Separation and covering for group based concatenation hierarchies

Thomas Place and Marc Zeitoun
Univ. Bordeaux, CNRS, Bordeaux INP, LaBRI, UMR 5800, F-33400, Talence, France
Email: {tplace,mz}@labri.fr

***Abstract*—Concatenation hierarchies are natural classifications of regular languages. All such hierarchies are built through the same construction process: one starts from an initial, specific class of languages (the basis) and builds new levels using two generic operations. Concatenation hierarchies have gathered a lot of interest since the early 70s, notably thanks to an alternate logical definition: each concatenation hierarchy can be defined as the quantification alternation hierarchy within a variant of first-order logic over words (while the hierarchies differ by their bases, the variants differ by their set of available predicates).**

Our goal is to understand these hierarchies. A typical approach is to look at two decision problems: membership and separation. In the paper we are interested in the latter, which is more general. For a class of languages $\mathcal{C}$, $\mathcal{C}$-separation takes two regular languages as input and asks whether there exists a third one in $\mathcal{C}$ including the first one and disjoint from the second one. Settling whether separation is decidable for the levels within a given concatenation hierarchy is among the most fundamental and challenging questions in formal language theory. In all prominent cases, it is open, or answered positively for low levels only. Recently, a breakthrough was made using a generic approach for a specific kind of hierarchies: those with a *finite* basis. In this case, separation is always decidable for levels 1/2, 1 and 3/2.

Our main theorem is similar but independent: we consider hierarchies with possibly infinite bases, but that contain only *group languages*. An example is the group hierarchy introduced by Pin and Margolis: its basis consists of all group languages. Another example is the quantifier alternation hierarchy of first-order logic with modular predicates FO$(<, \mathrm{MOD})$: its basis consists of the languages that count the length of words modulo some number. Using a generic approach, we show that for any such hierarchy, if separation is decidable for the basis, then it is decidable as well for levels 1/2, 1 and 3/2 (we actually solve a more general problem called covering). This complements the aforementioned result nicely: all bases considered in the literature are either finite or made of group languages. Thus, one may handle the lower levels of any prominent hierarchy in a *generic way*.

I. Introduction

Context. Concatenation hierarchies are natural classifications of regular languages. They were motivated by a celebrated theorem of Schützenberger [34], McNaughton and Papert [18], which is twofold. Its first part states that the class of regular languages that can be defined from singletons using Boolean operations and concatenation only (but *no* Kleene star) coincides with the class of languages that can be defined in first-order logic. In other words, *star-free* languages are exactly *first-order definable* ones. The second and more difficult part consists of an algorithm that inputs a regular language, and outputs whether it belongs to this class.

978-1-7281-3608-0/19/$31.00 ©2019 IEEE

Simultaneously, Brzozowski and Cohen [6] introduced the *dot-depth hierarchy*. It stratifies star-free languages in an infinitely increasing sequence of levels [7] spanning the whole class of star-free languages. Intuitively, a level in this hierarchy captures the number of alternations between complement and concatenation that are needed to express a star-free language. The dot-depth rose to prominence following the work of Thomas [39], who proved an exact correspondence with the quantifier alternation hierarchy of first-order logic: each level in the dot-depth hierarchy consists of all languages that can be defined with a prescribed number of quantifier blocks.

One of the most famous open problems in automata theory is to settle whether the membership problem is decidable for each individual level: is there an algorithm deciding whether an input regular language belongs to this level?

The literature about this problem is rich. We refer the reader to the surveys [29], [32], [22]. The bottom line is that after more than 45 years, little is known. Let us briefly survey the most important cornerstones in this line of research.

State of the art. The dot-depth is a particular instance of a *concatenation hierarchy*. All such hierarchies are built through a uniform construction scheme. Levels are numbered by integers 0,1,2... or half integers 1/2, 3/2, 5/2... Level 0 is a class of languages called the basis, which is the only parameter specific to the hierarchy. New levels are then built using two operations. The first is closure under marked concatenation $K, L \mapsto KaL$ (for some letter a) and union. This closure operator, when applied to level $n \in \mathbb{N}$, produces level $n + 1/2$. The second operator is closure under Boolean operations, which yields level $n + 1$ out of level $n + 1/2$.

For instance, the basis of the dot-depth hierarchy consists of 4 languages: the empty set, the singleton set consisting of the empty word, and their complements. Between 1971 and 2015, membership was shown to be decidable up to level 5/2 and for level 7/2 [2], [3], [15], [13], [33], [27], [28], [25], [26]. The state of the art is the same for the *Straubing-Thérien hierarchy* [36], [38], whose basis is the empty set and the set of all words. Actually the above results for the dot-depth were obtained by reduction to this hierarchy, via transfer theorems [37], [28].

A crucial point is that all recent results are based on more general problems than membership: separation and covering. For instance, separation for a class $\mathcal{C}$ takes as input *two* regular languages, and asks whether there exists a third language, which belongs to $\mathcal{C}$, contains the first language, and is disjoint from the second. Membership is the particular case of separation when the input consists of a regular language and its complement. Thus, deciding membership for a class $\mathcal{C}$ reduces to deciding separation for $\mathcal{C}$, which itself reduces to deciding covering for $\mathcal{C}$.

A generic result. Recently, we proved that *all* known results on both the Straubing-Thérien and the dot-depth hierarchies follow from a single one. This generic theorem [33], [26] states that in any hierarchy having a finite basis satisfying mild hypotheses, covering is decidable at levels 1/2, 1 and 3/2. All results stated above follow from this via simpler observations.

Establishing such generic results is of course desirable. This avoids developing specific arguments, pinpointing the key hypotheses needed for obtaining covering or separation algorithms, and this is amenable to generalization.

Contribution. In the literature, there is another important kind of concatenation hierarchy: those with a basis made exclusively of group languages. A *group language* is a language recognized by an automaton whose letters act as permutations on states. The most prominent representative is the *group hierarchy*, introduced by Pin and Margolis [17], whose basis consists of all group languages. Using algebraic techniques, it was shown that the levels 1/2 and 1 both have decidable membership in [24] and [14] respectively. Note that our aforementioned generic result does not apply here, since the class of group languages is infinite.

Our contribution is another ***generic result*** which nicely complements the one applying to finite bases. While both statements are similar, they are independent. Our main theorem states that *covering* is decidable at levels 1/2, 1 and 3/2 in *any* hierarchy whose basis consists of *group languages* only, provided that this basis has *decidable separation*.

Applications. Let us state three applications of this result. First, the class of all group languages is known to have decidable separation, and even decidable covering. This follows from a result of Ash [4], which was connected to separation and covering by Almeida [1]. Therefore, our generic statement applies to the group hierarchy of Pin and Margolis.

The other examples have a nice logical interpretation. Indeed, the result of Thomas [39] was generalized in [32]: *every* concatenation hierarchy corresponds, level by level, to the quantification alternation hierarchy within a variant of first-order logic over words. While the hierarchies differ by their bases, the variants differ by their set of available predicates.

This correspondence makes it possible to present our second example as the quantifier alternation hierarchy of first-order logic with modular predicates, $\mathrm{FO}(<, \mathrm{MOD})$. It corresponds to the concatenation hierarchy whose basis consists of the languages that count the length of words modulo some number. It is simple to show that this basis consists of group languages and has decidable separation. Hence, our generic theorem applies: it strengthens and unifies results of [10], [16], which dealt with membership at levels up to 3/2, and [40], which proved separation at level 1 with a combinatorial proof leading to a brute force algorithm, orthogonal to our techniques. Finally, when the basis consists of languages recognized by commutative groups, we obtain by [32], [12] the quantifier alternation hierarchy of first-order logic endowed with predicates counting the number of occurrences of a letter before a position, modulo some integer. This basis has decidable covering [11], [1], so our result applies to this hierarchy, which was not yet investigated.

Historical remarks. Our approach differs from earlier work on these hierarchies. When the original group hierarchy was introduced [17], the approach was purely algebraic. Rather than considering the levels of this hierarchy (*i.e.*, classes of languages) directly, the idea was to work with *varieties of finite monoids* associated to them through Eilenberg's correspondence [12]. Moreover, most results obtained so far rely on an alternate definition of the levels within a group based hierarchy, which involves an algebraic operation called *wreath product*. Indeed, it was shown by Pin [19], [20] that for any class $\mathcal{C}$ of group languages, every level in the hierarchy of basis $\mathcal{C}$ is the wreath product of the corresponding level in the *Straubing-Thérien hierarchy* with $\mathcal{C}$. This theorem instilled hope for transferring results from the Straubing-Thérien hierarchy to group based hierarchies. An overview of these techniques applied to the group hierarchy is presented in [14]. They were generalized to decide membership of level 1 [35] when $\mathcal{C}$ has decidable separation. However, an example of a class $\mathcal{C}$ with decidable membership and such that level 1 has undecidable membership was given in [5]. This result dashes the hope of transferring decidability of membership from the basis to higher levels: stronger hypotheses on the basis are needed.

Let us explain the reasons why we do not frame the paper in this setting. Since our objective is to gather understanding at the *language level*, this algebraic approach has several downsides:

- First, it requires a *heavy preliminary machinery*, useless for our purpose, such as the definition of varieties, Eilenberg's theorem, the definition wreath product, etc.
- Second, it requires to rely on *technical black-box results*, obfuscating the intuition on the language theoretic side.
- Third, it is well suited for *more specific classes* than ours: it permits to work only with varieties, which requires closure under inverse morphisms on the language side. This is useless, we are not bound by such restrictions.
- Finally, most algebraic papers investigating such hierarchies consider the *membership problem*: the key idea towards decidability boils down to obtaining equations, which is accurate, but costs an extremely technical framework. While this approach can work for membership [14], [35], it didn't prove to do so for separation and covering.

Here, we use an orthogonal, purely language theoretic approach: we work directly with the bare definition of levels 1/2, 1 and 3/2 in a group hierarchy, built from the basis using polynomial closure and Boolean closure. This approach bypasses the above downsides: on the one hand, it is simpler. On the other hand, we obtain more general results: the bases have to fulfill less hypotheses, and we solve more general problems than membership, and for higher levels than in the historical approach.

Organization. In Section II, we recall the basic notions that we need. The main theorem is stated in Section III. The framework that we use is recalled in Sections IV and VI. Finally, Sections V, VII and VIII are devoted to presenting the algorithms of the three parts of the main theorem: the decidability of covering at levels 1/2, 1 and 3/2, respectively. Due to space limitations, some proofs are omitted.

II. Preliminaries

A. Classes of languages

We fix an arbitrary finite alphabet A for the whole paper. As usual, A^* denotes the set of all words over A, including the empty word ε. We let $A^+ = A^* \setminus \{\varepsilon\}$. For $u, v \in A^*$, we write uv the word obtained by concatenating u and v. Finally, for $w \in A^*$, we write $|w| \in \mathbb{N}$ for its length.

A *language* is a subset of A^*. We denote the singleton language $\{u\}$ by u. One may lift the concatenation operation to languages: for $K, L \subseteq A^*$, we let $KL = \{uv \mid u \in K \text{ and } v \in L\}$. Additionally, we consider *marked concatenation*. Given $K, L \subseteq A^*$, a marked concatenation of K with L is a language of the form KaL for some letter $a \in A$.

A *class of languages* $\mathcal{C}$ is a set of languages. We shall work with robust classes satisfying standard closure properties:

- $\mathcal{C}$ is a *lattice* when it is closed under union and intersection, $\emptyset \in \mathcal{C}$ and $A^* \in \mathcal{C}$.
- A *Boolean algebra* is a lattice closed under complement.
- $\mathcal{C}$ is *quotient-closed* when for every $L \in \mathcal{C}$ and $w \in A^*$, the following two languages belong to $\mathcal{C}$:

$$\begin{aligned} w^{-1}L &\stackrel{\text{def}}{=} \{u \in A^* \mid wu \in L\}, \\ Lw^{-1} &\stackrel{\text{def}}{=} \{u \in A^* \mid uw \in L\}. \end{aligned}$$

All classes considered in the paper are (at least) quotient-closed lattices. Furthermore, they are included in the class of *regular languages*. These are the languages that can be equivalently defined by monadic second-order logic, finite automata or finite monoids. We use the definition based on monoids.

