



**HAL**  
open science

## Positive First-order Logic on Words

Denis Kuperberg

► **To cite this version:**

Denis Kuperberg. Positive First-order Logic on Words. 2021 36th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS), Jun 2021, Rome, France. pp.1-13, 10.1109/LICS52264.2021.9470602 . hal-03375096

**HAL Id: hal-03375096**

**<https://hal.science/hal-03375096v1>**

Submitted on 12 Oct 2021

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

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

# Positive First-order Logic on Words

Denis Kuperberg  
CNRS, LIP, ENS Lyon  
Email: denis.kuperberg@ens-lyon.fr

**Abstract**—We study  $\text{FO}^+$ , a fragment of first-order logic on finite words, where monadic predicates can only appear positively. We show that there is an  $\text{FO}$ -definable language that is monotone in monadic predicates but not definable in  $\text{FO}^+$ . This provides a simple proof that Lyndon’s preservation theorem fails on finite structures. We additionally show that given a regular language, it is undecidable whether it is definable in  $\text{FO}^+$ .

## I. INTRODUCTION

Preservation theorems in first-order logic ( $\text{FO}$ ) establish a link between semantic and syntactic properties [AG97], [Ros08]. We will be particularly interested here in Lyndon’s theorem [Lyn59], which states that if a first-order formula is monotone in a predicate  $P$  (semantic property), then it is equivalent to a formula that is positive in  $P$  (syntactic property). As for other preservation theorems, this result may not hold when restricting the class of structures considered. Whether Lyndon’s Theorem was true when restricted to finite structures was an open problem for 28 years. It was finally shown to fail on finite structures in [AG87] with a very difficult proof, using a large array of techniques from different fields of mathematics such as probability theory, topology, lattice theory, and analytic number theory. A simpler but still quite intricate proof of this fact was later given by [Sto95], using Ehrenfeucht-Fraïssé games on grid-like structures equipped with two binary predicates.

The goal of this paper is to further restrict the class of structures under consideration, by allowing only finite words. We will therefore work in this paper with the particular signature associated with finite words: one binary predicate (the total order), and a finite set of monadic predicates (encoding the alphabet). We will call  $\text{FO}^+$  the fragment of first-order logic on these models, that is syntactically positive in the monadic predicates. Our purpose is twofold:

- Find out whether Lyndon’s Theorem holds on finite words, and investigate the relation of this framework with the more general case of finite structures.
- From the point of view of language theory: study the natural fragment of  $\text{FO}^+$ -definable languages, in particular given a regular language, can we decide whether it is  $\text{FO}^+$ -definable?

Recall that  $\text{FO}$  on words is a well-studied logic defining a proper fragment of regular languages. This fragment has many equivalent characterizations: definable by star-free expressions, aperiodic monoids, LTL formulas,... [DG08], [Sch65], [Kam68], [MP71].

## Contributions

We define a semantic notion of monotone language on alphabets equipped with a partial order: the language is required to be closed under replacement of a letter by a bigger one. This generalizes the monotonicity condition on monadic predicates in the sense of Lyndon. The negation-free logic  $\text{FO}^+$  can only define monotone languages, and can be seen as a fragment of the standard  $\text{FO}$  logic on words, in the context of ordered alphabets.

Answering our first objective, we show that Lyndon’s Theorem fails on finite words, by building a regular language that is monotone and  $\text{FO}$ -definable, but not  $\text{FO}^+$ -definable. This proof uses a variant of Ehrenfeucht-Fraïssé games that characterizes  $\text{FO}^+$ -definability, introduced in [Sto95], and instantiated here on finite words. As a corollary, using suitable axiomatizations of finite words, we obtain the failure of Lyndon’s theorem on finite structures, in a much simpler way than in [AG87], [Sto95].

Finally, answering our second objective, we show that  $\text{FO}^+$ -definability is undecidable for regular languages. This result is obtained using a reduction from the Turing Machine Mortality problem [Hoo66]. To our knowledge, this is the first example of a natural<sup>1</sup> class of regular languages for which membership is undecidable.

