labeled by some P_i , $i \in [k]$, such that $P = P_1 \dots P_k$. There must exist $i, i' \in [k]$, $i \neq i'$, such that $P_i, P_{i'} \neq \epsilon$;
4. when P is a sequential poset and n is labeled by $\omega^{>1}$ and edged by $\text{out}(n) = n \rightarrow n_1$ then T_P has the form $T_P = (m, (T_{P_i})_{i \in \omega})$ with m labeled by (n, ω) and T_{P_i} is a path in D_e from n_1 labeled by some P_i , $i \in \omega$, such that $P = \prod_{i \in \omega} P_i$. There must exist $i, i' \in \omega$ such that $i \neq i'$ and $P_i, P_{i'} \neq \epsilon$;
5. the construction when P is a sequential poset and n is labeled by $-\omega^{>1}$ is symmetrical;
6. when P is a sequential poset and n is labeled by $\natural^{>1}$ and edged by $\text{out}(n) = n \rightarrow n_1$ then T_P has the form $T_P = (m, (T_{P_i})_{i \in \alpha})$ for some $\alpha \in \mathcal{O} \setminus \{0, 1\}$ with m labeled by (n, α) and T_{P_i} is a path in D_e from n_1 labeled by some P_i , $i \in \alpha$, such that $P = \prod_{i \in \alpha} P_i$. There must exist $i, i' \in \alpha$ such that $i \neq i'$ and $P_i, P_{i'} \neq \epsilon$;
7. the construction when P is a sequential poset and n is labeled by $-\natural^{>1}$ is symmetrical;
8. when P is a sequential poset and n is labeled by $\diamond^{>1}$ and edged by $\text{out}(n) = n \rightarrow n_1, n \rightarrow n_2$ then T_P has the form $T_P = (m, (T_{P_j})_{j \in J \cup \hat{J}^*})$ for some $J \in \mathcal{S} \setminus \{0, 1\}$ with m labeled by $(n, J \cup \hat{J}^*)$ and T_{P_j} is a path in D_e from n_1 when $j \in J$, from n_2 when $j \in \hat{J}^*$, and labeled by some P_j , $j \in J \cup \hat{J}^*$, such that $P = \prod_{j \in J \cup \hat{J}^*} P_j$. There must exist $j, j' \in J \cup \hat{J}^*$ such that $j \neq j'$ and $P_j, P_{j'} \neq \epsilon$;
9. when n is labeled by a Presburger formula $\rho(x_1, \dots, x_k)$ and edged by $\text{out}(n) = n \rightarrow n_1, \dots, n \rightarrow n_k$ then T_P has the form $T_P = (m, (T_{P_{i,j}})_{i \in [k], j \in [y_i]})$ with m labeled by some $(n, (y_1, \dots, y_k))$ with $(y_1, \dots, y_k) \in L(\rho)$, and $T_{P_{i,j}}$ is a path in D_e from n_i labeled by some non-empty $P_{i,j}$, $i \in [k]$, $j \in [y_i]$, such that $P = \prod_{i \in [k]} \prod_{j \in [y_i]} P_{i,j}$. Note that since the parallel product of posets commutes, the ordering of the sequence $(T_{P_{i,j}})_{i \in [k], j \in [y_i]}$ has no consequence.

In the last case, when T_P has the form $T_P = (m, (T_{P_{i,j}})_{i \in [k], j \in [y_i]})$ and $n \rightarrow n_t \in E_S(D_e)$ for some $t \in [k]$, we say that $P_{t,j}$ is marked by $n \rightarrow n_t$ in T_P , and that $n \rightarrow n_t$ starts $T_{P_{t,j}}$, for all $j \in [y_t]$. Marking is an hereditary notion: every factor of P_i marked by some special edge e in T_{P_i} is also considered marked by e in T_P . The class of paths of D_e from n labeled by P is denoted $\mathcal{R}_P(D_e, n)$.

Note that ϵ can be the label of a path only by Case 9 when $(0, \dots, 0) \in L(\rho)$ or when ρ is a closed tautology. Note also that Property PP ensures that all $P_{i,j}$ are non-empty in Case 9. Finally observe also that a path may have infinite height (with the usual definition of height on trees). Proposition 47 below shows that Definition 45 is well-founded.

Definition 46. Let T and T' be two paths in the D-graph of some >1 -expression. Then T' is a (resp. strict) sub-path of T if $T' \subseteq T$ (resp. $T' \subsetneq T$). A strict sub-path T' of T is direct when it is not a strict sub-path of strict sub-path of T .

Proposition 47. Definition 45 is well-founded.

Proof. Consider the class of all (T_P, P) formed of a path T_P and its label P partially ordered by the relation $(T_P, P) < (T_{P'}, P')$ if and only if T_P is a strict sub-path of $T_{P'}$. We claim that $<$ is a well-ordering. Indeed, assume by contradiction that there exists an infinite decreasing sequence $\dots < (T_{P_i}, P_i) < \dots < (T_{P_1}, P_1) < (T_{P_0}, P_0)$. We may assume that $T_{P_{i+1}}$ is a direct sub-path of T_{P_i} , for all i . Each T_{P_i} is a path from a node n_i . Thus, there is an edge $e_i : n_i \rightarrow n_{i+1}$ in D_e for each i . The sequence e_0, e_1, \dots is infinite and consecutive. Since a loop of consecutive edges in a D-graph must contain a special edge, there is among the e_i 's a special edge e that occurs infinitely often in the sequence. Let i_0, i_1, \dots be the sequence of all indexes i_j such that $e_{i_j} = e$. The construction of P_{i_j} from $P_{i_{j+1}}$ involves at least one parallel product by a non-empty labeled poset, as a consequence of Proposition 44. Since the destination of a special edge is not labeled by a Presburger formula, the construction of P_{i_j} from $P_{i_{j+1}}$ also involves a sequential product by a non-empty sequential labeled poset. Thus necessarily $r_X(P_{i_j}) > r_X(P_{i_{j+2}})$. The sequence $r_X(P_{i_0}) > r_X(P_{i_2}) > \dots$ is an infinitely decreasing sequence of X-ranks, which can not exist since the ordering of X-ranks is well-founded. \square

Definition 48. Let D be the D-graph of some rational expression. Let n be a node of D . The language $L(n)$ of n consists of all labels of paths from n in D . The language of D is $L(D) = L(r(D))$.

Note that when D is the D-graph of some rational expression and n is a node of D , if $a \in L(n)$ for some $a \in A$, then necessarily n is labeled by a , or n is labeled in \mathcal{P} and has a **child** labeled by a .

It is just verification to check that when they are applied during the construction of the D-graph of a rational expression, Algorithms 1 and 2 preserve the languages of the D-graphs as well as Properties PP, SS and DAG.

Property PP is used in particular in order to compute, during the construction of D_e from e , the Presburger formulæ that will appear later in the P-MSO formula built from D_e . Property SS ensures that $L(n)$ do not contain parallel posets when n is the destination of a special edge.

Remark 49. As a consequence of Proposition 44, when there is a path from $r(D_e)$ labeled by some P , if it contains two different sub-paths labeled by F_1 and F_2 both marked by the same special edge, then F_1 and F_2 are necessarily sequential posets (Property SS), and either

(C₁) $F_1 \cap F_2 = \emptyset$. Possibly, $F_1 F_2$ is a sequential factor of P ;

(C₂) one is strictly included into the other, wlog. $F_1 \subsetneq F_2$. In this case, there is some $x \in F_2 \setminus F_1$ such that x is incomparable to all the elements of F_1 .

The following proposition can be checked with an induction on f .

Proposition 50. For any rational expression f , $L(D_f) = L(f)$.

Example 51. Let L_1 and L_2 be the languages of respectively $e_1 = a \circ_\xi (a(\xi \parallel \xi))^{\ast\xi}$ (see also Example 20) and of $e_2 = b^\diamond$. Consider the D-graphs D_e, D_{e_1} and D_{e_2} of respectively $e = e_1 \diamond e_2$, e_1 and e_2 pictured on the left side of Figure 4. We let ζ denote the linear ordering of all integers (negative, 0 and positive) and by \mathfrak{B} the finite linear ordering $\{x_1, x_2, x_3\}$ such that $x_1 < x_2 < x_3$. Then $\mathfrak{B}^* = \{(\{x_1\}, \{x_2, x_3\}), (\{x_1, x_2\}, \{x_3\})\}$. Let $x_4 = (\{x_1\}, \{x_2, x_3\})$ and $x_5 = (\{x_1, x_2\}, \{x_3\})$. Then $\mathfrak{B} \cup \mathfrak{B}^* = \{x_1, \dots, x_5\}$ ordered by $x_1 < x_4 < x_2 < x_5 < x_3$. The same reasoning can be applied to infer the infinite set of elements and the ordering of $\zeta \cup \hat{\zeta}^*$. On the right side of the figure is pictured a path T of D_e from $r(D_e)$ labeled by $P = \prod_{j \in \mathfrak{B} \cup \mathfrak{B}^*} F_j$, where $F_{x_1} = a(a \parallel a \parallel a)$, $F_{x_4} = \epsilon$, $F_{x_2} = a$, $F_{x_5} = b^\zeta$ and $F_{x_3} = a(a \parallel a)$. Note that $F_{x_1}, F_{x_2}, F_{x_3} \in L(a \circ_\xi (a(\xi \parallel \xi))^{\ast\xi})$ and $F_{x_4}, F_{x_5} \in L(b^\diamond)$. Note also that $F_{x_5} = \prod_{j \in \zeta \cup \hat{\zeta}^*} F_{5,j}$ where $F_{5,j} = b$ when $j \in \zeta$ and $F_{5,j} = \epsilon$ when $j \in \hat{\zeta}^*$. Observe that the path T has a unique direct sub-path T' also labeled by P . Let $T'_{x_1}, T'_{x_4}, T'_{x_2}, T'_{x_5}, T'_{x_3}$ be the direct sub-paths of T' taken in order from the left to the right. Then observe that each T'_{x_i} is labeled by F_{x_i} .