Regular languages. A *semigroup* is a pair $(S, \cdot)$ where S is a set and "$\cdot$" is an associative binary operation on S (often called multiplication). It is standard to abuse terminology and make the binary operation implicit: one simply says that "S is a semigroup". A *monoid* M is a semigroup whose multiplication has a neutral element denoted by "1_M". Recall that an idempotent of a semigroup S is an element $e \in S$ such that $ee = e$. A standard result in semigroup theory states that when S is *finite*, there exists $\omega(S) \in \mathbb{N}$ (written ω when S is understood) such that s^ω is idempotent for every $s \in S$.

Clearly, A^* is a monoid whose multiplication is concatenation (ε is the neutral element). Hence, given a monoid M, we may consider morphisms $\alpha : A^* \to M$. We say that a language $L \subseteq A^*$ is *recognized* by such a morphism α when there exists $F \subseteq M$ such that $L = \alpha^{-1}(F)$. It is well-known that the regular languages are exactly those which can be recognized by a morphism $\alpha : A^* \to M$ where M is a *finite* monoid.

Group languages. A group is a monoid G such that every element $g \in G$ has an inverse $g^{-1} \in G$, *i.e.*, $gg^{-1} = g^{-1}g = 1_G$. We call "*group language*" a language L which is recognized by a morphism into a *finite group*. In the paper, we consider classes that are quotient-closed Boolean algebras of group languages (*i.e.*, containing group languages only).

Example 1. *The most immediate example of quotient-closed Boolean algebra of group languages (which is also the largest one) is the class of* all *group languages. We write it* GR.

Quotient-closed Boolean algebras of group languages are more general than the classes of group languages that are usually considered. Typically, publications on the topic consider *varieties* of group languages, which involve an additional closure property called "inverse morphic image" (see [19] for details). Let us present a class which is a quotient-closed Boolean algebra of group languages, but not a variety.

Example 2. *The class* MOD *containing the finite Boolean combinations of languages $\{w \in A^* \mid |w| = k \mod m\}$ with $k, m \in \mathbb{N}$ such that $k < m$, is a quotient-closed Boolean algebra of group languages. We use it as a running example.*

B. Separation and covering

In the paper, we use two decision problems to investigate specific classes of languages (all built from quotient-closed Boolean algebras of group languages): separation and covering. We define them here. The former is standard while the latter was introduced in [31]. Both of them are parametrized by an arbitrary class of languages $\mathcal{C}$. We start with separation.

Separation. Given three languages K, L_1, L_2, we say that K *separates* L_1 from L_2 if $L_1 \subseteq K$ and $L_2 \cap K = \emptyset$. Given a class of languages $\mathcal{C}$, we say that L_1 is *$\mathcal{C}$-separable* from L_2 if some language in $\mathcal{C}$ separates L_1 from L_2. Observe that when $\mathcal{C}$ is not closed under complement, the definition is not symmetrical: L_1 could be $\mathcal{C}$-separable from L_2 while L_2 is not $\mathcal{C}$-separable from L_1. The separation problem associated to a given class $\mathcal{C}$ is as follows:

INPUT: Two regular languages L_1 and L_2.
OUTPUT: Is L_1 $\mathcal{C}$-separable from L_2?

Remark 3. *Separation generalizes the classical membership problem which asks whether a single regular language belongs to $\mathcal{C}$. Indeed, $L \in \mathcal{C}$ if and only if L is $\mathcal{C}$-separable from $A^* \setminus L$.*

Covering. This more general problem was introduced in [31]. Given a language L, a *cover of* L is a **finite** set of languages $\mathbf{K}$ such that $L \subseteq \bigcup_{K \in \mathbf{K}} K$. Moreover, if $\mathcal{C}$ is a class, a $\mathcal{C}$-cover of L is a cover $\mathbf{K}$ of L such that all $K \in \mathbf{K}$ belong to $\mathcal{C}$.

Covering takes as input a language L_1 and a *finite set of languages* $\mathbf{L}_2$. A *separating cover* for the pair $(L_1, \mathbf{L}_2)$ is a cover $\mathbf{K}$ of L_1 such that for every $K \in \mathbf{K}$, there exists $L \in \mathbf{L}_2$ which satisfies $K \cap L = \emptyset$. Finally, given a class $\mathcal{C}$, we say that the pair $(L_1, \mathbf{L}_2)$ is $\mathcal{C}$-coverable when there exists a separating $\mathcal{C}$-cover. The $\mathcal{C}$-covering problem is now defined as follows:

INPUT: A regular language L_1 and
a finite set of regular languages $\mathbf{L}_2$.
OUTPUT: Is $(L_1, \mathbf{L}_2)$ $\mathcal{C}$-coverable?

It is straightforward to prove that covering generalizes separation (provided that the class $\mathcal{C}$ is a lattice) as stated in the following lemma (see Theorem 3.5 in [31] for the proof).

Lemma 4. *Let $\mathcal{C}$ be a lattice and $L_1, L_2 \subseteq A^*$. Then L_1 is $\mathcal{C}$-separable from L_2, if and only if $(L_1, \{L_2\})$ is $\mathcal{C}$-coverable.*

III. Concatenation hierarchies and main theorem

In this section, we present the particular classes that we investigate in the paper and outline our results.

A. Closure operations and concatenation hierarchies

We are interested in *concatenation hierarchies*. We briefly recall this notion (see [22], [32] for details). A concatenation hierarchy is an increasing sequence of classes of languages, which depends on a single parameter: an arbitrary quotient-closed Boolean algebra, called its *basis*. Once the basis is fixed, the construction is uniform. Languages are classified into levels: each new level is built by applying one of two generic operations to the previous one. Let us define these operations.

Boolean closure. Given a class $\mathcal{C}$, its *Boolean closure*, denoted by $Bool(\mathcal{C})$ is the least Boolean algebra containing $\mathcal{C}$. The following lemma is immediate from the definitions (it holds simply because quotients commute with Boolean operations).

Lemma 5. *Let $\mathcal{C}$ be a quotient-closed lattice. Then $Bool(\mathcal{C})$ is a quotient-closed Boolean algebra.*

Polynomial closure. Given a class $\mathcal{C}$, the *polynomial closure* of $\mathcal{C}$, denoted by $Pol(\mathcal{C})$, is the least class containing $\mathcal{C}$ which is closed under both union and marked concatenation: for every $K, L \in Pol(\mathcal{C})$ and $a \in A$, we have $K \cup L \in Pol(\mathcal{C})$ and $KaL \in Pol(\mathcal{C})$. While this is not obvious from the definition, when the input class $\mathcal{C}$ is a quotient-closed lattice, its polynomial closure $Pol(\mathcal{C})$ remains a quotient-closed lattice (the difficulty is to prove closure under intersection). This was proved by Arfi [2] when $\mathcal{C}$ is also closed under complement.

Theorem 6 ([2], [21], [32])**.** *Let $\mathcal{C}$ be a quotient-closed lattice. Then, $Pol(\mathcal{C})$ is a quotient-closed lattice closed under concatenation and marked concatenation.*

Concatenation hierarchies. We let $\mathcal{C}$ as a quotient-closed Boolean algebra. The *concatenation hierarchy of basis $\mathcal{C}$* classifies languages into levels of two kinds: full levels (denoted by 0, 1, 2,...) and half levels (denoted by 1/2, 3/2, 5/2,...):

- Level 0 is the basis $\mathcal{C}$.
- Each *half level* $n+\frac{1}{2}$, for $n \in \mathbb{N}$, is the *polynomial closure* of the previous full level, *i.e.*, of level n.
- Each *full level* $n+1$, for $n \in \mathbb{N}$, is the *Boolean closure* of the previous half level, *i.e.*, of level $n+\frac{1}{2}$.

In view of Lemma 5 and Theorem 6, it is immediate that every full level is a quotient-closed Boolean algebra while every half level is a quotient-closed lattice. Moreover, we have the following standard fact (for example, see [26, Lemma 3.4] for a proof).

Fact 7. *Let $\mathcal{D}$ be a level greater or equal to 1 in some concatenation hierarchy. Then $\{w\} \in \mathcal{D}$ for every $w \in A^*$.*

In this paper, we are mainly interested in levels 1/2, 1 and 3/2. By definition, they correspond to the classes $Pol(\mathcal{C})$, $Bool(Pol(\mathcal{C}))$ and $Pol(Bool(Pol(\mathcal{C})))$. For the sake of avoiding clutter, we shall write $BPol(\mathcal{C})$ for $Bool(Pol(\mathcal{C}))$ and $PBPol(\mathcal{C})$ for $Pol(Bool(Pol(\mathcal{C})))$. Our main theorem applies to those levels for every basis $\mathcal{C}$ which is a *quotient-closed Boolean algebra of group languages*. It is as follows.

Theorem 8 (Main result)**.** *Consider a concatenation hierarchy whose basis $\mathcal{C}$ contains only **group languages** and such that $\mathcal{C}$-separation is decidable. Then, covering and separation are decidable for levels 1/2, 1 and 3/2 of this hierarchy.*

Theorem 8 is generic and applies to all "group based" hierarchies. It complements similar results which were recently proved in [25], [26] (for levels 1/2 and 3/2) and [30], [33] (for level 1) for *finitely based* hierarchies.

Theorem 9 ([26], [33])**.** *Consider a concatenation hierarchy whose basis is **finite**. Then, covering and separation are decidable for levels 1/2, 1 and 3/2.*

All prominent concatenation hierarchies investigated in the literature have a basis which is either finite or made of group languages. Hence, when put together, the two above theorems can be used to handle the lower levels of all prominent concatenation hierarchies in a *generic way*.

Additionally, one may combine Theorem 8 with a result of [32] to get information on level 5/2 of group based hierarchies. It is shown in [32] that given an arbitrary hierarchy, if separation is decidable for some half level, then so is the *membership problem* for the next half level. Hence, we get the following corollary of Theorem 8.

Corollary 10. *Consider a concatenation hierarchy whose basis $\mathcal{C}$ contains only **group languages** and such that $\mathcal{C}$-separation is decidable. Then, given as input a regular language L, one may decide whether L belongs to level 5/2 of this hierarchy.*

The proof of Theorem 8 spans the remaining sections of the paper. Each level mentioned in the statement is handled independently. In Section IV, we recall a framework designed to handle the covering problem. It was originally introduced in [31] and we use it to obtain all results announced in Theorem 8. We apply it for level 1/2 in Section V. Handling levels 1 and 3/2 is much more involved and requires extending the framework. We do so in Section VI and handle these levels in Sections VII and VIII respectively.

However, let us first illustrate Theorem 8 by presenting two prominent concatenation hierarchies to which it applies.

B. Important applications of Theorem 8

We present two prominent examples of concatenation hierarchies whose bases are quotient-closed Boolean algebra of group languages (we already presented these bases in Section II).

Group hierarchy. This is the hierarchy of basis GR (the class of *all* group languages). It was originally introduced by Margolis and Pin [17]. To our knowledge, the following results were known for this hierarchy:

- Separation is decidable for the basis GR. This is a corollary of a theorem by Ash [4] which solves a longstanding

conjecture in algebra (Rhodes' type II conjecture). We refer the reader to [14] for background on this question.
- Membership is decidable for levels 1/2 and 1. The former follows from the decidability of separation for GR and a transfer result of [32]. The latter is proved in [14].

Since separation is decidable for the basis GR, we obtain the following result from Theorem 8 and Corollary 10.

Corollary 11. *Separation and covering are decidable for levels 1/2, 1 and 3/2 in the group hierarchy. Moreover, membership is decidable for level 5/2.*

The hierarchy of basis MOD. This example is quite important because it has a natural *logical* equivalent. It is known that *every* concatenation hierarchy corresponds to the *quantifier alternation hierarchy* within a particular variant of first-order logic over words (the variants differ by the set of predicates that are allowed in sentences). This was first observed by Thomas [39] for a particular example: the dot-depth hierarchy of Brzozowski and Cohen [6] (whose basis is the finite class $\{\emptyset, \{\varepsilon\}, A^+, A^*\}$). However, the ideas of Thomas can be generalized to all concatenation hierarchies (see [32]).

Of course, depending on the concatenation hierarchy, the variant of first-order logic one ends up with may or may not be natural. It turns out that for the hierarchy of basis MOD, we get a *standard* one: first-order logic with modular predicates ($\text{FO}(<, \text{MOD})$). One may view a finite word w as a logical structure made of a sequence of positions numbered from 0 to $|w| - 1$. Each position can be quantified and carries a label in A. We denote by $\text{FO}(<, \text{MOD})$ the variant of first-order logic equipped with the following predicates:

- For each $a \in A$, a unary predicate P_a selecting positions labeled with letter "a".
- A binary predicate "$<$" interpreted as the linear order.
- For each integers $0 \leq i < d$, a unary predicate MOD_i^d selecting positions that are congruent to i modulo d.
- A constant D_i^d, which holds for words whose length is congruent to i modulo d.

