Under consideration for publication in Math. Struct. in Comp. Science

Unique decomposition of homogeneous languages and application to isothetic regions

EMMANUEL HAUCOURT[1] and NICOLAS NININ[2]

emmanuel.haucourt@lix.polytechnique.fr nicolas.ninin@epfl.ch

[1] *LIX - UMR 7161, 1 rue Honoré d'Estienne d'Orves, Campus de l'École Polytechnique, 91120, Palaiseau, France*

[2] *Laboratory for Topology and Neuroscience, Brain Mind Institute, EPFL, École Polytechnique Fédérale de Lausanne, Switzerland*

Received 13 July 2021

A language is said to be *homogeneous* when all its words have the same length. Homogeneous languages thus form a monoid under concatenation. It becomes *freely* commutative under the simultaneous actions of every permutation group $\mathfrak{S}_n$ on the collection of homogeneous languages of length $n \in \mathbb{N}$. One recovers the *isothetic regions* from (Haucourt (2017)) by considering the alphabet of connected subsets of the space $|G|$, viz the geometric realization of a *finite* graph G. Factoring the geometric model of a *conservative* program amounts to parallelize it, and there exists an *efficient* factoring algorithm for isothetic regions. Yet, from the theoretical point of view, one wishes to go beyond the class of conservative programs, which implies relaxing the finiteness hypothesis on the graph G. Provided that the collections of n-dimensional isothetic regions over G (denoted by $\mathcal{R}_n|G|$) are co-unital distributive lattices, the prime decomposition of isothetic regions is given by an algorithm which is, unfortunately, *very inefficient*. Nevertheless, if the collections $\mathcal{R}_n|G|$ satisfy the stronger property of being Boolean algebras, then the efficient factoring algorithm is available again. We relate the algebraic properties of the collections $\mathcal{R}_n|G|$ to the geometric properties of the space $|G|$. On the way, the algebraic structure $\mathcal{R}_n|G|$ is proven to be the universal tensor product, in the category of semilattices with zero, of n copies of the algebraic structure $\mathcal{R}_1|G|$.

1. Introduction

1.1. *Motivation from concurrency theory*

Geometric models of concurrent programs explicitly appeared for the first time in (Coffman et al. (1971)) as *progress graphs*. Later (Carson and Reynolds Jr. (1987)) formalized the construction of such models for a fragment of the language introduced by (Dijkstra (1965)). That sublanguage exactly contains every program that consists of a parallel composition of sequential processes without loops nor branchings, the processes being synchronized by means of *semaphores*. Mathematically speaking, the progress graph of a program with n processes is a *cubical region* of dimension n, viz a subset of $\mathbb{R}^n$ that is covered by a finite family of *cubes* (*i.e.* products of intervals). The relevance of cubical regions lies in the unique decomposition property they come with: the decomposition of the model of a program is indeed related to its parallelization (Balabonski and Haucourt (2010)). The geometric models defined so far would be satisfactory were it not for the drastic restriction (no loops nor branchings) on the class of programs they apply to. Several mathematical concepts have been proposed to extend the range of application of geometric models: *higher dimensional automata* (van Glabbeek (1991); Pratt (2000)), *locally ordered spaces* (Fajstrup et al. (2006)), *d-spaces* (Grandis (2003, 2009)), *streams* (Krishnan (2009)). However, these objects were introduced with a view to applying methods inspired from algebraic topology (Fajstrup et al. (2016)). As a consequence, each of them induces a category containing the unit interval (actually an object which stands for it) as well as any object that can be obtained by gluing cubes. In particular, most objects of such a category are so far from the geometric model of any program that they can be regarded as pathological from the computer science point of view. The issue is addressed in (Haucourt (2017)) providing each *conservative* program with a geometric model which lies in a restricted class of locally ordered spaces. The elements of that class are called *isothetic regions* (Definition 2.63). A sequential process is said to be conservative when the amount of resources it holds *only* depends on the point of the control flow graph the instruction pointer stands on. By extension, a conservative program is a parallel composition of conservative processes. The notion of an isothetic region is obtained by replacing the intervals of $\mathbb{R}$ by the connected subsets of the geometric realization of a *finite* graph G in the definition of cubical regions (Definition 2.56). The graph G, sometimes referred to as the *underlying graph* of the isothetic region, is actually to be understood as the disjoint union of the control flow graphs of the processes of the conservative program to model. On one hand, any sequential process without loops is conservative so the class of conservative programs broadly extends that of programs considered in (Carson and Reynolds Jr. (1987)). On the other hand, the class of isothetic regions clearly contains that of cubical ones. To complete the generalization, it remains to check that the unique decomposition theorem for cubical regions remains valid for the isothetic ones: this is the main purpose of this paper. On the way, we provide a factoring algorithm that is much more efficient than the one given in (Balabonski and Haucourt (2010)). We also clarify the algebraic framework on which our unique decomposition result is grounded and exactly determine how far, and at which cost, the finiteness hypothesis on G can be relaxed.

1.2. *A survey of geometric models*

To serve our purpose, we briefly describe the semantics of a toy language and provide some examples to motivate the abstract development to come. More details can be found in (Haucourt (2017)), which is the main source for this section and the next one.

A program P consists of a tuple $(P_1, \ldots, P_n)$ of *sequential processes* running concurrently, and sharing a pool of resources (Haucourt, 2017, Definition 2.2). The *arity* of a resource $\mathtt{a}$ is the number of *occurrences* of $\mathtt{a}$ that are available in the pool at the beginning of any execution of the program. A resource of arity 1 is called a *mutex* for

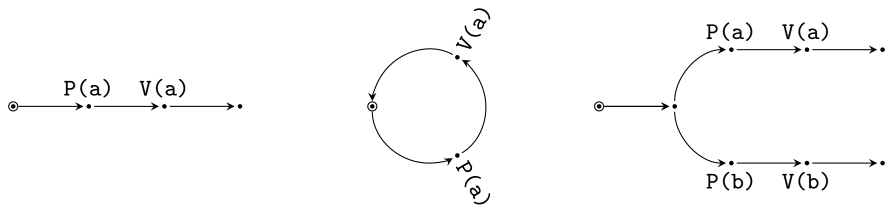

Fig. 1. Examples of processes

mutual exclusion. A process P_i takes an occurrence of the resource `a` by executing the instruction `P(a)`. If no occurrence of `a` is available in the pool of resources, the process is blocked until another process P_j releases one by executing the instruction `V(a)`. However, if P_j does not hold any occurrence of `a`, then the instruction `V(a)` is just ignored and the process P_i remains stalled. In particular, each process P_i is associated to a *wallet* containing the resource occurrences it *owns* at any time of an execution. This should be compared to the toy language described in (Dijkstra (1965)) whose semantics associates each resource with a mere *counter*.

The processes of a program $(P_1, \dots, P_n)$ are actually given as *automata* (Figure 1) so we can consider the tuple $(G_1, \dots, G_n)$ of their underlying directed graphs. Denoting the set of vertices (resp. arrows) of G_i by V_i (resp. A_i), the elements of the set

$$|G_i| \quad = \quad V_i \quad \cup \quad A_i \times]0,1[$$

can be seen as the possible positions of the *instruction pointer* of the process P_i. The instruction pointer of the program $(P_1, \dots, P_n)$ being the n-tuple of instruction pointers of its processes, an *overapproximation* of its valid *positions* is given by the product

$$|G_1| \times \cdots \times |G_n| \ .$$

In order to refine it, we would like to take into account the constraints imposed by the resource limitations. To do so, we assume that the amount of resources held by P_i at a given position *does not depend* on the execution trace that led to it[†]. More precisely, for each process P_i we suppose that we have a *potential function* F_i which provides the number of occurrences of any resource x held by P_i at position p_i. From there, we deduce the potential function F of the program: given a resource x and a point $p = (p_1, \dots, p_n)$ in $|G_1| \times \cdots \times |G_n|$, the value of the function $F(x, \_)$ at p is defined as

$$F(x,p) \quad = \quad \sum_{i=1}^{n} F_i(x, p_i) \ .$$

The point p is said to be *forbidden* when there exists a resource x whose arity is (strictly) less that $F(x,p)$. The *forbidden region* of a program is the set of its forbidden points; its *geometric model* $[\![P]\!]$ is the complement (in $|G_1| \times \cdots \times |G_n|$) of its forbidden region.

[†] By analogy with conservative forces in physics, such a process is said to be *conservative*.

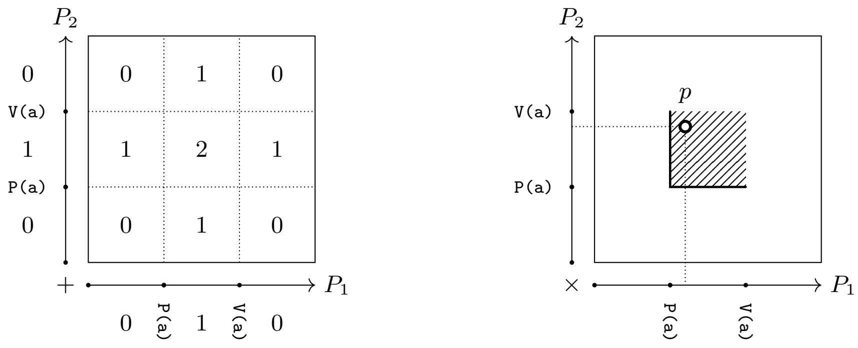

Fig. 2. The potential function of a program and its geometric model

Example 1.1. Suppose that `a` is a mutex, and consider a program (P_1, P_2) made of two copies of the process shown on the left part of Figure 1 — this process corresponds to the sequence of instructions `P(a);V(a)`. In particular, the underlying graphs G_1 and G_2 are equal to

so both sets $|G_1|$ and $|G_2|$ are, up to an obvious bijection, identified with the interval $[0, 3]$. The potential function of the program is shown on the left hand part of Figure 2. Any point p on the dashed square (Figure 2) corresponds to a position where the unique occurrence of the resource `a` is simultaneously held by P_1 and P_2, therefore it is forbidden. So the geometric model of the program under consideration is the complement (in $[0, 3]^2$) of the dashed square, that is to say $\left\{ (x, y) \in [0, 3]^2 \;\middle|\; x < 1 \text{ or } y < 1 \text{ or } x \geqslant 2 \text{ or } y \geqslant 2 \right\}$.

1.3. *Independent programs, geometric models, and homogeneous languages*

The *parallel composition* $P\,|\,Q$ of the programs $(P_1, \ldots, P_n)$ and $(Q_1, \ldots, Q_m)$ is defined as the program $(P_1, \ldots, P_n, Q_1, \ldots, Q_m)$. A resource x is said to be *used* by a process when the label of some of its vertices is the instruction `P(`x`)`. For example, all the processes on Figure 1 use the resource `a` while the resource `b` is only used by the third one. The programs P and Q are said to be *syntactically independent* when the sets of resources they respectively use are disjoint. In that case, one readily checks that P and Q are *model independent* in the sense that the following identity holds:

$$[\![P\,|\,Q]\!] \quad = \quad [\![P]\!] \times [\![Q]\!] \; .$$

The model independence of two programs P and Q is relevant in practice because it implies that they are *observationally independent*† in the sense that the result produced

† See (Haucourt, 2017, Definition 3.6).

by the execution of the parallel composition $P|Q$ only depends on the restrictions of the scheduling to the programs P and Q (Haucourt, 2017, Theorem 6.2).

Remark 1.2. Model independent programs may not be syntactically independent. In Example 1.1, assume that the arity of `a` is 2 instead of 1. The geometric model of the resulting program is $[0,3]^2$ though its processes (considered as programs) are *not* syntactically independent (as they both use the resource `a`).

In that context, we interpret the geometric model $[\![P]\!]$ of a program $(P_1, \ldots, P_n)$ as a *homogeneous*† *language* of length n over some alphabet Ω. In Example 1.1, the set Ω is the interval $[0,3]$: we will soon clarify the way Ω depends on the program to model.

The homogeneous languages over Ω form a (noncommutative) monoid $\mathcal{H}(\Omega)$ under concatenation (Section 2.1). Decomposing an element L of that monoid amounts to finding a partition $(I_1, \ldots, I_k)$ of the set $\{1, \ldots, n\}$ (n being the length of L) such that $L = L|_{I_1} * \cdots * L|_{I_k}$ (Definition 2.30). We can actually suppose that each element I_i of the partition is an interval, and that for $i < j$ we have $I_i < I_j$ (Section 2.3).

Example 1.3. Suppose that `a` and `b` are mutexes. The geometric model of the program `P(a).V(a)|P(b).V(b)|P(a).V(a)` has no non-trivial decomposition though the programs `P(a).V(a)|P(a).V(a)` and `P(b).V(b)` are model independent (since they are syntactically independent).

The issue raised in Example 1.3 is a direct consequence of the noncommutativity of the parallel composition operator we have defined. This defect is prohibitive in concurrency theory because the order in which the processes of a program are listed should neither alter its runtime behaviour nor its semantics. Yet, one readily checks that $[\![\pi \cdot P]\!] = \pi \cdot [\![P]\!]$ holds for any permutation π of the set $\{1, \ldots, n\}$. In other words, the action of the n^{th} symmetric group‡ $\mathfrak{S}_n$ on programs made of n processes is compatible with its action on n-dimensional geometric models (these actions are standard, yet they are recalled in Section 2.1). Since the programs P and $\pi \cdot P$ should be seen as identical, so should be their geometric models. The previous observation leads to introduce the *homogeneous monoid* over Ω, denoted by $\widetilde{\mathcal{H}}(\Omega)$, as the quotient of the monoid $\mathcal{H}(\Omega)$ by the congruence relating L and $\pi \cdot L$ for all $n \in \mathbb{N}$, all $\pi \in \mathfrak{S}_n$, and all $L \in \mathcal{H}(\Omega)$ of length n (Definition 2.4). The monoid $\widetilde{\mathcal{H}}(\Omega)$ is freely commutative (Corollary 2.34). Concretely, decomposing (the equivalence class of) a homogeneous language L amounts to finding a partition $X_1, \ldots, X_k$ of $\{1, \ldots, n\}$ such that

$$\pi \cdot L \quad = \quad (\pi \cdot L)|_{I_1} * \cdots * (\pi \cdot L)|_{I_k} \quad = \quad L|_{X_1} * \cdots * L|_{X_k}$$

where $I_1, \cdots, I_k$ is the unique partition of $\{1, \ldots, n\}$ into *intervals* such that $I_i < I_j$ when $i < j$ and there is a (necessarily unique) permutation π of $\{1, \ldots, n\}$ that induces an increasing bijection from X_i to I_i for all $i \in \{1, \ldots, k\}$ — see Section 2.3.

We now pay some attention to the alphabet Ω. Note that for any superset Ω' of Ω, any decomposition in $\mathcal{H}(\Omega')$ of an element of $\mathcal{H}(\Omega)$ only contains elements of $\mathcal{H}(\Omega)$. Clearly,

† A language whose words share the same length n is said to to be *homogeneous* of length n (Section 2.1).
‡ The n^{th} symmetric group is that of permutations of the set $\{1, \ldots, n\}$.

the same holds for $\tilde{\mathcal{H}}(\Omega)$ and $\tilde{\mathcal{H}}(\Omega')$. So, in view of factoring the geometric model of a program P, we can let Ω be the set $|G_1| \cup \cdots \cup |G_n|$ where $G_1, \ldots, G_n$ are the underlying graphs of the processes of P. Moreover, the graph over which a process is defined can be replaced by any isomorphic graph. So we can suppose that all the graphs G_i belongs to a *finite* family of graphs $\mathcal{G}$ in which two elements are either equal, or not isomorphic and disjoint in the sense that their sets of vertices (resp. arrows) are disjoint. We can even suppose that the elements of $\mathcal{G}$ are connected because any process has a distinguished vertex, called the *entry point*, from which any other vertex can be reached. It is then natural to let Ω be the (disjoint finite) union of the sets $|G|$ for $G \in \mathcal{G}$. For the programs studied in Examples 1.1 and 1.3, the family $\mathcal{G}$ is reduced to a single graph shown below.

$$\odot \longrightarrow \bullet \longrightarrow \bullet \longrightarrow \bullet$$

From the computer science point of view, the parallel decomposition of a program P into model independent subprograms is more interesting than the decomposition of its geometric model $\llbracket P \rrbracket$. Moreover, any decomposition of P readily induces a decomposition of $\llbracket P \rrbracket$. Conversely, for a given decomposition of the geometric model

$$\pi \cdot \llbracket P \rrbracket \quad = \quad \llbracket P \rrbracket|_{X_1} * \cdots * \llbracket P \rrbracket|_{X_k} \tag{D}$$

(where $X_1, \ldots, X_k$ is a partition of $\{1, \ldots, n\}$ and $\pi \in \mathfrak{S}_n$ its related permutation — see Section 2.3) we can ask whether we have $\llbracket P|_{X_i} \rrbracket = \llbracket P \rrbracket|_{X_i}$ for all $i \in \{1, \ldots, k\}$. If the answer is 'yes'†, then the decomposition of $\llbracket P \rrbracket$ lifts to a decomposition of P into model independent subprograms. As a consequence, all the decompositions of a program into model independent subprograms can be obtained from the prime decomposition (Definition 2.13) of its geometric model.

Yet, the benefit is limited in practice because the language we have associated to $\llbracket P \rrbracket$ is infinite. Nevertheless, this issue can be overcome by using a representation of the geometric models that takes advantage of the topology of Ω (Definition 2.56). More precisely, one replaces Ω by the collection $\mathcal{C}$ of all its connected subsets. Then, as we shall see, the geometric model of a program is an *isothetic region* (Definition 2.63), that is to say a *finite* union of $\mathcal{C}$-*blocks* (*i.e.* Cartesian products of elements of $\mathcal{C}$ — Definition 2.39).

Example 1.4. Considering the program `P(a).V(a)|P(a).V(a)` (Example 1.1), we have $\Omega = [0,3]$. It follows that $\mathcal{C}$ is the collection of all subintervals of $[0,3]$ and the geometric model is covered by finitely many rectangles:

$$[0,3] \times [0,1[\quad \cup \quad [0,3] \times [2,3] \quad \cup \quad [0,1[\times [0,3] \quad \cup \quad [2,3] \times [0,3] \ . \tag{C_1}$$

In that context, the geometric model of P can be seen as the homogeneous language

$$\{ \ \Omega I \ , \ \Omega J \ , \ I\Omega \ , \ J\Omega \ \}$$

(where Ω, I, and J respectively denote the intervals $[0,3]$, $[0,1[$, and $[2,3]$ — see Figure 3) to which we apply the generic factoring algorithm (Section 2.4).

† We conjecture that the answer is always 'yes' in the sense that when (D) holds, the geometric model of the subprogram $P|_{X_i}$ is indeed $\llbracket P \rrbracket|_{X_i}$ for every $i \in \{1, \ldots, k\}$.

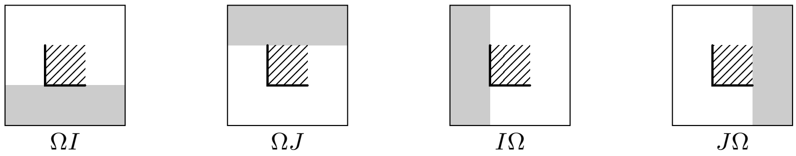

Fig. 3. The maximal rectangles around the square

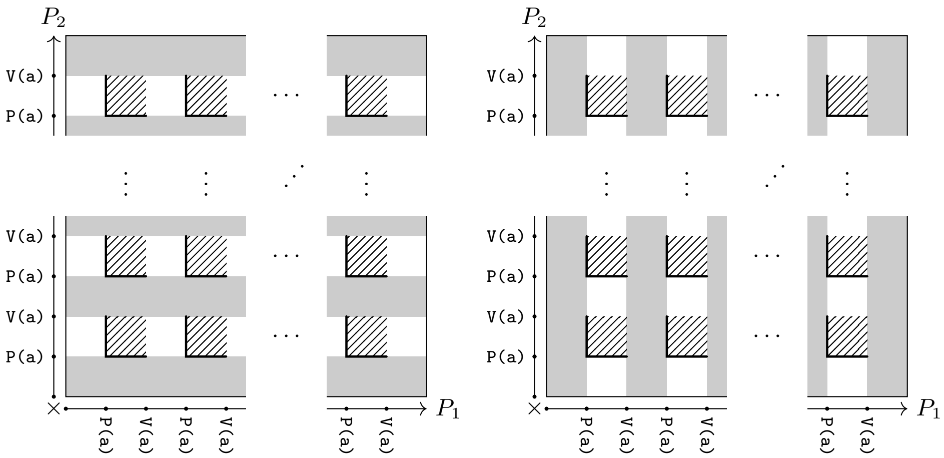

Fig. 4. Many maximal rectangles (*i.e.* $\mathcal{C}$-blocks), but only two maximal $\mathcal{C}'$-blocks

Interpreting the geometric model of a program P as a language over the alphabets Ω and $\mathcal{C}$ result in two distinct homogeneous languages. Nevertheless, their prime decompositions match provided that the $\mathcal{C}$-block *covering* from which the language over $\mathcal{C}$ is built *only* contains *maximal* $\mathcal{C}$-blocks (Definition 2.39) of the geometric model (Corollary 2.48). The covering (C_1) given in Example 1.4 is indeed the collection of all the maximal rectangles of the geometric model.

We can actually go further and replace the collection $\mathcal{C}$ by the collection $\mathcal{C}'$ of all the *finite unions* of elements of $\mathcal{C}$. Since the graphs over which processes are built are finite, the collection $\mathcal{C}'$ forms a Boolean subalgebra of the powerset $\wp(\Omega)$ (Theorem 3.12). Then applying Proposition 2.53, we deduce that Corollary 2.48 also applies to $\mathcal{C}'$-blocks. In other words, the prime decomposition of a geometric model seen as a language over Ω matches its prime decomposition as a language over $\mathcal{C}'$. This shift comes with two remarkable benefits. The first one is that the size of the language associated with a geometric model may be drastically reduced (Example 1.5).

Example 1.5. Consider the program made of two copies of the process

$$\underbrace{\texttt{P(a).V(a)} \ \ldots \ \texttt{P(a).V(a)}}_{N \text{ times}} \ .$$

Its forbidden region is the disjoint union of N^2 squares which forms a single maximal $\mathcal{C}'$-block. Its geometric model has $2(N+1)$ maximal rectangles (*i.e.* maximal $\mathcal{C}$-blocks) but only 2 maximal $\mathcal{C}'$-blocks (Figure 4). That situation generalizes to higher dimension n. Suppose that the arity of `a` is $n-1$ and consider a program made of n copies of the process introduced above. The forbidden region is made of N^n disjoint hypercubes which forms a single maximal $\mathcal{C}'$-block. Its geometric model has exactly $n(N+1)$ maximal hyperrectangles (*i.e.* maximal $\mathcal{C}$-blocks) but only n maximal $\mathcal{C}'$-blocks.