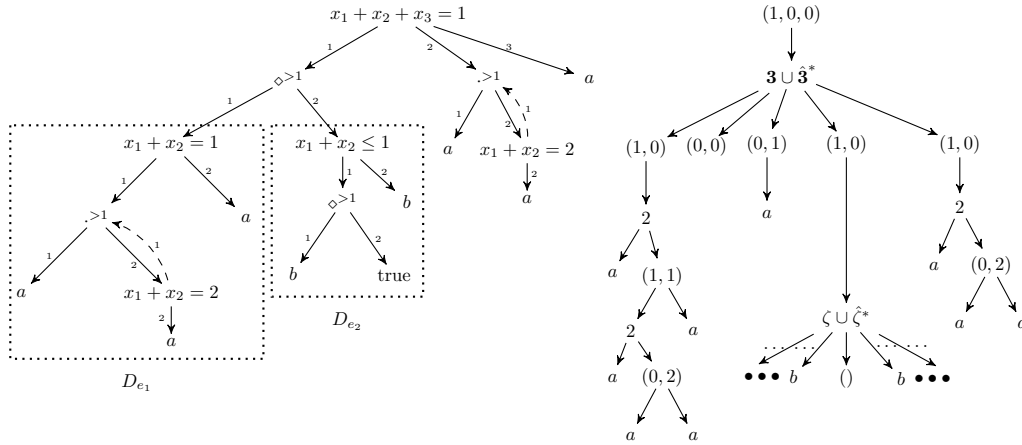


Figure 4: The D-graph D_e of $e = a \circ_\xi (a(\xi \parallel \xi))^{\ast\xi} \diamond b^\diamond$ and a path in D_e from $r(D_e)$. In order to lighten figures, nodes of paths are labeled by s instead of (n, s) where n is a node of D_e

Example 52. Let $e = ((a \parallel \xi^*)^{\ast\xi} + b(c \parallel d))^{\ast\xi}$ be a rational expression. The D-graph D_e of the >1 -expression of e and the poset $P = a \parallel (\xi(a \parallel (b(c \parallel d)b(c \parallel d))))$ of $L(D_e)$ are represented on the left side of Figure 5. Let n_1, \dots, n_{11} be the pre-order traversal of D_e without its special edges. We have $E_S(D_e) = \{n_4 \rightarrow n_2, n_4 \rightarrow n_3, n_4 \rightarrow n_7\}$. The leaves of D_e are $n_2, n_5, n_6, n_8, n_{10}, n_{11}$. Note that in $\text{out}(n_4) = n_4 \rightarrow n_2, n_4 \rightarrow n_3, n_4 \rightarrow n_2, n_4 \rightarrow n_3, n_4 \rightarrow n_5, n_4 \rightarrow n_7$ each of $n_4 \rightarrow n_2$ and $n_4 \rightarrow n_3$ occurs twice. On the right side of Figure 5 is pictured a path T_P in D_e from n_1 labeled by P . Its root is labeled by $(n_1, (1, 1, 0, 0))$ with $(1, 1, 0, 0) \in L(\rho_1)$. The path T_P has two direct sub-paths $T_{P_1} = (n_2, a)$ and $T_{P_2} = ((n_3, 2), T_{P_{2,1}}, T_{P_{2,2}})$ labeled respectively by $P_1 = a$ and $P_2 = P_{2,1}P_{2,2}$, with $P = P_1 \parallel P_2$, $T_{P_{2,1}}$ and $T_{P_{2,2}}$ direct sub-paths of T_{P_2} from n_4 and respectively labeled by $P_{2,1} = \xi$ and $P_{2,2} = a \parallel (b(c \parallel d)b(c \parallel d))$. In continuation, $T_{P_{2,1}}$ and $T_{P_{2,2}}$ have respectively the form $((n_4, (0, \dots, 0, 1, 0)), (n_5, \xi))$ and $((n_4, (1, 1, 0, \dots, 0)), T_{F_4}, T_{F_3})$ with $(0, \dots, 0, 1, 0), (1, 1, 0, \dots, 0) \in L(\rho_2)$ and $(n_5, \xi), T_{F_4}$ and T_{F_3} are from respectively n_5, n_2 and n_3 labeled by respectively $\xi, F_4 = a$ and $F_3 = b(c \parallel d)b(c \parallel d)$. Note that there is a path from n_4 , different from $T_{P_{2,2}}$ and with the same label. It differs only by its root which is labeled by $(n_4, (0, 0, 1, 1, 0, 0))$. Observe that T_{F_3} and T_{F_4} are started by respectively $n_4 \rightarrow n_3$ and $n_4 \rightarrow n_2$. Similarly, T_{F_1} and T_{F_2} are sub-paths of T_{F_3} from n_7 labeled by $F_1 = F_2 = b(c \parallel d)$ and started by $n_4 \rightarrow n_7$. Each F_i is marked by the edge that starts T_{F_i} , $i \in [4]$. As a consequence of Remark 49, the edge marking F_3 is necessarily different from those which mark F_1 and F_2 .

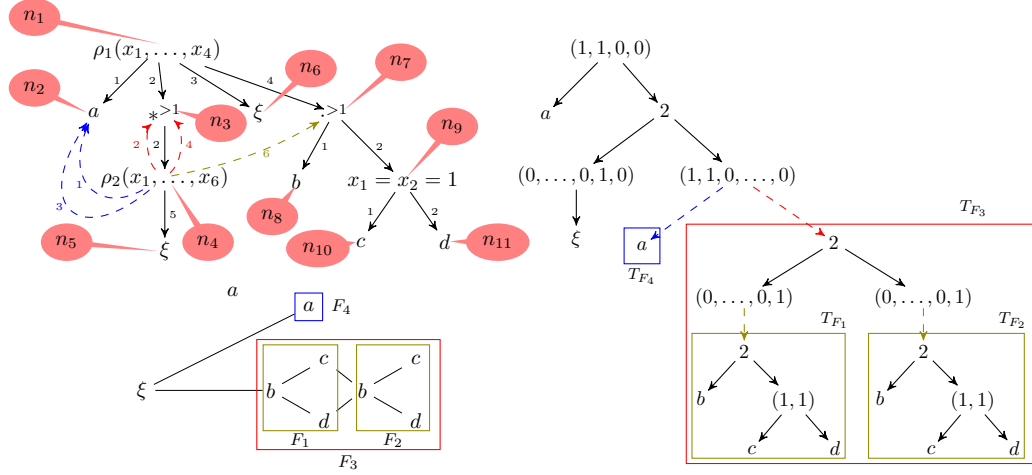


Figure 5: The D-graph D_e of the >1 -expression of $e = ((a \parallel \xi^*)^* \xi + (b(c \parallel d)))^* \xi$, a poset $P = a \parallel (\xi(a \parallel (b(c \parallel d)b(c \parallel d))))$ of $L(e)$ and a path in D_e from $r(D_e)$ labeled by P . Here $\rho_1(x_1, \dots, x_4) \equiv (x_1 = x_2 = 0 \wedge x_3 + x_4 = 1) \vee (x_1 \geq 1 \wedge x_2 + x_3 + x_4 \leq 1)$ and $\rho_2(x_1, \dots, x_6) \equiv (x_1 \geq 1 \wedge x_2 \leq 1 \wedge 3 \leq i \leq 6 x_i = 0) \vee (x_1 \geq 0 \wedge x_2 = 0 \wedge ((x_3 = x_4 = 0 \wedge x_5 + x_6 = 1) \vee (x_3 \geq 1 \wedge x_4 + x_5 + x_6 \leq 1)))$

6. Coloring

Let C be a non-empty finite set whose elements are named *colors* and $P \in SP^\diamond$. We let $F_s(P)$ denote the class of all sequential factors of P .

Definition 53. Let $P \in SP^\diamond$ and C be a non-empty finite set of colors. A s -coloring $\mathfrak{c}: F_s(P) \rightarrow C$ of P with C is a partial map from $F_s(P)$ to C . It is compatible if for any different $F, F' \in F_s(P)$ such that $\mathfrak{c}(F)$ and $\mathfrak{c}(F')$ are defined, either:

- $F \cap F' = \emptyset$ and, if $FF' \in F_s(P)$, then $\mathfrak{c}(F) \neq \mathfrak{c}(F')$;
- one is strictly included into the other, say $wlog F \subsetneq F'$, and if $\mathfrak{c}(F) = \mathfrak{c}(F')$ there is some $x \in F' \setminus F$ such that x is incomparable to all the elements of F .

In Sub-section 6.2 we show how compatible s -colorings can be encoded by means of MSO. The technique relies on another one, called *ms-coloring*, specific to some particular case of sequential factors (Sub-section 6.1). In Sub-section 6.3 we link compatible s -colorings and paths in D-graphs of rational expressions.

Before starting we need some additional definitions on posets.

A factor F of P is *sequentially maximal* if there is no R, S with at least one of them non-empty such that $R + F + S$ is a factor of P . We let $F_{ms}(P)$ denote the class of all elements of $F_s(P)$ that are sequentially maximal. A factor $F \in F_{ms}(P)$ is *direct* if there is no $P' \in F_{ms}(P) \setminus \{P, F\}$ such that $F \in F_{ms}(P')$. We let $DF_{ms}(P)$ denote the subclass of $F_{ms}(P)$ consisting of all strict direct factors of P .

Example 54. Let $P = (a \cup b) + c + (d \cup (e + (f \cup (g + h)))) \cup i$. Then:

- $a, a \cup b$ and $(a \cup b) + c$ are strict factors of P ;
- $a \cup b$ and $d \cup (e + (f \cup (g + h))) \cup i$ are parallel factors of P ;
- $(a \cup b) + c$ and $c + (d \cup (e + (f \cup (g + h)))) \cup i$ are sequential factors of P ;
- $g + h$ and $d \cup i$ are sequentially maximal factors of P ;
- $a, e + (f \cup (g + h))$ and P are direct sequentially maximal sequential factors of P .