A sentence of $\text{FO}(<, \text{MOD})$ defines a language (it consists of all words satisfying the sentence). Thus, $\text{FO}(<, \text{MOD})$ defines a *class of languages* which we also denote by $\text{FO}(<, \text{MOD})$.

We are not interested in $\text{FO}(<, \text{MOD})$ itself: we consider its quantifier alternation hierarchy. One may classify sentences of $\text{FO}(<, \text{MOD})$ by counting their number of quantifier alternations. Let $n \in \mathbb{N}$. A sentence is $\Sigma_n(<, \text{MOD})$, if it can be rewritten into a sentence in prenex normal form which has either, exactly n blocks of quantifiers starting with an $\exists$ or *strictly less* than n blocks of quantifiers. For example, consider the following sentence (already in prenex normal form)

$$\exists x_1 \exists x_2 \forall x_3 \exists x_4 \ \varphi(x_1, x_2, x_3, x_4) \quad \text{(with } \varphi \text{ quantifier-free)}$$

This sentence is $\Sigma_3(<, \text{MOD})$. In general, the negation of a $\Sigma_n(<, \text{MOD})$ sentence is not a $\Sigma_n(<, \text{MOD})$ sentence. Hence it is relevant to define $\mathcal{B}\Sigma_n(<, \text{MOD})$ sentences as the Boolean combinations of $\Sigma_n(<, \text{MOD})$ sentences. This gives a hierarchy of classes of languages: for every $n \in \mathbb{N}$, we have $\Sigma_n(<, \text{MOD}) \subseteq \mathcal{B}\Sigma_n(<, \text{MOD}) \subseteq \Sigma_{n+1}(<, \text{MOD})$.

This hierarchy has been widely investigated in the literature. For membership, it was known that the problem is decidable for $\mathcal{B}\Sigma_1(<, \text{MOD})$ (see [10]) as well as $\Sigma_1(<, \text{MOD})$ and $\Sigma_2(<, \text{MOD})$ (see [16]). Furthermore, it was recently shown that separation is decidable for $\mathcal{B}\Sigma_1(<, \text{MOD})$ [40]. We are able to reprove these results, generalize them to the covering problem and push them to $\Sigma_3(<, \text{MOD})$. Indeed, the generic correspondence of [32] implies that for every $n \in \mathbb{N}$, level n in the hierarchy of basis MOD corresponds to $\mathcal{B}\Sigma_n(<, \text{MOD})$ while level $n + \frac{1}{2}$ corresponds to $\Sigma_n(<, \text{MOD})$. Moreover, proving the decidability of MOD-separation is a simple exercise. Consequently, we obtain the following statement as a corollary of Theorem 8.

Corollary 12. *Separation and covering are decidable for the logics $\Sigma_1(<, \text{MOD})$,$\mathcal{B}\Sigma_1(<, \text{MOD})$ and $\Sigma_2(<, \text{MOD})$. Moreover, membership is decidable for $\Sigma_3(<, \text{MOD})$.*

IV. Framework

In this section, we present the framework which we shall use to obtain the three decidability results announced in Theorem 8. It was originally introduced in [31]. We refer the reader to this paper for details and the proofs of the statements.

A. Rating maps

The framework is based on an algebraic object called "rating map". They are morphisms of commutative and idempotent monoids. We write such monoids $(R, +)$: we call the binary operation "$+$" *addition* and denote the neutral element by 0_R. Being idempotent means that for all $r \in R$, we have $r + r = r$. For every commutative and idempotent monoid $(R, +)$, one may define a canonical ordering $\leq$ over R:

$$\text{For all } r, s \in R, \quad r \leq s \text{ when } r + s = s.$$

One may verify that $\leq$ is a partial order which is compatible with addition. Moreover, every morphism between commutative and idempotent monoids is increasing for this ordering.

Example 13. *For every set E, $(2^E, \cup)$ is an idempotent and commutative monoid. The neutral element is $\emptyset$ and the canonical ordering is inclusion.*

When dealing with subsets of a commutative and idempotent monoid $(R, +)$, we shall often apply a *downset operation*. Given $S \subseteq R$, we write:

$$\downarrow_R S = \{r \in R \mid r \leq s \text{ for some } s \in S\}.$$

We extend this notation to Cartesian products of arbitrary sets with R. Given some set X and $S \subseteq X \times R$, we write,

$$\downarrow_R S = \{(x, r) \in X \times R \mid \exists s \in R \text{ s.t. } r \leq s \text{ and } (x, s) \in S\}$$

We may now define rating maps. A rating map is a morphism $\rho : (2^{A^*}, \cup) \to (R, +)$ where $(R, +)$ is a *finite* idempotent and commutative monoid, called the *rating set of ρ*. That is, ρ is a map from 2^{A^*} to R satisfying the following properties:

(1) $\rho(\emptyset) = 0_R$.
(2) For all $K_1, K_2 \subseteq A^*$, $\rho(K_1 \cup K_2) = \rho(K_1) + \rho(K_2)$.

For the sake of improved readability, when applying a rating map ρ to a singleton set $\{w\}$, we shall write $\rho(w)$ for $\rho(\{w\})$. Additionally, we write $\rho_* : A^* \to R$ for the restriction of ρ to A^*: for every $w \in A^*$, we have $\rho_*(w) = \rho(w)$ (this notation is useful when referring to the language $\rho_*^{-1}(r) \subseteq A^*$, which consists of all words $w \in A^*$ such that $\rho(w) = r$).

Most of the theory makes sense for arbitrary rating maps. However, we shall often have to work with special rating maps satisfying additional properties. We define two kinds.

Nice rating maps. We say a rating map $\rho : 2^{A^*} \to R$ is nice when, for every language $K \subseteq A^*$, there exists finitely many words $w_1, \dots, w_n \in K$ such that $\rho(K) = \rho(w_1) + \cdots + \rho(w_k)$.

When a rating map $\rho : 2^{A^*} \to R$ is nice, it is characterized by the canonical map $\rho_* : A^* \to R$. Indeed, for $K \subseteq A^*$, we may consider the sum of all elements $\rho(w)$ for $w \in K$: while it may be infinite, this sum boils down to a finite one since R is commutative and idempotent. The hypothesis that ρ is nice implies that $\rho(K)$ is equal to this sum.

Multiplicative rating maps. A rating map $\rho : 2^{A^*} \to R$ is multiplicative when its rating set R has more structure: it needs to be an *idempotent semiring*. Moreover, ρ has to satisfy an additional property connecting this structure to language concatenation. Namely, it has to be a morphism of semirings.

A *semiring* is a tuple $(R, +, \cdot)$ where R is a set and "$+$" and "$\cdot$" are two binary operations called addition and multiplication, such that the following axioms are satisfied:

- $(R, +)$ is a commutative monoid.
- $(R, \cdot)$ is a monoid (the neutral element is denoted by 1_R).
- Multiplication distributes over addition. For $r, s, t \in R$, $r \cdot (s + t) = (r \cdot s) + (r \cdot t)$ and $(r + s) \cdot t = (r \cdot t) + (s \cdot t)$.
- The neutral element "0_R" of $(R, +)$ is a zero for $(R, \cdot)$: $0_R \cdot r = r \cdot 0_R = 0_R$ for every $r \in R$.

A semiring R is *idempotent* when $r + r = r$ for every $r \in R$, *i.e.*, when the additive monoid $(R, +)$ is idempotent (there is no additional constraint on the multiplicative monoid $(R, \cdot)$).

Example 14. *A key example of infinite idempotent semiring is the set 2^{A^*}. Union is the addition and language concatenation is the multiplication (with $\{\varepsilon\}$ as neutral element).*

Clearly, any finite idempotent semiring $(R, +, \cdot)$ is in particular a rating set: $(R, +)$ is an idempotent and commutative monoid. In particular, one may verify that the canonical ordering "$\leq$" on R is compatible with multiplication as well.

We may now define multiplicative rating maps: as expected they are semiring morphisms. Let $\rho : 2^{A^*} \to R$ be a rating map: $(R, +)$ is an idempotent commutative monoid and ρ is a morphism from $(2^{A^*}, \cup)$ to $(R, +)$. We say that ρ is multiplicative when the rating set R is equipped with a multiplication "$\cdot$" such that $(R, +, \cdot)$ is an idempotent semiring and ρ is also a monoid morphism from $(2^{A^*}, \cdot)$ to $(R, \cdot)$. That is, the two following additional axioms have to be satisfied:

(3) $\rho(\varepsilon) = 1_R$.

(4) For all $K_1, K_2 \subseteq A^*$, we have $\rho(K_1 K_2) = \rho(K_1) \cdot \rho(K_2)$.

Altogether, this exactly says that ρ must be a semiring morphism from $(2^{A^*}, \cup, \cdot)$ to $(R, +, \cdot)$.

Remark 15. *The rating maps which are both nice and multiplicative are finitely representable. Indeed, as we explained above, if a rating map $\rho : 2^{A^*} \to R$ is nice, it is characterized by the canonical map $\rho_* : A^* \to R$. When ρ is additionally multiplicative, ρ_* is finitely representable since it is a morphism into a finite monoid. Hence, we may speak about algorithms taking nice multiplicative rating maps as input.*

The rating maps which are not nice and multiplicative cannot be finitely represented in general. However, they remain crucial in the paper: while our main statements consider nice multiplicative rating maps, many proofs involve auxiliary rating maps which are neither nice nor multiplicative.

B. Optimal imprints

Now that we have rating maps, we turn to imprints. Consider a rating map $\rho : 2^{A^*} \to R$. Given any finite set of languages $\mathbf{K}$, we define the ρ-imprint of $\mathbf{K}$. Intuitively, when $\mathbf{K}$ is a cover of some language L, this object measures the "quality" of $\mathbf{K}$. The ρ-imprint *of* $\mathbf{K}$ is the following subset of R:

$$\mathcal{I}[\rho](\mathbf{K}) = \downarrow_R \{\rho(K) \mid K \in \mathbf{K}\}.$$

We may now define optimality. Consider an arbitrary rating map $\rho : 2^{A^*} \to R$ and a lattice $\mathcal{C}$. Given a language L, an optimal $\mathcal{C}$-cover of L for ρ is a $\mathcal{C}$-cover $\mathbf{K}$ of L which satisfies the following property:

$$\mathcal{I}[\rho](\mathbf{K}) \subseteq \mathcal{I}[\rho](\mathbf{K}') \quad \text{for every } \mathcal{C}\text{-cover } \mathbf{K}' \text{ of } L.$$

In general, there can be infinitely many optimal $\mathcal{C}$-covers for a given rating map ρ. The key point is that there always exists at least one (this requires the hypothesis that $\mathcal{C}$ is a lattice).

Lemma 16. *Let $\mathcal{C}$ be a lattice. For every language L and every rating map ρ, there exists an optimal $\mathcal{C}$-cover of L for ρ.*

Clearly, for a lattice $\mathcal{C}$, a language L and a rating map ρ, all optimal $\mathcal{C}$-covers of L for ρ have the same ρ-imprint. Hence, this unique ρ-imprint is a *canonical* object for $\mathcal{C}$, L and ρ. We call it the *$\mathcal{C}$-optimal ρ-imprint on* L and we write it $\mathcal{I}_{\mathcal{C}}[L, \rho]$:

$$\mathcal{I}_{\mathcal{C}}[L, \rho] = \mathcal{I}[\rho](\mathbf{K}) \quad \text{for any optimal } \mathcal{C}\text{-cover } \mathbf{K} \text{ of } L \text{ for } \rho.$$

An important special case is when $L = A^*$. In that case, we write $\mathcal{I}_{\mathcal{C}}[\rho]$ for $\mathcal{I}_{\mathcal{C}}[A^*, \rho]$. Finally, we have the following useful fact which is immediate from the definitions.

Fact 17. *Let ρ be a rating map and consider two languages H, L such that $H \subseteq L$. Then, $\mathcal{I}_{\mathcal{C}}[H, \rho] \subseteq \mathcal{I}_{\mathcal{C}}[L, \rho]$.*

C. Connection with covering

We may now connect these definitions to the covering problem. The key idea is that solving $\mathcal{C}$-covering boils down to computing $\mathcal{C}$-optimal imprints from input nice multiplicative rating maps. There are actually two statements (both taken from [31]). The first one is simpler but it only applies to classes $\mathcal{C}$ which are Boolean algebras while the second (more involved) one applies to all lattices. We start with the former.