Although we work in a specialized framework requiring a partial order on the alphabet, we believe that this is not an artificial construct, as it occurs naturally in several settings. On one hand, powerset alphabets – i.e. letters are sets of atomic predicates, and are naturally ordered by inclusion – are standard in verification and model theory. The link between powerset alphabets and model theory will be explicated in the paper. On the other hand, ordered letters can be used as an abstraction for factors of the input word that are naturally equipped with a partial order, for instance in quantitative generalizations of regular languages such as regular cost functions [Col12].

Some details and additional remarks can be found in the Appendix.

## Related works

### Monotone complexity

Positive fragments of first-order logic play a prominent role in complexity theory. Indeed, an active research program con-

<sup>1</sup>The concept of “natural” class is of course a bit informal here, but for instance we can think of it as classes inductively defined via a syntax. More generally, any class of regular languages not purposely defined to have an undecidable membership problem could be considered natural in this context.

sists in studying positive fragments of complexity classes. This includes for instance trying to lift equivalent characterizations of a class to their positive versions, or investigating whether a semantic and a syntactic definition of the positive variant of a class are equivalent. See [GS92] for an introduction to monotone complexity, and [LSS96], [Ste94] for examples of characterizations of the positive versions of the classes P and NP, in particular through extensions of first-order logic. The aforementioned paper [AG87], which was the first to show the failure of Lyndon’s theorem on finite structures, does so by reproving in particular an important result on monotone circuit complexity first proved in [FSS81]:  $\text{Monotone-AC}^0 \neq \text{Monotone} \cap \text{AC}^0$ .

### Membership in subclasses of regular languages

Related to our undecidability result, we can mention that there are syntactically defined classes of regular languages for which decidability of membership is an open problem. Such classes, also related to FO fragments, are the ones defined via quantifier-alternation: given a regular language, is it definable with an FO formula having at most  $k$  quantifier alternations? Recent works obtained decidability results for this question, but only for the first 3 levels of the quantifier alternation hierarchy [PZ19]. For higher levels, the problem remains open. Let us also mention the generalized star-height problem [PST89]: can a given regular language be defined in an extended regular expression (with complement allowed) with no nesting of Kleene star? In this case it is not even known whether all regular languages can be defined in this way.

### Quantitative extensions

First-order logic on words has been extended to quantitative settings, which naturally yields a negation-free syntax, because complementation becomes problematic in those settings. This is the case in the theory of regular cost functions [Col12], [KVB12], and in other quantitative extensions concerned with boundedness properties, such as MSO+U [Boj04] or Magnitude MSO [Col13]. We hope that the present work can shed a light on these extensions as well.

### Notations and prerequisites

If  $i, j \in \mathbb{N}$ , we note  $[i, j]$  the set  $\{i, i+1, \dots, j\}$ . If  $X$  is a set, we note  $\mathcal{P}(X)$  its powerset, i.e. the set of subsets of  $X$ . We will note  $A$  a finite alphabet throughout the paper. The set of finite words on  $A$  is  $A^*$ . The length of  $u \in A^*$  is denoted  $|u|$ . If  $L \subseteq A^*$  is a language, we will note  $\bar{L}$  its complement. We will note  $\text{dom}(u) = [0, |u| - 1]$  the set of positions of a word  $u$ . If  $u$  is a word and  $i \in \text{dom}(u)$ , we will note  $u[i]$  the letter at position  $i$ , and  $u[..i]$  the prefix of  $u$  up to position  $i$ . Similarly,  $u[i..j]$  is the infix of  $u$  from position  $i$  to  $j$  and  $u[i..]$  is the suffix of  $u$  starting in position  $i$ .

We will assume that the reader is familiar with the notion of regular languages of finite words, and with some ways to define such languages: finite automata (DFA for deterministic and NFA for non-deterministic), finite monoids, and first-order

logic. See e.g. [DG08] for an introduction of all the needed material.