6.1. ms-coloring

A *ms-coloring* of $P \in SP^\circ$ with C is a total map $\mathbf{c}: F_{ms}(P) \rightarrow C$. In general functions are not MSO-expressible, but a ms-coloring can be encoded using three sets X_c^w, X_c^s, X_c^p of elements of P , for each $c \in C$. For short let ms_C^X denote the sets $(X_c^w, X_c^s, X_c^p)_{c \in C}$. Also, given $c \in C$, $x \in P$ and $\alpha \in \{w, s, p\}$, we denote by $ms_C^X +_c^\alpha x$ the subsets $(X_{c'}^{\alpha w}, X_{c'}^{\alpha s}, X_{c'}^{\alpha p})_{c' \in C}$ of P defined by

$$X_{c'}^{\alpha\beta} = \begin{cases} X_{c'}^\beta \cup \{x\} & \text{when } \beta = \alpha \text{ and } c' = c; \\ X_{c'}^\beta & \text{otherwise.} \end{cases}$$

When $(y, x, x') \in P^3$ we denote by $ms_C^X +_c(y, x, x') = ms_C^X +_c^s y +_c^p x +_c^p x'$. Also, we denote by $ms_C^X +_c x = ms_C^X +_c^w x$ when $x \in P$.

The three following definitions lay the foundations on how ms_C^X is used to encode a ms-coloring of a poset.

Definition 55. Let $P \in SP^\circ$, ms_C^X some subsets of P , and let $c \in C$. Then $(y, x, x') \in X_c^s \times X_c^p \times X_c^p$ is directly bound by c for $F \in F_{ms}(P)$ in ms_C^X if the following conditions are true:

1. $F = F_1 + F_2 + F_3$, $F_2 = F_x \cup F_{x'}$, $x \in F_x, x' \in F_{x'}$ and $y \in F \setminus F_2$, for some $F_1, F_2, F_3, F_x, F_{x'}$;
2. there exists $z \in F \setminus F_2$ incomparable to (and distinct from) y ;
3. there is no $F' \in F_{ms}(F) \setminus \{F\}$ such that $F' = F'_1 + F'_2 + F'_3$, $F'_2 = F'_{2,1} \cup F'_{2,2}$ for some $F'_1, F'_2, F'_3, F'_{2,1}, F'_{2,2}$, such that one of the following conditions is true:
 - (a) there exist $x_1, x_2 \in X_c^p$ such that $x_1 \in F'_{2,1}, x_2 \in F'_{2,2}$ and $y \in F' \setminus F'_2$;
 - (b) there exist $y_1 \in X_c^s$ and $x_1 \in X_c^p$ such that $x \in F'_{2,1}, x_1 \in F'_{2,2}$ and $y_1 \in F' \setminus F'_2$;
 - (c) there exist $y_1 \in X_c^s$ and $x_1 \in X_c^p$ such that $x' \in F'_{2,1}, x_1 \in F'_{2,2}$ and $y_1 \in F' \setminus F'_2$.

In this case y , x , and x' are respectively directly s-bound, p-bound and p-bound by c for F in ms_C^X . In addition, we say that (y, x, x') (resp. y , x and x') is indirectly bound (resp. s-bound, p-bound and p-bound) by c in ms_C^X for all $F' \in F_{ms}(P) \setminus \{F\}$ such that $F \in F_{ms}(F')$.

Observe that if y, x and (y, x, x') are directly respectively s-bound, p-bound and bound by $c \in C$ for some $F \in F_{ms}(P)$ in ms_C^X then they are indirectly respectively s-bound, p-bound and bound in $(X_c^w \cap F', X_c^s \cap F', X_c^p \cap F')_{c \in C}$ for all $F' \in F_{ms}(P)$ strictly containing F . Also, it follows from Definition 55 that if (y, x, x') is directly bound by some $c \in C$ for P in ms_C^X then there is no $(y, s, s'), (t, x, s'), (t, s, x') \in X_c^s \times X_c^p \times X_c^p$ that is bound by c for some $F' \in F_{ms}(P) \setminus \{P\}$ in ms_C^X . An element of P^3 (resp. P) is *bound* (resp. *s-bound*, *p-bound*) by some $c \in C$ for some $F \in F_{ms}(P)$ in some ms_C^X if it is indifferently directly or indirectly bound (resp. s-bound, p-bound) by $c \in C$ for $F \in F_{ms}(P)$ in ms_C^X .

Definition 56. Let $P \in Seq$ of irreducible sequential factorization $\sum_{j \in J} P_j$ and ms_C^X some subsets of P . The set $\mathcal{C}_{ms_C^X}(P)$ of candidates for the ms-coloring of P is the smallest subset of $P \cup P^3$ such that:

- if there exist $j \in J$ such that $|P_j| = 1$ and $c \in C$ such that $P_j \subseteq X_c^w$ then $P_j \subseteq \mathcal{C}_{ms_C^X}(P)$;
- if there is no $j \in J$ such that $|P_j| = 1$ and there exist $c \in C$, $y, x, x' \in P$ such that (y, x, x') is directly bound by c for P in ms_C^X then $(y, x, x') \in \mathcal{C}_{ms_C^X}(P)$.

Definition 57. Let $P \in Seq$ of irreducible sequential factorization $\sum_{j \in J} P_j$ and ms_C^X some subsets of P . Then ms_C^X ms-colors P in $c \in C$, denoted $ms_C^X(P) = c$, if $\mathcal{C}_{ms_C^X}(P)$ is some singleton $\mathcal{C}_{ms_C^X}(P) = \{x\}$ and either $x \in X_c^w$ or $x \in X_c^s \times X_c^p \times X_c^p$. In this case we say that P is ms-colored in c by x . In addition ms_C^X completely ms-colors a poset $P \in SP^\circ$ if $ms_C^X(F)$ is defined for all $F \in F_{ms}(P)$. Finally, when $P \in SP^\circ$ and $\mathbf{c}: F_{ms}(P) \rightarrow C$ is a ms-coloring of P with C then the subsets ms_C^X of P encode \mathbf{c} if $ms_C^X(F) = \mathbf{c}(F)$ for all $F \in F_{ms}(P)$.

At the end of this sub-section we show how to encode the ms-coloring of some $P \in SP^\circ$ by some ms_C^X , by induction on $r_X(P)$. The ms-coloring in $c \in C$ of a factor of $F_{ms}(P)$ that has no candidate for its ms-coloring and has an element x comparable to all the others is easy, since it suffices to add x to X_c^w . Note that this does not influence the ms-coloring of other factors of $F_{ms}(P)$. However, adding an element of P into X_c^s or X_c^p in order to ms-color some factor of $F_{ms}(P)$ may influence the ms-coloring of other factors. This leads us to introduce the following definition.

Definition 58. Let $P \in SP^\circ$, ms_C^X some subsets of P , $F \in F_{ms}(P)$ and $c \in C$. Then $x \in P$ is s-free (resp. p-free) for (F, c) in ms_C^X if it is not s-bound (resp. p-bound) by c for F in $ms_C^X +_c^s x$ (resp. $ms_C^X +_c^p x$).

Observe that if $x \in P$ is α -free for some (F, c) with $F \in F_{ms}(P)$, $c \in C$ and $\alpha \in \{s, p\}$ then x is α -free for (F', c) in ms_C^X and in $(X_c^w \cap F', X_c^s \cap F', X_c^p \cap F')_{c \in C}$ for all $F' \in F_{ms}(F)$ such that $x \in F'$.

Example 59. Let $C = \{\mathbf{red}, \mathbf{blue}\}$ be a set of colors. Consider the poset $P \in SP^\circ$ of Figure 6. Let $P_1 = \{x_i : i \in [2; 16]\}$, $P_2 = \{x_i : i \in [17; 25]\}$ and $P_3 = \{x_{19}, x_{20}, x_{22}, x_{23}\}$. Then $P_1, P_2, P_3 \in F_{ms}(P)$. Let $\mathbf{c} : F_{ms}(P) \rightarrow C$ be a ms-coloring of P such that $\mathbf{c}(P_1) = \mathbf{red}$ and $\mathbf{c}(P_2) = \mathbf{c}(P_3) = \mathbf{blue}$. We define

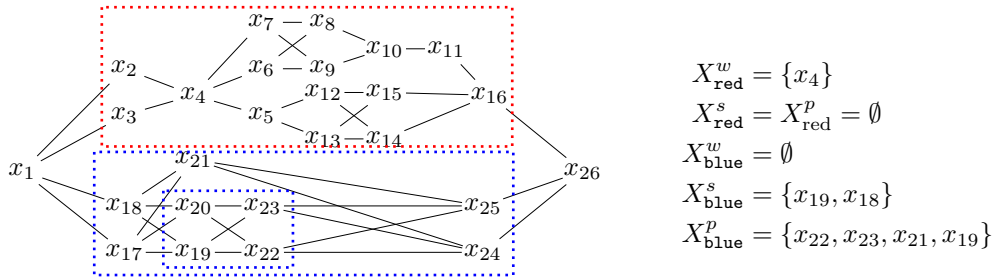


Figure 6: The Hasse diagram of a poset $P \in SP^\circ$ and a ms-coloring of $P_1, P_2, P_3 \in F_{ms}(P)$