Proposition 18. *Let $\mathcal{C}$ be a Boolean algebra. Assume that there exists an algorithm which computes $\mathcal{I}_{\mathcal{C}}[\rho]$ from an input nice multiplicative rating map ρ. Then, $\mathcal{C}$-covering is decidable.*

In order to handle all lattices, one needs to consider several optimal imprints simultaneously. This is formalized by the following additional object. Consider a lattice $\mathcal{C}$, a morphism $\alpha : A^* \to M$ into a finite monoid M and a rating map $\rho : 2^{A^*} \to R$, we define the following subset of $M \times R$:

$$\mathcal{P}^{\alpha}_{\mathcal{C}}[\rho] = \{(s,r) \in M \times R \mid r \in \mathcal{I}_{\mathcal{C}}\left[\alpha^{-1}(s), \rho\right]\}.$$

We call $\mathcal{P}^{\alpha}_{\mathcal{C}}[\rho]$ the *α-pointed $\mathcal{C}$-optimal ρ-imprint*. Clearly, $\mathcal{P}^{\alpha}_{\mathcal{C}}[\rho]$ encodes all the sets $\mathcal{I}_{\mathcal{C}}\left[\alpha^{-1}(s), \rho\right]$ for $s \in M$.

Proposition 19. *Let $\mathcal{C}$ be a lattice. Assume that there exists an algorithm which computes $\mathcal{P}^{\alpha}_{\mathcal{C}}[\rho]$ from an input morphism α and an input nice multiplicative rating map ρ. Then, $\mathcal{C}$-covering is decidable.*

V. Polynomial closure

We prove the first part of Theorem 8: for every quotient-closed Boolean algebra of group languages $\mathcal{C}$, if $\mathcal{C}$-separation is decidable, then so is $Pol(\mathcal{C})$-covering.

We use rating maps: for every morphism $\alpha : A^* \to M$ and nice multiplicative rating map $\rho : 2^{A^*} \to R$, we characterize $\mathcal{P}^{\alpha}_{Pol(\mathcal{C})}[\rho]$. When $\mathcal{C}$-separation is decidable, it is simple to deduce a least fixpoint algorithm for computing $\mathcal{P}^{\alpha}_{Pol(\mathcal{C})}[\rho]$. Proposition 19 then yields that $Pol(\mathcal{C})$-covering is decidable.

Remark 20. *Our characterization of $\mathcal{P}^{\alpha}_{Pol(\mathcal{C})}[\rho]$ does not actually require ρ to be nice. While useless for deciding $Pol(\mathcal{C})$-covering, this is significant. This will later be a requirement when handling $BPol(\mathcal{C})$ and $PBPol(\mathcal{C})$ (see remark 43).*

Before we can present the characterization, we require some terminology. We define a new notion for rating maps: optimal ε-approximations. They are specifically designed to handle covering for concatenation hierarchies whose bases are made of group languages (it is through this notion that our characterizations are parameterized by $\mathcal{C}$-separation).

A. Optimal ε-approximations

When handling $Pol(\mathcal{C})$ for a quotient-closed Boolean algebra of group languages $\mathcal{C}$ (as well as $BPol(\mathcal{C})$ and $PBPol(\mathcal{C})$ later), we shall often encounter optimal $\mathcal{C}$-covers of the singleton $\{\varepsilon\}$ for various rating maps. The definitions presented here are based on a single key idea: there always exists an optimal $\mathcal{C}$-cover of $\{\varepsilon\}$ which consists of a single language.

Remark 21. *The definitions make sense for any lattice $\mathcal{C}$. However, they are mainly relevant when $\mathcal{C}$ is a class of group languages. In practice, the other important lattices contain the singleton $\{\varepsilon\}$ and $\{\{\varepsilon\}\}$ is always an optimal $\mathcal{C}$-cover of $\{\varepsilon\}$.*

Let $\mathcal{C}$ be a lattice and $\rho : 2^{A^*} \to R$ be a rating map. A *$\mathcal{C}$-optimal ε-approximation for ρ* is a language $L \in \mathcal{C}$ such that $\varepsilon \in L$ and,

$$\rho(L) \leq \rho(L') \quad \text{for every } L' \in \mathcal{C} \text{ such that } \varepsilon \in L'.$$

Remark 22. *Since $\{\varepsilon\}$ is a singleton, one may show that L is a $\mathcal{C}$-optimal ε-approximation for ρ if and only if $\{L\}$ is an optimal $\mathcal{C}$-cover of $\{\varepsilon\}$ for ρ.*

As expected, there always exists a $\mathcal{C}$-optimal ε-approximation for any rating map ρ, provided that $\mathcal{C}$ is a lattice.

Lemma 23. *For any lattice $\mathcal{C}$ and any rating map $\rho : 2^{A^*} \to R$, there exists a $\mathcal{C}$-optimal ε-approximation for ρ.*

We complete the definition with a key remark. Clearly, all $\mathcal{C}$-optimal ε-approximations for ρ have the same image under ρ. This is a canonical object for $\mathcal{C}$ and ρ. We write it $\mathbb{i}_{\mathcal{C}}[\rho] \in R$:

$$\mathbb{i}_{\mathcal{C}}[\rho] = \rho(L) \quad \text{for any } \mathcal{C}\text{-optimal } \varepsilon\text{-approximation } L \text{ for } \rho.$$

We turn to a crucial property of ε-approximations: when ρ is a *nice rating map*, computing $\mathbb{i}_{\mathcal{C}}[\rho]$ boils down to separation.

Lemma 24. *Let $\mathcal{C}$ be a lattice and $\rho : 2^{A^*} \to R$ be a nice rating map. Then, $\mathbb{i}_{\mathcal{C}}[\rho]$ is the sum of all $r \in R$ such that $\{\varepsilon\}$ is not $\mathcal{C}$-separable from $\rho_*^{-1}(r)$.*

Lemma 24 has an important corollary. When $\rho : 2^{A^*} \to R$ is a nice multiplicative rating map, the languages $\rho_*^{-1}(r)$ are regular (they are recognized by ρ_*). Hence, given a lattice $\mathcal{C}$, if $\mathcal{C}$-separation is decidable, we may compute $\mathbb{i}_{\mathcal{C}}[\rho]$ from an input nice multiplicative rating map.

Corollary 25. *Consider a lattice $\mathcal{C}$ such that $\mathcal{C}$-separation is decidable. Then, given as input a nice multiplicative rating map $\rho : 2^{A^*} \to R$, one may compute the element $\mathbb{i}_{\mathcal{C}}[\rho] \in R$.*

Corollary 25 is crucial for our approach: this is exactly how the algorithms for $Pol(\mathcal{C})$-, $BPol(\mathcal{C})$- and $PBPol(\mathcal{C})$-covering announced in Theorem 8 are parametrized by $\mathcal{C}$-separation.

Let us point out that when considering a *specific* lattice $\mathcal{C}$, it is usually possible to write an algorithm for computing $\mathbb{i}_{\mathcal{C}}[\rho]$ from an input nice multiplicative rating map which *does not* involve separation. We present an example for the class MOD.

Lemma 26. *Let $\rho : 2^{A^*} \to R$ be a nice multiplicative rating map. Then, $\mathbb{i}_{\text{MOD}}[\rho] = (\rho(A))^{\omega} + \rho(\varepsilon)$.*

Finally, we have the following lemma, which considers the special case when the lattice $\mathcal{C}$ is a quotient-closed Boolean algebra. We shall need this lemma for proving the three main results of the paper.

Lemma 27. *Let $\mathcal{C}$ be a quotient-closed Boolean algebra and $\rho : 2^{A^*} \to R$ be a rating map. Then, $\mathbb{i}_{\mathcal{C}}[\rho] \in \mathcal{I}_{Pol(\mathcal{C})}\left[\{\varepsilon\}, \rho\right]$.*

Proof. Let $\mathbf{K}$ be an optimal $Pol(\mathcal{C})$-cover of $\{\varepsilon\}$. By definition, there exists $K \in \mathbf{K}$ such that $\varepsilon \in K$. Since $K \in Pol(\mathcal{C})$, it is immediate from a simple induction on the definition of polynomial closure that we have $H \in \mathcal{C}$ such that $\varepsilon \in H$ and $H \subseteq K$ (the key idea is that the marked concatenation of two languages cannot contain ε). By definition of $\mathbb{i}_{\mathcal{C}}[\rho]$, we have $\mathbb{i}_{\mathcal{C}}[\rho] \leq \rho(H)$. Consequently, $\mathbb{i}_{\mathcal{C}}[\rho] \leq \rho(K)$, which yields $\mathbb{i}_{\mathcal{C}}[\rho] \in \mathcal{I}[\rho](\mathbf{K})$. Since $\mathbf{K}$ is an optimal $Pol(\mathcal{C})$-cover of $\{\varepsilon\}$, we obtain $\mathbb{i}_{\mathcal{C}}[\rho] \in \mathcal{I}_{Pol(\mathcal{C})}\left[\{\varepsilon\}, \rho\right]$, as desired. $\square$

B. Characterization

We characterize $\mathcal{P}^{\alpha}_{Pol(\mathcal{C})}[\rho]$ when $\mathcal{C}$ is a quotient-closed Boolean algebra of group languages. Let $\mathcal{C}$ be such a class.

We first state the characterization. Consider some morphism $\alpha : A^* \to M$ and some multiplicative rating map $\rho : 2^{A^*} \to R$. Consider a subset $S \subseteq M \times R$. We say that S is $Pol(\mathcal{C})$-*complete* for α and ρ when it satisfies the following properties:

- ***Trivial elements***: For all $w \in A^*$, $(\alpha(w), \rho(w)) \in S$.
- ***Downset***: We have $\downarrow_R S = S$.
- ***Multiplication***: For all, $(s, q), (t, r) \in S$, $(st, qr) \in S$.
- ***$\mathcal{C}$-operation***: We have $(1_M, \mathbb{i}_{\mathcal{C}}[\rho]) \in S$.

We may now state the main theorem of the section: the least $Pol(\mathcal{C})$-complete subset of $M \times R$ with respect to inclusion is exactly $\mathcal{P}^{\alpha}_{Pol(\mathcal{C})}[\rho]$ (recall that $\mathcal{C}$ is required to be a quotient-closed Boolean algebra of group languages).

Theorem 28. *Consider a morphism $\alpha : A^* \to M$ and a multiplicative rating map $\rho : 2^{A^*} \to R$. Then, $\mathcal{P}^{\alpha}_{Pol(\mathcal{C})}[\rho]$ is the least $Pol(\mathcal{C})$-complete subset of $M \times R$.*

By Theorem 28, if $\mathcal{C}$-separation is decidable, then one may compute $\mathcal{P}^{\alpha}_{Pol(\mathcal{C})}[\rho]$ when given a morphism $\alpha : A^* \to M$ and a nice multiplicative rating map $\rho : 2^{A^*} \to R$ as input. Indeed, computing the least $Pol(\mathcal{C})$-complete subset of $M \times R$ is achieved by an obvious least fixpoint procedure. We are able to implement $\mathcal{C}$-operation, since we may compute $\mathbb{i}_{\mathcal{C}}[\rho] \in R$: by Corollary 25, this boils down to deciding $\mathcal{C}$-separation.

Together with Proposition 19, this yields that when $\mathcal{C}$-separation is decidable, so is $Pol(\mathcal{C})$-covering. Therefore, we get the first part of Theorem 8, regarding level 1/2. We conclude the section with the proof of Theorem 28.

C. Proof of Theorem 28

We fix a morphism $\alpha : A^* \to M$ and a multiplicative rating map $\rho : 2^{A^*} \to R$ for the proof. We need to show that $\mathcal{P}^{\alpha}_{Pol(\mathcal{C})}[\rho]$ is the least $Pol(\mathcal{C})$-complete subset of $M \times R$. We first prove that it is $Pol(\mathcal{C})$-complete. This means that the least fixpoint procedure computing $\mathcal{P}^{\alpha}_{Pol(\mathcal{C})}[\rho]$ is sound.