## II. MONOTONICITY ON WORDS

### A. Ordered alphabet

In this paper we will consider that the finite alphabet  $A$  is equipped with a partial order  $\leq_A$ . This partial order is naturally extended to words componentwise:  $a_1 a_2 \dots a_n \leq_A b_1 b_2 \dots b_m$  if  $n = m$  and for all  $i \in [1, n]$  we have  $a_i \leq_A b_i$ .

A special case that will be of interest here is when the alphabet is built as the powerset of a set  $P$  of *predicates*, i.e.  $A = \mathcal{P}(P)$ , and the order  $\leq_A$  is inclusion. We will call this a *powerset alphabet*.

Taking  $A = \mathcal{P}(P)$  is standard in settings such as verification and model theory, where several predicates can be considered independently of each other in some position.

Powerset alphabets constitute a particular case of ordered alphabets. The results obtained in this paper are valid for both the powerset case and the general case. Due to the nature of the results (existence of a counter-example and undecidability result), it is enough to show them in the particular case of powerset alphabets to cover both cases. Moreover, the powerset alphabet case allows us to directly establish a link with Lyndon’s theorem, which is stated in the framework of model theory. For these reasons, we will keep the more general notion of ordered alphabet for generic definitions, but we will prove our main results on powerset alphabets in order to directly obtain the stronger version of these results.

### B. Monotone languages

We fix  $A$  a finite ordered alphabet.

**Definition 1.** We say that a language  $L \subseteq A^*$  is *monotone* if for all  $u \leq_A v$ , if  $u \in L$  then  $v \in L$ .

**Example 2.** Let  $A = \{a, b\}$  with  $a \leq_A b$ . Then  $A^* b A^*$  is monotone but its complement  $a^*$  is not monotone.

**Definition 3.** Let  $L \subseteq A^*$ , the *monotone closure* of  $L$  is the language  $L^\uparrow = \{v \in A^* \mid \exists u \in L, u \leq_A v\}$ . It is the smallest monotone language containing  $L$ .

In particular, if  $a \in A$ , we will note  $a^\uparrow$  the set  $\{b \in A \mid a \leq_A b\}$ .

**Lemma 4.** Given an NFA  $\mathcal{A}$ , we can compute in time  $O(|\mathcal{A}| \cdot |\mathcal{A}|)$  an NFA  $\mathcal{A}^\uparrow$  for the monotone closure of  $L(\mathcal{A})$ .

*Proof.* We build an NFA  $\mathcal{A}^\uparrow$  from  $\mathcal{A}$ , by replacing every transition  $p \xrightarrow{a} q$  of  $\mathcal{A}$  by  $p \xrightarrow{a^\uparrow} q$ . We use here the standard convention where a transition  $p \xrightarrow{B} q$  with  $B \subseteq A$  stands for a set of transitions  $\{p \xrightarrow{b} q \mid b \in B\}$ . It is straightforward to verify that  $\mathcal{A}^\uparrow$  is an NFA for  $L^\uparrow$ : any run of  $\mathcal{A}^\uparrow$  on some word  $v$  can be mapped to a run of  $\mathcal{A}$  on some  $u \leq_A v$ .  $\square$

**Theorem 5.** Given a regular language  $L \subseteq A^*$ , it is decidable whether  $L$  is monotone. The problem is in P if  $L$  is given by a DFA and PSPACE-complete if  $L$  is given by an NFA.