ms_C^X that encodes $\mathbf{c}(P_i)$, $i \in [3]$. Note that in this example we encode only the restriction of \mathbf{c} to P_1, P_2, P_3 : the generalization of the method in order to encode the total map $\mathbf{c} : F_{ms}(P) \rightarrow C$ is easy. We start with all sets of ms_C^X empty. Observe that P_1 contains two elements comparable to all the other elements of P_1 : x_4 and x_{16} . Choose one of them, indifferently x_4 . Then add x_4 to $X_{\mathbf{red}}^w$. This suffices to encode that P_1 is ms-colored in \mathbf{red} , since there is no other candidate in ms_C^X for the ms-coloring of P_1 . Let us now encode that $\mathbf{c}(P_3) = \mathbf{blue}$. Note that since P_3 has no element comparable to all the others, we have to use a 3-tuple in P_3 . Choose for example (x_{19}, x_{22}, x_{23}) , and add x_{19} in $X_{\mathbf{blue}}^s$, x_{22}, x_{23} in $X_{\mathbf{blue}}^p$. Then (x_{19}, x_{22}, x_{23}) is directly bound by \mathbf{blue} for P_3 in ms_C^X . This encodes that P_3 is ms-colored in \mathbf{blue} , since there is no other $c \in C$, $(z, z', z'') \in X_c^s \times X_c^p \times X_c^p$ that can also encode a color for P_3 , and that this does not interfere with encoding of ms-coloring of other factors of $F_{ms}(P)$. Let us turn now to P_2 . It has no element comparable to all the others, so we have to choose a 3-tuple (y, y', y'') as we did for P_3 . Note that x_{19} is not a possible choice for y since it is already s-bound by \mathbf{blue} for P_3 in ms_C^X . Also, x_{20} is not a possible choice for y since it is not s-free for (P_3, \mathbf{blue}) and thus not s-free for (P_2, \mathbf{blue}) . Indeed, adding it in $X_{\mathbf{blue}}^s$ would make (x_{20}, x_{22}, x_{23}) another candidate for ms-coloring P_3 in blue and as a consequence, regardless of y' and y'' , (x_{20}, y', y'') could not be bound by \mathbf{blue} for P_2 in ms_C^X . Choose, for example, $y = x_{18}$ and $y' = x_{21}$. As we can not take $y'' \in \{x_{22}, x_{23}\}$ since they are already p-bound by \mathbf{blue} for P_3 in ms_C^X , then necessarily y'' is indifferently one of x_{20}, x_{19} , say x_{19} . Add y in $X_{\mathbf{blue}}^s$ and y', y'' in $X_{\mathbf{blue}}^p$ and this is done. Finally, $(X_{\mathbf{red}}^w, X_{\mathbf{red}}^s, X_{\mathbf{red}}^p) = (\{x_4\}, \emptyset, \emptyset)$ and $(X_{\mathbf{blue}}^w, X_{\mathbf{blue}}^s, X_{\mathbf{blue}}^p) = (\emptyset, \{x_{19}, x_{18}\}, \{x_{22}, x_{23}, x_{21}, x_{19}\})$.

Proposition 60. Let C be a non-empty finite set of colors and $P \in SP^\circ$. For any ms-coloring \mathbf{c} of P with C there exist some subsets ms_C^X of P that encode \mathbf{c} .

The remainder of this section is devoted to the proof of Proposition 60. The following lemma shows in particular that any poset P of Seq that is completely ms-colored by some ms_C^X contains, for all $c \in C$, at least a s-free and a p-free element for (P, c) in ms_C^X .

Lemma 61. *Let $P \in \text{Seq}$ completely ms-colored by some ms_C^X . If for all $c \in C$, all the elements of X_c^p (resp. X_c^s) are p-bound (resp. s-bound) by c for P in ms_C^X then there exist $x, y \in P$ such that for all $c \in C$, x is p-free and y is s-free for (P, c) in ms_C^X .*

Proof. We prove the lemma by induction on $r_X(P)$. If $r_X(P) = (0, i)$ for $i \in \{0, 1\}$ then P is either the singleton or a finite linear ordering. In both cases and for all $c \in C$, there is no $(y, x, x') \in P^3$ bound by c for P in ms_C^X . Thus, $\bigcup_{c \in C} X_c^s = \bigcup_{c \in C} X_c^p = \emptyset$ and for all $x \in P$, for all $c \in C$, x is naturally s-free and p-free

for (P, c) in ms_C^X . Otherwise, assume that the irreducible sequential factorization of P is $\sum_{j \in J} P_j$, for some $J \in \mathcal{S} \setminus \{0, 1\}$. Assume that P is ms-colored in e by $z \in X_e^w$. Then by Definitions 56–57, z is comparable to each element of $P \setminus \{z\}$. In this case, by Definition 58, z is s-free and p-free for (P, c) in ms_C^X , for all $c \in C$.

Otherwise, assume that P is ms-colored in e by $(y, x, x') \in X_e^s \times X_e^p \times X_e^p$. Then by Definition 57, $\mathcal{C}_{ms_C^X}(P) = \{(y, x, x')\}$ and by Definition 56 there is no $j \in J$ such that $|P_j| = 1$, and (y, x, x') is directly bound by e for P in ms_C^X . By Definitions 55, $P = P_1 + P_2 + P_3$, $P_2 = P_x \cup P_{x'} \cup P_{x''}$ for some $P_x, P_{x'} \in DF_{ms}(P)$ and $P_1, P_2, P_3, P_{x''}$ some factors of P , such that $x \in P_x, x' \in P_{x'}$ and $y \in P \setminus P_2$. Observe that there exists $P_y \in DF_{ms}(P)$ such that $y \in P_y$ and all the elements of P_y are comparable to those of P_2 . We have $r_X(P_y), r_X(P_x), r_X(P_{x'}) < r_X(P)$. Furthermore, x is not p-bound, x' is not p-bound and y is not s-bound by e for respectively $P_x, P_{x'}$ and P_y in ms_C^X , otherwise (y, x, x') cannot be directly bound by e for P in ms_C^X according to Definition 55. So, let $\overline{X_e^p} = X_e^p \cap P_y, \overline{X_e^s} = (X_e^s \cap P_y) \setminus \{y\}$ and $\overline{X_c^p} = X_c^p \cap P_y, \overline{X_c^s} = X_c^s \cap P_y$ for all $c \in C \setminus \{e\}$ and let $ms_C^{\overline{X}} = (X_c^w, \overline{X_c^s}, \overline{X_c^p})_{c \in C}$. Observe that P_y and $ms_C^{\overline{X}}$ satisfy the hypothesis of the lemma. Then by induction hypothesis, there exists $z \in P_y$ such that z is p-free for (P_y, c) in $ms_C^{\overline{X}}$, for all $c \in C$. In addition, for all $c \in C$, z is still p-free for (P_y, c) in ms_C^X . In fact, assume by contradiction that there exists $c' \in C$ such that z is not p-free for (P_y, c') in ms_C^X . Then by Definition 58, there exist $y_1 \in X_{c'}^s$ and $x_2 \in X_{c'}^p$ such that (y_1, x_2, z) is bound by c' for P_y in $ms_C^{X'} = ms_C^X +_{c'}^p z$. Note that as y is used in (y, x, x') to ms-color P in e then y is s-free for (P_y, e) in $ms_C^{\overline{X}}$. Thus z may not be p-free only when $y_1 = y$ and $c' = e$. Indeed, despite $y_1 \in \overline{X_{c'}^s}$ and $x_2 \in \overline{X_{c'}^p}$ when $y_1 \neq y$ or $c' \neq e$, (y_1, x_2, z) did not form a bound tuple since z was proved p-free for (P_y, c') in $ms_C^{\overline{X}}$.

By the lemma hypothesis, x_2 is p-bound by e for P in ms_C^X . Then there exist $y_2 \in X_e^s$ and $x'_2 \in X_e^p$ such that (y_2, x_2, x'_2) is bound by e for P in ms_C^X . In addition, (y_2, x_2, x'_2) is bound by e for P_y in ms_C^X since if $y_2, x'_2 \in P \setminus P_y$ then (y_2, x_2, x'_2) will be another candidate for ms-coloring P in e by ms_C^X which is not allowed by definition of the complete ms-coloring of P . Thus, (y, x_2, z) and (y_2, x_2, x'_2) are bound by e for P_y in respectively $ms_C^{X'}$ and ms_C^X . Assume that (y, x_2, z) and (y_2, x_2, x'_2) are directly bound by e for respectively P'_y in $ms_C^{X'}$ and P''_y in ms_C^X for some $P'_y, P''_y \in F_{ms}(P_y)$. Observe that (y_2, x_2, x'_2) is bound by e for P''_y in $ms_C^{\overline{X}}$ too since $y_2 \neq y$ (the contrary would contradict the fact that (y, x, x') is bound by e for P in ms_C^X). Observe also that $ms_C^{X'}$ contains ms_C^X . There are three cases:

1. if $P'_y \in F_{ms}(P''_y) \setminus \{P''_y\}$ then necessarily each of y, z and x_2 are incomparable to x'_2 . Then, naturally (y_2, z, x'_2) is also directly bound by e for P''_y in $ms_C^{\overline{X}} +_{c'}^p z$, which contradicts the fact that z is p-free for (P_y, e) in $ms_C^{\overline{X}}$;
2. in the case where $P'_y = P''_y$ then by Definition 55, $P'_y = P'_{y,1} + P'_{y,2} + P'_{y,3}$ such that $x_2, x'_2, z \in P'_{y,2}$, z and x'_2 are incomparable to x_2 and $y, y_2 \in P'_y \setminus P'_{y,2}$. In this case (y_2, x_2, z) is also directly bound by e for P'_y in $ms_C^{\overline{X}} +_{c'}^p z$, which contradicts the fact that z is p-free for (P_y, e) in $ms_C^{\overline{X}}$;
3. if $P''_y \in F_{ms}(P'_y) \setminus \{P'_y\}$ then (y, x_2, z) cannot be bound by e in $ms_C^{X'}$ by Definition 55.

Thus for all $c \in C$, z is still p-free for (P_y, c) in ms_C^X .

Let us prove that for all $c \in C$, z is still also p-free for (P, c) in ms_C^X . By contradiction suppose that there exists $c' \in C$ such that z is not p-free for (P, c') in ms_C^X . Then, there exist $y_1 \in X_{c'}^s$ and $x_2 \in X_{c'}^p$ such that (y_1, x_2, z) is bound by c' for P in $ms_C^{X'}$. As z is p-free for (P_y, c') in ms_C^X then $y_1, x_2 \notin P_y$. That means that (y_1, x_2, z) is directly bound for P in $ms_C^{X'}$. Then by Definition 55 there exist $P_{y_1}, P_{x_2} \in DF_{ms}(P)$ such

that $y_1 \in P_{y_1}$, $x_2 \in P_{x_2}$ and all the elements of P_y are comparable to those of P_{y_1} and incomparable to those of P_{x_2} . By hypothesis, y_1 and x_2 are respectively s-bound and p-bound by c' for P in ms_C^X . Then there exist $y_2 \in X_{c'}^s$ and $x_1, x'_1, x'_2 \in X_{c'}^p$ such that (y_1, x_1, x'_1) and (y_2, x_2, x'_2) are bound by c' for P in ms_C^X . By hypothesis P is ms-colored by (y, x, x') in ms_C^X . Then the complete ms-coloring of P by ms_C^X implies that (y_1, x_1, x'_1) and (y_2, x_2, x'_2) are bound by c' for respectively P_{y_1} and P_{x_2} in ms_C^X . Observe that $ms_C^{X'}$ contains ms_C^X and $P_{y_1}, P_{x_2} \in F_{ms}(P) \setminus \{P\}$. Thus by Definition 55, (y_1, x_2, z) cannot be bound by c' in $ms_C^{X'}$ which is a contradiction.