Soundness. We have to show that $\mathcal{P}^{\alpha}_{Pol(\mathcal{C})}[\rho]$ satisfies the four properties in the definition of $Pol(\mathcal{C})$-complete subsets. The first three (trivial elements, downset and multiplication) are generic: they are satisfied by $\mathcal{P}^{\alpha}_{\mathcal{D}}[\rho]$ for any quotient-closed lattice $\mathcal{D}$ (we refer the reader to [31] for the proof). Hence, they are satisfied by $\mathcal{P}^{\alpha}_{Pol(\mathcal{C})}[\rho]$ since $Pol(\mathcal{C})$ is a quotient-closed lattice by Theorem 6. We concentrate on $\mathcal{C}$-operation.

We show that $(1_M, \mathbb{i}_{\mathcal{C}}[\rho]) \in \mathcal{P}^{\alpha}_{Pol(\mathcal{C})}[\rho]$. In other words, we prove $\mathbb{i}_{\mathcal{C}}[\rho] \in \mathcal{I}_{Pol(\mathcal{C})}\left[\alpha^{-1}(1_M), \rho\right]$. By Lemma 27, we know that $\mathbb{i}_{\mathcal{C}}[\rho] \in \mathcal{I}_{Pol(\mathcal{C})}\left[\{\varepsilon\}, \rho\right]$. Moreover, $\{\varepsilon\} \subseteq \alpha^{-1}(1_M)$ since α is a morphism. Thus, Fact 17 yields that $\mathbb{i}_{\mathcal{C}}[\rho] \in \mathcal{I}_{Pol(\mathcal{C})}\left[\alpha^{-1}(1_M), \rho\right]$, finishing the soundness proof.

Completeness. We now prove that $\mathcal{P}^{\alpha}_{Pol(\mathcal{C})}[\rho]$ is included in every $Pol(\mathcal{C})$-complete subset of $M \times R$. This direction corresponds to completeness of the least fixpoint procedure computing $\mathcal{P}^{\alpha}_{Pol(\mathcal{C})}[\rho]$. For the proof, we fix $S \subseteq M \times R$ which is $Pol(\mathcal{C})$-complete subset and show that $\mathcal{P}^{\alpha}_{Pol(\mathcal{C})}[\rho] \subseteq S$. The argument is based on the following proposition

Proposition 29. *Let $t \in M$. There exists a $Pol(\mathcal{C})$-cover $\mathbf{K}_t$ of $\alpha^{-1}(t)$ such that for every $K \in \mathbf{K}_t$, we have $(t, \rho(K)) \in S$.*

Let us first use Proposition 29 to show that $\mathcal{P}^{\alpha}_{Pol(\mathcal{C})}[\rho] \subseteq S$. Let $(t, r) \in \mathcal{P}^{\alpha}_{Pol(\mathcal{C})}[\rho]$. We show that $(t, r) \in S$. By definition, we have $r \in \mathcal{I}_{Pol(\mathcal{C})}\left[\alpha^{-1}(t), \rho\right]$. Consider the $Pol(\mathcal{C})$-cover $\mathbf{K}_t$ of $\alpha^{-1}(t)$ given by Proposition 29. By hypothesis, we have $r \in \mathcal{I}[\rho](\mathbf{K}_t)$. Thus, we get $K \in \mathbf{K}_t$ such that $r \leq \rho(K)$. Finally, since $(t, \rho(K)) \in S$ and $S = \downarrow_R S$ (S is $Pol(\mathcal{C})$-complete), we obtain that $(t, r) \in S$, finishing the proof.

It remains to prove Proposition 29. This requires the following technical lemma, already observed in [9, Proposition 3.11], which follows from a simple pumping argument. Note that this result is important: this only point where we use the hypothesis that $\mathcal{C}$ contains only *group* languages.

Lemma 30. *Let H be a regular language and L be a group language such that $\varepsilon \in L$. There is a cover $\mathbf{K}$ of H such that for every language $K \in \mathbf{K}$, we have $n \in \mathbb{N}$ and $a_1, \dots, a_n \in A$ such that $K = La_1L \cdots a_nL$ and $a_1 \cdots a_n \in H$.*

Remark 31. *The hypothesis that H is regular in Lemma 30 is actually not needed: the result holds when H is arbitrary. This follows from a result by Bucher, Ehrenfeucht and Haussler [8] on well quasi-orders. However, the regular case is simpler to prove and suffices for what we need to do here.*

Proof. Since H is regular, we have a morphism $\gamma : A^* \to N$ into a finite monoid N recognizing H. Moreover, since L is a group language it is recognized by a morphism $\beta : A^* \to G$ where G is a finite group. Since $\varepsilon \in L$, we have $\beta^{-1}(1_G) \subseteq L$.

For every $w = a_1 \cdots a_n \in A^*$, we let $F_w \subseteq A^*$ be the language $F_w = La_1L \cdots a_nL$ ($F_\varepsilon = L$). Moreover, we let $F'_w \subseteq F_w$ be the language $F'_w = \beta^{-1}(1_G)a_1\beta^{-1}(1_G) \cdots a_n\beta^{-1}(1_G)$. Let $n = |N \times G|$. We show that the following is a cover of H:

$$\{F'_w \mid w \in H \text{ and } |w| \leq n\}.$$

Clearly, this implies that $\mathbf{K} = \{F_w \mid w \in H \text{ and } |w| \leq n\}$ is a cover of H, finishing the proof. We use the following fact.

Fact 32. *Let $w \in H$ such that $|w| > n$. Then, there exists $v \in H$ such that $|v| < |w|$ and $F'_w \subseteq F'_v$.*

We first use the fact to finish the main proof. Let $w \in H$. We have to find $K \in \{F'_w \mid w \in H \text{ and } |w| \leq n\}$ such that $w \in K$. Observe that $w \in F'_w$ since by definition, $\varepsilon \in \beta^{-1}(1_G)$. We may then apply Fact 32 repeatedly to build $v \in H$ such that $|v| \leq n$ and $F'_w \subseteq F'_v$. Consequently, we have $w \in F'_v$ and we may choose $K = F'_v$. It remains to prove Fact 32.

We fix $w \in H$ such that $|w| > n$. We let $w = a_1 \cdots a_{|w|}$, where each a_i is a letter. Clearly, $N \times G$ is a monoid for the componentwise multiplication. Let $\eta : A^* \to N \times G$ be the morphism defined by $\eta(u) = (\gamma(u), \beta(u))$ for every $u \in A^*$. Since $|w| > n = |N \times G|$, it follows from the pigeonhole principle that there exist i, j such that $1 \leq i < j \leq |w|$ and $\eta(a_1 \cdots a_i) = \eta(a_1 \cdots a_j)$. Define $v = a_1 \cdots a_i a_{j+1} \cdots a_{|w|}$. Clearly, we have $|v| < |w|$. Moreover, $\eta(w) = \eta(v)$, hence $\gamma(w) = \gamma(v)$. Consequently, $v \in H$, since $w \in H$ and H is recognized by γ. It remains to show that $F'_w \subseteq F'_v$.

Let $u \in F'_w$: we have $u = u_0 a_1 u_1 a_2 u_2 \cdots a_{|w|} u_{|w|}$ with $u_i \in \beta^{-1}(1_G)$ for every $i \leq |w|$. Let $x = u_i a_{i+1} u_{i+1} \cdots a_j u_j$. Clearly, $\beta(x) = \beta(a_{i+1} \cdots a_j)$ and we know by definition of i, j that $\beta(a_1 \cdots a_i) = \beta(a_1 \cdots a_j)$. Thus, $\beta(a_1 \cdots a_i) \cdot \beta(x) = \beta(a_1 \cdots a_i)$. Since G is a group, it follows that $\beta(x) = 1_G$. Consequently, $x \in \beta^{-1}(1_G)$. Moreover, we have,

$$\begin{array}{rcl} u & = & u_0 a_1 u_1 \cdots a_i x a_{j+1} u_{j+1} \cdots a_{|w|} u_{|w|} \\ v & = & a_1 \cdots a_i a_{j+1} \cdots a_{|u|} \end{array}$$

It follows that $u \in F'_v$, which concludes the proof. $\square$

For $t \in M$, we are now ready to construct the $Pol(\mathcal{C})$-cover $\mathbf{K}_t$ of Proposition 29. Let L be a $\mathcal{C}$-optimal ε-approximation for ρ: we have $L \in \mathcal{C}$, $\varepsilon \in L$ and $\rho(L) = \mathbb{i}_{\mathcal{C}}[\rho]$. Since $\alpha^{-1}(t)$ is a regular language and L is a group language (by hypothesis on $\mathcal{C}$), Lemma 30 yields a cover $\mathbf{K}_t$ of $\alpha^{-1}(t)$ such that every language $K \in \mathbf{K}_t$ is of the form $K = La_1L \cdots a_nL$ with $a_1, \dots, a_n \in A$ such that $a_1 \cdots a_n \in \alpha^{-1}(t)$.

Clearly, $\mathbf{K}_t$ is a $Pol(\mathcal{C})$-cover of $\alpha^{-1}(t)$ since $L \in \mathcal{C}$. It remains to prove that for every $K \in \mathbf{K}_t$, we have $(t, \rho(K)) \in S$. By definition of $\mathbf{K}_t$, there exists $w = a_1 \cdots a_n \in \alpha^{-1}(t)$ such that $K = La_1L \cdots a_nL$. By definition of $Pol(\mathcal{C})$-complete subsets, we have $(\alpha(a_i), \rho(a_i)) \in S$ for every $i \leq n$ (these are trivial elements). Moreover, we know from $Pol(\mathcal{C})$-operation and the definition of L as a $\mathcal{C}$-optimal ε-approximation for ρ that $(1_M, \rho(L)) = (1_M, \mathbb{i}_{\mathcal{C}}[\rho]) \in S$. It then follows from closure under multiplication that:

$$\left(1_M \cdot \prod_{1 \leq i \leq n} (\alpha(a_i) \cdot 1_M), \quad \rho(L) \cdot \prod_{1 \leq i \leq n} (\rho(a_i) \cdot \rho(L))\right) \in S.$$

Since α is a morphism and ρ is a multiplicative rating map (and therefore a morphism for multiplication), this exactly says that $(\alpha(w), \rho(K)) \in S$. Finally, $\alpha(w) = t$ by definition and we obtain as desired that $(t, \rho(K)) \in S$, finishing the proof. $\square$

VI. Extending the Framework

In this section, we introduce additional material about rating maps required to prove the remaining part of Theorem 8 (*i.e.*, concerning levels 1 and 3/2). Let us first overview the situation.

In Section IV, we proved that given an arbitrary lattice $\mathcal{D}$, deciding $\mathcal{D}$-covering boils down to computing $\mathcal{P}^{\alpha}_{\mathcal{D}}[\rho]$ from a morphism α and a nice multiplicative rating map ρ (see Proposition 19). In fact, if $\mathcal{D}$ is a Boolean algebra, we may even work with the simpler set $\mathcal{I}_{\mathcal{D}}[\rho]$ (see Proposition 18). This is how we handled $Pol(\mathcal{C})$ in the previous section: we showed that given a fixed quotient-closed Boolean algebra of group languages $\mathcal{C}$, if $\mathcal{C}$-separation is decidable, then $\mathcal{P}^{\alpha}_{Pol(\mathcal{C})}[\rho]$ can be computed via a least fixpoint algorithm. We shall use the same approach for the classes $BPol(\mathcal{C})$ and $PBPol(\mathcal{C})$: we obtain algorithms for computing $\mathcal{I}_{BPol(\mathcal{C})}[\rho]$ and $\mathcal{P}^{\alpha}_{PBPol(\mathcal{C})}[\rho]$. However, we do not compute these sets directly. Instead, we work with more involved sets, which *carry more information*. In this section, we introduce these other sets.

Remark 33. *This situation is not surprising. Typically, computing optimal imprints is achieved via fixpoint procedures (what we did for $Pol(\mathcal{C})$ in Section VII is a typical example). Hence, replacing the object that we truly want to compute by another one which carries more information makes sense.*

First, we present two constructions for building new rating maps out of already existing ones. They are taken from [33], where they are used as technical proof objects. Here, we use them in a more prominent way: they are central to the definitions of this section. We use these constructions to present refined variants of Proposition 18 and Proposition 19.