*Proof.* Notice that if  $\mathcal{A}$  is an NFA,  $L(\mathcal{A})$  is monotone if and only if  $L(\mathcal{A}^\dagger) \subseteq L(\mathcal{A})$ . This shows that the problem is in PSPACE in general, and that it is in P when  $\mathcal{A}$  is a DFA, since it reduces to checking emptiness of the intersection between  $\mathcal{A}^\dagger$  and the complement of  $\mathcal{A}$ . We show that the general problem is PSPACE-hard by reducing from NFA universality. Let  $\mathcal{A}$  be an NFA on an alphabet  $A$ . We build the ordered alphabet  $B = A \cup \{a, b\}$ , where  $a, b \notin A$ , and  $a \leq_B b$  is the only non-trivial inequality in  $B$ . We build an NFA  $\mathcal{B}$  of size polynomial in the size of  $\mathcal{A}$  and recognizing  $aA^* + bL(\mathcal{A})$ , using standard NFA constructions. We have that  $L(\mathcal{A}) = A^*$  if and only if  $L(\mathcal{B})$  is monotone, thereby completing the PSPACE-hardness reduction.  $\square$

### III. POSITIVE FIRST-ORDER LOGIC

#### A. Syntax and semantics

The main idea of positive FO, that we will note  $\text{FO}^+$ , is to guarantee via a syntactic restriction that it only defines monotone languages.

Notice that since monotone languages are not closed under complement (see Example 2), we cannot allow negation in the syntax of  $\text{FO}^+$ . This means we have to add dual versions of classical operators of first-order logic.

This naturally yields the following syntax for  $\text{FO}^+$ :

$$\varphi, \psi := a^\dagger(x) \mid x \leq y \mid x < y \mid \varphi \vee \psi \mid \varphi \wedge \psi \mid \exists x.\varphi \mid \forall x.\varphi$$

As usual, variables  $x, y, \dots$  range over the positions of the input word. The semantics is the same as classical FO on words, with the notable exception that  $a^\dagger(x)$  is true if and only if  $x$  is labelled by some  $b \in a^\dagger$ . Unlike classical FO, it is not possible to require that a position is labelled by a specific letter  $a$ , except when  $a^\dagger = \{a\}$ . This is necessary to guarantee that only monotone languages can be defined.

#### Formal semantics of $\text{FO}^+$

If  $\varphi$  is a formula with free variables  $\text{FV}(\varphi)$ , its semantics is a set  $\llbracket \varphi \rrbracket$  of pairs of the form  $(u, \alpha)$ , where  $u \in A^*$  and  $\alpha : \text{FV}(\varphi) \rightarrow \text{dom}(u)$  a valuation for the free variables. We write indistinctively  $u, \alpha \models \varphi$  or  $(u, \alpha) \in \llbracket \varphi \rrbracket$ , to signify that  $(u, \alpha)$  satisfies  $\varphi$ . If  $\text{FV}(\varphi) = \emptyset$ , we can simply write  $u \models \varphi$  instead of  $(u, \emptyset) \models \varphi$ . In this case, the language recognized by  $\varphi$  is  $\llbracket \varphi \rrbracket = \{u \in A^* \mid u \models \varphi\}$ .

We define  $\llbracket \varphi \rrbracket$  by induction on  $\varphi$ .

- $u, \alpha \models a^\dagger(x)$  if  $a \leq_A u[\alpha(x)]$ .
- $u, \alpha \models x \leq y$  if  $\alpha(x) \leq \alpha(y)$ .
- $u, \alpha \models x < y$  if  $\alpha(x) < \alpha(y)$ .
- $\llbracket \varphi \vee \psi \rrbracket = \llbracket \varphi \rrbracket \cup \llbracket \psi \rrbracket$ .
- $\llbracket \varphi \wedge \psi \rrbracket = \llbracket \varphi \rrbracket \cap \llbracket \psi \rrbracket$ .
- $u, \alpha \models \exists x.\varphi$  if there exists  $i \in \text{dom}(u)$  such that  $(u, \alpha[x \mapsto i]) \in \llbracket \varphi \rrbracket$ .
- $u, \alpha \models \forall x.\varphi$  if for all  $i \in \text{dom}(u)$ , we have  $(u, \alpha[x \mapsto i]) \in \llbracket \varphi \rrbracket$ .