The second benefit is an algorithm which factorizes the geometric model of a program $P = (P_1, \ldots, P_n)$ from the *maximal* $\mathcal{C}'$-block covering L of its forbidden region. For $i \in \{1, \ldots, n\}$, the i^{th} letter of $w \in L$ is a subset of Ω that is actually contained in $|G_i|$. So we define the *support* of $w \in L$ as the set $\text{supp}(w)$ of all the indices $i \in \{1, \ldots, n\}$ such that the i^{th} letter of w is *not* the entire $|G_i|$ but a proper subset. The finest partition of $\{1, \ldots, n\}$ that is coarser than the family $\{\text{supp}(w) \mid w \in L\}$ induces the prime decomposition of $[\![P]\!]$. The correctness of the algorithm derives from Theorem 2.50 and Remark 2.51.

Example 1.6. We want to factorize $[\![P]\!] = [0,3]^n \setminus [1,2[^n$ where P is the program made of n copies of the process `P(a).V(a)` and the arity of the resource `a` is $n-1$ (Example 1.5). The alphabet Ω is the interval $[0,3]$ while the collection $\mathcal{C}$ (resp. $\mathcal{C}'$) is that of (finite unions of) subintervals of $[0,3]$. Hence the $\mathcal{C}$-blocks are the hyperrectangles. The forbidden region is the hypercube $[1,2[^n$ seen as a word w whose support is $\{1, \ldots, n\}$. On one hand, we instantly deduce from our algorithm that $[\![P]\!]$ is irreducible. On the other hand, the geometric model $[\![P]\!]$ has n maximal $\mathcal{C}'$-blocks (Example 1.5) so the generic algorithm performs $n \cdot C_n^{\lfloor n/2 \rfloor}$ word comparisons to obtain the same result — see Section 2.4.

Example 1.7. Given $n \geqslant 2$, the n-philosophers program has exactly n processes

$$\begin{array}{l}
\mathtt{P(a_1).P(a_2).V(a_1).V(a_2)\,|} \\
\mathtt{P(a_2).P(a_3).V(a_2).V(a_3)\,|} \\
\qquad\qquad\vdots \\
\mathtt{P(a_{n-1}).P(a_n).V(a_{n-1}).V(a_n)\,|} \\
\mathtt{P(a_n).P(a_1).V(a_n).V(a_1)}
\end{array}$$

where all the resources $\mathtt{a}_1, \ldots, \mathtt{a}_n$ are mutexes. The alphabet Ω is the interval $[0,5]$ and the collections $\mathcal{C}$ and $\mathcal{C}'$ are defined as in Example 1.6. The forbidden region is covered by n maximal $\mathcal{C}'$-blocks, which are actually hyperrectangles. Each of them is the contribution of a single mutex. The hyperrectangle generated by $\mathtt{a}_1$ is $[1,3[\times[0,5]^{n-2} \times [2,4[$ and for $i \geqslant 2$, the one generated by $\mathtt{a}_i$ is $[0,5]^{i-2} \times [2,4[\times[1,3[\times[0,5]^{n-i}$. The resulting family of supports consists of the pairs $\{1,n\}$ and $\{i-1,i\}$ for $i \in \{2, \ldots, n\}$. It follows that the geometric model of the n-philosophers program is irreducible.

Example 1.8. (Balabonski and Haucourt (2010)). Let P be the program

```
P(a).P(c).V(c).V(a)|
P(b).P(c).V(c).V(b)|
P(a).P(c).V(c).V(a)|
P(b).P(c).V(c).V(b)
```

forbidden region generated by `a`	[1,4[× [0,5] × [1,4[× [0,5]	↦	{1,3}
forbidden region generated by `b`	[0,5] × [1,4[× [0,5] × [1,4[	↦	{2,4}
forbidden region generated by `c`	[0,5] × [2,3[× [2,3[× [2,3[[2,3[× [0,5] × [2,3[× [2,3[[2,3[× [2,3[× [0,5] × [2,3[[2,3[× [2,3[× [2,3[× [0,5]		these hypercubes are not maximal

Fig. 5. The forbidden region of the program from Example 1.8

where `a` and `b` are two mutexes, and the arity of the resource `c` is 2. The sets Ω, $\mathcal{C}$, and $\mathcal{C}'$ are the same as in Example 1.7 and the contribution of each resource to the forbidden region of P is detailed in Figure 5. In particular, the forbidden region generated by `c` is entirely contained in the union of the forbidden regions generated by `a` and `b`. One can indeed observe that because of the mutex `a`, only one of the processes P_1 and P_3 can hold an occurrence of `c`. The same remark applies to the processes P_2 and P_4 because of the mutex `b`. Finally, the forbidden region of the program only contains two maximal $\mathcal{C}'$-blocks (which are actually hyperrectangles) from which we obtain the partition $\{\{1,3\},\{2,4\}\}$ of $\{1,2,3,4\}$. The associated permutation π swaps 2 and 3 so the program $\pi \cdot P$ is the following one:

```
P(a).P(c).V(c).V(a)|
P(a).P(c).V(c).V(a)|
P(b).P(c).V(c).V(b)|
P(b).P(c).V(c).V(b)
```

Then we can check that we have $[\![\pi \cdot P]\!] = [\![P|_{\{1,3\}}]\!] * [\![P|_{\{2,4\}}]\!]$, which confirms that P can be decomposed in two model independent subprograms. By the way, expanding the following expression

$$[\![\pi \cdot P]\!] \quad = \quad \Big(\big(([0,1[\cup[4,5]) \times [0,5]\big) \quad \cup \quad \big([0,5] \times ([0,1[\cup[4,5])\big)\Big)^2$$

we note that the geometric model of P has 16 maximal $\mathcal{C}$-blocks and 4 maximal $\mathcal{C}'$-blocks.

According to the semantics of the instructions `P(_)` and `V(_)`, the forbidden region of a program represents the restrictions on parallel execution imposed by the programmer. The forbidden region of an ‘interesting’ concurrent program is thus expected to be ‘sparse’. As an illustration, those of the programs considered so far were much ‘simpler’ than the corresponding geometric models. In particular, the ratio between the number of maximal $\mathcal{C}'$-blocks of the forbidden region and that of maximal $\mathcal{C}'$-blocks of the geometric model could be taken as a relevant measure of the intricateness of a program that is due to concurrency. In the extreme case where the forbidden region is empty, that ratio drops to zero. The above observation also provides an extra motivation for using the factoring algorithm based on forbidden regions (Section 2.7) instead of the generic one (Section 2.4).

1.4. *Organization*

We abstract away the topological aspects of isothetic regions (Definition 2.63) from the problem of factoring them (Section 2). To this end, we introduce the notion of a *homogeneous language* (over an alphabet Ω) and prove Theorem 2.32, from which one easily deduces that the commutative monoid of (finite) homogeneous languages is free (Corollaries 2.34 and 2.36). In this context, *maximal block coverings* (Definition 2.42) provide the key ingredient to practical application of factoring. Indeed, Theorem 2.44 asserts that we can replace Ω by an alphabet $\mathcal{C} \subseteq \wp(\Omega) \setminus \{\emptyset\}$ containing all the singletons. Then, provided that the collection of isothetics regions satisfies some additional properties (Conditions (A_1) and (A_2)), shifting to the richer alphabet $\mathcal{C}$ allows more efficient representations of homogeneous languages (Corollary 2.48) and practical applications of the factoring algorithm described in Section 2.4. In particular, Conditions (A_1) and (A_2) are fulfilled as soon as the collection $\mathcal{C}$ is a Boolean subalgebra of the powerset $\wp(\Omega)$ (Proposition 2.53). When this is the case, we can apply the factoring algorithm from Section 2.7, which is proven to be much more efficient than the generic one (Section 2.4). In Section 2.8, we finally move to the practical applications mentioned in Section 1.1. In this context, the alphabet Ω is the underlying set of $|G|$, viz the *geometric realization* of a graph G, and the alphabet $\mathcal{C}$ is (related to) the collection of connected subsets of Ω (Definitions 2.56 and 2.63).

If the graph G is *finite*, then the collection of finite unions of connected subsets of Ω forms a Boolean subalgebra of the powerset $\wp(\Omega)$. Consequently, if we are only concerned with factoring the geometric models of conservative programs (Section 1.1), which provide a reasonable setting in practice, then we do not need to investigate any further. But going beyond that limitation requires to deal with *infinite* (control flow) graphs. In Section 3, we prove that for a given graph G, every collection $\mathcal{R}_n|G|$ (Definition 2.45) is a Boolean subalgebra of the powerset $\wp(|G|^n)$ if and only if $|G|$ has a *compactification* that is homeomorphic to the geometric realization of some finite graph (Theorem 3.12). Such a graph is said to be *almost finite* (Definition 3.14). In Section 4, we prove that Condition (A_2) is satisfied if and only if each connected component of G has a *spanning tree* containing all its arrows but finitely many ones (Theorem 4.10). Such a graph is said to be *almost a forest* (Definition 4.4).

Being almost finite (resp. almost a forest) is equivalent to the fact that the co-unital subsemilattice of $\wp(|G|)$ generated by the connected subsets of $|G|$, viz $\mathcal{R}_1|G|$, is a Boolean subalgebra (resp. a co-unital distributive sublattice) of $\wp(|G|)$ (Theorems 3.12 and 4.10). These results lead to the study of the algebraic structures of the collections $\mathcal{R}_n|G|$, in view of which we provide a short survey on universal tensor products (Section 5.2) with a special emphasis on the case of co-unital semilattices (Section 5.3). Researchers dealing with Boolean algebras usually define the tensor product of their favourite objects of study as the ordinary tensor product of the related Boolean rings seen as idempotent commutative† $\mathbb{F}_2$-algebras: they implicitly refer to the correspondence between Boolean algebras and Boolean rings (Section 5.1). Yet, our humble contribution consists in proving

† In fact, any $\mathbb{F}_2$-algebra is commutative.

that it can also be defined as a universal tensor product in the category of co-unital semilattices (Section 5.4). The latter approach is indeed better fitted to the proof that if G is almost finite (resp. almost a forest), then for each $n \in \mathbb{N}$, the Boolean algebra (resp. co-unital distributive lattice) $\mathcal{R}_n|G|$ is the universal tensor product of n copies of $\mathcal{R}_1|G|$ in the category of co-unital semilattices (Proposition 5.14).

2. The free commutative monoid of homogeneous languages over a set Ω

The disjoint union of the powersets $\wp(\Omega^n) \setminus \{\emptyset\}$ for $n \in \mathbb{N}$ can be turned into a free commutative monoid. The corresponding unique factorisation property is related to the parallelization of *conservative* programs (Haucourt (2017)). The present section generalizes results from (Balabonski and Haucourt (2010)), (Fajstrup et al., 2016, p.101–104), and (Ninin, 2017, Chapter 2).

2.1. *Homogeneous languages*

A *language* is a collection of finite sequences over a set Ω which are, in that context, respectively called *words* and *alphabet*. A language is said to be *homogeneous* when all the words it contains have the same length, which is then the *length of the language*. Concatenation of words, denoted by $*$, extends to languages in the obvious way:

$$L * L' := \{w * w' \mid w \in L \ ; \ w' \in L'\} \ .$$

In particular, the concatenation of two homogeneous languages L and L' of length n and n' is a homogeneous language of length $n + n'$. The empty language is absorbing in the sense that for any language L we have $\emptyset * L = \emptyset = L * \emptyset$, so its length is conventionally defined as $-\infty$. With respect to our purpose, the empty language is discarded.

Definition 2.1. The set of *nonempty* homogeneous languages over the alphabet Ω forms the (noncommutative) monoid $\mathcal{H}$, or $\mathcal{H}(\Omega)$ if we need to be explicit about the alphabet. Its neutral element is $\{\varepsilon\}$ (*i.e.* the singleton containing the empty word).

Definition 2.2. A monoid M is said to be *left-cancellative* when $xy = xz$ implies $y = z$ for all $x, y, z \in M$. *Right-cancellative* monoids are defined accordingly. A monoid is *cancellative* when it is both left-cancellative and right-cancellative.

Lemma 2.3. The monoid $\mathcal{H}$ is cancellative.

Proof. Suppose that we have $L * L' = L * L''$ in $\mathcal{H}$. In particular L is not empty so it contains a word w. Considering the words of $L * L'$ and that of $L * L''$ admitting w as a prefix, we deduce that $\{w\} * L' = \{w\} * L''$, and therefore $L' = L''$. $\square$

Given a word w of length n and a permutation π of the set $\{1, \dots, n\}$, the word $\pi \cdot w$ is obtained by moving the k^{th} letter of the word w to position $\pi(k)$. Dually, the k^{th} letter of the word $\pi \cdot w$ is the one that was at position $\pi^{-1}(k)$ in the word w. Hence the n^{th} *symmetric group* $\mathfrak{S}_n$ acts on the left of the set of words of length n as follows (words of length n are considered as mapping defined on the set $\{1,\ldots,n\}$):

$$\pi \cdot w \quad := \quad w \circ \pi^{-1} \quad = \quad (w_{\pi^{-1}(1)} \cdots w_{\pi^{-1}(n)}) \; .$$

By extension, the group $\mathfrak{S}_n$ also acts on the left of the collection of homogeneous languages of length n by defining $\pi \cdot L$ as $\{\pi \cdot w \mid w \in L\}$. Two homogeneous languages are said to be equivalent, denoted by $L \sim L'$, when they have the same length n and there exists $\pi \in \mathfrak{S}_n$ such that $L' = \pi \cdot L$. We define the *juxtaposition* $\pi \otimes \pi' \in \mathfrak{S}_{n+n'}$ of the permutations $\pi \in \mathfrak{S}_n$ and $\pi' \in \mathfrak{S}_{n'}$ as follows:

$$\pi \otimes \pi'(k) := \left\{ \begin{array}{ccc} \pi(k) & \text{if} & 1 \leqslant k \leqslant n \\ \big(\pi'(k-n)\big) + n & \text{if} & n+1 \leqslant k \leqslant n+n' \; . \end{array} \right.$$

The juxtaposition operator satisfies the following Godement exchange law

$$(\pi \cdot L) * (\pi' \cdot L') \quad = \quad (\pi \otimes \pi') \cdot (L * L') \; ,$$

which ensures that the relation $\sim$ is a congruence over $\mathcal{H}$.

Definition 2.4. The *homogeneous monoid* over Ω, denoted by $\tilde{\mathcal{H}}$, is the quotient of the monoid $\mathcal{H}$ by the congruence $\sim$. We denote by $\bar{L}$ the image of the homogeneous language L by the quotient morphism $q : \mathcal{H} \to \tilde{\mathcal{H}}$.

Remark 2.5. If the alphabet Ω is a singleton (resp. the empty set) then the homogeneous monoid $\tilde{\mathcal{H}}$ is isomorphic to $(\mathbb{N}, +, 0)$ (resp. to the null monoid).

Remark 2.6. If the alphabet Ω contains at least two elements, then the canonical morphism from the abelianization of $\mathcal{H}$ to $\tilde{\mathcal{H}}$ is not an isomorphism. Indeed, the language $\{aab, baa\}$ is irreducible in the abelianization of $\mathcal{H}$ but not in $\tilde{\mathcal{H}}$ since letting π be the permutation swapping 1 and 2, we have $\pi \cdot \{aab, baa\} = \{aab, aba\} = \{a\} * \{ab, ba\}$.

2.2. *Free commutative monoids*

We gather the basic facts about free commutative monoids that are needed in the sequel of the section. The existence of unique decompositions is related to the subtle distinction between prime and irreducible elements in a commutative monoid. In this section we denote the neutral element of a monoid by ε. More details can be found in the first chapter of (Geroldinger and Halter-Koch (2006)). However, we insist that in the latter, monoids are defined as cancellative semigroups (Geroldinger and Halter-Koch, 2006, p.xiii), which is rather unusual.

Definition 2.7. A *unit* of a commutative monoid is an element u for which there exists an element u' such that $uu' = \varepsilon$. A commutative monoid is said to be *reduced* when it has no unit but its neutral element. One says that d *divides* x when there exists x' such that $x = dx'$, this situation being denoted by $d|x$. The elements x and y are said to be *equivalent* when $y = ux$ for some unit u.

Example 2.8. The monoid $\mathbb{Z}^\times = (\mathbb{Z} - \{0\}, \times, 1)$ is not reduced but the monoid $\mathbb{N}^\times = (\mathbb{N} - \{0\}, \times, 1)$ is. One can always *reduce* a commutative monoid by identifying its equivalent elements, which amounts to identifying all its units with the neutral element.

Definition 2.9. A nonunit element is said to be *irreducible*[†] when it can only be divided, up to equivalence, by ε and itself. A nonunit element is said to be *prime* when it divides a or b as soon as it divides their product. Denote by $I(M)$ and $P(M)$ the set of irreducible elements and the set of prime elements of a commutative monoid M.

Example 2.10. In the monoids $\mathbb{N}^\times$ and $\mathbb{Z}^\times$, an element is prime iff it is irreducible.

Definition 2.11. We say that $G \subseteq M$ generates M when any element of M is, up to equivalence, a product of elements of G. By convention, the empty product is ε.

Example 2.12. Let the *support* of a mapping from X to $\mathbb{N}$ be the subset of X on which it is nonzero. The collection of all the mappings with finite support is denoted by $F(X)$. It is a commutative monoid under pointwise addition, the null map being the neutral element. One readily checks that $F(X)$ is reduced, cancellative, and generated by the mappings $g_x : X \to \mathbb{N}$ defined by $g_x(x) = 1$ and $g_x(y) = 0$ for $y \neq x$. As elements of $F(X)$, the mappings g_x are prime and irreducible. In other words, a reduced commutative monoid is free iff any of its elements x can be written as a product $i_1 * \cdots * i_n$ of irreducible elements in a unique way up to reordering of the terms. The construction $X \mapsto F(X)$ extends to a functor $F : \mathbf{Set} \to \mathbf{CMon}$ which is left adjoint to the forgetful one.

Definition 2.13. A commutative monoid of the form $F(X)$ is said to be *free*. For each $x \in F(X)$, the tuple $(i_1, \ldots, i_n)$ is called the *prime decomposition* of x (Example 2.12).

Proposition 2.14. A commutative monoid M is free if and only if it is reduced, cancellative, and the sets $P(M)$ and $I(M)$ are equal and generate M.

Proof. Let X be the set of irreducible elements of M. The map ϕ which sends $f \in F(X)$ to the product of the elements $x^{f(x)}$ is well-defined because M is commutative, and onto M by hypothesis. Suppose that we have $a_1 \cdots a_n = b_1 \cdots b_m$ with $a_i, b_j \in X$ for all $i \in \{1, \ldots, n\}$ and all $j \in \{1, \ldots, m\}$. By hypothesis a_1 is prime and we can suppose that it divides b_1. Since b_1 is irreducible and the monoid M is reduced, we have $a_1 = b_1$. Therefore we have $a_2 \cdots a_n = b_2 \cdots b_m$ because M is cancellative. Suppose that $n \geqslant m$. By an obvious induction we end up with $a_{n-m} \cdots a_n = \varepsilon$ from which we deduce that $a_i = \varepsilon$ for all $i \in \{n - m, \ldots, n\}$ because M is reduced. So the map ϕ is one-to-one. $\square$

Example 2.15. The monoids $(\mathbb{N}, +, 0)$ and $(\mathbb{N}\backslash\{0\}, \times, 1)$ are freely commutative.

Example 2.16. The monoid $(\mathbb{R}_+, +, 0)$ has neither prime nor irreducible elements. Indeed, for every real number $x > 0$ we have $x \geqslant \frac{x}{2} + \frac{x}{2}$ though $x \ngeqslant \frac{x}{2}$.

Example 2.17. The semilattice $(\mathbb{N}, \vee, 0)$ is commutative monoid, in which every nonzero element is prime but not irreducible. Note that in the semilattice $(\{0, 1\}, \vee, 0)$ the element 1 is prime and irreducible. Also note that semilattices are reduced but not cancellative.

Example 2.18. In $\mathbb{Z}_6$ one notes that 2 is a prime but not irreducible since $2 = 2 \cdot 4$ (mod 6) and neither 2 nor 4 are unit as they are zero divisors (Hungerford, 2003, p.136).

[†] Other notions of irreducibility are listed in (Anderson and Valdes-Leon, 1997, p.199-200). However, if the monoid under consideration is reduced and cancellative, they coincide.

Example 2.19. The set $\{a + b\sqrt{10} \mid a, b \in \mathbb{Z},\ a \neq 0 \text{ or } b \neq 0\}$ forms a submonoid of the multiplicative monoid of nonzero real numbers, in which the elements 2, 3 and $4 \pm \sqrt{10}$ are irreducible but not prime. (Hungerford, 2003, p.140).

Example 2.20. (Nakayama and Hashimoto (1950)). In the semiring $\mathbb{N}[X]$, the polynomial $X^5 + X^4 + X^3 + X^2 + X + 1$ has two incompatible decompositions

$$(X + 1)(X^4 + X^2 + 1) = (X^3 + 1)(X^2 + X + 1)$$

from which we deduce that neither $X^3 + 1$ nor $X^4 + X^2 + 1$ is prime in $\mathbb{N}[X]$. In the meantime, their prime decompositions in the unique factorization domain $\mathbb{Z}[X]$, see (Hungerford, 2003, Remark, p.138), are given by

$$X^3 + 1 = (X + 1)(X^2 - X + 1) \quad \text{and} \quad X^4 + X^2 + 1 = (X^2 + X + 1)(X^2 - X + 1)\ .$$

It follows that both $X^3 + 1$ and $X^4 + X^2 + 1$ are irreducible in $\mathbb{N}[X]$.

Definition 2.21. A commutative monoid M is said to be *graded* when there exists a morphism of monoid $d : M \to \mathbb{N}$ such that $d^{-1}(\{0\})$ is the set of units of M.

Remark 2.22. A free commutative monoid is graded since each of its elements can be associated with the number of terms (with their multiplicity) of its prime decomposition. Conversely, a graded monoid is not far from being freely commutative.

Proposition 2.23. For every graded monoid M, $I(M)$ generates M and contains $P(M)$.

Proof. Given $x_1, \ldots, x_n$ non-unit elements of M we have

$$d(x_1 \cdots x_n) = d(x_1) + \cdots + d(x_n) \geqslant n$$

because M is graded. It is therefore generated by its irreducible elements. Suppose that $p = a \cdot b$ is prime. So we can suppose p divides a, and then we have

$$d(a) + d(b) = d(a \cdot b) = d(p) \leqslant d(a)\ .$$

Therefore $d(b) = 0$ from which we deduce that b is a unit of M. □

Corollary 2.24. A commutative monoid M is free iff it is reduced, cancellative, graded, and all its irreducible elements are prime.