Now, let $\overline{X_e^s} = X_e^s \cap P_x$, $\overline{X_e^p} = (X_e^p \cap P_x) \setminus \{x\}$ and $\overline{X_c^p} = X_c^p \cap P_x$, $\overline{X_c^s} = X_c^s \cap P_x$ for all $c \in C \setminus \{e\}$ and let $ms_C^{\overline{X}} = (X_c^w, \overline{X_c^s}, \overline{X_c^p})_{c \in C}$. Observe that P_x and $ms_C^{\overline{X}}$ satisfy the hypothesis of the lemma. Then by induction hypothesis, for all $c \in C$, there exists $z \in P_x$ such that z is s-free for (P_x, c) in $ms_C^{\overline{X}}$. The proof that z is still s-free for (P, c) , for all $c \in C$, in ms_C^X uses similar arguments. \square

Using the previous lemma, we show that the ms-coloring of a poset $P \in Seq$ can be encoded without changing the ms-coloring of its sequentially maximal strict sequential factors.

Lemma 62. *Let $P \in Seq$ and ms_C^X some subsets of P . If the following conditions are true:*

- $\mathcal{C}_{ms_C^X}(P) = \emptyset$;
- for all $P' \in F_{ms}(P) \setminus \{P\}$, P' is completely ms-colored by ms_C^X ;
- for all $c \in C$, all the elements of X_c^p (resp. X_c^s) are p-bound (resp. s-bound) by c for P in ms_C^X ,

then for all $e \in C$, there exist $z \in P \cup P^3$, $ms_C^{X'} = ms_C^X +_e z$, such that:

- for all $P' \in F_{ms}(P) \setminus \{P\}$, $x \in P^3$ and $c \in C$, x is bound by c in ms_C^X for P' if and only if x is bound by c in $ms_C^{X'}$ for P' ;
- z ms-colors P in e .

Proof. Assume we want to ms-color P in $e \in C$. If $P = \{x\}$ then each set of ms_C^X is empty. Define $ms_C^{X'}$ by all sets empty except $X_e^w = \{x\}$. In this case, $\mathcal{C}_{ms_C^{X'}}(P) = \{x\}$. If $|P| > 1$, assume that its irreducible sequential factorization is $P = \sum_{j \in J} P_j$, for some $J \in \mathcal{S} \setminus \{0, 1\}$. If there exists $j \in J$ such that $P_j = \{x\}$ then it suffices to set $ms_C^{X'} = ms_C^X +_e x$ to reach the lemma. Assume now that the irreducible sequential factorization of P is $P = \sum_{j \in J} P_j$, for some $J \in \mathcal{S} \setminus \{0, 1\}$ and each P_j is a parallel poset. Let $j, j' \in J$ such that $j \neq j'$. Then $P_j = P_{j,1} \cup P_{j,2} \cup P_{j,3}$ and $P_{j'} = P_{j',1} \cup P_{j',2} \cup P_{j',3}$ such that $P_{j,1}, P_{j,2}, P_{j',1}, P_{j',2} \in DF_{ms}(P)$. By hypothesis, each of $P_{j,1}, P_{j,2}, P_{j',1}, P_{j',2}$ are completely ms-colored by ms_C^X . In addition, all the elements of $P_{k,t} \cap X_c^p$ and $P_{k,t} \cap X_c^s$ are respectively p-bound and s-bound by c for $P_{k,t}$ in ms_C^X , $k \in \{j, j'\}$, $t \in [2]$, for all $c \in C$. By Lemma 61 there exist $y_{j,1} \in P_{j,1}$ s-free, $x_{j',1} \in P_{j',1}$ p-free and $x_{j',2} \in P_{j',2}$ p-free for respectively $(P_{j,1}, e)$, $(P_{j',1}, e)$ and $(P_{j',2}, e)$ in ms_C^X . In addition, one can prove, by contradiction and using the same arguments of Lemma 61's proof, that $y_{j,1}, x_{j',1}$ and $x_{j',2}$ are still respectively s-free, p-free and p-free for (P, e) in ms_C^X . So, by setting $ms_C^{X'} = ms_C^X +_e (y_{j,1}, x_{j',1}, x_{j',2})$, then $(y_{j,1}, x_{j',1}, x_{j',2})$ is directly bound by e for P in $ms_C^{X'}$ and $(y_{j,1}, x_{j',1}, x_{j',2})$ ms-colors P in e . Furthermore, since $y_{j,1}, x_{j',1}$ and $x_{j',2}$ are respectively s-free, p-free, p-free for (P, e) in ms_C^X then for all $P' \in F_{ms}(P) \setminus \{P\}$, $x \in P^3$ and $c \in C$, x is bound by c in $ms_C^{X'}$ for P' if and only if x is bound by c in ms_C^X for P' . \square

We are now ready to prove Proposition 60.

Proof of Proposition 60. The case of the empty poset is trivial, so assume $P \in SP^{\diamond+}$. We build, by induction on $r_X(P)$, a ms-coloring $ms_C^X = (X_c^w, X_c^s, X_c^p)_{c \in C}$ encoding \mathbf{c} , making sure that for all $c \in C$, X_c^p and X_c^s contain only respectively p-bound and s-bound elements for P in ms_C^X . We start with all sets of ms_C^X empty.

The first case is when $r_X(P) = (0, 0)$. In this case, P is the singleton $\{x\}$. Then, set $X_e^w = \{x\}$, where $e = \mathbf{c}(P)$, and all the other sets composing ms_C^X to \emptyset and this is done.

Otherwise, assume that $P \in Seq$ and its irreducible sequential factorization is $\sum_{j \in J} P_j$, for some $J \in \mathcal{S} \setminus \{0, 1\}$. For all $j \in J$, let $P_j = \bigcup_{i \in [k_j]} P_{j,i}$ be the irreducible parallel factorisation of P_j . Note that each $r_X(P_{j,i}) < r_X(P)$ for all $P_{j,i}$. Then by induction hypothesis, there exists $ms_C^{X_{j,i}}$ satisfying the hypothesis of the proposition for each $P_{j,i} \in DF_{ms}(P)$. Let $ms_C^{X'}$ be the union, set by set, of the different $ms_C^{X_{j,i}}$, for all $j \in J, i \in [k_j], k_j > 1$. By construction of $ms_C^{X'}$, for all $c \in C$, each element x of X_c^p (resp. X_c^s) is p-bound (resp. s-bound) by c for some element $F_x \in F_{ms}(P) \setminus \{P\}$ in $ms_C^{X'}$. This ensures that $C_{ms_C^{X'}}(P) = \emptyset$ and $ms_C^{X'}(F) = c(F)$ for all $F \in F_{ms}(P) \setminus \{P\}$. It suffices to apply Lemma 62 to conclude.

Finally, assume that P is of the form $\bigcup_{j \in [k]} P_j$ for some $k > 1$ where each $P_j \in DF_{ms}(P)$. Since each $P_j \in Seq$, building a ms-coloring $ms_C^{X_j}$ satisfying the hypothesis of the proposition for each P_j can be achieved using the same arguments that those used in the previous paragraph. We define ms_C^X as the union, set by set, of the different $ms_C^{X_j}$ s, for all $j \in [k]$. \square

We proved that any ms-coloring \mathfrak{c} of P with C can be encoded by some ms_C^X as above. Furthermore, we claim that there exist MSO formulæ:

- $ms_C^X(F) = c$ which is satisfied if and only if $ms_C^X(F) = c$ (cf. Definition 57), assuming $F \in F_{ms}(P)$. It expresses that
 - if F has at least one element comparable to all the others, then there is a unique $(e, x) \in C \times F$ such that x is comparable to all the other elements of F , and $x \in X_e^w$. If such (e, x) exists then $e = c$;
 - or F has no element comparable to all the others, and there is a unique $(e, x) \in C \times F^3$ such that x is directly bound by e for F in ms_C^X (cf. Definition 55). If such (e, x) exists then $e = c$.
- $ms\text{-Coloring}(P, ms_C^X)$ which is satisfied if and only if ms_C^X is a complete ms-coloring of P . It expresses that for each $F \in F_{ms}(P)$ there exists $c \in C$ such that $ms_C^X(F) = c$.

These formulæ are obtained by a direct translation into MSO of Definitions 55–57.

6.2. Encoding a compatible s-coloring with MSO

We encode a compatible s-coloring by means of MSO with a set X_c^v for each $c \in C$, and a complete ms-coloring of P with 2^C encoded by some $ms_{2^C}^X$. For short we let s_C^X denote the sets $((X_c^v)_{c \in C}, ms_{2^C}^X)$.

Definition 63. *Let $P \in SP^{\circ+}$, $F \in F_s(P)$ of sequentially irreducible factorisation $F = \sum_{j \in J} F_j$, X_c^v a subset of P for each $c \in C$, and $ms_{2^C}^X$ a complete ms-coloring of P . Then s_C^X s-colors F in $c \in C$, denoted $s_C^X(F) = c$, if for all $j \in J$,*

1. $F_j \in X_c^v$ when $|F_j| = 1$;
2. when $|F_j| > 1$, $c \in ms_{2^C}^X(F_{j,i})$ for all $F_{j,i}$ such that $F_j = \bigcup_{i \in [n_j]} F_{j,i}$ is the irreducible parallel factorisation of F_j ;
3. there is no F', F'' not both empty such that $F' + F + F'' \in F_s(P)$ and $s_C^X(F' + F + F'') = c$.

Finally, when $P \in SP^\diamond$ and $\mathfrak{c}: F_s(P) \rightarrow C$ is a s-coloring of P with C then the subsets s_C^X of P encode \mathfrak{c} if $s_C^X(F) = \mathfrak{c}(F)$ for all $F \in F_s(P)$.

As a consequence of Condition 3 of Definition 63, observe that when F and F' are two sequential factors of P such that FF' is also a sequential factor of P , there is no s_C^X encoding a s-coloring such that $s_C^X(F) = s_C^X(F')$.