Here the valuation  $\alpha[x \mapsto i]$  maps  $y$  to  $\begin{cases} i & \text{if } y = x \\ \alpha(y) & \text{if } y \neq x \end{cases}$ .

**Example 6.** On alphabet  $A = \{a, b, c\}$  with  $a \leq_A b$ .

- $\forall x.a^\dagger(x)$  recognizes  $\{a, b\}^*$ .
- $\exists x.b^\dagger(x)$  recognizes  $A^*bA^*$ .

**Remark 7.** In the powerset alphabet framework where  $A = \mathcal{P}(P)$ , we can naturally view  $\text{FO}^+$  as the negation-free fragment of first-order logic, by having atomic predicates  $a^\dagger(x)$  range directly over  $P$  instead of  $A = \mathcal{P}(P)$ . We can then drop the  $a^\dagger$  notation, as predicates from  $P$  are considered independently of each other. This way,  $p(x)$  will be true if and only if the letter  $S \in A$  labelling  $x$  contains  $p$ . A letter predicate  $S^\dagger(x)$  in the former syntax can then be expressed by  $\bigwedge_{p \in S} p(x)$ , so  $\text{FO}^+$  based on predicates from  $P$  is indeed equivalent to  $\text{FO}^+$  based on  $A$ . We will take this convention when working on powerset alphabets.

**Example 8.** Let  $A = \mathcal{P}(P)$  with  $P = \{a, b\}$ . The formula  $\exists x, y. x \leq y \wedge a(x) \wedge b(y)$  recognizes  $A^*\{a, b\}A^* + A^*\{a\}A^*\{b\}A^*$ .

#### B. Properties of $\text{FO}^+$

**Lemma 9.** Assume the order on  $A$  is trivial, i.e. no two distinct letters are comparable. Then all languages are monotone, and any FO-definable language is  $\text{FO}^+$ -definable.

*Proof.* The fact that all languages are monotone in this case follows from the fact that for two words  $u, v$  we have  $u \leq_A v$  if and only if  $u = v$ .

If  $L$  is definable by an FO formula  $\varphi$ , we can build an  $\text{FO}^+$  formula  $\psi$  from  $\varphi$  by pushing negations to the leaves using the usual rewritings such as  $\neg(\varphi \wedge \psi) = \neg\varphi \vee \neg\psi$  and  $\neg(\exists x.\varphi) = \forall x.\neg\varphi$ . For all letter  $a \in A$  and variable  $x$ , we then replace all occurrences of  $\neg a(x)$  by  $\bigvee_{b \neq a} b(x)$ . Finally, the negation of  $x \leq y$  (resp.  $x < y$ ) can be written  $y < x$  (resp.  $y \leq x$ ).  $\square$

**Lemma 10.** The logic  $\text{FO}^+$  can only define monotone languages.

*Proof.* By induction on formulas, see Appendix A for details.  $\square$

It is natural to ask whether the converse of Lemma 10 holds: if a language is FO-definable and monotone, then is it necessarily  $\text{FO}^+$ -definable? This will be the purpose of Section IV.

#### C. Ordered Ehrenfeucht-Fraïssé games

We will explain here how  $\text{FO}^+$ -definability can be captured by an ordered variant of Ehrenfeucht-Fraïssé games, that we will call  $\text{EF}^+$ -games.

This notion was defined in [Sto95] for general structures, we will instantiate it here on words.

We define the  $n$ -round  $\text{EF}^+$ -game on two words  $u, v \in A^*$ , noted  $\text{EF}_n^+(u, v)$ . This game is played between two players, Spoiler and Duplicator.

If  $k \in \mathbb{N}$ , a  $k$ -position of the game is of the form  $(u, \alpha, v, \beta)$ , where  $\alpha : [1, k] \rightarrow \text{dom}(u)$  and  $\beta : [1, k] \rightarrow \text{dom}(v)$  are valuations for  $k$  variables in  $u$  and  $v$  respectively. We can

think of  $\alpha$  and  $\beta$  as giving the position of  $k$  previously placed tokens in  $u$  and  $v$ .