Example 2.25. The monoid $M =< a, b \mid ab = ba;\ aab = abb >$ is commutative, reduced, graded, and satisfies $I(M) = P(M) = \{a, b\}$, but it is not cancellative because $aa \neq ab$.

Example 2.26. The monoid $\tilde{\mathcal{H}}$ (Definition 2.4) is reduced and graded by the length of homogeneous languages. We denote by $\tilde{\mathcal{H}}_n$ the subset of the elements of length n.

Any submonoid of a graded monoid is graded. Yet, a submonoid of a free commutative monoid might not be free e.g. define $\alpha = x + 2y$, $\beta = 2x + y$, and $\gamma = x + y$ so the submonoid of $\mathbb{N}^{\{x,y\}}$ generated by α, β, and γ satisfies $\alpha + \beta = 3\gamma$.

Definition 2.27 (Grillet (1969a)). A submonoid M' of M is said to be *consistent* when for all $x, y \in M$, if $x \cdot y \in M'$ then both x and y belong to M'.

Lemma 2.28. Any consistent submonoid M' of a free commutative monoid M is free, and the prime decomposition of an element of M' is its prime decomposition in M.

Proof. Let M' be a consistent submonoid of a free commutative monoid M. Then M' is graded. Let p be an irreducible element of M', it is also irreducible in M by consistency of M'. Moreover, if p divides $x \cdot y$ with $x, y \in M'$ then we can suppose that $x = p \cdot x'$ with $x' \in M$. Since M' is consistent, x' belongs to M'. We conclude by Corollary 2.24. □

2.3. *Partitions of the set* $\{1, \ldots, n\}$ *and factorizations of an element of* $\tilde{\mathcal{H}}_n$

We establish a result (Theorem 2.32) which generalizes (Ninin, 2017, Proposition 2.4.1) and completes the proof of (Balabonski and Haucourt, 2010, Theorem 1). The latter indeed contains the fact that every irreducible element of $\mathcal{H}$ is prime, but not that $\mathcal{H}$ is cancellative. Given the natural numbers a and b we denote by $[a : b]$ the set of natural numbers x such that $a \leqslant x \leqslant b$. An *interval* of $\mathbb{N}$ is a subset I such that $[a : b] \subseteq I$ for all $a, b \in I$. Given the intervals I and J we denote by $I < J$ the fact that every element of I is strictly less than any element of J. For each partition $(X_1, \ldots, X_k)$ of the interval $[1 : n]$ there exists a unique permutation $\pi \in \mathfrak{S}_n$ such that:

1 for all $i \in [1 : k]$, the set $\pi(X_i)$ is an interval (denoted by I_i),
2 for all $i, i' \in [1 : k]$, if $i < i'$ then $I_i < I_{i'}$, and
3 for all $i \in [1 : k]$, the restricted mapping $\pi : X_i \to I_i$ is increasing.

Remark 2.29. Define the *index* i_x of an element $x \in [1 : n]$ as the unique element i_x of $[1 : k]$ such that $x \in X_{i_x}$. Also define the *rank* r_x of x as the number of elements of X_{i_x} that are less or equal than x. The permutation π is the unique element of $\mathfrak{S}_n$ such that for all $x, y \in [1 : n]$, if (i_x, r_x) is less of equal than (i_y, r_y) in the lexicographic order, then $\pi(x) \leqslant \pi(y)$. Even more explicitly, we have $\pi(x) = n_x + r_x$ where n_x is the number of elements of $X_1 \cup \cdots \cup X_{i_x - 1}$, with the convention that $n_x = 0$ when $i_x = 1$.

Definition 2.30. Given a word w of length n and a subset $X = \{x_1 < \cdots < x_k\}$ of $\{1, \ldots, n\}$, the *restriction* of w to X, which we denote by $w|_X$, is the word of length k whose i^{th} letter, for $i \in [1 : k]$, is the x_i^{th} letter of w. By extension, the restriction to X of a homogeneous language L of length n, which we denote by $L|_X$, is the set of words $w|_X$ for $w \in L$. One checks that for all $\pi \in \mathfrak{S}_n$ that is increasing on X we have

$$(\pi \cdot w)|_X \quad = \quad w|_{\pi^{-1}(X)} \qquad \text{and} \qquad (\pi \cdot L)|_X \quad = \quad L|_{\pi^{-1}(X)} \ .$$

Remark 2.31. The language $\pi \cdot L$ is always included in the language $L|_{X_1} * \cdots * L|_{X_k}$. The partition $(X_1, \ldots, X_k)$ induces a factorization of $\bar{L} \in \tilde{\mathcal{H}}_n$ when the other inclusion is also satisfied, that is to say when the following equality holds in $\mathcal{H}$:

$$\pi \cdot L \quad = \quad L|_{X_1} * \cdots * L|_{X_k} \ . \tag{1}$$

The other way round, given a permutation $\pi \in \mathfrak{S}_n$ such that $\pi \cdot L = L_1 * \cdots * L_k$ in $\mathcal{H}$, which amounts to a factorization of $\bar{L}$, one recovers the partition $(X_1, \ldots, X_k)$ by noting that $X_i = \pi^{-1}([d_{i-1} + 1 : d_i])$ with $d_0 = 0$ and $d_i = d_{i-1} + \mathit{length}(X_i)$ for all $i \in [1 : k]$. Note that the permutation related to the partition $(X_1, \ldots, X_k)$ may differ from π since the latter is not required to be increasing on each element X_i of the partition. Thus each factorization $H_1 * \cdots * H_k$ of $\bar{L}$ is induced by some partition $(X_1, \ldots, X_k)$ of $[1:n]$ such that $\overline{L|}_{X_i} = H_i$ for all $i \in [1:k]$.

The next result is the technical cornerstone of the proof that $\tilde{\mathcal{H}}$ is freely commutative, its proof is based on ideas from (Ninin (2017)).

Theorem 2.32. Let L be a homogeneous language of length n. If X and Y are two subsets of $[1:n]$ such that $\overline{L|}_X$ and $\overline{L|}_Y$ both divide $\overline{L}$, then the equality

$$\overline{L|}_X \quad = \quad \overline{L|}_{X \,\cap\, Y} * \overline{L|}_{X \,\cap\, ([1:n]\setminus Y)}$$

holds in the homogeneous monoid $\tilde{\mathcal{H}}$ (Definition 2.4).

Proof. Denote the cardinals of $X \cap Y$ and $X \cap ([1:n] \setminus Y)$ by A and B, and let $\{\alpha_1 < \cdots < \alpha_A\}$ and $\{\beta_1 < \cdots < \beta_B\}$ be their enumerations. Let π and π' be the permutations associated with the partitions $(X, [1:n] \setminus X)$ and $(Y, [1:n] \setminus Y)$. By definition of π, for each $i \in [1:\mathrm{card}(X)]$, $\pi^{-1}(i)$ belongs to X. If it also belongs to Y, then there is a unique $j \in [1:A]$ such that $\pi^{-1}(i) = \alpha_j$. In that case we define $\pi''(i) = j$. The same way, if $\pi^{-1}(i)$ belongs to $[1:n] \setminus Y$, then there is a unique $j \in [1:B]$ such that $\pi^{-1}(i) = \beta_j$. In that case we define $\pi''(i) = A + j$. So π'' is the permutation of $[1:\mathrm{card}(X)]$ associated to the partition of $[1:\mathrm{card}(X)]$ induced by the partition $(Y, [1:n] \setminus Y)$ of $[1:n]$. From Remark 2.31 we deduce the following inclusion

$$\pi'' \cdot L|_X \quad \subseteq \quad L|_{X \,\cap\, Y} * L|_{X \,\cap\, ([1:n]\setminus Y)} \ .$$

Now we prove the opposite one. Since $\overline{L|}_X$ and $\overline{L|}_Y$ both divide $\overline{L}$, we have

$$\pi \cdot L \quad = \quad L|_X * L|_{[1:n]\setminus X} \qquad \text{and} \qquad \pi' \cdot L \quad = \quad L|_Y * L|_{[1:n]\setminus Y}$$

from which we deduce that

$$L|_X * L|_{[1:n]\setminus X} \quad = \quad (\pi \circ \pi'^{-1}) \cdot (L|_Y * L|_{[1:n]\setminus Y}) \ .$$

It follows that

$$L|_X \quad = \quad \big((\pi \circ \pi'^{-1}) \cdot (L|_Y * L|_{[1:n]\setminus Y})\big)\,|_{[1:\mathrm{card}(X)]} \ .$$

We will conclude by proving the following inclusion:

$$\big((\pi \circ \pi'^{-1}) \cdot (L|_Y * L|_{[1:n]\setminus Y})\big)\,|_{[1:\mathrm{card}(X)]} \quad \supseteq \quad \pi''^{-1} \cdot \Big(L|_{X \,\cap\, Y} * L|_{X \,\cap\, ([1:n]\setminus Y)}\Big) \ .$$

Let w be a word of length $\mathrm{card}(X)$ of the form

$$w^{(1)}|_{X \cap Y} * w^{(2)}|_{X \cap ([1:n]\setminus Y)}$$

for some words $w^{(1)}$ and $w^{(2)}$ of L. We will prove, letter by letter, that the words $\pi''^{-1} \cdot w$ and $(\pi \circ \pi'^{-1}) \cdot (w^{(1)}|_Y * w^{(2)}|_{[1:n]\setminus Y})$ are equal. Let i be an element of $[1:\mathrm{card}(X)]$. By definition of π, the index $\pi^{-1}(i)$ belongs to X. Moreover, if $\pi^{-1}(i) \in Y$, then the index $\pi' \circ \pi^{-1}(i)$ belongs to $[1:\mathrm{card}(Y)]$ and there is a unique $j \in [1:A]$ such that $\pi^{-1}(i) = \alpha_j$; otherwise it belongs to $[\mathrm{card}(Y)+1:n]$ and there is a unique $j \in [1:B]$

such that $\pi^{-1}(i) = \beta_j$. In the first case the i^{th} letter of the word

$$(\pi \circ \pi'^{-1}) \cdot (w^{(1)}|_Y * w^{(2)}|_{[1:n]\setminus Y})$$

is the $\pi'(\alpha_j)^{th}$ letter of the word

$$w^{(1)}|_Y * w^{(2)}|_{[1:n]\setminus Y}$$

which is also the j^{th} letter of the word $w^{(1)}|_Y$ or equivalently the α_j^{th} letter of the word $w^{(1)}$. Moreover, by definition of π'', we have $\pi''(i) = j$, which is the rank of α_j in the set $X \cap Y$. So the $\pi''(i)^{th}$ letter of the word w is the j^{th} letter of the word $w^{(1)}|_{X\cap Y}$ which is also the α_j^{th} letter of the word $w^{(1)}$. The second case (*i.e.* $\pi^{-1}(i) \in [1:n] \setminus Y$) is treated the same way. □

Corollary 2.33. Let L be a homogeneous language of length n, and X be a subset of $[1:n]$ such that $\overline{L|_X}$ is irreducible and divides $\bar{L}$. If $(X_1, \ldots, X_k)$ is a partition of $[1:n]$ inducing a factorization of $\bar{L}$, then there exists some $i \in [1:k]$ such that $X \subseteq X_i$.

Proof. By induction on k with the help of Theorem 2.32. □

Corollary 2.34. The monoid $\tilde{\mathcal{H}}$ is freely commutative.

Proof. From Example 2.26 we already know that $\tilde{\mathcal{H}}$ is reduced and graded. As a consequence of Corollary 2.24, proving that $\mathcal{H}$ is freely commutative amounts to proving that it is cancellative and that its irreducible elements are prime.

Suppose that we have $H * H_1 = H * H_2$ in $\tilde{\mathcal{H}}$. Equivalently, we have L, L_1, and L_2, representatives of H, H_1, and H_2 such that $L * L_1 = \pi \cdot (L * L_2)$ where n is the length of L, m is that of L_1 and L_2, and π is a permutation of $[1:n+m]$ (Remark 2.31). We can suppose that π is increasing on $\pi^{-1}([1:n])$ so that we have

$$L \quad = \quad (L * L_1)|_{[1:n]} \quad = \quad (\pi \cdot (L * L_2))|_{[1:n]} \quad = \quad (L * L_2)|_{\pi^{-1}([1:n])}$$

the second equality being due to the final remark in Definition 2.30. We also have $L = (L * L_2)|_{[1:n]}$ so the sets $\pi^{-1}([1:n])$ and $[1:n]$ are either equal or disjoint (Theorem 2.32). In the first case, the permutation π can be written as $\pi_1 \otimes \pi_2$ where π_1 and π_2 are permutations of $[1:n]$ and $[1:m]$ respectively (Section 2.1) so we have

$$L * L_1 \quad = \quad \pi \cdot (L * L_2) \quad = \quad (\pi_1 \otimes \pi_2) \cdot (L * L_2) \quad = \quad (\pi_1 \cdot L) * (\pi_2 \cdot L_2)\ .$$

Since $\mathcal{H}$ is right-cancellative (Lemma 2.3), we deduce that $L_1 = \pi_2 \cdot L_2$. The second case, we consider a permutation π' that leaves each element of $[n+1 : n+m] \setminus \pi^{-1}([1:n])$ unchanged and maps $[1:n]$ onto $\pi^{-1}([1:n])$ in a way that

$$(\pi' \circ \pi) \cdot (L * L_2) \quad = \quad \pi \cdot (L * L_2)\ .$$

Then we apply the first case noting that $\pi' \circ \pi([1:n]) = [1:n]$.

Let L be a homogeneous language of length n such that $\bar{L} = H_1 * H_2$ and suppose that some irreducible element H of $\tilde{\mathcal{H}}$ divides $H_1 * H_2$. There exists some subset X of $[1:n]$ such that $\overline{L|_X} = H$ while the factorization $H_1 * H_2$ is related to a partition (X_1, X_2) of $[1:n]$. We conclude that $X \subseteq X_1$ or $X \subseteq X_2$ (Corollary 2.33), in other words that H divides H_i for some $i \in \{1, 2\}$. □

Even if we take term reordering into account, there may be many partitions leading to the same decomposition. For example, all the partitions $\{\{1\},\{2,3\}\}$, $\{\{2\},\{1,3\}\}$, and $\{\{3\},\{1,2\}\}$ lead to the same decomposition when $\bar{L} = H_1 * H_2$ where the length of H_1 is 1 and $H_2 = H_1 * H_1$ in $\tilde{\mathcal{H}}$. The phenomenon vanishes when the elements of the decomposition are irreducible.

Corollary 2.35. Up to term reordering, there is a unique partition of $[1:n]$ corresponding to the prime decomposition of $\bar{L}$.

Proof. Two such partitions have the same number of elements, the result follows from an immediate induction on that number together with Corollary 2.33. □

By definition, an element of $\tilde{\mathcal{H}}$ is an equivalence class whose elements are sets of the same cardinal. Therefore we can define the cardinal of an element of $\tilde{\mathcal{H}}$ as the cardinal of any of its representatives. In particular, the collection $\tilde{\mathcal{H}}^f$ of finite elements of $\tilde{\mathcal{H}}$ forms a consistent submonoid of $\tilde{\mathcal{H}}$ because for all nonempty languages L and L', the language $L \cdot L'$ is finite if and only if the languages L and L' are so. The next result is an immediate consequence of Corollary 2.34 and Lemma 2.28.

Corollary 2.36. The monoid $\tilde{\mathcal{H}}^f$ is freely commutative.

2.4. *The generic factoring algorithm*

Assuming that the range of homogeneous languages L under consideration is restricted to finite ones, whether the identity (1) from Remark 2.31 holds is decided by an obvious algorithm that performs $\mathrm{card}(L)$ tests of equality between words. Whether an element of $\tilde{\mathcal{H}}^f$ of length n is prime can thus be decided by checking all the subsets $X \subseteq [1:n]$ whose cardinal is at most $\lfloor n/2 \rfloor$ (the greatest natural number less of equal than the half of n). This is the algorithm given in (Balabonski and Haucourt (2010)). Denoting the number of such subsets by $C_n^{\lfloor n/2 \rfloor}$, the algorithm requires at most $\mathrm{card}(L) \times C_n^{\lfloor n/2 \rfloor}$ word comparisons. Its complexity is thus more than exponential in n.

2.5. *Homogeneous monoids inherit preorders from alphabets*

Suppose that the alphabet Ω comes with a preorder $\preccurlyeq_\Omega$ and denote by $\preccurlyeq_\Omega^n$ the product preorder on Ω^n. Given two homogeneous languages L and L' of the same length n, write $L \preccurlyeq_{\mathcal{H}} L'$ when for each $w \in L$ there exists $w' \in L'$ such that $w \preccurlyeq_\Omega^n w'$. In that case, for all $\pi \in \mathfrak{S}_n$ we also have $\pi \cdot L \preccurlyeq_{\mathcal{H}} \pi \cdot L'$. Then given two elements H and H' of $\tilde{\mathcal{H}}$, write $H \preccurlyeq_{\tilde{\mathcal{H}}} H'$ when there exist representatives L and L' of H and H' such that $L \preccurlyeq_{\mathcal{H}} L'$.

Proposition 2.37. The map q from Definition 2.4 induces a morphism of preordered monoids† from $(\mathcal{H}, \preccurlyeq_{\mathcal{H}})$ onto $(\tilde{\mathcal{H}}, \preccurlyeq_{\tilde{\mathcal{H}}})$. Moreover, if $\preccurlyeq_{\mathcal{H}}$ is antisymmetric, then so is $\preccurlyeq_{\tilde{\mathcal{H}}}$.

† Given a preorder $\preccurlyeq$ on a monoid M, we say that $(M, \preccurlyeq)$ is a *preordered monoid* when for all elements a, b, c, d of M, if $a \preccurlyeq b$ and $c \preccurlyeq d$ then $ac \preccurlyeq bd$. The *morphisms of preordered monoids* are the preorder-preserving morphisms of monoids.

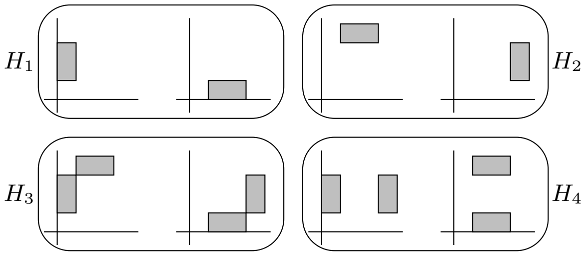

Fig. 6. Two elements with two non comparable minimal upper bounds

Proof. The morphism of monoids q is obviously onto and readily preserves the binary relations in the sense that if $L \preccurlyeq_{\mathcal{H}} L'$, then $q(L) \preccurlyeq_{\tilde{\mathcal{H}}} q(L')$. We deduce that if $H_1 \preccurlyeq_{\tilde{\mathcal{H}}} H_1'$ and $H_2 \preccurlyeq_{\tilde{\mathcal{H}}} H_2'$, then $H_1 * H_2 \preccurlyeq_{\tilde{\mathcal{H}}} H_1' * H_2'$. It remains to check that the relation $\preccurlyeq_{\tilde{\mathcal{H}}}$ is a preorder. It is obviously reflexive. If $H_1 \preccurlyeq_{\tilde{\mathcal{H}}} H_2 \preccurlyeq_{\tilde{\mathcal{H}}} H_3$ then we have $L_1 \in H_1$, $L_2, L_2' \in H_2$, and $L_3 \in H_3$ such that $L_1 \preccurlyeq_{\mathcal{H}} L_2$ and $L_2' \preccurlyeq_{\mathcal{H}} L_3$. By definition, there exists some permutation π such that $\pi \cdot L_2 = L_2'$, from which we deduce that $\pi \cdot L_1 \preccurlyeq_{\mathcal{H}} \pi \cdot L_2 \preccurlyeq_{\mathcal{H}} L_3$, and therefore that $H_1 \preccurlyeq_{\tilde{\mathcal{H}}} H_3$. Now suppose that $H \preccurlyeq_{\tilde{\mathcal{H}}} H'$ and $H' \preccurlyeq_{\tilde{\mathcal{H}}} H$ and that $\preccurlyeq_{\mathcal{H}}$ is antisymmetric. Then we have $L \in H$, $L' \in H'$ and some permutation π such that $L \preccurlyeq_{\mathcal{H}} L' \preccurlyeq_{\mathcal{H}} \pi \cdot L$. An immediate induction provides the following sequence, where $m \in \mathbb{N}$:

$$L \preccurlyeq_{\mathcal{H}} \pi \cdot L \preccurlyeq_{\mathcal{H}} \cdots \preccurlyeq_{\mathcal{H}} \pi^m \cdot L \preccurlyeq_{\mathcal{H}} \cdots$$

By finiteness of the group of permutations of a finite set, we have $\pi^m = \mathrm{id}$ for some $m \in \mathbb{N} \setminus \{0\}$, hence $L = L'$ because $\preccurlyeq_{\mathcal{H}}$ is antisymmetric. □

Remark 2.38. The preorder $\preccurlyeq_{\tilde{\mathcal{H}}}$ does not, in general, inherit the properties from $\preccurlyeq_{\Omega}$. For example, let the alphabet Ω be the complete lattice of subintervals of $\mathbb{R}$ ordered by inclusion. The frames on Figure 6 are extensive descriptions of the elements H_1, H_2, H_3 and H_4 of $\tilde{\mathcal{H}}$. Formally we have $H_1 = \{[0,1] \times [1,3], [1,3] \times [0,1]\}$ and $H_2 = \{[1,3] \times [3,4], [3,4] \times [1,3]\}$. Then observe that H_3 and H_4, described below, are non comparable minimal upper bounds of H_1 and H_2. The preorder $\preccurlyeq_{\tilde{\mathcal{H}}}$ thus lack binary least upper bounds.

$$H_3 = \{[0,1] \times [1,3] \cup [1,3] \times [3,4], [1,3] \times [0,1] \cup [3,4] \times [1,3]\}$$

$$H_4 = \{[0,1] \times [1,3] \cup [3,4] \times [1,3], [1,3] \times [0,1] \cup [1,3] \times [3,4]\}$$

2.6. *Blocks of words*

Let $\mathcal{C}$ be a collection of subsets of Ω containing all the singletons so that in some sense, it contains Ω. We denote the inclusion relation on $\mathcal{C}$ by $\subseteq_{\mathcal{C}}$ while $=_{\Omega}$ stands for the discrete order on Ω. Any point of Ω^n can be seen as a word on the alphabet Ω so a (nonempty) subset of Ω^n can be seen as a (nonempty) homogeneous language of length n over Ω.