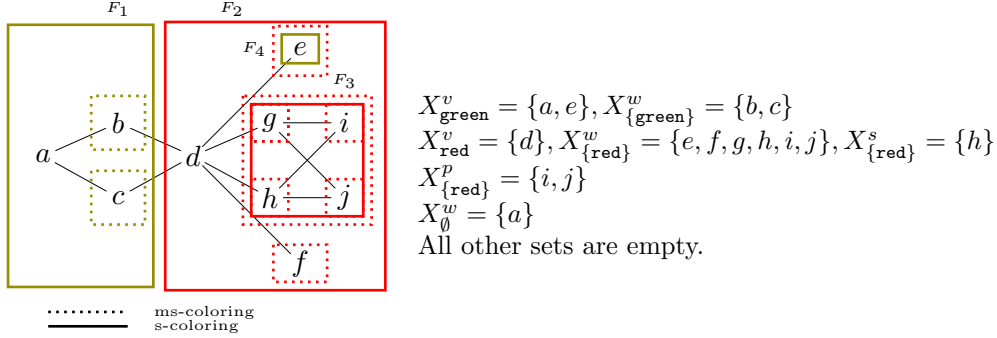


Figure 7: The Hasse diagram of a poset $P \in SP^\circ$ and a s-coloring s_C^X of P .

Example 64. Let $C = \{\text{red}, \text{green}\}$ be a set of colors and $P = a + (b \cup c) + d + (e \cup ((g \cup h) + (i \cup j)) \cup f)$. Let $F_1 = a + (b \cup c)$, $F_2 = P \setminus F_1$, $F_3 = (g \cup h) + (i \cup j)$ and $F_4 = e$. Then $F_1, F_2, F_3, F_4 \in F_s(P)$. Note that $F_3 \subsetneq F_2$ and that there exists $x \in F_2 \setminus F_3$ such that x is incomparable to any $y \in F_3$. Let also $\mathbf{c}: F_s(P) \rightarrow C$ be the s-coloring defined by $\mathbf{c}(F_1) = \mathbf{c}(F_4) = \text{green}$ and $\mathbf{c}(F_2) = \mathbf{c}(F_3) = \text{red}$. Note that \mathbf{c} is compatible. Figure 7 represents P with an encoding s_C^X of \mathbf{c} .

The set $F_{ms}(P)$ consists of the singleton posets b, c, e, g, h, i, j, f , of F_3 and of P . We have $ms_{2C}^X(b) = ms_{2C}^X(c) = \{\text{green}\}$ and $ms_{2C}^X(e) = ms_{2C}^X(g) = ms_{2C}^X(h) = ms_{2C}^X(i) = ms_{2C}^X(j) = ms_{2C}^X(f) = \{\text{red}\}$. We also have $ms_{2C}^X(F_3) = \{\text{red}\}$ since $\mathcal{C}_{ms_{2C}^X}(F_3) = \{(h, i, j)\}$, $(h, i, j) \in X_{\{\text{red}\}}^s \times X_{\{\text{red}\}}^p \times X_{\{\text{red}\}}^p$ and there is no decomposition of F_3 into $F_3 = X + Y + Z$ for some X, Y, Z such that Y is a singleton. As $\mathcal{C}_{ms_{2C}^X}(P) = \{a\}$ and $a \in X_\emptyset^w$ we have $ms_{2C}^X(P) = \emptyset$, and thus P is completely ms-colored by ms_{2C}^X .

We have $s_C^X(F_1) = \text{green}$ since $a \in X_{\text{green}}^v$, $d \notin X_{\text{green}}^v$, $DF_{ms}(F_1) = \{\{b\}, \{c\}\}$, and $\text{green} \in ms_{2C}^X(F)$ for all $F \in DF_{ms}(F_1)$. We have $s_C^X(F_2) = \text{red}$ since $d \in X_{\text{red}}^v$, $\text{red} \notin ms_{2C}^X(b)$, $\text{red} \notin ms_{2C}^X(c)$, $DF_{ms}(F_2) = \{\{f\}, F_3, F_4\}$ and $\text{red} \in ms_{2C}^X(F)$ for all $F \in DF_{ms}(F_2)$. We have $s_C^X(F_3) = \text{red}$ since $DF_{ms}(F_3) = \{\{g\}, \{h\}, \{i\}, \{j\}\}$, $\text{red} \in ms_{2C}^X(F)$ for all $F \in DF_{ms}(F_3)$, and there is no X, Y non both empty such that $X + F_3 + Y$ is a factor of P . Finally $s_C^X(F_4) = \text{green}$ since $e \in X_{\text{green}}^v$ and there is no X, Y non both empty such that $X + e + Y$ is a factor of P .

Proposition 65. Let $P \in SP^\circ$, C be a non-empty finite set of colors and $\mathbf{c}: F_s(P) \rightarrow C$ a compatible s-coloring of P . There exist some subsets s_C^X of P that encode \mathbf{c} .

Proof. From \mathbf{c} define a ms-coloring $\mathbf{c}': F_{ms}(P) \rightarrow 2^C$ as follows: for all $F \in F_{ms}(P)$,

$$\mathbf{c}'(F) = \{c : F \in DF_{ms}(F') \text{ for some } F' \in F_s(P) \text{ such that } \mathbf{c}(F') = c\}$$

According to Proposition 60, \mathbf{c}' is encoded by some ms_{2C}^X . In order to define an encoding $s_C^X = ((X_c^v)_{c \in C}, ms_{2C}^X)$ of \mathbf{c} it suffices now to define X_c^v as follows, for each $c \in C$: $x \in X_c^v$ if and only if there exists $F \in F_s(P)$ such that $\mathbf{c}(F) = c$ and x is comparable to all the elements of $F \setminus \{x\}$. We verify that $s_C^X(F) = \mathbf{c}(F)$ for all $F \in F_s(P)$. Assume first that $\mathbf{c}(F) = c$ for some $F \in F_s(P)$ and $c \in C$, and let $F = \sum_{j \in J} F_j$ be its irreducible sequential factorisation. Let $j \in J$ such that F_j consists of a unique element x , which is thus comparable to all the elements of $F \setminus \{x\}$. By definition $x \in X_c^v$. Thus Item 1 of Definition 63 is verified. Similar arguments apply when F_j is not a singleton: Item 2 of Definition 63 is also verified. Since there is no $F' \in F_s(P)$ such that $\mathbf{c}(F') = c$ and FF' or $F'F \in F_s(P)$, and there is no X, Y not both empty such that $\mathbf{c}(XFY) = c$, then Item 3 of Definition 63 is verified, and thus $s_C^X(F) = c$. Similar arguments are used to show that if $s_C^X(F) = c$ then $\mathbf{c}(F) = c$. \square

We proved that any compatible s-coloring \mathbf{c} of P with C can be encoded by some s_C^X as above. Furthermore, we claim that there exist MSO formulæ:

- $s_C^X(F) = c$ which is satisfied if and only if s_C^X s-colors F in $c \in C$, assuming $F \in F_s(P)$. It is obtained by a direct translation into MSO of Definition 63;
- $\mathbf{s}\text{-Coloring}(P, s_C^X)$ which is satisfied if and only if $\mathbf{ms}\text{-Coloring}(P, ms_{2^C}^X)$ is true and for each non-empty parallel factor F of P there exists $c \in 2^C$ such that $ms_{2^C}^X(F') = c$, for all $F' \in DF_{ms}(F)$.

6.3. Compatible s-colorings and D-graphs

s-coloring is a key argument in the transformation of the D-graph of a rational expression into a P-MSO formula (Section 7). We use s-coloring in order to associate a special edge to some $F \in F_s(P)$ as follows.

Proposition 66. *Let D_e be the D-graph of a rational expression e , $C = \mathbb{B} \times E_S(D_e)$, n a node of D_e . Assume that there is a path T_P from n labeled by some $P \in SP^\circ(A)$. There exist a compatible s-coloring of P with C and its encoding s_C^X with MSO such that $\pi_2(s_C^X(F)) = f \in E_S(D_e)$ if and only if F is marked by f in T_P , for any $F \in F_s(P)$.*

The encoding for s-colorings does not allow $s_C^X(F) = s_C^X(F')$ when $F, F', FF' \in F_s(P)$. Alternation of the booleans in C is used when we need a s-coloring to associate the same special edge to F and F' .

Proof. Let $\mathbf{c}: F_s(P) \rightarrow C$ be a s-coloring of P such that, for all $F \in F_s(P)$, $\mathbf{c}(F) = (b, f)$ if and only if the following conditions are verified:

- F is marked by f in T_P ;
- for all $F' \in F_s(P)$, if $FF' \in F_s(P)$ and $\mathbf{c}(F') = (b', f)$ then $b \neq b'$.

Obviously \mathbf{c} verifies the specification of the proposition (see also Remark 49), and the existence of s_C^X comes from Proposition 65. \square

Example 67. *Let us continue Example 52. Denote by b_l, c_l, d_l (resp. b_r, c_r, d_r) the leftmost (resp. rightmost, cf. Figure 5) elements of P labeled by b, c and d . Let $C = \mathbb{B} \times E_S(D_e)$. By Proposition 66 there exists a compatible s-coloring and its encoding $s_C^X = ((X_c^v)_{c \in C}, ms_{2^C}^X)$ such that $s_C^X(F_1) = (b_1, n_4 \rightarrow n_7)$, $s_C^X(F_2) = (b_2, n_4 \rightarrow n_7)$, $s_C^X(F_3) = (b_3, n_4 \rightarrow n_3)$, $s_C^X(F_4) = (b_4, n_4 \rightarrow n_2)$ for some $b_1, b_2, b_3, b_4 \in \mathbb{B}$, and $s_C^X(F)$ is undefined when $F \neq F_1, F_2, F_3, F_4$. Note that b_1 must be different from b_2 since $F_1 F_2 \in F_s(P)$ and $\pi_2(s_C^X(F_1)) = \pi_2(s_C^X(F_2)) = n_4 \rightarrow n_7$. Following the techniques of Section 6 we have*

- $X_{(b_1, n_4 \rightarrow n_7)}^v = \{b_l\}$;
- $X_{(b_2, n_4 \rightarrow n_7)}^v = \{b_r\}$;
- $X_{(b_3, n_4 \rightarrow n_3)}^v = \{b_l, b_r\}$;
- $X_{(b_4, n_4 \rightarrow n_2)}^v = F_4$;
- $X_c^v = \emptyset$ for all $c \in C \setminus \{(b_1, n_4 \rightarrow n_7), (b_2, n_4 \rightarrow n_7), (b_3, n_4 \rightarrow n_3), (b_4, n_4 \rightarrow n_2)\}$;
- $ms_{2^C}^X(c_l) = ms_{2^C}^X(d_l) = \{(b_3, n_4 \rightarrow n_3), (b_1, n_4 \rightarrow n_7)\}$;
- $ms_{2^C}^X(c_r) = ms_{2^C}^X(d_r) = \{(b_3, n_4 \rightarrow n_3), (b_2, n_4 \rightarrow n_7)\}$;
- $ms_{2^C}^X(F) = \emptyset$ when $F \in F_{ms}(P) \setminus \{c_l, d_l, c_r, d_r\}$.