A  $k$ -position  $(u, \alpha, v, \beta)$  is *valid* if for all  $i \in [1, k]$ , we have  $u[\alpha(i)] \leq_A v[\beta(i)]$ , and for all  $i, j \in [1, k]$ ,  $\alpha(i) \leq \alpha(j)$  if and only if  $\beta(i) \leq \beta(j)$ .

Notice the difference with usual EF-games: here we do not ask that tokens placed in the same round have same label, but that the label in  $u$  is  $\leq_A$ -smaller than the label in  $v$ . This feature is intended to capture  $\text{FO}^+$  instead of  $\text{FO}$ .

The game starts from the 0-position  $(u, \emptyset, v, \emptyset)$ .

At each round, starting from a  $k$ -position  $(u, \alpha, v, \beta)$ , the game is played as follows. If  $(u, \alpha, v, \beta)$  is not valid, then Spoiler wins. Otherwise, if  $k = n$ , then Duplicator wins. Otherwise, Spoiler chooses a position in one of the two words, and places token number  $k + 1$  on it. Duplicator answers by placing token number  $k + 1$  on a position of the other word. Let us call  $\alpha'$  and  $\beta'$  the extensions of  $\alpha$  and  $\beta$  with these new tokens. If  $(u, \alpha', v, \beta')$  is not a valid  $(k + 1)$ -position, then Spoiler immediately wins the game, otherwise, the game moves to the next round with  $(k + 1)$ -position  $(u, \alpha', v, \beta')$ .

We will note  $u \preceq_n v$  when Duplicator has a winning strategy in  $\text{EF}_n^+(u, v)$ .

**Definition 11.** The quantifier rank of a formula  $\varphi$ , noted  $\text{qr}(\varphi)$  is its number of nested quantifiers. It can be defined by induction in the following way: if  $\varphi$  is atomic then  $\text{qr}(\varphi) = 0$ , otherwise,  $\text{qr}(\varphi \wedge \psi) = \text{qr}(\varphi \vee \psi) = \max(\text{qr}(\varphi), \text{qr}(\psi))$  and  $\text{qr}(\exists x.\varphi) = \text{qr}(\forall x.\varphi) = \text{qr}(\varphi) + 1$ .

The following Theorem shows the link between the  $n$ -round  $\text{EF}^+$  game and formulas of rank at most  $n$ .

**Theorem 12** ([Sto95, Thm 2.4]). *We have  $u \preceq_n v$  if and only if for all formulas  $\varphi$  of  $\text{FO}^+$  with  $\text{qr}(\varphi) \leq n$ , we have  $(u \models \varphi) \Rightarrow (v \models \varphi)$ .*

Since the proof of Theorem 12 does not appear in [Sto95], it can be found in Appendix B for completeness.

Let us now see how we can use  $\text{EF}^+$  games to characterize  $\text{FO}^+$ -definability.

**Corollary 13.** *A language  $L$  is not  $\text{FO}^+$ -definable if and only if for all  $n \in \mathbb{N}$ , there exists  $(u, v) \in L \times \overline{L}$  such that  $u \preceq_n v$ .*

*Proof.*  $\Leftarrow$ : Let  $n \in \mathbb{N}$ , there exists  $(u, v) \in L \times \overline{L}$  such that  $u \preceq_n v$ . By Theorem 12, any formula of quantifier rank  $n$  accepting  $u$  must accept  $v$ , so no formula of quantifier rank  $n$  recognizes  $L$ . This is true for all  $n \in \mathbb{N}$ , so  $L$  is not  $\text{FO}^+$ -definable.