Definition 2.39. A *block*, or $\mathcal{C}$-*block* if one needs to emphasize on the dependency, is a finite Cartesian product of nonempty elements of $\mathcal{C}$. Blocks are thus identified with words over the alphabet of nonempty elements of $\mathcal{C}$. A block contained in $L \subseteq \Omega^n$ is said to be a block of L. Such a block is said to be *maximal* when no block of L strictly contains it. A collection of blocks whose union is L is called a *block covering* of L.

Remark 2.40. The restriction of the relation $\preccurlyeq_{\mathcal{H}\Omega}$ to $\mathcal{H}_n\Omega$, where $\preccurlyeq_{\mathcal{H}\Omega}$ is defined from $=_\Omega$ at the beginning of Section 2.5, is the inclusion relation between subsets of Ω^n. Given two collections of sets $\mathcal{F}$ and $\mathcal{F}'$, we write $\mathcal{F} \preccurlyeq \mathcal{F}'$ when every element of $\mathcal{F}$ is included in some element of $\mathcal{F}'$. The relation $\preccurlyeq$ is called the *covering preorder*. In particular, the restriction to $\mathcal{H}_n\mathcal{C}$ of the relation $\preccurlyeq_{\mathcal{H}\mathcal{C}}$ defined from $\subseteq_\mathcal{C}$ at the beginning of Section 2.5 is the covering preorder over the collection of n-dimensional block coverings. Hence we will write $\subseteq$ and $\preccurlyeq$ instead of $\preccurlyeq_{\mathcal{H}\Omega}$ and $\preccurlyeq_{\mathcal{H}\mathcal{C}}$. With the notation introduced in Section 2.5 we will also write $\tilde{\subseteq}$ and $\tilde{\preccurlyeq}$ instead of $\preccurlyeq_{\tilde{\mathcal{H}}\Omega}$ and $\preccurlyeq_{\tilde{\mathcal{H}}\mathcal{C}}$.

Remark 2.41. The composition law of the monoid $\mathcal{H}\Omega$ (*i.e.* the concatenation of languages) readily corresponds to the usual Cartesian product of sets. From the standard logical equivalence

$$A \times B \quad \subseteq \quad C \times D \qquad \text{if and only if} \qquad A \subseteq C \text{ and } B \subseteq D \tag{E1}$$

which holds for all sets A, B, C, and D, we deduce that the two monoids $(\mathcal{H}\Omega, \subseteq)$ and $(\mathcal{H}\mathcal{C}, \preccurlyeq)$ are preordered, and also that the map γ sending each element of $\mathcal{H}\mathcal{C}$ to the union of its elements is a morphism of preordered monoids. As a consequence of equivalence (E1) we obtain, for all $L, L' \in \mathcal{H}\Omega$, the following equalities in $\mathcal{H}\mathcal{C}$:

$$\big\{\text{blocks of } L \times L'\big\} \quad = \quad \big\{\text{blocks of } L\big\} * \big\{\text{blocks of } L'\big\} \tag{E2}$$

$$\big\{\text{max. blocks of } L \times L'\big\} \quad = \quad \big\{\text{max. blocks of } L\big\} * \big\{\text{max. blocks of } L'\big\} \; . \tag{E3}$$

Because $\mathcal{C}$ contains all the singletons, each element L of $\mathcal{H}\Omega$ is covered by its blocks. Hence the map sending each element of $\mathcal{H}\Omega$ to its collection of blocks is both a morphism of preordered monoids (Equality (E2)) and the right adjoint to γ (*i.e.* $\gamma \circ \alpha = \mathrm{id}$ and $\mathrm{id} \preccurlyeq \alpha \circ \gamma$).

In the spirit of Galois connections, the right adjoint to γ should provide an abstraction of L that is simpler than L itself. It is not the case if L is abstracted by its whole collection of blocks since the latter contains every singleton contained in L. In view of Equality (E3) it would be natural to take the collection of maximal blocks of L as an abstraction of it. However, without taking precautions, it might be that L has no maximal block. The generic example of that situation is the subset $\mathbb{Z}$ of $\mathbb{R}$ together with the collection $\mathcal{C}$ of finite unions of intervals. To address that issue, we make an additional assumption about the $\mathcal{C}$-blocks: for all $n \in \mathbb{N}$ and all $L \subseteq \Omega^n$,

$$\text{every block of } L \text{ is included in some maximal block of } L. \tag{A'_1}$$

Definition 2.42. Since every singleton is a block, Condition (A'_1) implies that the set of all maximal blocks of $L \in \mathcal{H}\Omega$ covers L. This set is called the *maximal block covering* of L and it is denoted by $\alpha(L)$. Note that $\alpha(L) = \emptyset$ if and only if $L = \emptyset$.

Proposition 2.43. The map $\alpha : \mathcal{H}\Omega \to \mathcal{HC}$ sending L to its maximal block covering is both a morphism of preordered monoids and a right adjoint to the morphism γ which sends a family of blocks to the union of its elements (Remark 2.41). The morphism α induces an isomorphism onto its image, its converse being induced by γ.

Proof. We have $\gamma \circ \alpha = \mathrm{id}$ by Definition 2.42 and $\mathrm{id} \preccurlyeq \alpha \circ \gamma$ by Condition (A'_1). The map α is a morphism of monoids by Equality (E3), and it is preorder preserving by definition of the covering preorder and Equivalence (E1). The last statement is obvious. □

We now prove that the Galois connection and the isomorphism described in Proposition 2.43 are compatible with the quotient maps q_Ω and $q_\mathcal{C}$ from Definition 2.4. Given a homogeneous language $L \in \mathcal{H}_n\mathcal{C}$ and a permutation $\pi \in \mathfrak{S}_n$, one readily has $\pi \cdot \gamma(L) = \gamma(\pi \cdot L)$. Due to the universal property of the quotient map $q_\mathcal{C}$, there exists a unique map $\tilde{\gamma} : \tilde{\mathcal{H}}\mathcal{C} \to \tilde{\mathcal{H}}\Omega$ such that $q_\Omega \circ \gamma = \tilde{\gamma} \circ q_\mathcal{C}$. Given a subset $X \subseteq \Omega^n$ (*i.e.* $L \in \mathcal{H}_n\Omega$) and a permutation $\pi \in \mathfrak{S}_n$, if B is a maximal block of L, then $\pi \cdot B$ is a maximal block of $\pi \cdot L$. Due to the universal property of the quotient map q_Ω, there exists a unique map $\tilde{\alpha} : \tilde{\mathcal{H}}\Omega \to \tilde{\mathcal{H}}\mathcal{C}$ such that $q_\mathcal{C} \circ \alpha = \tilde{\alpha} \circ q_\Omega$.

Theorem 2.44. Both mappings $\tilde{\gamma}$ and $\tilde{\alpha}$ are morphisms of preordered monoids inducing a Galois connection $\tilde{\gamma} \dashv \tilde{\alpha}$. The image of $\tilde{\alpha}$, denoted by $\mathrm{img}(\tilde{\alpha})$, is a consistent submonoid of $\tilde{\mathcal{H}}\mathcal{C}$. The induced maps $\tilde{\gamma} : (\mathrm{img}(\tilde{\alpha}), \tilde{\preccurlyeq}) \to (\tilde{\mathcal{H}}\Omega, \tilde{\subseteq})$ and $\tilde{\alpha} : (\tilde{\mathcal{H}}\Omega, \tilde{\subseteq}) \to (\mathrm{img}(\tilde{\alpha}), \tilde{\preccurlyeq})$ are isomorphisms of preordered (freely commutative) monoids, inverse of each other.

Proof. Both mappings q_Ω and $q_\mathcal{C}$ are surjective morphisms of preordered monoids (Proposition 2.37), so we deduce that both mappings $\tilde{\alpha}$ and $\tilde{\gamma}$ are morphisms of preordered (commutative) monoids from the fact that both γ and α are such morphisms. The mapping α is one-to-one and the mapping γ is onto because $\gamma \circ \alpha = \mathrm{id}$. For the same reason, the restriction of γ to $\mathrm{img}(\alpha)$ is one-to-one, hence bijective. From the following relations (where all the involved maps has to be understood replacing $\mathcal{HC}$ and $\tilde{\mathcal{H}}\mathcal{C}$ by $\mathrm{img}(\alpha)$ and $\mathrm{img}(\tilde{\alpha})$ respectively),

$$
\begin{array}{rclcrcl}
q_\Omega \circ \gamma & = & \tilde{\gamma} \circ q_\mathcal{C} & \qquad & q_\mathcal{C} \circ \alpha & = & \tilde{\alpha} \circ q_\Omega \\
\alpha \circ \gamma & = & \mathrm{id} & & \gamma \circ \alpha & = & \mathrm{id}
\end{array}
$$

we deduce that $\tilde{\alpha} \circ \tilde{\gamma} \circ q_\mathcal{C} = q_\mathcal{C}$ and $\tilde{\gamma} \circ \tilde{\alpha} \circ q_\Omega = q_\Omega$. Therefore $\tilde{\alpha} \circ \tilde{\gamma} = \mathrm{id}$ and $\tilde{\gamma} \circ \tilde{\alpha} = \mathrm{id}$ because both $q_\mathcal{C}$ and q_Ω are onto. Hence $\tilde{\alpha}$ and $\tilde{\gamma}$ are isomorphisms of monoids between $\mathrm{img}(\tilde{\alpha})$ and $\tilde{\mathcal{H}}\Omega$ which is free (Corollary 2.34). To prove the consistency of $\mathrm{img}(\tilde{\alpha})$ it remains to prove that $\tilde{\alpha}$ sends irreducible elements of $\tilde{\mathcal{H}}\Omega$ to irreducible elements of $\tilde{\mathcal{H}}\mathcal{C}$. Let I be an irreducible element of $\tilde{\mathcal{H}}\Omega$ and suppose that $\tilde{\alpha}(I) = H * H'$ in $\tilde{\mathcal{H}}\mathcal{C}$. We have $I = \tilde{\gamma} \circ \tilde{\alpha}(I) = \tilde{\gamma}(H) * \tilde{\gamma}(H')$ and without loss of generality, we can suppose that $\tilde{\gamma}(H')$ is the unit of the monoid $\tilde{\mathcal{H}}\Omega$. Therefore H' is the unit of the monoid $\mathcal{HC}$ because the units of $\tilde{\mathcal{H}}\Omega$ and $\tilde{\mathcal{H}}\mathcal{C}$ are the only elements of zero length. Hence $\tilde{\alpha}(I)$ is irreducible in $\mathcal{HC}$, and the submonoid $\mathrm{img}(\tilde{\alpha})$ is consistent in $\mathcal{HC}$. □

Definition 2.45. The image of $\mathcal{H}^f\mathcal{C}$ under γ is exactly the collection of elements of $\mathcal{H}\Omega$ having a finite block covering, we denote it by $\mathcal{R}\Omega$. It forms a consistent submonoid of

$\mathcal{H}\Omega$. For all $n \in \mathbb{N}$, we denote by $\mathcal{R}_n\Omega$ the collection of elements of $\mathcal{R}\Omega$ of length n. Then observe that $\mathcal{R}_n\Omega$ forms a *semilattice with zero* (or *join-semilattice with zero*), that is to say, from the order theoretic point of view, a poset with a least element 0 in which every pair $\{a, b\}$ has a least upper bound $a \vee b$.

Remark 2.46. Let $\mathcal{C}'$ be the least subset of $\wp(\Omega)$ containing all the elements of $\mathcal{C}$ and closed under binary union. Observe that we have $\gamma(\mathcal{H}^f\mathcal{C}) = \gamma(\mathcal{H}^f\mathcal{C}')$. Hence we can always replace the alphabet $\mathcal{C}$ by the alphabet $\mathcal{C}'$ in Definition 2.45.

The maximal block covering of L contains far less elements than the whole collection of blocks of L. As an extreme example, suppose that Ω is an infinite set and let L be Ω^n. In view of the applications described in Section 1.3, we focus on subsets of Ω^n admitting a finite block covering. However, we have to be careful that the finite block covering property is misleading: a subset of Ω^n satisfying it might have infinitely many maximal blocks. For example, let $\mathcal{C}$ be the collection of intervals of $\mathbb{R}$ of length at most 1 and let L be the hypercube $[0, 2]^n$. One may argue that the preceding example is pathological, but we will see that ill behaved ones also arise in a very natural setting (Section 4). For the moment, we just make another assumption about $\mathcal{C}$-blocks: for all $n \in \mathbb{N}$ and all $L \subseteq \Omega^n$,

$$\text{if } L \text{ has a finite block covering, then it has finitely many maximal blocks.} \qquad (A_2)$$

We also introduce a weaker (yet more relevant) form of Condition (A_1'): for all $n \in \mathbb{N}$ and all $L \subseteq \Omega^n$ *with a finite block covering*:

$$\text{every block of } L \text{ is included in some maximal block of } L. \qquad (A_1)$$

Proposition 2.47. If the collection of blocks satisfies Conditions (A_1) and (A_2) then the maps α and γ from Proposition 2.43 can be restricted to† the preordered monoids $(\mathcal{R}\Omega, \subseteq)$ and $(\mathcal{H}^f\mathcal{C}, \preccurlyeq)$. Those restrictions induce isomorphisms of preordered monoids between $(\mathcal{R}\Omega, \subseteq)$ and its image under α, and a Galois connection $\gamma|_{\mathcal{H}^f\mathcal{C}} \dashv \alpha|_{\mathcal{R}\Omega}$.

Proof. Condition (A_2) exactly states that the restriction of α to $\mathcal{R}\Omega$ is well-defined. The remaining statements immediately derive from Definition 2.42 and Proposition 2.43 which readily adapts to $\mathcal{R}\Omega$. □

Corollary 2.48. If the collection of blocks satisfies Conditions (A_1) and (A_2), then the morphisms of preordered monoids $\tilde{\gamma}$ and $\tilde{\alpha}$ from Theorem 2.44 can be restricted to $\tilde{\mathcal{R}}\Omega$ and $\tilde{\mathcal{H}}^f\mathcal{C}$. These restrictions induce an isomorphism of preordered monoids between $\tilde{\mathcal{R}}\Omega$ and its image under $\tilde{\alpha}$, and a Galois connection $\tilde{\gamma}|_{\tilde{\mathcal{H}}^f\mathcal{C}} \dashv \tilde{\alpha}|_{\tilde{\mathcal{R}}\Omega}$.

Proof. The proof of Theorem 2.44 remains valid for the restrictions $\gamma|_{\tilde{\mathcal{H}}^f\mathcal{C}}$ and $\alpha|_{\tilde{\mathcal{R}}\Omega}$ except that it is built on Proposition 2.47 instead of Proposition 2.43. □

† Strictly speaking, if Condition (A_1') fails, then the map α from Proposition 2.43 is not well-defined.

2.7. *A less generic but drastically more efficient factoring algorithm*

From Corollary 2.48, we obtain the prime decomposition of a (possibly) infinite language L from that of the finite language $\alpha(L)$. The latter decomposition is given by the algorithm from Section 2.4 whose complexity is really poor. However, under the assumption that $\mathcal{C}$ is a Boolean subalgebra of the powerset $\wp(\Omega)$, there exists a drastically more efficient factoring algorithm (Ninin, 2017, Chapter 2). The subsets of Ω^n are identified with the elements of $\mathcal{H}_n\Omega$ (Section 2.6). The *complement* of a subset L of Ω^n is denoted by L^c. In this section, $(I_1,\dots,I_k)$ is a partition of $[1:n]$ into intervals such that $I_i < I_j$ for all $i,j \in [1:k]$ such that $i<j$. We denote the cardinal of I_i by d_i.

Lemma 2.49. For $i \in [1:k]$, let L_i be a subset of Ω^{d_i}. If B is a block of $\Omega^n \setminus L_1\times\cdots\times L_k$, then there exists $i\in[1:k]$ such that

$$\Omega^{d_1+\cdots+d_{i-1}} \times B|_{I_i} \times \Omega^{d_{i+1}+\cdots+d_k} \quad\subseteq\quad \Omega^n \setminus L_1\times\cdots\times L_k\ . \tag{2}$$

Proof. There exists $i\in[1:k]$ such that $B|_{I_i} \subseteq \Omega^{d_i}\setminus L_i$. Otherwise, for each $i\in[1:k]$ we choose a word w_i in $B|_{I_i}\cap L_i$, and form the concatenation $w_1 * \cdots * w_k$ which belongs to $B\cap L_1\times\cdots\times L_k$ thus leading to a contradiction. Hence we have inclusion (2). □

As a Boolean subalgebra of $\wp(\Omega)$, the collection $\mathcal{C}$ contains the element Ω, so it makes sense to define the *support* of a block $B\subseteq\Omega^n$ as the set

$$\mathrm{supp}(B) \quad=\quad \big\{x\in[1:n] \bigm| B|_{\{x\}} \neq \Omega\big\}\ .$$

Theorem 2.50 (Ninin (2017)). Let $L\in\mathcal{H}_n\Omega$ and $\mathcal{F}$ be a block covering of L^c. The finest partition $(X_1,\dots,X_k)$ of $[1:n]$ that is coarser than the family $\{\mathrm{supp}(B)\mid B\in\mathcal{F}\}$ induces a decomposition of $\bar{L}$. Moreover, if $\mathcal{F}$ only contains maximal blocks of L^c, then it induces the prime decomposition of $\bar{L}$.

Proof. By Remark 2.31, it suffices to prove that the inclusion

$$L|_{X_1} * \cdots * L|_{X_k} \quad\subseteq\quad \pi\cdot L \tag{3}$$

holds with π being the permutation associated to the partition $(X_1,\dots,X_k)$. This amounts to proving that any word $w\in\Omega^n$ such that $w|_{X_i}$ belongs to $L|_{X_i}$ for all $i\in[1:k]$ is an element of L. Let d_i be the cardinal of X_i (*i.e.* the length of $L|_{X_i}$) and let $\mathcal{F}_i$ be the set $\big\{B\in\mathcal{F} \bigm| \mathrm{supp}(B)\subseteq X_i\big\}$. First we check the following inclusion:

$$L|_{X_i} \quad\subseteq\quad \Omega^{d_i}\setminus\bigcup_{B\in\mathcal{F}_i} B|_{X_i}\ . \tag{4}$$

The mapping that sends $w\in\Omega^n$ to $w|_{X_i}\in\Omega^{d_i}$ is onto, so the following inclusion holds for any $B\in\mathcal{F}_i$:

$$\Omega^{d_i}\setminus B|_{X_i} \quad\subseteq\quad (\Omega^n\setminus B)\,|_{X_i}\ . \tag{5}$$

Conversely, if w belongs to $\Omega^n\setminus B$, there exists some $x\in[1:n]$ such that the x^{th} letter of w does not belongs to $B|_{\{x\}}$. Since $B\in\mathcal{F}_i$, the index x lies in X_i, and therefore $\omega|_{X_i}$ belongs to $\Omega^{d_i}\setminus B|_{X_i}$. Hence inclusion (5) is actually an equality:

$$\Omega^{d_i}\setminus B|_{X_i} \quad=\quad (\Omega^n\setminus B)\,|_{X_i}\ . \tag{6}$$

By unfolding definitions and applying the standard relation between direct image and set intersection, we have

$$L|_{X_i} \quad = \quad \left(\Big(\bigcup_{B\in\mathcal{F}} B\Big)^c \right)\Big|_{X_i} \quad = \quad \left(\bigcap_{B\in\mathcal{F}} B^c \right)\Big|_{X_i} \quad \subseteq \quad \bigcap_{B\in\mathcal{F}} (B^c)|_{X_i} \ .$$

Observe that if $B \in \mathcal{F}_j$ for $i \neq j$, then $(B^c)|_{X_i} = \Omega^{d_i}$ so we can restrict the indexing set to $\mathcal{F}_i$ and apply relation (6) :

$$\bigcap_{B\in\mathcal{F}} (B^c)|_{X_i} \quad = \quad \bigcap_{B\in\mathcal{F}_i} (B^c)|_{X_i} \quad = \quad \bigcap_{B\in\mathcal{F}_i} \Omega^{d_i} \setminus B|_{X_i} \quad = \quad \Omega^{d_i} \setminus \bigcup_{B\in\mathcal{F}_i} B|_{X_i} \ .$$

We have proven relation (4). Then let $w \in \Omega^n$ be so that for all $i \in [1:k]$, the extracted word $w|_{X_i}$ belongs to $X|_{X_i}$. Given $B \in \mathcal{F}$ there exists some $i \in [1:k]$ such that $B \in \mathcal{F}_i$ hence $w \notin B$. Since the word w does not belong to any element of the covering $\mathcal{F}$ of L^c, it belongs to L. We have proven relation (3). Now suppose that $\mathcal{F}$ only contains maximal blocks of L^c (possibly not all of them) and let $(X'_1, \ldots, X'_{k'})$ be the partition associated to the prime decomposition of L (Corollary 2.35). Let π' be the corresponding permutation. For a given $B \in \mathcal{F}$, the block $\pi' \cdot B$ is included in $\Omega^n \setminus L|_{X'_1} * \cdots * L|_{X'_{k'}}$ therefore by Lemma 2.49, there exists $j \in [1:k']$ such that

$$\Omega^{d'_1+\cdots+d'_{j-1}} \times (\pi' \cdot B)|_{X'_j} \times \Omega^{d'_{j+1}+\cdots+d'_{k'}} \quad \subseteq \quad \Omega^n \setminus L|_{X'_1} * \cdots * L|_{X'_{k'}} \ .$$

Since $\pi' \cdot B$ is maximal, it is actually equal to $\Omega^{d'_1+\cdots+d'_{j-1}} \times (\pi' \cdot B)|_{X'_j} \times \Omega^{d'_{j+1}+\cdots+d'_{k'}}$, hence the support of B is included in X'_j. The finest partition compatible with all the supports of the elements of $\mathcal{F}$, namely $(X_1, \ldots, X_k)$, is thus finer than the partition $(X'_1, \ldots, X'_{k'})$. Moreover, according to the first part of the proof, the partition $(X_1, \ldots, X_k)$ induces a factorization of L. It follows from Corollary 2.35 that the partitions $(X_1, \ldots, X_k)$ and $(X'_1, \ldots, X'_{k'})$ are the same up to term reordering. $\square$

The decomposition of L that is associated to a finite block covering of L^c is thus provided by a mere *union-find* algorithm (Cormen et al., 2009, Chapter 21). The algorithm is illustrated is Examples 1.6, 1.7, and 1.8 In these examples, the alphabets Ω and $\mathcal{C}$ are respectively a compact interval of $\mathbb{R}$ and the collection of finite unions its subintervals.