A. Nested rating maps

We present two constructions. The first one involves two objects: a lattice $\mathcal{D}$ and a rating map $\rho : 2^{A^*} \to R$. We build a new map $\xi_{\mathcal{D}}[\rho]$ whose rating set is 2^R. We let:

$$\begin{array}{rccl} \xi_{\mathcal{D}}[\rho]: & (2^{A^*}, \cup) & \to & (2^R, \cup) \\ & K & \mapsto & \mathcal{I}_{\mathcal{D}}[K, \rho] \end{array}$$

The second construction takes an additional object as input: a map $\alpha : A^* \to M$ where M is some arbitrary finite set (in practice, α will be a monoid morphism). We build a new map $\zeta^{\alpha}_{\mathcal{D}}[\rho]$ whose rating set is $2^{M \times R}$:

$$\begin{array}{rccl} \zeta^{\alpha}_{\mathcal{D}}[\rho]: & (2^{A^*}, \cup) & \to & (2^{M \times R}, \cup) \\ & K & \mapsto & \{(s, r) \mid r \in \mathcal{I}_{\mathcal{D}}\left[\alpha^{-1}(s) \cap K, \rho\right]\} \end{array}$$

It is straightforward to verify that these two maps are in fact rating maps (a proof is available in [33], see Proposition 6.2). We state this result in the following proposition.

Proposition 34. *Given a lattice $\mathcal{D}$, a map $\alpha : A^* \to M$ and a rating map $\rho : 2^{A^*} \to R$, $\xi_{\mathcal{D}}[\rho]$ and $\zeta^{\alpha}_{\mathcal{D}}[\rho]$ are rating maps.*

A key point is that the rating maps $\xi_{\mathcal{D}}[\rho]$ and $\zeta^{\alpha}_{\mathcal{D}}[\rho]$ are **not** nice in general, even when the original rating map ρ is (we refer the reader to [33] for a counterexample).

Another question is whether they are multiplicative. Clearly, when $\alpha : A^* \to M$ is a monoid morphism and $\rho : 2^{A^*} \to R$ is a multiplicative rating map, we may lift the multiplications of M and R to the sets 2^R and $2^{M \times R}$ in the natural way. This makes $(2^R, \cup, \cdot)$ and $(2^{M \times R}, \cup, \cdot)$ semirings. Unfortunately, $\xi_{\mathcal{D}}[\rho]$ and $\zeta^{\alpha}_{\mathcal{D}}[\rho]$ are **not** multiplicative, even under these hypotheses. However, they behave almost as multiplicative rating maps when the class $\mathcal{D}$ satisfies appropriate properties.

Consider an arbitrary lattice $\mathcal{G}$ and a rating map $\rho : 2^{A^*} \to R$ whose rating set is a semiring $(R, +, \cdot)$ (but ρ need not be multiplicative). We say that ρ is *$\mathcal{G}$-multiplicative* when there exists an endomorphism μ_ρ of $(R, +)$ such that:

1) For every $q, r, s \in R$, $\mu_\rho(q \mu_\rho(r) s) = \mu_\rho(qrs)$.
2) $1_R \leq \rho(\varepsilon)$.
3) For every $H, K \in \mathcal{G}$ and $a \in A$, we have,

$$\begin{array}{rcl} \rho(H) & = & \mu_\rho(\rho(H)) \\ \rho(HK) & = & \mu_\rho\big(\rho(H) \cdot \rho(K)\big) \\ \rho(HaK) & = & \mu_\rho\big(\rho(H) \cdot \rho(a) \cdot \rho(K)\big) \end{array}$$

When $\mathcal{G}$ is the class of all languages (*i.e.*, $\mathcal{G} = 2^{A^*}$), we say that ρ is *quasi-multiplicative*. In practice, we shall always assume implicitly that the endomorphism μ_ρ is fixed.

Remark 35. *A true multiplicative rating map is always quasi-multiplicative. Indeed, in this case, it suffices to choose μ_ρ as the identity; $\mu_\rho(r) = r$ for all $r \in R$.*

We have the two following lemmas, which apply to $\zeta^\alpha_{\mathcal{D}}[\rho]$ and $\xi_{\mathcal{D}}[\rho]$ respectively (see [33, Lemmas 6.7 and 6.8] for proofs). The statements are rather *ad hoc*: they are designed to accommodate the situations that we shall encounter. We use $\zeta^\alpha_{\mathcal{D}}[\rho]$ in the case when $\mathcal{D} = Pol(\mathcal{C})$ while we use $\xi_{\mathcal{D}}[\rho]$ in the case when $\mathcal{D} = BPol(\mathcal{C})$.

Lemma 36. *Let $\mathcal{D}$ be a quotient-closed lattice closed under concatenation, $\alpha : A^* \to M$ be a morphism and $\rho : 2^{A^*} \to R$ be a multiplicative rating map. Then, $\zeta^\alpha_{\mathcal{D}}[\rho]$ is quasi-multiplicative, and the endomorphism $\mu_{\zeta^\alpha_{\mathcal{D}}[\rho]}$ of $(2^{M\times R}, \cup)$ is:*

$$\mu_{\zeta^\alpha_{\mathcal{D}}[\rho]}(T) = \downarrow_R T \quad \textit{for all } T \in 2^{M\times R}.$$

Lemma 37. *Let $\mathcal{G}$ be a quotient-closed lattice closed under concatenation and* **marked** *concatenation and let $\mathcal{D}$ = $Bool(\mathcal{G})$. Let $\rho : 2^{A^*} \to R$ a multiplicative rating map. Then, $\xi_{\mathcal{D}}[\rho]$ is $\mathcal{G}$-multiplicative and the endomorphism $\mu_{\xi_{\mathcal{D}}[\rho]}$ of $(2^R, \cup)$ is:*

$$\mu_{\xi_{\mathcal{D}}[\rho]}(T) = \downarrow_R T \quad \textit{for all } T \in 2^R.$$

B. Application to classes of group languages

We may now present the refinements of Proposition 18 and Proposition 19 announced at the beginning. They are designed to handle the classes investigated in the paper. Here, we discuss and prove the variant for Boolean algebras and simply state the one for lattices.

By Proposition 18, given a Boolean algebra $\mathcal{D}$, deciding $\mathcal{D}$-covering boils down to computing $\mathcal{I}_{\mathcal{D}}[\rho] \subseteq R$ from a nice multiplicative rating map $\rho : 2^{A^*} \to R$. It turns out that when $\mathcal{D}$ is a full level in some concatenation hierarchy whose basis $\mathcal{C}$ is made of group languages, one may instead compute a specific subset of $\mathcal{I}_{\mathcal{D}}[\rho]$.

By definition, $\mathcal{I}_{\mathcal{D}}[\rho] = \mathcal{I}_{\mathcal{D}}[A^*, \rho] = \xi_{\mathcal{D}}[\rho](A^*)$. We may consider the element $\mathbb{i}_{\mathcal{C}}[\xi_{\mathcal{D}}[\rho]] \in 2^R$ which is equal to $\xi_{\mathcal{D}}[\rho](L)$ for a $\mathcal{C}$-optimal ε-approximation L for $\xi_{\mathcal{D}}[\rho]$. Hence, we have,

$$\mathbb{i}_{\mathcal{C}}[\xi_{\mathcal{D}}[\rho]] = \xi_{\mathcal{D}}[\rho](L) \subseteq \xi_{\mathcal{D}}[\rho](A^*) = \mathcal{I}_{\mathcal{D}}[\rho].$$

It turns out that because of our hypothesis on $\mathcal{D}$, one may compute $\mathcal{I}_{\mathcal{D}}[\rho]$ from its subset $\mathbb{i}_{\mathcal{C}}[\xi_{\mathcal{D}}[\rho]]$ (see below for the proof). This yields the following corollary of Proposition 18.

Proposition 38. *Let $\mathcal{C}$ be a quotient-closed Boolean algebra of group languages and let $\mathcal{D}$ be a strictly positive full level within the concatenation hierarchy of basis $\mathcal{C}$. Assume that there exists an algorithm computing $\mathbb{i}_{\mathcal{C}}[\xi_{\mathcal{D}}[\rho]]$ from an input nice multiplicative rating map ρ. Then, $\mathcal{D}$-covering is decidable.*

Before we prove Proposition 38, let us present the generalized variant of Proposition 19 which considers all lattices instead of just Boolean algebras.

Proposition 39. *Let $\mathcal{C}$ be a quotient-closed Boolean algebra of group languages and let $\mathcal{D}$ be a half level within the concatenation hierarchy of basis $\mathcal{C}$. Assume that there exists an algorithm computing $\mathbb{i}_{\mathcal{C}}[\zeta^\alpha_{\mathcal{D}}[\rho]]$ from a morphism α and a nice multiplicative rating map ρ. Then, $\mathcal{D}$-covering is decidable.*

C. Proof of Proposition 38

We fix $\mathcal{C}$ and $\mathcal{D}$ as in the statement of the proposition. We show that for an arbitrary nice multiplicative rating map ρ, one may compute $\mathcal{I}_{\mathcal{D}}[\rho]$ from its subset $\mathbb{i}_{\mathcal{C}}[\xi_{\mathcal{D}}[\rho]]$. The result is then an immediate corollary of Proposition 18.

Fix a nice multiplicative rating map $\rho : 2^{A^*} \to R$ for the proof. We let $S \subseteq R$ be the least subset of R that satisfies the following properties:

1) We have $\mathbb{i}_{\mathcal{C}}[\xi_{\mathcal{D}}[\rho]] \subseteq S$.
2) We have $\rho(w) \in S$ for every $w \in A^*$.
3) We have $\downarrow_R S = S$.
4) For every $r, r' \in S$, we have $rr' \in S$.

Clearly, one may compute S from ρ and $\mathbb{i}_{\mathcal{C}}[\xi_{\mathcal{D}}[\rho]]$. Hence, Proposition 38 is now immediate from the following lemma.

Lemma 40. *We have $S = \mathcal{I}_{\mathcal{D}}[\rho]$.*

We prove Lemma 40. Let L be a $\mathcal{C}$-optimal ε-approximation for $\xi_{\mathcal{D}}[\rho]$: we have $L \in \mathcal{C}$, $\varepsilon \in L$ and $\mathbb{i}_{\mathcal{C}}[\xi_{\mathcal{D}}[\rho]] = \xi_{\mathcal{D}}[\rho](L)$.

We first show that $S \subseteq \mathcal{I}_{\mathcal{D}}[\rho]$. This amounts to proving that $\mathcal{I}_{\mathcal{D}}[\rho]$ satisfies the four properties in the definition of S. That $\mathbb{i}_{\mathcal{C}}[\xi_{\mathcal{D}}[\rho]] \subseteq \mathcal{I}_{\mathcal{D}}[\rho]$ is immediate since $\mathbb{i}_{\mathcal{C}}[\xi_{\mathcal{D}}[\rho]] = \xi_{\mathcal{D}}[\rho](L)$, $\mathcal{I}_{\mathcal{D}}[\rho] = \xi_{\mathcal{D}}[\rho](A^*)$, $L \subseteq A^*$ and $\xi_{\mathcal{D}}[\rho]$ is a rating map. The other three properties are generic: they are satisfied whenever ρ is a multiplicative rating map and $\mathcal{D}$ is a quotient-closed Boolean algebra (which is the case here as $\mathcal{D}$ is a full level within the hierarchy of basis $\mathcal{C}$). We refer the reader to [31, Lemma 6.11] for the proof.

We turn to the converse inclusion: $\mathcal{I}_{\mathcal{D}}[\rho] \subseteq S$. Consider $r \in \mathcal{I}_{\mathcal{D}}[\rho]$, we show that $r \in S$. The argument is based on the following lemma, which is where we use the hypothesis that $\mathcal{C}$ is a class of group languages (which implies that $L \in \mathcal{C}$ is recognized by a finite group).