$\Rightarrow$  (contrapositive): Assume there exists  $n \in \mathbb{N}$  such that for all  $(u, v) \in L \times \overline{L}$ ,  $u \not\preceq_n v$ . By Theorem 12, this means that for all  $(u, v) \in L \times \overline{L}$ , there exists a formula  $\varphi_{u,v}$  of quantifier rank  $n$  accepting  $u$  but not  $v$ . Since there are finitely many  $\text{FO}^+$  formulas of rank  $n$  up to logical equivalence [Lib04, Lem 3.13], the set of formulas  $F = \{\varphi_{u,v} \mid (u, v) \in L \times \overline{L}\}$  can be chosen finite. We define  $\psi = \bigvee_{u \in L} \bigwedge_{v \notin L} \varphi_{u,v}$ , where the conjunctions and disjunction are finite since  $F$  is finite. For all  $u \in L$ ,  $u \models \bigwedge_{v \notin L} \varphi_{u,v}$  hence  $u \models \psi$ , and conversely,

a word satisfying  $\psi$  must satisfy some  $\bigwedge_{v \notin L} \varphi_{u,v}$ , so it cannot be in  $\overline{L}$ .  $\square$

#### IV. SYNTAX VERSUS SEMANTICS

We will now answer the natural question posed in Section III-B: is any  $\text{FO}$ -definable monotone language  $\text{FO}^+$ -definable?

##### A. A counter-example language

This section is dedicated to the proof of the following Theorem:

**Theorem 14.** *There is an  $\text{FO}$ -definable monotone language  $K$  on a powerset alphabet that is not  $\text{FO}^+$ -definable.*

Let  $P = \{a, b, c\}$  and  $A = \mathcal{P}(P)$ , ordered by inclusion.

We will note  $\binom{a}{b}$ ,  $\binom{b}{c}$ ,  $\binom{c}{a}$  for the letters  $\{a, b\}$ ,  $\{b, c\}$ ,  $\{a, c\}$  respectively, and  $\top$  for  $\{a, b, c\}$ . If  $x \in P$  we will often note  $x$  instead of  $\{x\}$  to lighten notations.

We now define the desired language by:

$$K := (a^\uparrow b^\uparrow c^\uparrow)^* + A^* \top A^*.$$

We claim that  $K$  satisfies the requirements of Theorem 14.

Notice that the second disjunct  $A^* \top A^*$  could be omitted if we were to consider only the alphabet  $A \setminus \{\top\}$ . When sticking with a powerset alphabet, this disjunct is necessary to obtain an  $\text{FO}$ -definable language. Indeed, if we just define  $K_0 = (a^\uparrow b^\uparrow c^\uparrow)^*$  on alphabet  $A$ , we have  $K_0 \cap (\top^*) = (\top \top \top)^*$ . Since  $\top^*$  is  $\text{FO}$ -definable but  $(\top \top \top)^*$  is not (see [DG08]), and  $\text{FO}$ -definable languages are closed under intersection, we have that  $K_0$  is not  $\text{FO}$ -definable.

**Lemma 15.**  *$K$  is monotone and  $\text{FO}$ -definable.*

*Proof.* The fact that  $K$  is monotone is straightforward from its definition, as the union of two monotone languages.

To show that  $K$  is  $\text{FO}$ -definable, we can use the classical characterizations of first-order definable languages [DG08]. Here we will verify that its minimal automaton  $\mathcal{A}$  is counter-free, that is, no word induces a non-trivial cycle in  $\mathcal{A}$ .

The minimal DFA  $\mathcal{A}$  recognizing  $K$  is depicted in Figure 1. We note  $\neg a = \{\emptyset, \{b\}, \{c\}, \{b, c\}\}$  the sub-alphabet of  $A$  of letters not containing  $a$ , similarly for  $\neg b$  and  $\neg c$ . The edges going to rejecting state  $\perp$  are grayed and dashed, and the ones going to accepting sink state  $q_\top$  are grayed, for readability. We also note  $a' = a^\uparrow \setminus \{\top\} = \{\{a\}, \binom{a}{b}, \binom{a}{c}\}$ , and similarly for  $b', c'$ .