Remark 2.51. In practical cases, the subsets L of Ω^n that we want to factor are included in the product $\Omega_1 \times \cdots \times \Omega_n$ where for all $x, y \in \{1, \ldots, n\}$, the subsets Ω_x and Ω_y of Ω are disjoint or equal — see Section 1.3. The *padding* and *unpadding* maps

$$B \quad \mapsto \quad (B|_1 \cup \Omega_1^c) \times \cdots \times (B|_x \cup \Omega_x^c) \times \cdots \times (B|_n \cup \Omega_n^c) \qquad \text{(Padding)}$$

$$B \quad \mapsto \quad (B|_1 \cap \Omega_1) \times \cdots \times (B|_x \cap \Omega_x) \times \cdots \times (B|_n \cap \Omega_n) \qquad \text{(Unpadding)}$$

(where for all $x \in \{1, \ldots, n\}$, Ω_x^c and $B|_x$ respectively stands for $\Omega \setminus \Omega_x$ and the restriction of B to $\{x\}$ (Definition 2.30)) turn any block covering of $\Omega_1 \times \cdots \times \Omega_n \setminus L$ into a block covering of $\Omega^n \setminus L$, and vice-versa. These maps preserve maximal block coverings and Theorem 2.50 remains valid defining the *support* of a block $B \subseteq \Omega_1 \times \cdots \times \Omega_n$ as $\{x \in [1:n] \mid B|_x \neq \Omega_x\}$ provided that we consider block coverings of $\Omega_1 \times \cdots \times \Omega_n \setminus L$.

We conclude this section proving that under the assumption that $\mathcal{C}$ is a Boolean subalgebra of the powerset $\wp(\Omega)$, Conditions (A_1) and (A_2) are both satisfied.

Lemma 2.52. Let $\mathcal{F}$ and $\mathcal{F}'$ be finite block coverings of two subsets L and L' of Ω^n containing all the maximal blocks of L and L' respectively. Assume that any block of L (resp. L') is included in a maximal block of L (resp. L'). Then the collection

$$\mathcal{F}'' \quad = \quad \left\{ B \cap B' \;\middle|\; B \in \mathcal{F};\ B' \in \mathcal{F}' \right\}$$

is a finite block covering of $L \cap L'$ containing all its maximal blocks. Moreover, any block of $L \cap L'$ is included in a maximal block.

Proof. A block B'' of $L \cap L'$ is contained in a maximal block B of L and in a maximal block B' of L'. Then $B \cap B'$ is a block of $L \cap L'$ because $\mathcal{C}$ is stable under intersection. Consequently, if B'' is maximal, then it is equal to $B \cap B'$. Since any block of $L \cap L'$ is included in some element of $\mathcal{F}''$, which is finite because so are $\mathcal{F}$ and $\mathcal{F}'$, every block of $L \cap L'$ is included in a maximal block of $L \cap L'$. $\square$

Proposition 2.53. If the collection $\mathcal{C}$ is a Boolean subalgebra of the powerset $\wp(\Omega)$, then the collection of blocks satisfies both Conditions (A_1) and (A_2).

Proof. If B is a block, then by Lemma 2.49, the set B^c has finitely many maximal blocks (they are of the form (2) with $\Omega \setminus B|_{I_i}$ instead of $B|_{I_i}$) and any block of B^c is contained in one of them. Now let $\mathcal{F}$ be a finite block covering of L^c containing all its maximal blocks and such that any block of L^c is included in a maximal one. From de Morgan's law we have $(L \cup B)^c = L^c \cap B^c$. Then denoting the collection of maximal blocks of B^c by $\mathcal{F}'$, we can apply Lemma 2.52 to the block coverings $\mathcal{F}$ and $\mathcal{F}'$ of L^c and B^c. By induction, we have proven that if the subset L of Ω^n has a finite block covering, then its complement has finitely many maximal blocks and each of its blocks is included in a maximal one. The conclusion follows from the usual relation $(L^c)^c = L$. $\square$

2.8. *Application to isothetic regions over the geometric realization of a finite graph*

In practice, Ω is related to the geometric model of a *conservative* program — see Section 1.2 and (Haucourt (2017)). In particular it is a *continuum* (Nadler Jr. (1992)), viz a connected compact metric space, and the powerset $\wp(\Omega)$ is not a tractable Boolean algebra. In view of Sections 1.3, 2.6 and 2.7 we are in search for a convenient substitute for it. All the basic facts we need concerning topology can be found in (Munkres (2000)).

Definition 2.54. A *graph* is an ordered pair (∂^-, ∂^+) of maps from its set A of *arrows* to its set V of *vertices*.

$$A \underset{\partial^+}{\overset{\partial^-}{\rightrightarrows}} V$$

The vertices $\partial^- a$ and $\partial^+ a$ are the *source* and the *target* of the arrow a. An arrow a is said to be *v-ingoing* (resp. *v-outgoing*) when $\partial^+ a = v$ (resp. $\partial^- a = v$), and it is said to be *v-adjacent* when it is v-ingoing or v-outgoing. A graph in which each vertex has finitely many adjacent arrows is said to be *locally finite*. The *size* of a graph is the

$$\cdots \quad \longrightarrow\bullet\longrightarrow\bullet\longrightarrow\bullet\longrightarrow\bullet\longrightarrow\bullet\longrightarrow\bullet\longrightarrow\bullet\longrightarrow \quad \cdots$$

Fig. 7. The totally ordered set $\mathbb{Z}$ seen as a graph.

cardinal of the set $V \cup A$, it may be infinite. A *graph morphism* from $(\partial^-, \partial^+ : A \rightrightarrows V)$ to $(\partial^-, \partial^+ : A' \rightrightarrows V')$ is pair of maps $(f : A \to A', g : V \to V')$ such that $\partial^- \circ f = g \circ \partial^-$ and $\partial^+ \circ f = g \circ \partial^+$. Graphs and their morphisms form the category **Grph**.

Example 2.55. The *n-cycle*, for $n \geqslant 1$, is the graph whose vertices are the elements of the set $\{0, \ldots, n-1\}$ with an arrow from k to $k+1$ modulo n. A graph that is isomorphic to some n-cycle with $n \geqslant 1$ is said to be *cyclic*. The case $n = 0$ also makes sense because the ordered set of integers $(\mathbb{Z}, \leqslant)$ can be seen as a graph whose set of vertices is $\mathbb{Z}$ and that of arrows is $\{(n, n+1) \mid n \in \mathbb{Z}\}$ (Figure 7). When the context is clear, this graph is denoted by $\mathbb{Z}$. A *chain* is a graph that is isomorphic to some connected subgraph of $\mathbb{Z}$. It is said to be *proper* when it is not isomorphic to $\mathbb{Z}$, and *trivial* when it has only one vertex.

Definition 2.56. The *geometric realization* of a graph $G : A \rightrightarrows V$ is the union of the sets V and $A\times]0,1[$ equipped with the greatest topology that makes the following maps continuous, with α ranging through the set of arrows of the graph.

$$\begin{array}{ccc} [0,1] & \to & V \sqcup A\times]0,1[\\ t & \mapsto & \begin{cases} \partial^-\alpha & \text{if } t = 0 \\ (\alpha, t) & \text{if } 0 < t < 1 \\ \partial^+\alpha & \text{if } t = 1 \end{cases} \end{array}$$

Example 2.57. The geometric realization of an n-cycle with $n \geqslant 1$ is homeomorphic to the unit circle, that of a chain is either homeomorphic to $\mathbb{R}$, $\mathbb{R}_+$, $[0,1]$ or $\{0\}$.

Remark 2.58. A *star* is the colimit of a diagram in **Top** made of inclusion maps $f_k : \{0\} \hookrightarrow \mathbb{R}_+ \times \{k\}$ with k ranging in some cardinal κ. Therefore a star is, up to homeomorphism, entirely determined by the cardinal κ which is, by definition, the *degree* of the star. One checks that every point p of the geometric realization $|G|$ of a graph G admits a basis of neighbourhoods whose elements are stars. In particular, the space $|G|$ is *locally simply connected.* Moreover, any star that belongs to such a basis of neighbourhood has the same degree κ, which only depends on the point p. By definition, the cardinal κ is the degree of p, it is denoted by $\deg_G(p)$. As one expects, if $p \in V$, then $\deg_G(p)$ matches the degree of p as a vertex of G, namely $\mathrm{card}\{p\text{-ingoing arrows}\} + \mathrm{card}\{p\text{-outgoing arrows}\}$; otherwise p belongs to $A\times]0,1[$ and its degree is 2. The next four lemmas are standard.

Lemma 2.59. There is a canonical bijection between the connected components of G and that of $|G|$. Hence the latter is connected iff so is the former.

Lemma 2.60. The geometric realization of a graph is Hausdorff.

Lemma 2.61. A graph is finite iff its geometric realization is compact.

Lemma 2.62. A graph is locally finite if and only if its geometric realization is locally compact if and only if its geometric realization is metrizable.

Definition 2.63. When Ω is the geometric realization of some graph G, we let $\mathcal{C}$ be the collection of its connected subsets. Then, the elements of $\mathcal{R}|G|$ (Definition 2.45) are called the *isothetic regions*† over the graph G. In particular, for all $X \in \mathcal{R}_1|G|$, the collection $\alpha_1(X)$ is that of connected components of X. For each isothetic region X, we denote by $\mathcal{R}|G||_X$ (or just $\mathcal{R}X$) the collection of isothetic regions contained in X.

Any increasing union‡ of blocks is a block because any increasing union of connected sets is connected. Therefore, by the Hausdorff maximal principle, the collection of blocks based on $\mathcal{C}$ always satisfies Condition (A_1'). Moreover, it is proven in (Haucourt (2017)) that if the graph G is *finite*, then the collection $\mathcal{R}_1|G|$ is a Boolean subalgebra of $\wp(|G|)$. Hence by Proposition 2.53 we know that $\mathcal{R}_1|G|$ satisfies both Conditions (A_1) and (A_2). On the way we check that the maximal $\mathcal{C}$-blocks of an isothetic region X are the maximal $\mathcal{C}$-blocks of the maximal $\mathcal{R}_1|G|$-blocks of X so the collection $\mathcal{C}$ also satisfies Condition (A_2). Following Remark 2.46, the collection $\mathcal{R}_1|G|$ can thus be chosen as the alphabet $\mathcal{C}$ in Sections 2.6 and 2.7 (it is denoted by $\mathcal{C}'$ in Section 1.3). This substitution has a practical interest: if $X_1, \ldots, X_n$ are 1-dimensional cubical regions, viz disjoint unions of intervals, and for $i \in [1 : n]$ the cubical region X_i has c_i connected components, then the n-dimensional cubical region $X_1 \times \cdots \times X_n$ has $c_1 \times \cdots \times c_n$ maximal $\mathcal{C}$-blocks but only one maximal $\mathcal{R}_1|G|$-block — see Example 1.5.

3. Isothetic regions whose subregions form a Boolean algebra

From the final paragraph of Section 2.8, we know that the isothetic regions over a finite graph provide an ideal framework to apply the results from Sections 2.6 and 2.7. Isothetic regions based on finite graphs cover all the geometric models of conservative programs (Haucourt (2017)). Nevertheless, for theoretical purposes, we would like to go beyond that limitation. Therefore we give an exact characterization, in topological terms, of the class of graphs G such that $\mathcal{R}_1|G|$ form a Boolean subalgebra of $\wp(|G|)$ (Theorem 3.12), which is the right property to apply the efficient factoring algorithm (Section 2.7).

3.1. *Freudenthal extension of a topological space*

We will need to extend certain topological spaces in a way that we now explain. Given K_0 and K_1 compact closed subsets of a space X with $K_0 \supseteq K_1$, each nonempty connected component C_0 of $X \setminus K_0$ is contained in a unique connected component C_1 of $X \setminus K_1$. Therefore we have a functor $\mathcal{K}_X$ sending each closed compact subset of X (ordered by reversed inclusion) to the set of nonempty connected components of its complement in X. The above construction is illustrated on Figure 8.

† The term 'isothetic region' is borrowed from (Preparata and Shamos, 1985, p.329).
‡ The union of a family of sets is said to be *increasing* when that family is a $\subseteq$-chain.

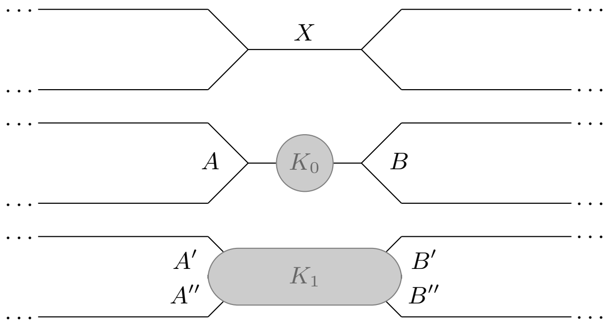

Fig. 8. $K_0 \subseteq K_1$; $A', A'' \mapsto A$; $B', B'' \mapsto B$

Definition 3.1. Following the terminology from (Diestel and Kühn (2003)) the elements of $\lim \mathcal{K}_X$, the (inverse) limit of $\mathcal{K}_X$ in **Set**, are called the *directions* of X. They indeed correspond to the inclusion-reversing maps d from $\mathcal{K}(X)$ to the collection of nonempty subspaces of X such that $d(K)$ is a connected component of $X \setminus K$. The set $X \sqcup \lim \mathcal{K}_X$ is equipped with the topology whose open subsets are those U such that $U \setminus \lim \mathcal{K}_X$ is open in X and for all directions $d \in U$, there exists some compact closed subset $K \subseteq X$ such that $d(K) \subseteq U$. The resulting topological space is denoted by $\mathcal{D}X$ and called the *Freudenthal extension* of X. In the case where $\mathcal{D}X$ is actually compact, it is called the *Freudenthal compactification* of X.

Remark 3.2. According to Definition 3.1, $(X \backslash K) \cup \lim \mathcal{K}_X$ is a neighbourhood of $\lim \mathcal{K}_X$ for all closed compact set K of X. This is why $\mathcal{K}(X)$ only contains closed compact subsets instead of all. Since we mainly consider Hausdorff spaces, the distinction vanishes.

Remark 3.3. Assuming that d and K are respectively a direction and a compact closed subset of X, the closure (in X) of $d(K)$ is not compact. Otherwise $K \cup \mathrm{clo}(d(K))$ would be compact closed and for d is order reversing we would have $d(K \cup \mathrm{clo}(d(K))) \subseteq d(K) \subseteq \mathrm{clo}(d(K))$. But we would also have $d(K \cup \mathrm{clo}(d(K))) \subseteq X \setminus \mathrm{clo}(d(K))$ by definition of a direction. As a consequence, if the connected components of X are compact, then $\lim \mathcal{K}_X = \emptyset$ and $\mathcal{D}X = X$. In particular the Freudenthal extension of a topological space may not be compact.

Example 3.4 (The real line). The Freudenthal compactification of $\mathbb{R}$ is $[0,1]$. Note that $|\mathbb{Z}| = |\{\cdots < -2 < -1 < 0 < 1 < 2 < \cdots\}| \cong \mathbb{R}$ and $|\{0 < 1\}| \cong [0,1]$.

Example 3.5 (The infinite comb). The *infinite comb* is the graph G depicted on Figure 9. Formally, its set of vertices is $\mathbb{Z} \times \{0,1\}$ with one arrow from $(n,0)$ to $(n+1,0)$ (resp. $(n,1)$) for all $n \in \mathbb{Z}$. Its geometric realization can be embedded into the plane.

$$|G| \quad \cong \quad \left(\mathbb{R} \times \{0\} \quad \cup \quad \bigcup_{n \in \mathbb{Z}} \{n\} \times [0,1] \right) \quad \subseteq \quad \mathbb{R}^2$$

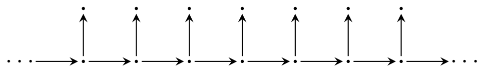

Fig. 9. The infinite comb.

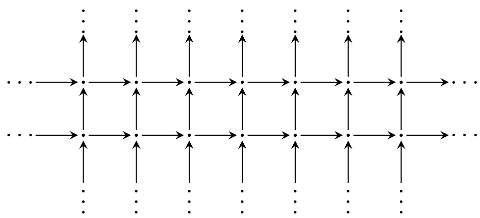

Fig. 10. The infinite grid.

Consequently, the Freudenthal compactification of the (geometric realization of the) *infinite comb* has two directions.

Example 3.6 (One-point compactification). The Freudenthal compactification of $\mathbb{R}^n$ for $n \geqslant 2$ is the $n+1$-dimensional sphere. More generally, if X is a locally compact Hausdorff space in which the complement of any compact subset is connected, then the Freudenthal compactification of X is its Alexandroff compactification (*i.e.* a single point ∞ is added at infinity).

Example 3.7 (The infinite grid). The *infinite grid* is the graph G depicted on Figure 10. Formally, its set of vertices is $\mathbb{Z} \times \{0,1\}$ with one arrow from (a,b) to (c,d) iff $a \leqslant c$, $b \leqslant d$, and $(c-a)+(d-b)=1$. Its geometric realization can be embedded into the plane.

$$|G| \quad \cong \quad \bigcup_{n \in \mathbb{Z}} \big(\, \mathbb{R} \times \{n\} \; \cup \; \{n\} \times \mathbb{R} \, \big) \quad \subseteq \quad \mathbb{R}^2$$

As a consequence of Example 3.6, the Freudenthal compactification of the (geometric realization of the) *infinite grid* has a single direction. It is not locally simply connected since no neighborhood of ∞ is simply connected. In particular, by Remark 2.58, it cannot be the realization of a graph.

Lemma 3.8. Let X be a topological space every compact closed subset K of which is contained in a compact closed subset K' whose complement $X \setminus K'$ is nonempty connected. Then X has a single direction.

Proof. Let f and g be two directions of X, we have $f(K') = g(K')$ because $X \setminus K'$ has a single nonempty connected component. Consequently the connected component of $X \setminus K$ that contains $f(K')$ is also the one that contains $g(K')$. In other words $f(K) = g(K)$. $\square$

Example 3.9. Let p such that $p \notin \mathbb{N}$. Consider the graph G whose set of vertices is $\mathbb{N} \sqcup \{p\}$ with one arrow from p to each element of $\mathbb{N}$. The geometric realization of G is $p \;\sqcup\; \mathbb{N} \times]0,1]$ with the obvious topology. Any compact subspace of $|G|$ is contained in

graph G	$\mathcal{R}_1 G$ stable under complement	counterexample
finite graph	yes	–
$\mathbb{Z}$	yes	–
star with finitely many branches	yes	–
star with infinitely many branches	no	center of the star
infinite comb	no	line
infinite grid	no	all the vertical lines connected by a horizontal one

Fig. 11. Stability of $\mathcal{R}_1 G$ under complement.

$K = p \sqcup S{\times}]0,1]$ for some finite subset S of $\mathbb{N}$. The connected components of $|G| \setminus K$ are the branches $\{n\}{\times}]0,1]$, for $n \in \mathbb{N} \setminus S$, and they are relatively compact in $|G|$. By Remark 3.3, the geometric realization of G has no direction, hence $\mathcal{D}X = X$.

Example 3.10. Let p such that $p \notin \mathbb{N} \times \mathbb{N}$. Consider the graph G whose set of vertices is $\mathbb{N} \sqcup \{p\}$ with one arrow from p to $(n,0)$ and one arrow from (n,k) to $(n,k+1)$ for all $n,k \in \mathbb{N}$. The value n identifies a branch while k indicates how many vertices one finds between the vertices (n,k) and p. The geometric realization of G is $p \sqcup \mathbb{N}{\times}]0,+\infty[$ with the obvious topology. Any compact subspace of $|G|$ is contained in $K = p \sqcup S{\times}]0,r]$ for some finite subset S of $\mathbb{N}$ and $r > 0$. The connected components of $|G| \setminus K$ are of the form $\{n\}{\times}]r,+\infty[$ with r being zero exactly when for $n \in \mathbb{N} \setminus S$. They are *not* relatively compact in $|G|$. Hence each branch of G induces a direction. In fact $\mathcal{D}|G|$ is homemorphic with the geometric realization of the graph described in Example 3.9.

Lemma 3.11. The Freudenthal extension preserves disjoint unions.

3.2. *Graphs such that the collection $\mathcal{R}_1|G|$ forms a Boolean algebra*

The collection $\mathcal{R}_1|G|$ is always stable under binary union. By De Morgan's law, proving that $\mathcal{R}_1|G|$ is a Boolean subalgebra of $\wp(|G|)$ amounts to prove that it is also stable under complement. In Figure 11, we have summarized the properties of $\mathcal{R}_1|G|$ for some graphs G. The following theorem generalizes Lemma 2.61, is the main result of this section.

Theorem 3.12. Given a graph G, the following are equivalent:

1 The collection $\mathcal{R}_1|G|$ is a Boolean subalgebra of $\wp(|G|)$.
2 For all $n \in \mathbb{N}$, the collection $\mathcal{R}_n|G|$ is a Boolean subalgebra of $\wp(|G|^n)$.
3 The graph G has finitely many connected components, all its vertices have finitely

many adjacent arrows, and the degree of all but finitely many vertices is 2. In other words the following sum is finite.

$$\sum_{v \text{ vertex}} |\deg_G(v) - 2| + \#\{\text{connected components}\}$$

4 The graph G can be obtained as a coequalizer of $D \rightrightarrows L$ with D being finite and discrete, and L being a finite disjoint union of points, segments, and half-lines.

5 The Freudenthal compactification of $|G|$ is homeomorphic to the geometric realization of some finite graph.

When the preceding statements are satisfied, the number of directions of $|G|$ is the number of half-lines appearing in L.

Proof. We prove that the first assertion implies the second one. If G has infinitely many connected components then $\mathcal{R}_1 G$ has no greatest element. From now on G is assumed to be connected.

Suppose some vertex v has infinitely many adjacent arrows and let $S \subseteq |G|$ be an open star centered in v. If $|G| \setminus S$ has infinitely many connected components, then S is a connected subset of $|G|$ whose complement does not belong to $\mathcal{R}_1 G$. Otherwise $(|G| \setminus S) \cup \{v\}$ is a finite union of connected components of $|G|$ whose complement, which is the union of infinitely many pairwise disjoint segments, does not belong to $\mathcal{R}_1 G$.

Suppose that there are infinitely many vertices whose number of adjacent arrows is not 2 and let T be a spanning tree of G (*i.e.* a connected subgraph of G containing all the vertices of G such that T loses its connectedness if one removes a single arrow from it). The set theoretic difference $|G| - |T|$ is thus a disconnected union of segments $B \times]0,1[$ with B being a set of arrows of G. If B is infinite then we have a connected component of $|G|$ whose complement in $|G|$ does not belong to $\mathcal{R}_1|G|$. Assume that B is finite. It follows that all the vertices v but finitely many ones have the same neighborhood in T than in G. In particular T has infinitely many vertices whose number of adjacent arrows is not 2. From a general fact about trees we deduce that T has infinitely many vertices whose number of adjacent arrows is at least 3. By an easy induction we build a linear subgraph L of T containing infinitely many vertices with (at least) 3 adjacent arrows. The subgraph $L \subseteq T$ is connected and $|T| - |L|$ has infinitely many connected components (at least one for each vertex of L with at least 3 neighbors in T). We note that any connected component of $|G| - |L|$ is actually obtained as the union of connected components of $|T| - |L|$ related by "bridges", that is to say $B' \times]0,1[$ for a subset $B' \subseteq B$. As B was assumed to be finite $|G| - |L|$ has infinitely many connected components. Whether B is finite or not, the collection $\mathcal{R}_1|G|$ is not a Boolean subalgebra of $2^{|G|}$.