Proposition 66 states that the s-coloring \mathbf{c} of P issued from marks in a path can be encoded with MSO using some s_C^X .

7. From D-graphs to P-MSO

Let e be a rational expression, D_e the D-graph of its >1 -expression and $C = \mathbb{B} \times E_S(D_e)$. We are going to parse D_e in order to compute a P-MSO sentence ϕ_{D_e} such that P is a model for ϕ_{D_e} if and only if $P \in L(D_e)$, for all $P \in SP^\circ(A)$. For each node n of D_e , we define a P-MSO formula ϕ_n accordingly to the definition of a path from n . The formula ϕ_n depends on a second-order parameter X , and we want ϕ_n to be satisfied if and only if X is a factor of P labeling a path T_X from n . The formula $\phi_n(X)$ depends on the label of n and its edging out(n) = $n \rightarrow n_1, \dots, n \rightarrow n_k$. If n is labeled by some letter $a \in A$, then $\phi_n(X) \equiv |X| = 1 \wedge \forall x (x \in X \rightarrow a(x))$. In the case where n is labeled by $\cdot >^1$ then n has two **childs** n_1 and n_2 , and $\phi_n(X)$ expresses that there exists a partition of X into non-empty X_1, X_2 such that $X_1 < X_2$ and $X_i, i \in [2]$, satisfies $\phi_{n_i}(X_i)$. This construction for $\cdot >^1$ is P-MSO definable. The cases of other labels are mere adaptations of the case of linear orderings [5], except for Presburger formulæ. Indeed, recall that by construction, D_e without its special edges has a structure of acyclic graph (Property DAG), and that only nodes labeled in \mathcal{P} may be sources of special edges, which may cause circular dependencies between the ϕ_n s. Recall also that a sub-path T_R is started by a special edge e if and only if its label R is marked by e in T_R (Definition 45), and that R is necessarily a sequential poset because of Property SS. Circular dependencies between the ϕ_n s are avoided as follows. We encode by means of MSO the marking of labels of sub-paths using s-coloring: we assume R to be the label of a sub-path marked by some special edge e if and only if R is s-colored by (b, e) for some boolean b . Proposition 66 states that this can always be done. When n is the source of a special edge $n \rightarrow m$, instead of making ϕ_n dependant of ϕ_m for testing that a sequential factor $F \in L(m)$, we make ϕ_n dependant of a formula that checks if F is s-colored by $(b, n \rightarrow m)$ for some boolean b . This supposes that it is known that every sequential factor F s-colored by $(b, n \rightarrow m)$ satisfies ϕ_m .

Formally, set $C = \mathbb{B} \times E_S(D_e)$. When the label of a node n is some $\rho(x_1, \dots, x_k)$ and out(n) = $n \rightarrow n_1, \dots, n \rightarrow n_k$, set

$$\phi_n(X) \equiv \mathcal{Q}(X, \chi_1, \dots, \chi_k, \rho(x_1, \dots, x_k))$$

where

$$\chi_i \equiv \begin{cases} \forall Y (\forall y y \in Y) \rightarrow \phi_{n_i}(Y) & \text{when } n \rightarrow n_i \in E_N(D_e); \\ \forall Y (\forall y y \in Y) \rightarrow \bigvee_{b \in \mathbb{B}} s_C^X(Y) = (b, n \rightarrow n_i) & \text{when } n \rightarrow n_i \in E_S(D_e). \end{cases}$$

Informally speaking, the sentence ϕ_{D_e} encodes that $P \in L(D_e)$ if and only if there is a path T_P in D_e from $r(D_e)$ and labeled by P ; in order to avoid circular dependencies it uses an encoding s_C^X of a compatible s-coloring of P such that $\pi_2(s_C^X(F)) = f$ if and only if F is marked by f in T_P , for any sequential factor F of P . It guarantees that when F is marked by $f = n \rightarrow m$ then F satisfies $\phi_m(F)$. Formally

$$\begin{aligned} \phi_{D_e} \equiv & \exists R \exists s_C^X (\forall x x \in R) \wedge \mathbf{s}\text{-Coloring}(R, s_C^X) \wedge \phi_{r(D_e)}(R) \\ & \wedge \left(\bigwedge_{n \rightarrow m \in E_S(D_e)} \forall F (\mathbf{F}_s(F, R) \wedge \bigvee_{b \in \mathbb{B}} s_C^X(F) = (b, n \rightarrow m)) \rightarrow \phi_m(F) \right) \\ & \bigvee_{e \in L(e)} \forall X (\forall x x \in X) \rightarrow |X| = 0 \end{aligned}$$

Example 68. Let $e = a \circ_\xi (a(\xi \parallel \xi))^{\ast \xi}$ be the rational expression e_1 of Example 51. In this example we detail the construction of the P-MSO formula ϕ_{D_e} from D_e , using the techniques of Section 7. The D-graph D_e and the poset $P = a(a \parallel a(a \parallel a \parallel a)) \parallel a(a \parallel a)$ of $L(D_e)$ are represented on the left side of Figure 8. Let n_1, \dots, n_6 be the pre-order traversal of D_e without its special edges. We have $E_S(D_e) = \{n_4 \rightarrow n_2\}$, $C = \mathbb{B} \times E_S(D_e)$ and

- $\phi_{n_1}(X) \equiv \mathcal{Q}(X, \overline{\phi_{n_2}(Y)}, \overline{\phi_{n_6}(Y)}, x_1 + x_2 = 1)$;
- $\phi_{n_2}(X) \equiv \exists X_1, X_2, X = X_1 + X_2 \wedge \phi_{n_3}(X_1) \wedge \phi_{n_4}(X_2)$;
- $\phi_{n_3}(X) \equiv \phi_{n_5}(X) \equiv \phi_{n_6}(X) \equiv |X| = 1 \wedge \forall x (x \in X \rightarrow a(x))$;
- $\phi_{n_4}(X) \equiv \mathcal{Q}(X, \overline{\bigvee_{b \in \mathbb{B}} s_C^X(Y) = (b, n_4 \rightarrow n_2)}, \overline{\phi_{n_5}(Y)}, x_1 + x_2 = 2)$

where $\overline{\psi(Y)} \equiv \forall Y (\forall y y \in Y) \rightarrow \psi(Y)$. Then

$$\begin{aligned} \phi_{D_e} &\equiv \exists R \exists s_C^X (\forall x x \in R) \wedge \mathbf{s}\text{-Coloring}(R, s_C^X) \wedge \phi_{n_1}(R) \\ &\wedge (\forall F (F_s(F, R) \wedge \bigvee_{b \in \mathbb{B}} s_C^X(F) = (b, n_4 \rightarrow n_2)) \rightarrow \phi_{n_2}(F)) \end{aligned}$$

On the right side of Figure 8 is pictured a path T_P in D_e from n_1 labeled by P . The root of T_P is labeled by $(n_1, (1, 0))$ with $(1, 0) \in L(x_1 + x_2 = 1)$. Then T_P has a unique direct sub-path T'_P from n_2 also labeled by P . Thus $P \in L(n_2)$. Observe that $T'_P = (m, T_{P_1}, T_{P_2})$ where m is labeled by $(n_2, 2)$, and T_{P_1} and T_{P_2} are some paths in D_e from respectively n_3 and n_4 labeled respectively by $P_1 = a$ and $P_2 = a \parallel a(a \parallel a(a \parallel a(a \parallel a))) \parallel a(a \parallel a)$. Hence, $P = P_1 P_2$. Let $F_i, i \in [4]$, be the labels of the sub-paths $T_{F_i}, i \in [4]$, of T_P started by the special edge $n_4 \rightarrow n_2$, as in Figure 8. A factor of P is marked by $n_4 \rightarrow n_2$ in T_P if and only if it is one of the F_i s. Therefore, observe that there is some $x \in F_i \setminus F_j$ which is incomparable to all the elements of F_j when $F_j \subsetneq F_i$, as mentioned in Item (C₂) of Remark 49. By Proposition 66, this marking can be expressed by means of MSO by some encoding s_C^X of a compatible s -coloring. A factor of P marked in T_P satisfies $\bigvee_{b \in \mathbb{B}} s_C^X(Y) = (b, n_4 \rightarrow n_2)$, for some $b \in \mathbb{B}$, if and only if it is one of the F_i s. The maximum elements of P_2 satisfy ϕ_{n_5} , P_1 satisfies ϕ_{n_3} and P_2 satisfies ϕ_{n_4} . Thus P satisfies ϕ_{n_2} as well as ϕ_{n_1} and is a model for ϕ_{D_e} .