Lemma 41. *There exist $\ell \in \mathbb{N}$ and ℓ letters $a_1, \dots, a_\ell \in A$ such that $r \in \xi_{\mathcal{D}}[\rho](La_1L \cdots a_\ell L)$.*

Proof. For every word $w = a_1 \cdots a_\ell \in A^*$, we let H_w be the language $H_w = La_1L \cdots a_\ell L$ ($H_\varepsilon = L$). We have to find $w \in A^*$ such that $r \in \xi_{\mathcal{D}}[\rho](H_w)$. Clearly A^* is a regular language and L is a group language by hypothesis. Therefore, Lemma 30 implies that there exists a *finite* language U such that $A^* \subseteq \bigcup_{w\in U} H_w$. Since $\xi_{\mathcal{D}}[\rho]$ is a rating map, this implies:

$$\xi_{\mathcal{D}}[\rho](A^*) = \bigcup_{w\in U} \xi_{\mathcal{D}}[\rho](H_w).$$

Moreover, by hypothesis, we have $r \in \mathcal{I}_{\mathcal{D}}[\rho] = \mathcal{I}_{\mathcal{D}}[A^*, \rho] = \xi_{\mathcal{D}}[\rho](A^*)$. Thus, $r \in \xi_{\mathcal{D}}[\rho](H_w)$ for some $w \in A^*$ and the lemma is proved. ☐

By definition $\mathcal{D}$ is a **strictly positive** full level within the hierarchy of basis $\mathcal{C}$. Therefore, $\mathcal{D} = Bool(\mathcal{G})$ where $\mathcal{G}$ is the preceding half level in the hierarchy. In particular, $\mathcal{G}$ is a quotient-closed lattice closed under concatenation and marked concatenation by Lemma 5 and Theorem 6. Consequently, we obtain from Lemma 37 that $\xi_{\mathcal{D}}[\rho]$ is $\mathcal{G}$-multiplicative and the associated endomorphism of $(2^R, \cup)$ is defined by $\mu_{\xi_{\mathcal{D}}[\rho]}(T) = \downarrow_R T$ for all $T \in 2^R$. Moreover, we have

$L \in \mathcal{C} \subseteq \mathcal{G}$ by definition. Hence, by Axioms 1 and 3 in the definition of $\mathcal{G}$-multiplicative rating maps, the hypothesis that $r \in \xi_{\mathcal{D}}[\rho](La_1L \cdots a_\ell L)$ given by Lemma 41 implies that,

$$r \in \downarrow_R(\xi_{\mathcal{D}}[\rho](L) \cdot \xi_{\mathcal{D}}[\rho](a_1) \cdot \xi_{\mathcal{D}}[\rho](L) \cdots \xi_{\mathcal{D}}[\rho](a_\ell) \cdot \xi_{\mathcal{D}}[\rho](L))$$

By definition of L, we have $\xi_{\mathcal{D}}[\rho](L) = \mathbb{i}_{\mathcal{C}}[\xi_{\mathcal{D}}[\rho]]$. Moreover, we have the following fact.

Fact 42. *For every $a \in A$, we have $\xi_{\mathcal{D}}[\rho](a) \subseteq \downarrow_R\{\rho(a)\}$.*

Proof. Fact 7 yields that $\{a\} \in \mathcal{D}$ ($\mathcal{D}$ is a level $n \geq 1$ in the hierarchy of basis $\mathcal{C}$). Hence, $\{\{a\}\}$ is a $\mathcal{D}$-cover of $\{a\}$. Thus, $\xi_{\mathcal{D}}[\rho](a) = \mathcal{I}_{\mathcal{D}}[\{a\}, \rho] \subseteq \mathcal{I}[\rho](\{a\}) = \downarrow_R\{\rho(a)\}$, which finishes the proof. $\square$

Altogether, we obtain:

$$r \in \downarrow_R(\mathbb{i}_{\mathcal{C}}[\xi_{\mathcal{D}}[\rho]] \cdot \{\rho(a_1)\} \cdot \mathbb{i}_{\mathcal{C}}[\xi_{\mathcal{D}}[\rho]] \cdots \{\rho(a_\ell)\} \cdot \mathbb{i}_{\mathcal{C}}[\xi_{\mathcal{D}}[\rho]]).$$

By definition of S, this implies $r \in S$, concluding the proof.

VII. Boolean polynomial closure

We turn to the second part of Theorem 8: for a quotient-closed Boolean algebra of group languages $\mathcal{C}$, if $\mathcal{C}$-separation is decidable, then so is $BPol(\mathcal{C})$-covering. We fix $\mathcal{C}$ for the section.

We rely on rating maps and use the notions introduced in Section VI. For any *nice* multiplicative rating map $\rho : 2^{A^*} \to R$, we characterize $\mathbb{i}_{\mathcal{C}}[\xi_{BPol(\mathcal{C})}[\rho]]$ as the greatest subset of R satisfying specific properties. When $\mathcal{C}$-separation is decidable, this yields a fixpoint algorithm for computing $\mathbb{i}_{\mathcal{C}}[\xi_{BPol(\mathcal{C})}[\rho]]$ from ρ. Consequently, we get the decidability of $BPol(\mathcal{C})$-covering from Proposition 38.

Remark 43. *The conditions for applying our characterizations are more restrictive than those we had for $Pol(\mathcal{C})$ in the previous section. We require ρ to be nice: while we are able to handle $Pol(\mathcal{C})$ for arbitrary multiplicative rating maps, we are restricted to nice ones for $BPol(\mathcal{C})$. This is irrelevant for the decidability of covering: considering nice multiplicative rating maps suffices. However, this is a key point: the proof of the $BPol(\mathcal{C})$-characterization involves handling the simpler class $Pol(\mathcal{C})$ for auxiliary $Pol(\mathcal{C})$-multiplicative rating maps which are **not** nice. This is why we are not able to get results for higher levels in concatenation hierarchies: our knowledge about level 1/2 is stronger than the decidability of covering, and we are unable to replicate it for levels 1 and 3/2 (the situation is the same for $PBPol(\mathcal{C})$).*

We first present the characterization of $\mathbb{i}_{\mathcal{C}}[\xi_{BPol(\mathcal{C})}[\rho]]$. For every multiplicative rating map $\rho : 2^{A^*} \to R$, we shall define the $BPol(\mathcal{C})$*-complete subsets of R for ρ*. Our theorem states that when ρ is nice, $\mathbb{i}_{\mathcal{C}}[\xi_{BPol(\mathcal{C})}[\rho]]$ is the *greatest* such subset.

We fix the multiplicative rating map $\rho : 2^{A^*} \to R$ for the definition. We successively lift the multiplication "$\cdot$" of R to the sets 2^R, $R \times 2^R$ and $2^{R\times 2^R}$ in the natural way. This makes $(2^{R\times 2^R}, \cup, \cdot)$ an idempotent semiring. For every $S \subseteq R$, we use it as the rating set of a nice multiplicative rating map,

$$\eta_{\rho,S} : (2^{A^*}, \cup, \cdot) \to (2^{R\times 2^R}, \cup, \cdot).$$

Since we are defining a nice multiplicative rating map, it suffices to specify the evaluation of letters. For $a \in A$, we let,

$$\eta_{\rho,S}(a) = \{(\rho(a), \quad S \cdot \{\rho(a)\} \cdot S)\}.$$

By definition, we have $\mathbb{i}_{\mathcal{C}}[\eta_{\rho,S}] \subseteq R \times 2^R$ for every $S \subseteq R$.

We now define the $BPol(\mathcal{C})$-complete subsets. We say that $S \subseteq R$ is $BPol(\mathcal{C})$-complete for ρ if for every $s \in S$, there exist $(r_1, U_1), \dots, (r_k, U_k) \in \mathbb{i}_{\mathcal{C}}[\eta_{\rho,S}]$ such that,

$$\begin{array}{c} s \leq r_1 + \cdots + r_k \text{ and,} \\ r_1 + \cdots + r_k \in \downarrow_R U_i \text{ for every } i \leq k \end{array} \tag{1}$$

We are ready to state the main theorem of this section: when ρ is nice, the greatest $BPol(\mathcal{C})$-complete subset of R (with respect to inclusion) is exactly $\mathbb{i}_{\mathcal{C}}[\xi_{BPol(\mathcal{C})}[\rho]]$.

Theorem 44. *Let $\rho : 2^{A^*} \to R$ be a nice multiplicative rating map. Then, $\mathbb{i}_{\mathcal{C}}[\xi_{BPol(\mathcal{C})}[\rho]]$ is the greatest $BPol(\mathcal{C})$-complete subset of R for ρ.*

Due to lack of space, we omit the proof of Theorem 44. It is rather involved and exploits two results for rating maps that were designed in [33] to handle Boolean closure in a general context, which are completed with arguments specific to our setting. Here, we only discuss the applications of Theorem 44.

Provided that $\mathcal{C}$-separation is decidable, Theorem 44 yields a greatest fixpoint procedure for computing $\mathbb{i}_{\mathcal{C}}[\xi_{BPol(\mathcal{C})}[\rho]]$ from an input nice multiplicative rating map $\rho : 2^{A^*} \to R$. Indeed, consider the following sequence of subsets $S_0 \supseteq S_1 \supseteq S_2 \cdots$. We let $S_0 = R$ and for $i \geq 1$, S_i contains all $s \in S_{i-1}$ such that there exists $(r_1, U_1), \dots, (r_k, U_k) \in \mathbb{i}_{\mathcal{C}}[\eta_{\rho,S_{i-1}}]$ satisfying (1):

$$\begin{array}{c} s \leq r_1 + \cdots + r_k \text{ and,} \\ r_1 + \cdots + r_k \in \downarrow_R U_i \text{ for all } i \leq k \end{array}$$

Clearly, computing S_i from S_{i-1} boils down to computing $\mathbb{i}_{\mathcal{C}}[\eta_{\rho,S_{i-1}}]$. By Corollary 25, this is possible since we have an algorithm for $\mathcal{C}$-separation ($\eta_{\rho,S_{i-1}}$ is a **nice** multiplicative rating map that we may compute from S_{i-1}).

Finally, since R is finite, the sequence $S_0 \supseteq S_1 \supseteq S_2 \cdots$ stabilizes at some point: there exists some $i \in \mathbb{N}$ such that $S_i = S_j$ for all $j \geq i$. One may verify that S_i is the greatest $BPol(\mathcal{C})$-complete subset of R and we may compute it. Theorem 44 then states that $S_i = \mathbb{i}_{\mathcal{C}}[\xi_{BPol(\mathcal{C})}[\rho]]$.

By Proposition 38, a procedure for computing $\mathbb{i}_{\mathcal{C}}[\xi_{BPol(\mathcal{C})}[\rho]]$ yields an algorithm for $BPol(\mathcal{C})$-covering. Hence, we get the part of Theorem 8 regarding $BPol(\mathcal{C})$ (*i.e.*, level 1) as a corollary: when $\mathcal{C}$-separation is decidable, so is $BPol(\mathcal{C})$-covering (as well as $BPol(\mathcal{C})$-separation by Lemma 4).

VIII. Nested polynomial closure

This section is devoted to the final part of Theorem 8: for every quotient-closed Boolean algebra of group languages $\mathcal{C}$, if $\mathcal{C}$-separation is decidable, then so is $PBPol(\mathcal{C})$-covering. As usual, we fix $\mathcal{C}$ for the section.

We rely on the framework outlined in Sections IV and VI. For every morphism $\alpha : A^* \to M$ and every *nice* multiplicative rating map $\rho : 2^{A^*} \to R$, we characterize $\mathbb{i}_{\mathcal{C}}[\zeta^{\alpha}_{PBPol(\mathcal{C})}[\rho]]$ as the least subset of $M \times R$ satisfying specific properties. When

$\mathcal{C}$-separation is decidable, this yields a least fixpoint algorithm for computing $\mathbb{i}_{\mathcal{C}}[\zeta^{\alpha}_{PBPol(\mathcal{C})}[\rho]]$ from ρ. Consequently, we get the decidability of $PBPol(\mathcal{C})$-covering by Proposition 39.

We start by presenting the characterization. Consider a morphism $\alpha : A^* \to M$ and a multiplicative rating map $\rho : 2^{A^*} \to R$. We define a notion of $PBPol(\mathcal{C})$-complete subset of $M \times R$. Our main theorem then states that when ρ is nice, the least such subset is exactly $\mathbb{i}_{\mathcal{C}}[\zeta^{\alpha}_{PBPol(\mathcal{C})}[\rho]]$. The definition depends on auxiliary nice multiplicative rating maps. We first present them.

We successively lift the multiplications of M and R to the sets $2^{M\times R}$, $R \times 2^{M\times R}$ and $2^{R\times 2^{M\times R}}$ in the natural way. This makes $(2^{R\times 2^{M\times R}}, \cup, \cdot)$ an idempotent semiring. For every $S \subseteq M \times R$, we define a nice multiplicative rating map:

$$\eta_{\alpha,\rho,S} : (2^{A^*}, \cup, \cdot) \to (2^{R\times 2^{M\times R}}, \cup, \cdot).$$

Since we are defining a nice multiplicative rating map, it suffices to specify the evaluation of letters. For $a \in A$, we let,

$$\eta_{\alpha,\rho,S}(a) = \{(\rho(a), \quad S \cdot \{(\alpha(a), \rho(a))\} \cdot S)\}$$

Observe that by definition, we have $\mathbb{i}_{\mathcal{C}}[\eta_{\alpha,\rho,S}] \subseteq R \times 2^{M\times R}$.