Conversely since G has finitely many connected components, the collection $\mathcal{R}_1 G$ is a Boolean subalgebra of $\wp(|G|)$ iff $\mathcal{R}_1 C$ is a Boolean subalgebra of $\wp(|C|)$ for all connected components of G. We can thus suppose that G is connected. We write $|G|$ as $D \cup D^{\mathrm{c}}$ with

$$D = \{x \in |G| \mid x \text{ admits a neighbourhood that is not isomorphic to } \mathbb{R}\}$$

First remark that D is a discrete subspace of $|G|$. Also, an element of $|G|$ belongs to D iff its degree is not 2. Therefore, by hypothesis, D is finite. Since G is a connected graph, $|G|$ is a connected space. On the contrary D^c (*i.e.* the complement of D in $|G|$) is a disconnected union of copies of $]0,1[$ (according to our description of $|G|$). Let us consider their boundaries in $|G|$. If some of them has an empty boundary, then it is both open and closed, and disconnected from its complement in $|G|$. Therefore $|G| \cong \mathbb{R}$. Now suppose that the boundary of any connected component of D^c contains at least one element. This element belongs to D. For each $v \in D$ and each connected component C of D^c whose boundary contains v, the degree of v is augmented by at least 1 (if $C \cup \{v\} \cong \mathbb{R}_+$) and at most 2 (if $C \cup \{v\} \cong \mathbb{S}^1$). Hence D^c has finitely many connected components, let us say $C_1, \ldots, C_n$. In order to conclude, remark that a finite union of connected components of $|G|$ can be written as a disjoint union of the following form, where $D' \subseteq D$ and X_k is a finite union of intervals of C_k.

$$D' \cup X_1 \cup \cdots \cup X_n$$

Suppose the second point is satisfied. Then consider the graph G' obtained from G as follows: G and G' have the same set of isolated vertices (*i.e.* those with null degree), they share their set of arrows and we have $\partial^- \alpha' = \partial^+ \alpha$ in G' iff the same holds in G and $\deg \partial^+ \alpha = 2$ (in G). As a consequence the degree of a vertex in G' does not exceed 2 so G' is a disjoint union of linear graphs. In particular we have a canonical morphism from G' to G which is entirely defined by the fact that it is the identity map on arrows and vertex of zero degree. Moreover, the number of connected components of G' is finite because it is less than

$$\#\big\{\text{linear connected components of } G\big\} \;+ \sum_{\substack{v \text{ vertex s.t.} \\ \deg_G(v) \neq 2}} \deg_G(v) \;.$$

Some connected component of G' may be a circle or a line, yet both can be obtained as the coequalizer of the form $\{0,1\} \rightrightarrows L''$ with L'' being a segment in the former case, and the disjoint union of two half lines in the latter.

Conversely, consider a graph G obtained as the coequalizer of f and g as in the fourth assertion. The coequalizer morphism induces a map from the connected components of L' onto the connected components of G so there are finitely many of them. Hence the vertices of G are the classes of the least equivalence relation over the vertices of L' that contains the binary relation

$$\big\{(f(x), g(x)) \;\big|\; x \text{ vertex of } L'\big\} \;.$$

Since D is finite there are finitely many classes that are not reduced to a singleton (hence G has finitely many vertices whose degree differs from 2), and each class is finite (hence the degree of each vertex of G is finite).

Suppose that some (and then all) of the first three statements is (are) satisfied. From Lemma 2.62 we know that $|G|$ is locally compact. As a left adjoint, the realization functor preserves the coequalizer given by the third statement so any copy of the half-line in L gives rise to a copy of $\mathbb{R}_+$. Let n be the number of copies of the half-line in L. Given $k \in \mathbb{N}$ consider G_k the coequalizer of $D \rightrightarrows L_k$ with L_k being obtained from L by keeping the k-length initial segment of every half-line of K. Then the sequence

$$|G_0| \subset \cdots \subset |G_k| \subset |G_{k+1}| \subset \cdots$$

forms an exhaustion[†] of $|G|$, and for k sufficiently large, $|G| \setminus |G_k|$ is homeomorphic to n copies of the real line $\mathbb{R}$. It follows that for k sufficiently large, the Freudenthal extension of $|G|$ is homeomorphic to $|G_k|$.

Suppose that the fifth statement is satisfied. Let ϕ be some embedding of $|G|$ into $|G'| \cong \mathcal{F}|G|$ where G' is a finite graph. Following Remark 2.58, one has $\deg(x) = \deg(\phi(x))$ so the degrees of all the points of $|G|$ (and thus all the vertices of G) are finite. Moreover ϕ induces an embedding of the discrete subspace $\{x \in |G| \mid \deg(x) \neq 2\}$ into the discrete subspace $\{x \in |G'| \mid \deg(x) \neq 2\}$ which is finite. Therefore G has finitely many vertices with a degree that differs from 2. By Lemma 3.11, we conclude that $|G|$ (and therefore G) has finitely many connected components. $\square$

Remark 3.13. The finiteness hypothesis on D cannot be dropped from the statement of the fourth point of Theorem 3.12. Consider indeed the case where D is the set of vertices of a line L. Then let the first morphism of the coequalizer diagram be the inclusion of D in L while the second one sends all the elements of D to a single point. The coequalizer is an infinite bouquet of circles, viz an infinite graph with a single vertex.

Definition 3.14. A graph satisfying one (and then all) of the characterizing properties from Theorem 3.12 is said to be *almost finite*. In particular, all the almost finite graphs are countable, yet the converse is obviously false (Examples 3.5 and 3.7).

Theorem 3.12 can be slightly extended.

Definition 3.15. A *generalized Boolean algebra* is a co-unital distributive lattice with a binary operator $\backslash$ satisfying $(x\backslash y) \vee (x \wedge y) = x$ and $(x\backslash y) \wedge y = 0$ (Figure 13).

Remark 3.16. A lattice is said to be *relatively complemented* when for all $a \sqsubseteq b \sqsubseteq c$ there exists some d such that $b \wedge d = a$ and $b \vee d = c$ (Birkhoff, 1967, p.16). If the lattice is distributive, then such an element d is unique due to the fact that in any distributive lattice: if $c \wedge x = c \wedge y$ and $c \vee x = c \vee y$, then $x = y$ (Birkhoff, 1967, Thm.10, p.12). It follows that the generalized Boolean algebras are exactly the relatively complemented distributive lattices with zero. Given a generalized Boolean algebra, the element d is given by $(c\backslash b) \vee a$. Conversely, in a relatively complemented distributive lattice with zero, one has $0 \sqsubseteq x \wedge y \sqsubseteq x$ and $x\backslash y$ is the unique z such that $(x \wedge y) \vee z = x$ and $x \wedge y \wedge z = 0$.

One checks that the difference operator distributes over the meet and the join operators, and that De Morgan's laws have their counterparts for generalized Boolean algebras.

Lemma 3.17. Let $x, y_1, \ldots, y_n$ be elements of a distributive lattice.

— (distributivity) If all the differences $y_i\backslash x$ exist then the differences $(y_1 \vee \cdots \vee y_n)\backslash x$ and $(y_1 \wedge \cdots \wedge y_n)\backslash x$ also exist and they are respectively equal to $(y_1\backslash x) \vee \cdots \vee (y_n\backslash x)$ and $(y_1\backslash x) \wedge \cdots \wedge (y_n\backslash x)$.

[†] An exhaustion is a $\subset$-increasing family of compact subsets that covers the whole space

— (De Morgan laws) If all the differences $x\backslash y_i$ exist then the differences $x\backslash(y_1 \vee \cdots \vee y_n)$ and $x\backslash(y_1 \wedge \cdots \wedge y_n)$ also exist and they are respectively equal to $(x\backslash y_1)\wedge\cdots\wedge(x\backslash y_n)$ and $(x\backslash y_1) \vee \cdots \vee (x\backslash y_n)$.

Corollary 3.18. All the connected components of the graph G are almost finite if and only if $\mathcal{R}_1|G|$ is a *generalized* Boolean algebra.

Proof. Given two isothetic regions X and Y of dimension 1, the difference $X \setminus Y$ involves only finitely many connected components of $|G|$ so it still has finitely many connected components. Conversely, for each connected component G' of the graph G, the Boolean operations on the elements of $\mathcal{R}_1|G'|$ are obtained by taking the operations on subsets of $|G|$ relatively to $|G'|$. □

4. Isothetic regions whose subregions form a distributive lattice with zero

Theorem 3.12 and Corollary 3.18 imply that the factoring algorithm from Section 2.7 only applies to isothetic regions based on disjoint unions of *almost* finite graphs. Such a framework does not allow us to treat the geometric models of *nonconservative* programs. From the theoretical point of view, it remains interesting to determine the class of isothetic regions to which the generic factoring algorithm applies (Section 2.4). So following Section 2.6, we aim at characterizing the graphs whose associated collection of isothetic regions satisfies Condition (A_2), which amounts to knowing whether an isothetic region, viz a finite union of blocks, has finitely many maximal blocks. If we do not make any assumption on the graph G, there may be counter-examples.

Remark 4.1. Consider the case where the geometric realization of G is $\mathbb{R}$ and let A and B be the squares $[0,2]^2$ and $[1,3]^2$. Both A and B are maximal blocks of $A \cup B$, but they are not the only ones. Indeed, the rectangles $[1,2]\times[0,3]$ and $[0,3]\times[1,2]$ are two other maximal blocks of $A\cup B$, and there is no other. There may be infinitely many such unexpected maximal blocks (Example 4.2).

Example 4.2. Let G be the graph with two vertices $\{\bot,\top\}$ and infinitely many arrows from $\bot$ to $\top$: let us say that $G(\bot,\top) = \mathbb{N}$. Note that the subset $\mathbb{N}\times]0,1[$ of $|G|$ has infinitely many connected components and denote it by E. For the sake of readability, for all $n \in \mathbb{N}$, denote the subset $\{n\}\times]0,1[$ of $|G|$ by $]0,1[_n$. By extension, we denote the subsets $\{\bot\}\cup(\{n\}\times]0,1[)$, $(\{n\}\times]0,1[)\cup\{\top\}$, and $\{\bot\}\cup(\{n\}\times]0,1[)\cup\{\top\}$ by $[0,1[_n$, $]0,1]_n$, and $[0,1]_n$. Let I and I' be the sets $\{\bot\}\cup E$ and $E\cup\{\top\}$, and let J and J' be the sets $[0,1[_0$ and $]0,1]_0$. Then define A and B as $I\times J$ and $I'\times J'$, and note that both A and B are maximal blocks of $A\cup B$. The intersection of the first projections of A and B is E. Then for all $n\in\mathbb{N}$, the set $]0,1[_n\times[0,1]_0$ is a maximal block of $A\cup B$. There are even examples with locally finite graphs (Example 4.3).

Example 4.3. The *infinite ladder* is the graph G whose set of vertices is $\mathbb{N}\times\{\bot,\top\}$ with a single arrow from (n,ε) to $(n+1,\varepsilon)$ for $\varepsilon\in\{\bot,\top\}$ and a single arrow from $(n,\bot)$ to $(n,\top)$ for all $n\in\mathbb{N}$ (Figure 12). The *bottom vertices* are, by definition, those of the form $(n,\bot)$. The *top vertices* are defined dually. Denote by $G_\bot$ the subgraph of G containing

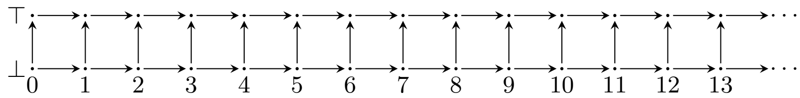

Fig. 12. The infinite ladder

all the bottom vertices together with all the arrows between them. The subgraph $G_\top$ is defined dually. The geometric realization of $G_\bot$, which is homeomorphic to $\mathbb{R}_+$, is seen as the subspace of $|G|$ gathering all the bottom vertices and all the segments $\{\alpha\}\times]0,1[$ such that α is an arrow of $G_\bot$. The subspace $|G| \setminus (|G_\bot| \cup |G_\top|)$, which we denote by E, is made of infinitely many disjoint copies of $]0,1[$. Then Example 4.2 is adapted letting I and I' be $|G_\bot| \cup E$ and $E \cup |G_\top|$.

4.1. *Spanning trees and bouquets of circles*

A point p of a connected topological space X is said to be *separating* when $X \setminus \{p\}$ is no longer connected. An *arc* of a topological space is a subspace that is homeomorphic to the compact unit segment. Such an arc is said to *relate* its extremities (*i.e.* its two non-separating points). A *tree* is a subset of the geometric realization of a graph in which any two distinct points are related by a unique arc. Such a subset is also said to be a *subtree* of the geometric realization of the graph. A connected graph G is a tree in the above sense exactly when for any arrow a of G, the space $|G| \setminus \{a\}\times]0,1[$ is disconnected. In other words the notion of a tree we have just defined generalizes the usual one. A *spanning tree* of a graph is a subtree containing all the vertices of the graph.

Definition 4.4. We say that a connected graph is *almost a tree* when it has a spanning tree containing all its arrows but finitely many. We say that it is *almost a forest* when each of its connected components is almost a tree.

Remark 4.5. Extending the geometric model construction to nonconservative programs requires to deal with isothetic regions based on locally finite trees. Roughly speaking, such trees appear as unfolded control flow graphs, therefore in most cases, they are infinite.

Definition 4.6. A *bouquet of circles* is a graph with a single vertex denoted by $*$. A *topological bouquet of circles* is the geometric realization of a bouquet of circles.

Remark 4.7. Given a bouquet of circles G, the intersection of the isothetic regions

$$\{*\} \cup \bigcup_{\substack{\alpha \\ \text{arrow} \\ \text{of } G}} \{\alpha\}\times]0,\tfrac{1}{2}] \qquad \text{and} \qquad \{*\} \cup \bigcup_{\substack{\alpha \\ \text{arrow} \\ \text{of } G}} \{\alpha\}\times[\tfrac{1}{2},1[$$

is the set $\{*\} \cup \{(\alpha, \frac{1}{2}) \mid \alpha \text{ arrow of } G\}$, which is finite iff the graph G is finite.

This section relies on a standard result from algebraic topology stating that a graph is almost a tree if and only if it is *homotopy equivalent* to a finite bouquet of circles. It is obtained by collapsing a spanning tree of G — see (Munkres, 2000, Theorem 84.7, p.511) or (Hatcher, 2002, Proposition 1A.2, p.84). In particular, *all* the spanning trees of a graph that is almost a tree contain all but finitely many arrows of that graph. The next result is also standard.

Lemma 4.8. The collection of subtrees of a tree is stable under intersection.

Proof. Let T_0 and T_1 be two subtrees of a tree T. Let x and y be two distinct points of $T_0 \cap T_1$. There is an arc α_0 in T_0 and an arc α_1 in T_1 that both relate x and y. In particular α_0 and α_1 are two arcs of the tree T joining x and y, hence they are equal. □

Remark 4.9. Generalizing Remark 4.7, if T is a spanning tree of a connected graph G, then the intersection of the connected sets

$$T \quad\cup\quad (G \setminus T)\times]0, \tfrac{1}{2}] \qquad \text{and} \qquad T \quad\cup \bigcup_{\alpha\in G\setminus T} \{\alpha\} \times [\tfrac{1}{2}, 1[$$

is the set $T \cup \big\{\langle\alpha, \{\frac{1}{2}\}\rangle \bigm| \alpha \in G \setminus T\big\}$ which finite iff the set of arrows $G \setminus T$ is finite.

4.2. *Characterizing isothetic regions with finitely many maximal blocks*

Theorem 4.10. Given a graph G, the following assertions are equivalent:

1 The collection of maximal blocks of any isothetic region over $|G|$ is finite (*i.e.* Condition (A_2) from Section 2.6 is satisfied).
2 For all $n \in \mathbb{N}$, the collection $\mathcal{R}_n|G|$ is stable under intersection (*i.e.* it is a sublattice of that of subsets of $|G|^n$).
3 The collection $\mathcal{R}_1|G|$ is stable under intersection (*i.e.* it is a sublattice of that of subsets of $|G|$).
4 The graph G is almost a forest.

Proof. It is easy to check that the third point implies the second one. The intersection of two n-dimensional isothetic regions X and Y that are respectively covered by the finite families of blocks $\{A_1, \ldots, A_N\}$ and $\{B_1, \ldots, B_M\}$, with $N, M \in \mathbb{N}$, can be written as follows:

$$X \cap Y \quad = \quad \bigcup_{i=1}^{N} \bigcup_{j=1}^{M} A_i \cap B_j .$$

Moreover we have the following standard equality

$$A_i \cap B_j \quad = \quad (\mathrm{proj}_1 A_i \cap \mathrm{proj}_1 B_j) \times \cdots \times (\mathrm{proj}_n A_i \cap \mathrm{proj}_n B_j)$$

from which we deduce that $A_i \cap B_j$ is also an isothetic region. Indeed, by hypothesis, each component of the product here above has finitely many connected components.

Proving that the second point implies the first one is slightly more technical. Let $\langle A', B'\rangle, \langle A_1, B_1\rangle, \ldots, \langle A_N, B_N\rangle$, with $N \in \mathbb{N}$, be elements of the set $\mathcal{R}X \times \mathcal{R}Y$ (Definition 2.63). Suppose that $A' \times B'$ is contained in the union $A_1 \times B_1 \cup \cdots \cup A_N \times B_N$. Given $a \in A'$, we denote by S_a the set of indices $i \in \{1, \ldots, N\}$ such that $a \in A_i$. The "horizontal" slice $\{a\} \times B'$ is thus contained in the union of products $A_i \times B_i$ for $i \in S_a$. In particular the element a belongs to the intersection

$$\bigcap \{A_i \bigm| i \in S_a\}$$

while the set B' is included in the union

$$\bigcup\{B_i \mid i \in S_a\}.$$

Applying the preceding reasoning to all the elements of A, we obtain a finite collection $S_1, \ldots, S_M$ of subsets of $\{1, \ldots, N\}$ such that

$$A' \quad \subseteq \quad \underbrace{\bigcup_{m=1}^{M} \bigcap_{i \in S_m} A_i \times B_i}_{=U} \qquad \text{and} \qquad B' \quad \subseteq \quad \underbrace{\bigcap_{m=1}^{M} \bigcup_{i \in S_m} A_i \times B_i}_{=V} \ . \tag{7}$$

We have proven that $A' \times B'$ is contained in the union $A_1 \times B_1 \cup \cdots \cup A_N \times B_N$ iff there exist finitely many finite subsets $S_1, \ldots, S_M$ satisfying the relations (7). From the hypothesis on the collections $\mathcal{R}_n G$ for $n \in \mathbb{N}$, we readily deduce that both $\mathcal{R}X$ and $\mathcal{R}Y$ are co -unital distributive lattices whose join and meet are given by the set-theoretic union and intersection. Now let $\mathcal{F}$ be the family of all the products $U \times V$ obtained from relations (7) letting $\{S_1, \ldots, S_M\}$ range through the family of all the finite collections of finite subsets of $\{1, \ldots, N\}$. The elements of $\mathcal{F}$ are isothetic regions because $\mathcal{R}X$ and $\mathcal{R}Y$ are stable under intersection. The family $\mathcal{F}$ is finite and by construction, a product $A' \times B'$ with $A' \in \mathcal{R}X$ and $B' \in \mathcal{R}Y$ is included in the isothetic region $A_1 \times B_1 \cup \cdots \cup A_N \times B_N$ if and only if it is included in some element $U \times V$ of $\mathcal{F}$. Therefore the collection of maximal isothetic region of the form $A' \times B'$ where $A' \in \mathcal{R}X$ and $B' \in \mathcal{R}Y$ is finite. Assuming that all the elements of $\mathcal{R}X$ and $\mathcal{R}Y$ have finitely many maximal blocks we deduce from Equality (E3) (Remark 2.41) that any element of $\mathcal{R}(X \times Y)$ has finitely many maximal blocks. By definition, any one-dimensional isothetic region has finitely many maximal blocks, which are actually its connected components. Then letting X and Y be $\mathcal{R}_1 G$ and $\mathcal{R}_n G$, we deduce the first point from an immediate induction on the dimension n.

The proof that the first point implies the third one is a generalization of Examples 4.2 and 4.3: let X and Y be one-dimensional isothetic regions whose intersection has infinitely many connected components. We can suppose that both X and Y are connected. In particular, the union $X \cup Y$ is connected because the intersection $X \cap Y$ is not empty. We conclude noting that for all connected components C of the intersection $X \cap Y$, the block $C \times (X \cup Y)$ is maximal in the isothetic region $X^2 \cup Y^2$. We prove that the fourth point implies the third one. Let T be a spanning tree of a graph G and assume that the set of arrows $G \setminus T$, which we denote by A, is finite. Let X and Y be two connected subsets of $|G|$. Then both $X \setminus A \times]0,1[$ and $Y \setminus A \times]0,1[$ are finite disconnected unions of trees. Moreover both $X \cap A \times]0,1[$ and $Y \cap A \times]0,1[$ are finite disconnected unions of intervals. It follows from Lemma 4.8 that the intersection $X \cap Y$ has finitely many connected components. The converse is given by Remark 4.9. $\square$

5. Higher dimensional isothetic regions from 1-dimensional ones

The set theoretic operations on higher dimensional isothetic regions are derived from those on one-dimensional isothetic regions and the following standard identities where

A, B, C and D are subsets of Ω with respect to which complements are taken:

$$\left.\begin{array}{rcl} A\times(B\cup C) & = & (A\times B)\cup(A\times C) \\ (A\cup B)\times C & = & (A\times C)\cup(B\times C) \end{array}\right\} \qquad (\times \text{ distributes over } \cup)$$

$$(A\times B)\cap(C\times D) \quad = \quad (A\cap C)\times(B\cap D) \qquad (\{\times,\cap\}\text{ - exchange law})$$

$$(A\times B)^c \quad = \quad (A^c\times\Omega)\cup(\Omega\times B^c)\ . \qquad (\times\text{ and complement})$$

The purpose of this section is to formalize this remark by expressing the algebraic structure of n-dimensional isothetic regions as a tensor product of n copies of the algebraic structure of 1-dimensional isothetic regions.