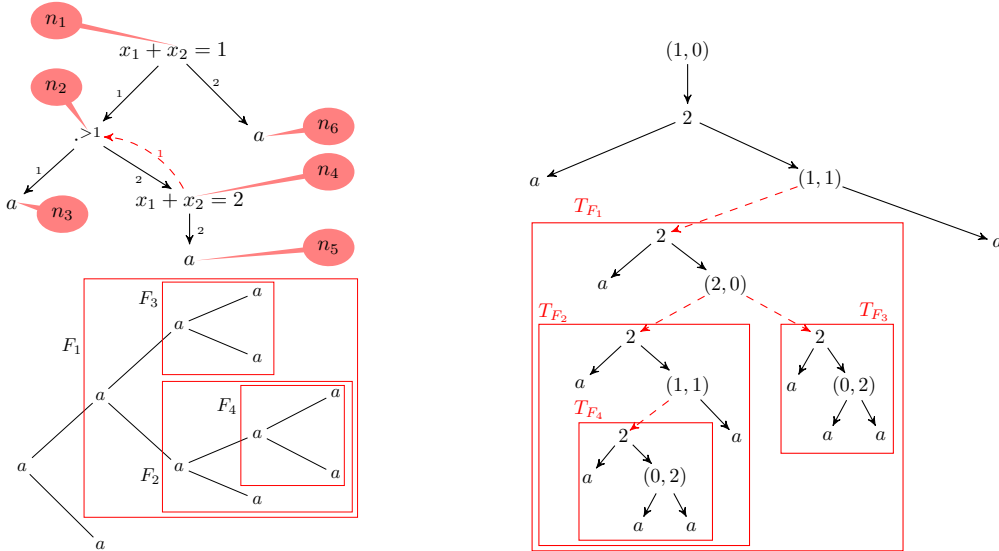


Figure 8: The D-graph D_e of the >1 -expression of $e = a \circ_\xi (a(\xi \parallel \xi))^{\ast\xi}$, a poset $P = a(a \parallel a(a \parallel a(a \parallel a))) \parallel a(a \parallel a)$ of $L(e)$ and a path in D_e from n_1 labeled by P

8. From P-MSO to rational expressions

The transformation of a P-MSO formula φ into a rational expression e such that $L(\varphi) = L(e)$ is based on several known results and uses only well-known techniques. Thus we give here only the main arguments, and refer the reader to the bibliography for the details. Automata over posets of $SP^\circ(A)$ effectively equivalent to rational expressions were introduced in [6]. Melting the techniques of the translation of a MSO formula to an automaton over countable and scattered linear orderings [5], and those of the translation of a P-MSO formula to an automaton over finite N-free posets [3], we build by induction on φ a branching automaton \mathcal{A}_φ such that $L(\mathcal{A}_\varphi) = L(\varphi)$. The only inductive step which can not be directly deduced from [5] or [3] is the construction of $\mathcal{A}_{\neg\varphi'}$ from $\mathcal{A}_{\varphi'}$. This case is a direct consequence of Theorem 18.

9. Conclusion

Links between automata theory and formal logic have been initiated by Büchi in his earlier works, in particular in order to use automata as a basis for decision procedures for MSO. The results presented in this paper have the same use. The construction of a branching automaton \mathcal{A}_φ from a formula φ of P-MSO such that $L(\mathcal{A}_\varphi) = L(\varphi)$ is effective. Furthermore, it is decidable whether $L(\mathcal{A}_\varphi) = \emptyset$ or not. As a consequence:

Theorem 69. *Let A be an alphabet. The P-MSO theory of $SP^\circ(A)$ is decidable.*

Since the pioneer works of Büchi, links between automata and logic have been used in many ways, such as the characterization of fragments of MSO and classification of languages of words, trees and other structures. Our work is a step in that direction for languages of transfinite posets. **It could be continued for example by a study of the relative expressive power of logics weaker than P-MSO. This is usually done, for logic over finite structures (such as finite words), by a change of the predicates, or by restricting the number of variables or the alternation of quantifiers. For linear orderings, it is also interesting to restrict over the interpretation of second-order variables, for example, over their cardinality or orderings. This has been done in [13] when sets are restricted to singletons (see also [7, 8]), ordinals, or scattered linear orderings, with algebraic Schützenberger-like characterizations of such logics.**

Acknowledgements

We would like to thank the anonymous referees for their comments on this work.

References

- [1] A. Amrane and N. Bedon. Logic and rational languages of scattered and countable series-parallel posets. In *Proceedings of LATA'19*, volume 11417 of *Lect. Notes in Comput. Sci.* Springer, 2019.
- [2] N. Bedon. Logic and bounded-width rational languages of posets over countable scattered linear orderings. In Sergei Artemov and Anil Nerode, editors, *Logical Foundations of Computer Science*, volume 5407 of *Lect. Notes in Comput. Sci.*, pages 61–75. Springer-Verlag, January 2009.
- [3] N. Bedon. Logic and branching automata. *Logical Meth. in Comput. Sci.*, 11(4:2):1–38, 2015.
- [4] N. Bedon. Complementation of branching automata for scattered and countable N-free posets. *Int. J. Found. Comput. Sci.*, 19(25):769–799, 2018.
- [5] N. Bedon, A. Bès, O. Carton, and C. Rispal. Logic and rational languages of words indexed by linear orderings. *Theory of Comp. Syst.*, 46(4):737–760, May 2010.
- [6] N. Bedon and C. Rispal. Series-parallel languages on scattered and countable posets. *Theoret. Comput. Sci.*, 412(22):2356–2369, 2011.
- [7] Nicolas Bedon and Chloé Rispal. Schützenberger and Eilenberg theorems for words on linear orderings. *Journal of Computer and System Sciences*, 78(2):517–536, March 2012.
- [8] Alexis Bès and Olivier Carton. Algebraic characterization of FO for scattered linear orderings. In *Proc. of CSL'11, LIPIcs*, volume 12, pages 67–81, 2011.
- [9] V. Bruyère and O. Carton. Automata on linear orderings. *J. Comput. System Sci.*, 73(1):1–24, 2007.
- [10] J. R. Büchi. Weak second-order arithmetic and finite automata. *Zeit. Math. Logik. Grund. Math.*, 6:66–92, 1960.
- [11] J. R. Büchi. On a decision method in the restricted second-order arithmetic. In *Proc. Intern. Congr. on Logic, Methodology and Philosophy of Science, Berkeley 1960*, pages 1–11. Stanford University Press, 1962.
- [12] J. R. Büchi. Transfinite automata recursions and weak second order theory of ordinals. In *Proc. Int. Congress Logic, Methodology, and Philosophy of Science (1964)*, pages 2–23. North-Holland Publ. Co., 1965.
- [13] Thomas Colcombet and A. V. Sreejith. Limited set quantifiers over countable linear orderings. In Magnús M. Halldórsson, Kazuo Iwama, Naoki Kobayashi, and Bettina Speckmann, editors, *Automata, Languages, and Programming - 42nd International Colloquium, ICALP 2015, Kyoto, Japan, July 6-10, 2015, Proceedings, Part II*, volume 9135 of *Lecture Notes in Computer Science*, pages 146–158. Springer, 2015.
- [14] J. Doner. Tree acceptors and some of their applications. *Journal of Computer and System Sciences*, 4(5):406–451, 1970.
- [15] S. Eilenberg and M-P. Schützenberger. Rational sets in commutative monoids. *J. of Algebra*, 13(2):173–191, 1969.
- [16] C. C. Elgot. Decision problems of finite automata design and related arithmetics. *Trans. Amer. Math. Soc.*, 98:21–51, 1961.
- [17] S. Ginsburg and E. H. Spanier. Semigroups, Presburger formulas, and languages. *Pacific J. of Mathematics*, 16(2):285–296, 1966.
- [18] J. Grabowski. On partial languages. *Fundam. Inform.*, 4(1):427–498, 1981.
- [19] F. Hausdorff. Grundzüge einer theorie der geordneten mengen. *Mathematische Annalen*, 65(4):435–505, 1908.

- [20] S. C. Kleene. Representation of events in nerve nets and finite automata. In *Automata studies*, pages 3–41. Princeton University Press, 1956.
- [21] D. Kuske. Infinite series-parallel posets: logic and languages. In *ICALP 2000*, volume 1853 of *Lect. Notes in Comput. Sci.*, pages 648–662. Springer-Verlag, 2000.
- [22] D. Kuske. Towards a language theory for infinite N-free pomsets. *Theoret. Comput. Sci.*, 299:347–386, 2003.
- [23] K. Lodaya and P. Weil. A Kleene iteration for parallelism. In *Foundations of Software Technology and Theoretical Computer Science*, pages 355–366, 1998.
- [24] K. Lodaya and P. Weil. Series-parallel posets: algebra, automata and languages. In *STACS’98*, volume 1373 of *Lect. Notes in Comput. Sci.*, pages 555–565. Springer-Verlag, 1998.
- [25] K. Lodaya and P. Weil. Series-parallel languages and the bounded-width property. *Theoret. Comput. Sci.*, 237(1–2):347–380, 2000.
- [26] K. Lodaya and P. Weil. Rationality in algebras with a series operation. *Inform. Comput.*, 171:269–293, 2001.
- [27] R.J. Parikh. On context-free languages. *J. ACM*, 13(4):570–581, October 1966.
- [28] M. Presburger. Über die vollständigkeit eines gewissen systems der arithmetik ganzer zahlen, in welchem die addition als einzige operation hervortritt. In *Proc. Sprawozdaniez I Kongresu Matematykw Krajow Slowiaskich, Warsaw*, pages 92–101, 1930. English translation: On the completeness of certain system of arithmetic of whole numbers in which addition occurs as the only operation. *Hist. Philos. Logic*, 12:92–101, 1991.
- [29] M. O. Rabin. Decidability of second-order theories and automata on infinite trees. *Trans. Amer. Math. Soc.*, 141:1–5, 1969.
- [30] J. G. Rosenstein. *Linear Orderings*. Academic Press, 1982.
- [31] J. W. Thatcher and J. B. Wright. Generalized finite automata theory with an application to a decision problem of second-order logic. *Mathematical System Theory*, 2(1):57–81, 1968.
- [32] W. Thomas. Languages, automata, and logic. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, volume III, pages 389–455. Springer-Verlag, 1997.
- [33] B. A. Trakhtenbrot. Finite automata and monadic second order logic. *Siberian Math.*, 3:101–131, 1962. (Russian). Translation in AMS Transl. 59 (1966), 23-55.
- [34] J. Valdes, R. E. Tarjan, and E. L. Lawler. The recognition of series parallel digraphs. *SIAM J. Comput.*, 11:298–313, 1982.
- [35] J. Winkowski. An algebraic approach to concurrence. In *MFCS’79*, volume 74 of *Lect. Notes in Comput. Sci.*, pages 523–532. Springer Verlag, 1979.