We may now define $PBPol(\mathcal{C})$-complete subsets. Consider $S \subseteq M \times R$. We say that S is $PBPol(\mathcal{C})$-complete for α and ρ when the following conditions are satisfied:

- ***Downset.*** We have $\downarrow_R S \subseteq S$.
- ***Multiplication.*** We have $S \cdot S \subseteq S$.
- $\mathcal{C}$***-operation.*** For all $(r, T) \in \mathbb{i}_{\mathcal{C}}[\eta_{\alpha,\rho,S}]$, we have $T \subseteq S$.
- $PBPol(\mathcal{C})$***-operation.*** For all $(r, T) \in \mathbb{i}_{\mathcal{C}}[\eta_{\alpha,\rho,S}]$ and every idempotent $(e, f) \in \downarrow_R T \subseteq M \times R$, we have:

$$(e, f \cdot (1_R + r) \cdot f) \in S.$$

We may now state the main theorem of this section. When ρ is nice, $\mathbb{i}_{\mathcal{C}}[\zeta^{\alpha}_{PBPol(\mathcal{C})}[\rho]]$ is the least $PBPol(\mathcal{C})$-complete subset of $M \times R$ (with respect to inclusion).

Theorem 45. *Fix a morphism $\alpha : A^* \to M$ and a nice multiplicative rating map $\rho : 2^{A^*} \to R$. Then, $\mathbb{i}_{\mathcal{C}}[\zeta^{\alpha}_{PBPol(\mathcal{C})}[\rho]]$ is the least $PBPol(\mathcal{C})$-complete subset of $M \times R$.*

The proof of Theorem 45 is omitted. Similarly to most results of this kind, it involves two directions: one needs to show that $\mathbb{i}_{\mathcal{C}}[\zeta^{\alpha}_{PBPol(\mathcal{C})}[\rho]]$ is $PBPol(\mathcal{C})$-complete and then that it is included in every other $PBPol(\mathcal{C})$-complete subset. The former direction is proved directly. The latter one requires applying the main theorem of [26] (for classes $PBPol(\mathcal{D})$ with $\mathcal{D}$ a **finite** quotient-closed Boolean algebra).

As expected, when $\mathcal{C}$-separation is decidable, Theorem 45 yields a least fixpoint procedure for computing $\mathbb{i}_{\mathcal{C}}[\zeta^{\alpha}_{PBPol(\mathcal{C})}[\rho]]$ from a morphism $\alpha : A^* \to M$ and a nice multiplicative rating map $\rho : 2^{A^*} \to R$. One starts from the empty set $\emptyset \subseteq R$ and saturates it with the four operations in the definition of $PBPol(\mathcal{C})$-complete subsets. Clearly, they may be implemented. This is immediate for downset and multiplication. Moreover, we are able to implement $\mathcal{C}$-operation and $PBPol(\mathcal{C})$-operation by Corollary 25 since $\mathcal{C}$-separation is decidable. Eventually, the computation reaches a fixpoint and it is straightforward to verify that this set is the least $PBPol(\mathcal{C})$-complete subset of $M \times R$, *i.e.*, $\mathbb{i}_{\mathcal{C}}[\zeta^{\alpha}_{PBPol(\mathcal{C})}[\rho]]$ by Theorem 45.

By Proposition 39, we know that a procedure for computing $\mathbb{i}_{\mathcal{C}}[\zeta^{\alpha}_{PBPol(\mathcal{C})}[\rho]]$ yields an algorithm for $PBPol(\mathcal{C})$-covering. Hence, we get the $PBPol(\mathcal{C})$ part of Theorem 8 as a corollary: when $\mathcal{C}$-separation is decidable, so is $PBPol(\mathcal{C})$-covering (as well as $PBPol(\mathcal{C})$-separation by Lemma 4).

IX. Conclusion

We proved that for any quotient-closed Boolean algebra of *group languages* $\mathcal{C}$, if $\mathcal{C}$-separation is decidable, then so are separation and covering for levels 1/2, 1 and 3/2 in the concatenation hierarchy of basis $\mathcal{C}$. A corollary is that these levels enjoy decidable membership, as well as level 5/2. This result nicely complements analogous statements that apply to *finitely based* concatenation hierarchies [33], [26].

These results can be instantiated for several classical bases. First, one may consider the basis made of all group languages. This yields the so-called "group hierarchy" of Pin and Margolis [17], for which decidability of membership was known only up to level 1. Another application is the quantifier alternation hierarchy of first-order logic with modular predicates. It corresponds to the basis consisting of languages counting the length of words modulo some number (separation was shown decidable at level 1 in [40] with specific techniques). A third example is the basis consisting of all languages counting the number of occurrences of letters modulo some number, which are exactly languages recognized by finite commutative groups. In all cases, we get decidability of separation and covering for levels 1/2, 1 and 3/2 (and membership for level 5/2).

There are natural follow-up questions to this work. The most immediate one is whether our results may be pushed to higher levels. This is difficult. There are no known generic results for the levels above 2, even for finitely based hierarchies. Actually, there is one hierarchy for which covering is know to be decidable up to level 5/2: the Straubing-Thérien hierarchy (its basis is $\{\emptyset, A^*\}$). Unfortunately, this is based on a specific property of this hierarchy: its levels 2 and 5/2 are also levels 1 and 3/2 in another finitely based hierarchy [23]. We do not have a similar property for arbitrary group based hierarchies.

One can also investigate other closure operations, such as the star-free closure $\mathcal{C} \mapsto \mathrm{SF}(\mathcal{C})$, which is the union of all levels in the hierarchy of basis $\mathcal{C}$ (*i.e.*, $\mathrm{SF}(\mathcal{C})$ is the least class containing $\mathcal{C}$, closed under Boolean operations and marked concatenation). If $\mathcal{C}$ is a quotient-closed Boolean algebra of group languages and $\mathcal{C}$-separation is decidable, is $\mathrm{SF}(\mathcal{C})$-covering decidable?

Acknowledgments

Partly funded by the DeLTA project (ANR-16-CE40-0007), and, for the first author, by IUF, Institut Universitaire de France.

References

[1] J. Almeida, "Some algorithmic problems for pseudovarieties," *Publicationes Mathematicae Debrecen*, vol. 54, pp. 531–552, 1999.

[2] M. Arfi, "Polynomial operations on rational languages," in *STACS'87*, ser. Lect. Notes Comp. Sci., 1987, pp. 198–206.

[3] ——, "Opérations polynomiales et hiérarchies de concaténation," *Theoretical Computer Science*, vol. 91, no. 1, pp. 71–84, 1991.

[4] C. J. Ash, "Inevitable graphs: a proof of the type II conjecture and some related decision procedures," *IJAC*, vol. 1, no. 1, pp. 127–146, 1991.

[5] K. Auinger and B. Steinberg, "On the extension problem for partial permutations," *Proc. Amer. Math. Soc.*, vol. 131, no. 9, pp. 2693–2703, 2003. [Online]. Available: https://doi.org/10.1090/S0002-9939-03-06860-6

[6] J. A. Brzozowski and R. S. Cohen, "Dot-depth of star-free events," *J. Comp. Sys. Sci.*, vol. 5, no. 1, pp. 1–16, 1971.

[7] J. A. Brzozowski and R. Knast, "The dot-depth hierarchy of star-free languages is infinite," *J. Comp. Sys. Sci.*, vol. 16, no. 1, pp. 37–55, 1978.

[8] W. Bucher, A. Ehrenfeucht, and D. Haussler, "On total regulators generated by derivation relations," *Theoretical Computer Science*, vol. 40, pp. 131–148, 1985.

[9] A. Cano, G. Guaiana, and J. Éric Pin, "Regular languages and partial commutations," *Information and Computation*, vol. 230, pp. 76 – 96, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0890540113000771

[10] L. Chaubard, J. Pin, and H. Straubing, "First order formulas with modular predicates," in *LICS'06*, 2006.

[11] M. Delgado, "Abelian pointlikes of a monoid," *Semigroup Forum*, vol. 56, pp. 339–361, 1998.

[12] S. Eilenberg, *Automata, languages, and machines*. Academic Press, 1976, vol. B.

[13] C. Glaßer and H. Schmitz, "Languages of dot-depth 3/2," in *STACS'00*. Springer, 2000, pp. 555–566.

[14] K. Henckell, S. W. Margolis, J. Pin, and J. Rhodes, "Ash's type II theorem, profinite topology and malcev products: Part I," *IJAC*, vol. 1, no. 4, pp. 411–436, 1991.

[15] R. Knast, "A semigroup characterization of dot-depth one languages," *RAIRO TIA*, vol. 17, no. 4, pp. 321–330, 1983.

[16] M. Kufleitner and T. Walter, "One quantifier alternation in first-order logic with modular predicates," *RAIRO TIA*, vol. 49, no. 1, 2015.

[17] S. W. Margolis and J. Pin, "Products of group languages," in *FCT'85*. Springer, 1985.

[18] R. McNaughton and S. A. Papert, *Counter-Free Automata*. MIT Press, 1971.

[19] J.-É. Pin, "Bridges for concatenation hierarchies," in *ICALP'98*. Springer, 1998, pp. 431–442.

[20] J.-E. Pin, "Algebraic tools for the concatenation product," *Theoretical Computer Science*, vol. 292, pp. 317–342, 2003.

[21] J.-É. Pin, "An explicit formula for the intersection of two polynomials of regular languages," in *DLT 2013*. Springer, 2013, pp. 31–45.

[22] ——, "The dot-depth hierarchy, 45 years later," in *The Role of Theory in Computer Science. Essays Dedicated to Janusz Brzozowski*, 2017.

[23] J.-É. Pin and H. Straubing, "Monoids of upper triangular boolean matrices," in *Semigroups. Structure and Universal Algebraic Problems*. North-Holland, 1985, vol. 39, pp. 259–272.

[24] J.-E. Pin and P. Weil, "Semidirect products of ordered semigroups," *Communications in Algebra*, vol. 30, pp. 149–169, 2002.

[25] T. Place, "Separating regular languages with two quantifiers alternations," in *LICS'15*. IEEE Computer Society, 2015, pp. 202–213.

[26] ——, "Separating regular languages with two quantifier alternations," *Logical Methods in Computer Science*, vol. 14, no. 4, 2018.

[27] T. Place and M. Zeitoun, "Going higher in the first-order quantifier alternation hierarchy on words," in *ICALP'14*, 2014, pp. 342–353.

[28] ——, "Separation and the successor relation," in *STACS'15*. Springer, 2015, pp. 662–675.

[29] ——, "The tale of the quantifier alternation hierarchy of first-order logic over words," *SIGLOG news*, vol. 2, no. 3, pp. 4–17, 2015.

[30] ——, "Separation for dot-depth two," in *32th Annual ACM/IEEE Symposium on Logic in Computer Science*, ser. LICS'17, 2017.

[31] ——, "The covering problem," *Logical Methods in Computer Science*, vol. 14, no. 3, 2018.

[32] ——, "Generic results for concatenation hierarchies," *Theory of Computing Systems (ToCS)*, 2018, selected papers from CSR'17.

[33] ——, "Separation for dot-depth two," 2018, in preparation, preprint at http://www.labri.fr/ zeitoun/research/pdf/boolpol-full.pdf.

[34] M. P. Schützenberger, "On finite monoids having only trivial subgroups," *Information and Control*, vol. 8, no. 2, pp. 190–194, 1965.

[35] B. Steinberg, "Inevitable graphs and profinite topologies: Some solutions to algorithmic problems in monoid and automata theory, stemming from group theory," *International Journal of Algebra and Computation*, vol. 11, no. 01, pp. 25–71, 2001.

[36] H. Straubing, "A generalization of the Schützenberger product of finite monoids," *Theoret. Comp. Sci.*, vol. 13, no. 2, pp. 137–150, 1981.

[37] ——, "Finite semigroup varieties of the form $\mathrm{V} * \mathrm{D}$," *Journal of Pure and Applied Algebra*, 1985.

[38] D. Thérien, "Classification of finite monoids: The language approach," *Theoretical Computer Science*, vol. 14, no. 2, pp. 195–208, 1981.

[39] W. Thomas, "Classifying regular events in symbolic logic," *J. Comp. Sys. Sci.*, vol. 25, no. 3, pp. 360–376, 1982.

[40] G. Zetzsche, "Separability by piecewise testable languages and downward closures beyond subwords," in *LICS'18*, 2018.