5.1. *The standard approach to tensor product of Boolean algebras*

We expand the short paragraph from (Pierce, 1989, p.840) explaining the relation between tensors product and coproducts of Boolean algebras. A ring R is said to be *idempotent* when $x^2 = x$ holds for all its elements. An algebra is said to be *idempotent* when so is its underlying ring. An idempotent ring R is said to be *Boolean* when it is unital (*i.e.* the product has a neutral element, denoted by 1). Any Boolean ring $(R, +, \cdot, 0, 1)$ is turned into a Boolean algebra by setting $x^c = x+1$, $x\wedge y = x\cdot y$, and $x\vee y = x+y+x\cdot y$. Conversely, any *Boolean algebra* $(A, \vee, \wedge, 0, 1, (\_)^c)$ is turned into a Boolean ring $(A, +, \cdot, 0, 1)$ by setting $a + b = (a \vee b) \wedge (a \wedge b)^c$ and $a \cdot b = a \wedge b$. Those constructions extend to an isomorphism between the category of Boolean algebras and that of Boolean rings, denoted by $\mathbf{BR}$ — see (Johnstone, 1982, p.4-7) or (Givant and Halmos, 2009, p.1-20)

$$\mathbf{BA} \underset{\mathcal{A}}{\overset{\mathcal{R}}{\rightleftarrows}} \mathbf{BR}\ .$$

The way Boolean rings are related to $\mathbb{F}_2$-algebras comes from the specific features implied by idempotency: any idempotent ring has characteristic 2 (*i.e.* $x + x = 0$ holds for all its elements), and any ring of characteristic 2 is commutative (Givant and Halmos, 2009, p.3). Consequently, any idempotent ring can be seen as an idempotent $\mathbb{F}_2$-algebra in a unique way: the unary operators $0_{\mathbb{F}_2}\cdot(\_)$ and $1_{\mathbb{F}_2}\cdot(\_)$ are respectively interpreted as the null map and the identity map. It follows that the category of idempotent $\mathbb{F}_2$-algebras can be identified with that of idempotent rings, which we denote by $\mathbf{IR}$. The finite fields $\mathbb{F}_{2^q}$ with $q \geqslant 2$ provide examples of $\mathbb{F}_2$-algebras that are not idempotent. The *ordinary tensor product* of two commutative algebras over a given field (Lang, 2002, p.629-631) consists of the tensor product of their underlying vector spaces (Lang, 2002, p.601-603) endowed with the unique bilinear product that extends the $\{\otimes, *\}$-*exchange law*

$$(x * x') \otimes (y * y') \quad = \quad (x\otimes y) * (x' \otimes y')\ . \qquad \text{(EL)}$$

In particular, one easily checks that if both components of the ordinary tensor product have a unit, then the pure tensor $1 \otimes 1$ is its unit. The ordinary tensor product of two commutative unital algebras over a given field is actually their coproduct (Lang, 2002, Proposition 6.1, p.630), which initially motivates the construction. In the present context, we are mostly interested in the fact that it also preserves idempotency.

Lemma 5.1. The ordinary tensor product of two idempotent $\mathbb{F}_2$-algebras is idempotent.

Proof. Let A and B be two idempotent $\mathbb{F}_2$-algebras. By definition of the ordinary tensor product and idempotency of A and B, we have the following equalities

$$(a \otimes b) * (a \otimes b) \quad = \quad (a * a) \otimes (b * b) \quad = \quad a \otimes b$$

for all $a \in A$ and all $b \in B$. Any element of the ordinary tensor product is a linear combination of elements of the form $(a \otimes b)$ with coefficients taken in $\mathbb{F}_2$ because the ordinary tensor product of algebras over a field is built on the tensor product of their underlying vector spaces. One can also note that $*$ distributes over $+$ and that any product of pure tensors is a pure tensor because of the $\{\otimes, *\}$-exchange law (EL). The quadratic map $x \mapsto x^2$ being linear in characteristic 2 we deduce that $A \otimes_{\mathbb{F}_2} B$ is idempotent. $\square$

Following Lemma 5.1 and the relation between Boolean rings and Boolean algebras, the *tensor product* of the Boolean algebras A_0 and A_1 is defined as $\mathcal{A}(\mathcal{R}(A_0) \otimes_{\mathbb{F}_2} \mathcal{R}(A_1))$.

Remark 5.2. Coproducts of Boolean algebras can also be obtained through the *Stone duality* (Johnstone, 1982, p.71) providing an isomorphism between the category of Boolean algebras and the dual of the category of compact Hausdorff totally disconnected spaces, or *Stone spaces*. In particular, a *finite* Boolean algebra is related to a *finite* Stone space, in other word to some natural number n. Since a duality exchanges products and coproducts, the monoid of (isomorphism classes) of non-degenerate finite Boolean algebras equipped with the tensor product is actually the multiplicative monoid $(\mathbb{N} \setminus \{0\}, \times, 1)$. This simple remark is a special case of Ketonen's theorem: every countable, commutative semigroup can be embedded in the commutative monoid of (isomorphism classes of) countable Boolean algebras under tensor product. We warn the reader that following the terminology introduced by (Sikorski (1950)), in the original paper (Ketonen (1978)) *tensor product* of Boolean algebras are called *direct product* of Boolean algebras.

5.2. *Universal tensor product*

The terminology used in this section is borrowed from (Borceux, 1994, Chapter 3). A *signature* is a mapping α from a set Θ to $\mathbb{N}$. Each element $\theta \in \Theta$ should be thought of as an operator and $\alpha(\theta)$ as its *arity* (*i.e.* the number of arguments of θ). An *interpretation* of the signature is a set X together with a mapping $\theta_X : X^{\alpha(\theta)} \to X$ for each $\theta \in \Theta$. Given two interpretations X and Y of the same signature, a *morphism of interpretations* from X to Y is a mapping $f : X \to Y$ such that for all $\theta \in \Theta$ and for all $(x_1, \ldots, x_{\alpha(\theta)}) \in X^{\alpha(\theta)}$ the following equality holds:

$$f(\theta_X(x_1, \ldots, x_{\alpha(\theta)})) \quad = \quad \theta_Y(f(x_1), \ldots, f(x_{\alpha(\theta)})) \ .$$

An *algebraic theory* $\mathbb{T}$ is a signature together with a collection of axioms of the form

$$\forall x_1 \ldots \forall x_n \qquad \Phi(x_1, \ldots, x_n) \quad = \quad \Psi(x_1, \ldots, x_n)$$

where Φ and Ψ are terms built on the operators of Θ and whose free variables are in $\{x_1, \ldots, x_n\}$. A *model* of the theory is an interpretation of its signature satisfying

Structure	Signature	Axioms	Category
semilattice	$\vee$	commutative idempotent semigroup	**SL**
semilattice with zero	$\vee$, 0	commutative idempotent monoid	$\mathbf{SL_0}$
lattice	$\vee$, $\wedge$	two semilattices with $x \vee (x \wedge y) = x$ and $x \wedge (x \vee y) = x$	**Lat**
distributive lattice	$\vee$, $\wedge$	lattice in which $\wedge$ distributes over $\vee$	**DL**
distributive lattice with zero	$\vee$, 0, $\wedge$	distributive lattice in which $\vee$ has a neutral element	$\mathbf{DL_0}$
generalized Boolean algebra	$\vee$, 0, $\wedge$, $\backslash$	distributive lattice with zero s.t. $(x \backslash y) \vee (x \wedge y) = x$ and $(x \backslash y) \wedge y = 0$	**GBA**
bounded distributive lattice	$\vee$, 0, $\wedge$, 1	distributive lattice in which both $\vee$ and $\wedge$ have a neutral element	$\mathbf{DL_{01}}$
Boolean algebra	$\vee$, 0, $\wedge$, 1, ${\_}^{c}$	bounded distributive lattice s.t. $x^{c} \wedge x = 0$ and $x^{c} \vee x = 1$	**BA**
	$\vee$, 0, $\wedge$, 1, $\backslash$	generalized Boolean algebra with unit	

Fig. 13. Some extensions of the algebraic theory of semilattices

all its axioms. A *morphism of models* is just a morphism of interpretations between models of the theory. The models of an algebraic theory $\mathbb{T}$ and their morphisms form the complete and cocomplete category $\mathbf{Mdl}_{\mathbb{T}}$ (Borceux, 1994, Theorem 3.4.5, p.138). Most of the objects considered in algebra (*e.g.* semigroups, monoids, groups, rings, modules, algebras, and all their commutative variants) are models of some algebraic theory. Fields provide a noteworthy exception, the reason being that the zero element of a field has no inverse. Figure 13 summarizes the list of theories we will have to deal with. We often write *co-unital distributive lattice* instead of *distributive lattice with zero*.

Definition 5.3. Given A, B and X three models of the same theory, a *bimorphism* from $A \times B$ to X is a mapping $f : A \times B \to X$ such that for all $a \in A$ and for all $b \in B$ the mappings $f(a, \_) : B \to X$ and $f(\_\,, b) : A \to X$ are morphisms. The composite $f \circ g$ of a bimorphism $f : A \times B \to X$ and a morphism $g : X \to Y$ is again a bimorphism. As a consequence, there is a functor $\mathrm{Bim}(A, B)$ from the category of models of the theory to **Set** sending X to the set of bimorphisms from $A \times B$ to X.

An important result about algebraic theories is that the functor $\mathrm{Bim}(A, B)$ is *representable* (Borceux, 1994, Th.3.10.3 p.167-171). In other words there is a (necessarily unique) model $A \otimes B$ such that the functor $\mathrm{Bim}(A, B)$ is isomorphic to $\mathbf{Mdl}_{\mathbb{T}}(A \otimes B, \_)$. It amounts to say that there is a bimorphism $T : A \times B \to A \otimes B$ such that for every bimorphism $F : A \times B \to X$ there is a unique morphism $h \in \mathbf{Mdl}_{\mathbb{T}}(A \otimes B, X)$ such that

$F = h \circ T$. Following the common usage, the elements of the image of T are called the *pure tensors* and for all $(a, b) \in A \times B$, we write $a \otimes b$ instead of $T(a, b)$.

Definition 5.4. The bimorphism T is the *tensor product* of A and B in $\mathbf{Mdl}_{\mathbb{T}}$. We also write *universal tensor product* in the event that another notion of tensor product be under consideration for the objects of $\mathbf{Mdl}_{\mathbb{T}}$ (*e.g.* commutative algebras).

Since the notion of bimorphism dramatically depends on the underlying algebraic category (*i.e.* on the theory modelled by its objects) so does the tensor product.

Example 5.5. Let A and B be two monoids and $f : A \times B \to X$ be a bimorphism of monoids. Since morphisms preserve neutral elements we have $f(\varepsilon_A, b) = f(a, \varepsilon_B) = \varepsilon_X$ for all $a \in A$ and all $b \in B$. If $g : A \times B \to X$ is just a bimorphism of semigroups one may have $a \in A$ and $b \in B$ such that $g(\varepsilon_A, b) \neq g(a, \varepsilon_B)$. See also (Grillet, 1969a, Theorem 2.3). The tensor product of (commutative) semigroups have been introduced and studied by Grillet (1969a,b) with a view to homological algebra.

5.3. *Description of the universal tensor product of semilattices with zero*

A monoid in which the identity $xx = x$ holds for any element x is said to be *idempotent.* A semilattice with zero (or co-unital semilattice) can be seen as an idempotent commutative monoid. Semilattices with zero and their morphisms (*i.e.* the mappings that preserve the join operator $\vee$ and the least element 0) form the category $\mathbf{SL_0}$ (Figure 13). The tensor product in $\mathbf{SL_0}$ is thus a special instance of the construction described in Section 5.2. Given two co-unital semilattices A and B, it is obtained as the quotient of the co-unital semilattice of finite subsets of $A \times B$ by the least congruence $\sim$ satisfying $\{\langle a, b\rangle, \langle a', b\rangle\} \sim \{\langle a \vee a', b\rangle\}$, $\{\langle a, b\rangle, \langle a, b'\rangle\} \sim \{\langle a, b \vee b'\rangle\}$, and $\{\langle \emptyset, b\rangle\} \sim \{\langle a, \emptyset\rangle\} \sim \emptyset$ for all $a, a' \in A$ and all $b, b' \in B$ (Fraser, 1976, Theorem 2.3).

Remark 5.6. The tensor product $A \otimes_{\mathbf{SL}} B$ of two semilattices A and B, where $\mathbf{SL}$ denotes the category of semilattices, can be described as the quotient of the semilattice of *nonempty* finite subsets of $A \times B$ by the least congruence $\sim$ satisfying $\{\langle a, b\rangle, \langle a', b\rangle\} \sim \{\langle a \vee a', b\rangle\}$, and $\{\langle a, b\rangle, \langle a, b'\rangle\} \sim \{\langle a, b \vee b'\rangle\}$ for all $a, a' \in A$ and all $b, b' \in B$. In particular, even if A admits a zero element 0_A, the pure tensors $\langle 0_A, b\rangle$ and $\langle 0_A, b'\rangle$ are *not* identified. In the subsequent section, we consider fields of sets, and the pure tensors $\langle a, b\rangle$ are meant to represent set-theoretic cartesian products $a \times b$. In that context, when one of the sets a and b is empty, it is natural to identify $\langle a, b\rangle$ with the empty set. This is why we consider tensor products in $\mathbf{SL_0}$ instead of $\mathbf{SL}$. Yet, both are related by the natural isomorphism

$$A \otimes_{\mathbf{SL_0}} B \quad \cong \quad \big(A \setminus \{0_A\} \otimes_{\mathbf{SL}} B \setminus \{0_B\}\big)^o$$

where A and B are co-unital semilattices, and $(\_)^o : \mathbf{SL} \to \mathbf{SL_0}$ is the functor which adds a zero element to every semilattice. It is a special case of (Grillet, 1969a, Theorem 6.2).

The next example emphasizes how the presence of a zero is important in the behaviour of universal tensor products, it also illustrates Remark 5.6.

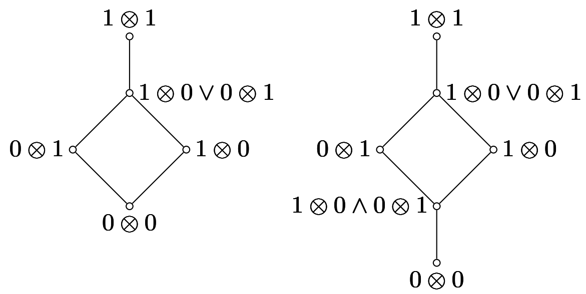

Fig. 14. The Hasse diagrams of $B \otimes_{\mathbf{SL}} B$ and $B \otimes_{\mathbf{DL}} B$

Example 5.7 (Haucourt and Ninin (2014)). Let B be the Boolean algebra $\{0, 1\}$ with $x \wedge y = \min(x, y)$, $x \vee y = \max(x, y)$, and $x^c = x + 1 \mod 2$. We determine $B \otimes B$ in $\mathbf{SL_0}$, $\mathbf{DL_0}$, $\mathbf{SL}$, and $\mathbf{DL}$. Let X be a co-unital semilattice and $T : B \times X \to X$ be the bimorphism of $\mathbf{SL_0}$ defined by $T(0, \_) = 0$ and $T(1, x) = x$ for all $x \in X$. Any bimorphism F of $\mathbf{SL_0}$ satisfies $F(0, \_) = 0$ therefore $f = F(1, \_)$ is the only morphism of $\mathbf{SL_0}$ satisfying the equation $F = f \circ T$. In other words $B \otimes_{\mathbf{SL_0}} X \cong X$. We now prove that the tensor product $B \otimes_{\mathbf{SL}} B$ is the bounded distributive lattice C whose corresponding Hasse diagram is depicted on Figure 14. Observe that any order-preserving map defined over B also preserves binary joins and meets, from which we deduce that a set map $F : \{0, 1\}^2 \to X$ is a bimorphism of $\mathbf{SL}$ iff it satisfies the relations

$$F(0,0) \sqsubseteq F(0,1) \sqsubseteq F(1,1) \qquad \text{and} \qquad F(0,0) \sqsubseteq F(1,0) \sqsubseteq F(1,1) \ .$$

Let us check that there is a unique $h \in \mathbf{SL}(C, X)$ satisfying $F = h \circ T$. Firstly, we have $h(a \otimes b) = F(a, b)$ for all $(a, b) \in \{0, 1\}^2$. Because h is a morphism of $\mathbf{SL}$ it preserves binary join, therefore it satisfies

$$h(0 \otimes 1 \vee 1 \otimes 0) \quad = \quad h(0 \otimes 1) \vee h(1 \otimes 0) \quad = \quad F(0,1) \vee F(1,0)$$

so h is uniquely defined. Checking that h is indeed a morphism of $\mathbf{SL}$ is a routine verification based on the previous relation and the fact that F is a bimorphism. Note that h might not preserve existing meets. A similar reasoning proves that the tensor product $B \otimes B$ in $\mathbf{DL}$ is the bounded distributive lattice whose corresponding Hasse diagram is depicted on Figure 14.

Remark 5.8. Grillet (1969a) presents the tensor product of two semigroups A and B as the quotient of the semigroup of *nonempty* words on the set $A \times B$ under the quotient generated by the relations

$$(a, bb') \sim (a, b)(a, b') \quad \text{and} \quad (aa', b) \sim (a, b)(a', b)$$

for all $a, a' \in A$ and all $b, b' \in B$. Denoting the category of (commutative) semigroups by ($\mathbf{CSG}$) $\mathbf{SG}$, we have the chain of inclusion functors

$$\mathbf{SL_0} \hookrightarrow \mathbf{SL} \hookrightarrow \mathbf{CSG} \hookrightarrow \mathbf{SG} \ .$$

Each of them has a left adjoint which is easily described in terms of presentations of semigroups. In order to relate the presentation of the tensor product of semigroups (Grillet, 1969a, Theorem 2.1) to that co-unital semilattices, it would suffice to know how these left adjoint functors behave with respect to tensor products. For all semigroups A and B we have $C(A \otimes_{\mathbf{SG}} B) \cong C(A) \otimes_{\mathbf{CSG}} C(B)$ where C is the left adjoint to the inclusion functor $\mathbf{CSG} \hookrightarrow \mathbf{SG}$ (Grillet, 1969b, Theorem 1.1). Moreover we have $A^o \otimes_{\mathbf{SL_0}} B^o \cong \big(A \otimes_{\mathbf{SL}} B\big)^o$ for all semilattices A and B (Remark 5.6). To conclude, we would need a similar result for the left adjoint $E : \mathbf{CSG} \to \mathbf{SL}$. One step in that direction is given by (Grillet, 1969b, p.282): if A is a commutative semigroup, then $E(A)$ is isomorphic to $1 \otimes_{\mathbf{CSG}} A$. Because the tensor product in $\mathbf{CSG}$ is commutative and associative[†] (Grillet, 1969b, Proposition 2.1), we have the following isomorphisms:

$$\begin{aligned} E(A \otimes_{\mathbf{CSG}} B) &\cong 1 \otimes_{\mathbf{CSG}} (A \otimes_{\mathbf{CSG}} B) \\ &\cong (1 \otimes_{\mathbf{CSG}} 1) \otimes_{\mathbf{CSG}} (A \otimes_{\mathbf{CSG}} B) \\ &\cong (1 \otimes_{\mathbf{CSG}} A) \otimes_{\mathbf{CSG}} (1 \otimes_{\mathbf{CSG}} B) \\ &\cong E(A) \otimes_{\mathbf{CSG}} E(B)\ . \end{aligned}$$

Applying the universal properties of the tensor products in $\mathbf{CSG}$ and $\mathbf{SL}$ we deduce that $E(A) \otimes_{\mathbf{CSG}} E(B) \cong E(A) \otimes_{\mathbf{SL}} E(B)$. In the end we obtain the following isomorphism:

$$A \otimes_{\mathbf{SL_0}} B \quad\cong\quad \Big(E\big(C(A \setminus \{0_A\} \otimes_{\mathbf{SG}} B \setminus \{0_B\})\big)\Big)^o\ .$$

5.4. *Universal tensor products of (generalized) Boolean algebras in the category* $\mathbf{SL_0}$

The purpose of this section is to prove that the universal tensor product of two Boolean algebras in the category $\mathbf{SL_0}$ is their ordinary tensor product (as defined in Section 5.1). In spite of the large amount of resources dealing with tensor product of (semi)lattices and the even much larger amount of resources dealing with Boolean algebras, we have not been able to find this result in the literature, so we write down the details. Yet, most of the content is already known and we give a reference each time we have been able to find some. Lemmas 5.9, 5.10, and 5.11 are drawn from (Fraser (1976)) which actually deals with tensor product in $\mathbf{SL}$, but following Remark 5.6, one readily adapts the needed results to the tensor product in $\mathbf{SL_0}$. The first lemma provides a simple characterization of the partial order associated with the tensor product. The second one is a generalized form of exchange law between the pure tensors and the meet operator of $A \otimes_{\mathbf{SL_0}} B$, which is indeed a co-unital distributive lattice. The last one relates universal tensor products of bounded distributive lattices in $\mathbf{SL_0}$ to their coproduct in $\mathbf{DL_{01}}$.

Lemma 5.9 (Fraser (1976), Theorem 2.5). Let A, B be distributive lattices with zero. For $n \in \mathbb{N}$, given $a, a_1, \ldots, a_n \in A \setminus \{0_A\}$ and $b, b_1, \ldots, b_n \in B \setminus \{0_B\}$, we have

$$a \otimes_{\mathbf{SL_0}} b \quad\sqsubseteq\quad \bigvee_{i=1}^{n} a_i \otimes_{\mathbf{SL_0}} b_i$$

[†] However, the tensor in **SG** is not associative (Grillet, 1969a, Example 2.3).

iff there are finitely many subsets $S_1, \ldots, S_m$ of $\{1, \ldots, n\}$ such that

$$a \sqsubseteq \bigvee_{j=1}^{m} \bigwedge_{i \in S_j} a_i \quad \text{and} \quad b \sqsubseteq \bigwedge_{j=1}^{m} \bigvee_{i \in S_j} b_i \ .$$

Lemma 5.10 (Fraser (1976), Theorem 2.6). The universal tensor product in $\mathbf{SL_0}$ of two co-unital distributive lattices is a co-unital distributive lattice in which the relation (GEL) holds for all $a_1, \ldots, a_m, c_1, \ldots, c_n \in A$ and all $b_1, \ldots, b_m, d_1, \ldots, d_n \in B$.

$$\left(\bigvee_{i=1}^{m} a_i \otimes_{\mathbf{SL_0}} b_i\right) \wedge \left(\bigvee_{j=1}^{n} c_j \otimes_{\mathbf{SL_0}} d_j\right) \quad = \quad \bigvee_{i=1}^{m} \bigvee_{j=1}^{n} (a_i \wedge c_j) \otimes_{\mathbf{SL_0}} (b_i \wedge d_j) \qquad \text{(GEL)}$$

Lemma 5.11 (Fraser (1976), Theorem 3.3). The universal tensor product (in $\mathbf{SL_0}$) of two bounded distributive lattices is their coproduct (in $\mathbf{DL_{01}}$).

The arguments from Example 5.7 also prove that $B \otimes_{\mathbf{DL_0}} X \cong X$ holds for every co -unital distributive lattices X. However, the universal tensor product of two co-unital distributive lattices in $\mathbf{SL_0}$ differ, in general, from their universal tensor product in $\mathbf{DL_0}$ (Fraser, 1976b, p.182). We extend the main result of (Haucourt and Ninin (2014)).

Proposition 5.12. The class of (generalized) Boolean algebras is closed under universal tensor product in the category of co-unital semilattices.

Proof. Suppose that A and B are two (generalized) Boolean algebras. By Lemma 5.10 we know that the tensor product of two (generalized) Boolean algebras in $\mathbf{SL_0}$ is a distributive lattice with zero. Given $a \in A$ and $b \in B$ we have the equalities

$$a \otimes b \quad = \quad (0 \otimes b) \vee (a \otimes b) \quad = \quad (0 \otimes 0) \vee (0 \otimes b) \vee (a \otimes b) \quad = \quad (0 \otimes 0) \vee (a \otimes b)$$

which proves that $0 \otimes 0$ is the zero of $A \otimes B$. The $\{\otimes, \wedge\}$-exchange law is given by Lemma 5.10. In particular, if both A and B have a unit, then its unit is $1_A \otimes 1_B$. To prove that the tensor product $A \otimes B$ is a (generalized) Boolean algebra, it remains to check that the difference operator is well-defined. First we prove that the difference between pure tensors exists and satisfies the relation

$$(a \otimes b) \backslash (c \otimes d) \quad = \quad \big((a \backslash c) \otimes b\big) \vee \big(a \otimes (b \backslash d)\big) \ .$$

The general case will follow from Lemma 3.17. We evaluate the expression

$$(a \backslash c \otimes b) \qquad \vee \qquad \underbrace{(a \otimes b \backslash d)}_{(a \backslash c \,\otimes\, b \backslash d) \,\vee\, (a \wedge c \,\otimes\, b \backslash d)} \qquad \vee \qquad \underbrace{(a \otimes b \wedge c \otimes d)}_{(a \wedge c \,\otimes\, b \wedge d)} \ .$$

The third term is rewritten applying the exchange law. The mapping $x \mapsto x \otimes b \backslash d$ preserves joins and a can be written as $a \backslash c \vee (a \wedge c)$ so the second term can be rewritten as above. By the same arguments we gather the first two terms and the last two ones in the expression

$$\underbrace{(a \backslash c \otimes b) \quad \vee \quad (a \backslash c \otimes b \backslash d)}_{a \backslash c \otimes b} \quad \vee \quad \underbrace{(a \wedge c \otimes b \backslash d) \quad \vee \quad (a \wedge c \otimes b \wedge d)}_{a \wedge c \otimes b} \ .$$

The original expression thus boils down to $(a\backslash c \otimes b) \vee (a \wedge c \otimes b) = a \otimes b$. We also need to evaluate the expression

$$\big((a\backslash c \otimes b)\ \vee (a \otimes b\backslash d)\big)\ \wedge\ \underbrace{(a \otimes b \wedge c \otimes d)}_{(a\wedge c\ \otimes\ b\wedge d)}\ .$$

By distributivity of $\wedge$ over $\vee$ and the exchange law it can be expressed as follows:

$$\big(\underbrace{(a\backslash c \wedge a \wedge c)}_{0} \otimes b \wedge d\big)\ \ \vee\ \ \big(a \wedge c \otimes \underbrace{(b\backslash d \wedge b \wedge d)}_{0}\big)\ .$$

Therefore it is reduced to $(0 \otimes b \wedge d)\ \vee\ (a \wedge c \otimes 0)$. Since tensor products are taken in $\mathbf{SL_0}$ the mappings $x \mapsto x \otimes b \wedge d$ and $y \mapsto a \wedge c \otimes y$ preserve zero, so the last expression is zero. We emphasize that the last argument does not hold in $\mathbf{SL}$ (Remark 5.6). It is the only place in the proof where this subtlety indeed matters, however, as shown by Example 5.7, it is crucial. $\square$

Corollary 5.13. The universal tensor product in $\mathbf{SL_0}$ of two Boolean algebras is their coproduct in $\mathbf{BA}$.

Proof. Given the Boolean algebras A and B, the universal tensor product $A \otimes_{\mathbf{SL_0}} B$ is a Boolean algebra (Proposition 5.12). From Lemma 5.11 we deduce that $A \otimes_{\mathbf{SL_0}} B$ is the coproduct of A and B in the category $\mathbf{DL_{01}}$. Since $\mathbf{BA}$ is a full subcategory of $\mathbf{DL_{01}}$, the universal tensor product $A \otimes_{\mathbf{SL_0}} B$ is also the coproduct of A and B in $\mathbf{BA}$. $\square$

5.5. *Application to isothetic regions*

Let X and Y be two isothetic regions over the graph G. Given two finite subsets C and C' of $\mathcal{R}X \times \mathcal{R}Y$ such that $C \sim C'$ (Section 5.3), we readily have the equality

$$\bigcup_{(A,B)\in C} A \times B \qquad = \qquad \bigcup_{(A',B')\in C'} A' \times B'\ .$$

Consequently, by definition of $\mathcal{R}(X \times Y)$, the map $\Phi_{\mathrm{X,Y}}$ which sends a finite subset C of $\mathcal{R}X \times \mathcal{R}Y$ to the set theoretic union of its elements induces a morphism of co-unital semilattices from $\mathcal{R}X \otimes_{\mathbf{SL_0}} \mathcal{R}Y$ onto $\mathcal{R}(X \times Y)$.

Proposition 5.14. If G is almost a forest (resp. almost finite), then $\Phi_{X,Y}$ is an isomorphism of co-unital distributive lattices (resp. Boolean algebras).

Proof. The map $\Phi_{X,Y}$ is one-to-one (hence an isomorphism) precisely when for all $A' \in \mathcal{R}X$ and all $B' \in \mathcal{R}Y$, the inclusion

$$A' \times B' \qquad \subseteq \qquad \bigcup_{(A,B)\in C} A \times B$$

implies that $C \sim C \cup \{(A', B')\}$. Assuming that $C = \{\langle A_1, B_1\rangle, \ldots, \langle A_N, B_N\rangle\}$ and with the notation from Section 5.3, the latter is equivalent to the relation

$$A' \otimes B' \quad \sqsubseteq \quad \bigvee_{i=1}^{N} A_i \otimes B_i\ . \tag{8}$$

From the proof that the second point of Theorem 4.10 implies the first one, we deduce that there exists a finite collection $S_1, \ldots, S_M$ of subsets of $\{1, \ldots, N\}$ such that the inclusions (7) are satisfied. The relation (8) then follows from Lemma 5.9. Moreover, with the help of Lemma 5.10, a routine computation proves that the mapping $\Phi_{X,Y}$ is also meet preserving, therefore it is an isomorphism of co-unital distributive lattices. In particular, assuming that G is almost finite (Definition 3.14), the source and the target of the morphism $\Phi_{X,Y}$ are Boolean algebras. Since we readily have $\Phi_{X,Y}(X \times Y) = X \otimes Y$, the mapping $\Phi_{X,Y}$ is an isomorphism of Boolean algebras. □

6. Future work

So far we have considered n-dimensional isothetic regions as *subsets* of $|G|^n$ for some fixed graph G. Yet, the mathematical structure of $|G|^n$ is much richer than that of a mere set, so any isothetic region inherits from that structure. Moreover, these *enriched* isothetic regions have *invariants* that are worth to study. In all the cases that we consider here, the mathematical objects lie in a Cartesian (resp. co-Cartesian) category and the way they are built preserves Cartesian (resp. co-Cartesian) products.

Locally ordered length metric spaces. Each isothetic region X is a sub-locally ordered space $S(X)$ of $|G|^n$ for a unique $n \in \mathbb{N}$, the dimension of X (Haucourt (2017)). Given a tuple $X_1, \ldots, X_k$ of isothetic regions over $|G|$, the induced locally ordered space $S(X_1 * \cdots * X_k)$ is isomorphic to the product of locally ordered spaces $S(X_1) \times \cdots \times S(X_k)$ because the Cartesian product (in any category) is associative. Indeed, the later assertion guarantees that the map which sends an element of $X_1 \times \cdots \times X_k$ (*i.e.* a k-word whose i^{th} letter is a n_i-word on $|G|$) to the concatenation of its letters is actually an isomorphism of locally ordered spaces. It follows that any factorization of an isothetic region in the sense of Section 2.1 induces a factorization of the corresponding locally ordered space. So we wonder if the prime decomposition of an isothetic region implies the prime decomposition of the corresponding locally ordered space. To deal with that question, we lean on two facts. The first one is that the underlying topological space of $S(X)$ is actually a *length metric space* in a natural way (Haucourt (2017)). The second one is a unique decomposition theorem for *geodesic metric spaces* of finite affine rank (Foertsch and Lytchak (2008)). A *length space* is a metric space in which the distance between two points is the infimum of the lengths of the paths joining them (Bridson and Haefliger, 1999, Chap.I.3). Length spaces raise a technical problem: in general, a subspace of a length space is not a length space. That issue is addressed by a standard construction: any metric space (X, d) is associated with a length space by replacing the induced metrics by the so-called *intrinsic metrics*, defined for all $x, y \in X$ by

$$d'(x, y) = \inf\{\ell(\alpha) \mid \alpha \text{ paths joining } x \text{ and } y\} \ .$$

If no such path exists (*i.e.* when the points do not lie in the same path-connected component) the distance between them is conventionally defined as infinite. More specifically, a length space is said to be *geodesic* when any couple of its points are connected by a path whose length is the distance between these points (Bridson and Haefliger, 1999, p.9). The geometric realization of any graph is known to admit a canonical metric structure assuming that any arrow is of length 1. Such metric spaces are called *metric graphs* (Bridson and Haefliger, 1999, Chap.1) and their are geodesic Polish† spaces. Given two metric spaces (X, d_X) and (Y, d_Y) the topological space $X \times Y$ is equipped with the *max-metric* $d_{X \times Y} = \max(d_X, d_Y)$. If both Y and Y are geodesic (Polish) spaces, then so is the max-metric space $X \times Y$. In particular $|G|^n$ is a geodesic Polish space. Consequently, every isothetic region X is a separable length space as a subspace of some geodesic Polish space of the form $|G|^n$. For such spaces, the affine rank of X boils down to the greatest $d \in \mathbb{N}$ such that the hypercube $[0,1]^d$ can be embedded in it, therefore it is bounded by n. However, the metric spaces thus associated to isothetic regions may not be geodesic (*e.g.* the punctured Euclidean plane $\mathbb{R}^2 \setminus \{0\}$). Moreover, due to the Hopf-Rinow Theorem (Bridson and Haefliger, 1999, p.35), that situation is very likely to happen when the isothetic region under consideration is not a closed subset of $|G|^n$, which is the case for many of the geometric models of concurrent programs (Haucourt (2017)). Therefore, we would first need to prove that the unique decomposition theorem from (Foertsch and Lytchak (2008)) remains valid for length spaces instead of geodesic ones.

Finite Connected Loop-Free Categories. A category $\mathcal{C}$ is said to be *loop-free* when for all objects x and y of $\mathcal{C}$, if both homsets $\mathcal{C}(x,y)$ and $\mathcal{C}(y,x)$ are nonempty, then $x = y$ and $\mathcal{C}(x,x) = \{\mathrm{id}_x\}$. The collection of isomorphism classes of nonempty finite connected loop-free categories is a countable commutative monoid $\mathcal{M}$ under Cartesian product. It is reduced and graded by the number of morphisms of its elements. Theorem 6.1 actually generalizes a result on finite posets which is itself a direct consequence of a refinement theorem for products of posets (Hashimoto (1951)).

Theorem 6.1 (Balabonski (2007)). The monoid $\mathcal{M}$ is freely commutative.

Remark 6.2. The connectedness hypothesis cannot be dropped. Indeed we obtain a counter-example by interpreting the product and the sum of monomials as the Cartesian product and the coproduct of categories, and then by substituting the category $\{\cdot \to \cdot\}$ to X in Hashimoto's polynomial $X^5 + X^4 + X^3 + X^2 + X + 1$ (Example 2.20).

The preceding digression is related to isothetic regions through their *fundamental categories* (Fajstrup et al., 2016, Section 4.2.3) and their *categories of components* (Fajstrup et al., 2016, Chapter 6), the latter being properly defined only when the isothetic region under consideration is loop-free:

Definition 6.3. An isothetic region X is said to be *loop-free* (resp. *diconnected*) when its fundamental category $\overrightarrow{\pi_1}X$ is loop-free (resp. connected).

Remark 6.4. An isothetic region has a natural structure of topological space and it is rather easy to prove that if it is connected, then it is also arc-connected. However it

† complete separable metric

might not be diconnected as it can be observed on the 2-dimensional cubical region

$$X \quad = \quad]0,1[\times]1,2[\quad \cup \quad \{(1,1)\} \quad \cup \quad]1,2[\times]0,1[\ .$$

Indeed, one can always relate a point of the square $]0,1[\times]1,2[$ to a point of the square $]1,2[\times]0,1[$ via a continuous path which necessarily goes trough the point $(1,1)$. Yet, no such path can be written as a *finite* concatenation of increasing paths and decreasing paths on X (which inherits from the product order on $\mathbb{R}^2$). The fundamental category of X is indeed the disjoint union of two copies of the posetal category $(]0,1[,\leqslant)\times(]0,1[,\leqslant)$ and a copy of the terminal category†.

The category of components is an invariant actually defined for any loop-free category $\mathcal{C}$, and it is denoted by $\overrightarrow{\pi_0}\mathcal{C}$. The category of components of a loop-free isothetic region X is therefore defined as the category of components of its fundamental category, namely $\overrightarrow{\pi_0}(\overrightarrow{\pi_1}X)$. It is expected to be finite even though this has not been formally proven yet.

One easily checks that given two equivalent isothetic regions (considered as homogeneous languages (Sections 2.1 and 2.6)), one is loop-free (resp. diconnected) iff so is the other. One also easily checks that a Cartesian product of *non-empty* isothetic regions is loop-free (resp. diconnected) iff so are all the operands of the product. In other words the collection of diconnected loop-free isothetic regions forms a *pure* submonoid of that of isothetic regions. As a consequence we have:

Proposition 6.5. The commutative monoid of nonempty diconnected loop-free isothetic regions is free.

The fundamental category construction preserves Cartesian products for the same reasons that the fundamental groupoid construction does (Brown (2006)). The category of components construction also preserves binary Cartesian product but it is not so obvious (Haucourt, 2016, Section 8.5). Then we have the morphism of commutative monoids

$$\{\text{diconnected loop-free regions}\} \xrightarrow{\overrightarrow{\pi_0}\circ\overrightarrow{\pi_1}} \{\text{finite connected loop-free categories}\} \ .$$

In addition, the category of components of a loop-free category $\mathcal{C}$ is reduced to a single morphism iff $\mathcal{C}$ is isomorphic to a lattice (Fajstrup et al., 2016, p.112). Therefore the kernel of the morphism $\overrightarrow{\pi_0}\circ\overrightarrow{\pi_1}$ is not null; it is natural to try to characterize it:

Conjecture 6.6. If X is a prime connected loop-free region whose fundamental category is not isomorphic to a lattice, then $\overrightarrow{\pi_0}(\overrightarrow{\pi_1}X)$ is prime.

A partial answer is given by (Ninin, 2017, Chapter 3) for a restricted collection of cubical regions X and for a canonical overapproximation of $\overrightarrow{\pi_0}(\overrightarrow{\pi_1}X)$.

† A category is said to be *posetal* when it is loop-free and when there is at most one morphism from an object to another. The *terminal* category has a single object and a single morphism.

References

Anderson, D. F. and Valdes-Leon, S. Factorization in Commutative Rings with Zero Divisors, II. In Anderson, D. (ed.), *Factorization in Integral Domains*, Lecture Notes in Pure and Applied Mathematics, pp. 197–219. Marcel Dekker, New York, 1997.

Allen, F. E. Control Flow Analysis. In *Proceedings of a Symposium on Compiler Optimization*, pages 1–19, New York, NY, USA, 1970. ACM.

Balabonski, T. Concurrence, géométrie, et factorisation de catégories. Master's thesis, École Normale Supérieure de Lyon, 2007.

Balabonski, T. and Haucourt, E. A Geometric Approach to the problem of Unique Decomposition of Processes. In *Concurrency Theory 21th International Conference*, vol. 6269 of *Lecture Notes in Computer Science*, pages 132–146. Springer, 2010.

Birkhoff, G. *Lattice Theory*, vol. 25 of *Colloquium Publications.* American Mathematical Society, 3rd edition, 1967. Providence, Rhode Island, USA.

Borceux, F. *Handbook of Categorical Algebra, II. Categories and Structures*, vol. 51 of *Encyclopedia of Math. and its App.* Cambridge University Press, Cambridge, 1994.

Bridson, M. R. and Haefliger, A. *Metric Spaces of Non-Positive Curvature*, vol. 319 of *Grundlehren der mathematischen Wissenschaften.* Springer-Verlag, Berlin, 1999.

Brown, R. *Topology and Groupoids.* BookSurge Publishing, 2006.

Carson, S. D. and Reynolds Jr., P. F. The Geometry of Semaphore Programs. *ACM Transactions on Programming Languages and Systems*, 9(1):25–53, 1987.

Coffman, E. G., Elphick, M., and Shoshani, A. System Deadlocks. *ACM Comput. Surv.*, 3(2):67–78, 1971.

Cormen, T. H., Leiserson, C. E., Rivest, R. L., and Stein, C. *Introduction to Algorithms.* MIT Press, USA, 3rd edition, 2009.

Diestel, R. and Kühn, D. Graph-theoretical versus topological ends of graphs. *Journal of Combinatorial Theory*, 87:197–206, 2003.

Dijkstra, E. W. Cooperating sequential processes. Technical report, Technological University, Eindhoven, The Netherlands, 1965. Reprinted in F. Genuys Ed. 1968. *Programming Languages*, Academic Press, New York, 43-112. *Article 1.*

Fajstrup, L., Goubault, É., and Raußen, M. Algebraic Topology and Concurrency. *Theoretical Computer Science*, 357(1):241–278, 2006. Extended version of report R-99-2008, Department of Mathematical Sciences, Aalborg university, DK-9220 Aalborg Øst. 1999.

Fajstrup, L., Goubault, É., Haucourt, E., Mimram, S., and Raußen, M. *Directed Algebraic Topology and Concurrency.* Springer, Switzerland, 2016.

Foertsch, T. and Lytchak, A. The De Rham Decomposition Theorem for Metric Spaces. *Geometric and Functional Analysis*, 18:120–143, 2008.

Fraser, G. A. The semilattice tensor product of distributive lattices. *Transactions of the American Mathematical Society*, 217, 1976.

Fraser, G. A. Tensor products of semilattices and distributive lattices. *Semigroup Forum*, 13:178–184, 1976b.

Geroldinger, A. and Halter-Koch, F. *Non-Unique Factorizations: Algebraic, Combinatorial, and Analytic Theory.* Pure and Applied Mathematics. Chapman & Hall, 2006. Boca Raton, Florida, USA.

Givant, S. and Halmos, P. *Introduction to Boolean Algebras.* Undergraduate Texts in Mathematics. Springer, New York, USA, 2009.

Grandis, M. Directed Homotopy Theory, I. The Fundamental Category. *Cahiers de Topologie et Géométrie Différentielle Catégoriques*, 44(4):281–316, 2003.

Grandis, M. *Directed Algebraic Topology : Models of Non-Reversible Worlds*, vol. 13 of *New Mathematical Monographs.* Cambridge University Press, Cambridge, 2009.

Grillet, P. A. The Tensor Product of Semigroups. *Transactions of the American Mathematical Society*, 138:267–280, 1969a.

Grillet, P. A. The Tensor Product of Commutative Semigroups. *Transactions of the American Mathematical Society*, 138:281–293, 1969b.

Hashimoto, J. On Direct Product Decomposition of Partially Ordered Sets. *Annals of Mathematics*, 54(2):315–318, 1951.

Hatcher, A. *Algebraic Topology.* Cambridge University Press, Cambridge, 2002.

Haucourt, E. Some Invariants of Directed Topology towards a Theoretical Base for a Static Analyzer Dealing with Fine-Grain Concurrency, 2016. Available on HAL.

Haucourt, E. The geometry of conservative programs. *Mathematical Structures in Computer Science*, to appear (online since October 2017).

Haucourt, E. and Ninin, N. The Boolean Algebra of Cubical Areas as a Tensor Product in the Category of Semilattices with Zero. In *7th Interaction and Concurrency Experience (ICE 2014)*, Electronic Proceedings in Theoretical Computer Science, 2014.

Hungerford, T. W. *Algebra.* Graduate Texts in Math. Springer-Verlag, New York, 2003.

Johnstone, P. T. *Stone Spaces*, vol. 3 of *Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 1982.

Ketonen, J. The Structure of Countable Boolean Algebras. *Annals of Mathematics*, 108 (1):41–89, 1978.

Krishnan, S. A Convenient Category of Locally Preordered Spaces. *Applied Categorical Structures*, 17(5):445–466, 2009.

Lang, S. *Algebra.* Graduate Texts in Mathematics. Springer, New York, 3rd ed., 2002.

Munkres, J. R. *Topology.* Prentice-Hall, USA, 2nd ed., 2000.

Nadler Jr., S. B. *Continuum Theory, An Introduction*, vol. 158 of *Monographs and Textbooks in Pure and Applied Mathematics.* Marcel Dekker, New York, USA, 1992.

Nakayama, T. and Hashimoto, J. On a problem of G. Birkoff. *Proceeding of the American Mathematical Society*, 1:141–142, 1950.

Ninin, N. *Factorisation des régions cubiques et applications à la concurrence.* PhD thesis, Université Paris 11 Orsay, 2017.

Pierce, R. S. Countable Boolean algebras. In *Handbook of Boolean Algebras vol.3*, pages 775–876. Elsevier, Amsterdam, The Netherlands, 1989.

Pratt, V. Higher dimensional automata revisited. *Mathematical Structures in Computer Science*, 10(4):525–548, 2000.

Preparata, F. P. and Shamos, M. I. *Computational Geometry: An Introduction.* Texts and Monographs in Computer Science. Springer-Verlag, New York, 2nd printing, 1985.

Sikorski, R. The Cartesian Product of Boolean Algeras. *Fundamenta Matematicae*, 37 (1):25–54, 1950.

van Glabbeek, R. J. Bisimulations for Higher Dimensional Automata.
URL `http://theory.stanford.edu/~rvg/hda`