To show that  $K$  is  $\text{FO}$ -definable, it suffices to show that  $\mathcal{A}$  is counter-free, i.e. that there is no word  $u \in A^*$ , distinct states  $p, q$  of  $\mathcal{A}$ , and integer  $k$ , such that  $p \xrightarrow{u} q$  and  $q \xrightarrow{u^k} p$ . Assume for contradiction that such  $u, p, q, k$  exist. Since the only non-trivial strongly connected component in  $\mathcal{A}$  is  $\{q_a, q_b, q_c\}$ , these states are the only candidates for  $p, q$ . Since  $p, q$  are distinct, it means  $|u|$  is not a multiple of 3, and  $u$  induces a 3-cycle, either  $q_a \xrightarrow{u} q_b \xrightarrow{u} q_c \xrightarrow{u} q_a$  if  $|u| \equiv 1 \pmod{3}$  or in the reverse order if  $|u| \equiv 2 \pmod{3}$ . Thus, the first letter of  $u$  can be read from all states from  $\{q_a, q_b, q_c\}$ , while staying in this component.

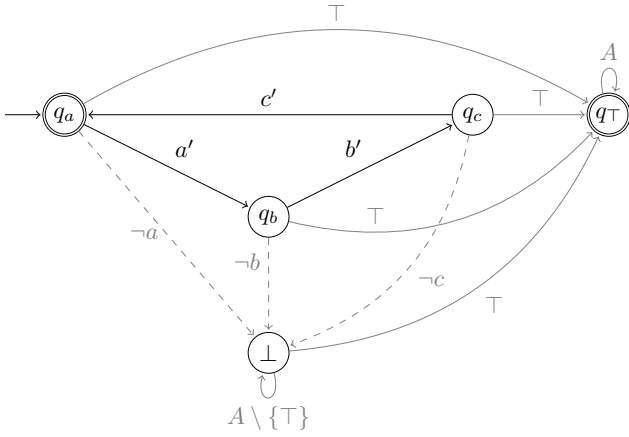


Fig. 1. The minimal DFA  $\mathcal{A}$  of  $K$

Such a letter does not exist, so we reach a contradiction. The DFA  $\mathcal{A}$  is counter-free, so  $K$  is FO-definable [DG08].  $\square$

**Remark 16.** In order to get a better intuition on the language  $K$ , we can also build its 21-element syntactic monoid and verify its aperiodicity. This is done in Appendix C. In addition, it is useful to understand how an FO formula can describe the language  $K$ , as we will later build on this understanding in Section V. We describe the behaviour of such a formula in Appendix D.

**Lemma 17.**  $K$  is not  $\text{FO}^+$ -definable.

*Proof.* We establish this using Corollary 13. Let  $n \in \mathbb{N}$ , and  $N = 2^n$ . We define  $u = (abc)^N$  and  $v = \left[ \begin{pmatrix} a \\ b \\ c \end{pmatrix} \begin{pmatrix} b \\ c \\ a \end{pmatrix} \right]^{N-1} \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ . Notice that  $u \in K$ , and  $v \notin K$  because  $|v| \equiv 2 \pmod 3$ , and  $v$  does not contain  $\top$ . By Corollary 13, it suffices to prove that  $u \preceq_n v$  to conclude. We give a strategy for Duplicator in  $\text{EF}_n^+(u, v)$ . The strategy is an adaptation from the classical strategy showing that  $(aa)^*$  is not FO-definable [Lib04]. We consider that at the beginning, tokens *first*, *last* are placed on the first and last positions on  $u$ , and *first'*, *last'* on the first and last position of  $v$ . The strategy of Duplicator during the game is then as follows: every time Spoiler places a token in one of the words, Duplicator answers in the other by replicating the closest distance (and direction) to an existing token. This strategy is illustrated in Figure 2, where move  $i$  of Spoiler (resp. Duplicator) is represented by  $\boxed{i}$  (resp.  $\circledast i$ ).

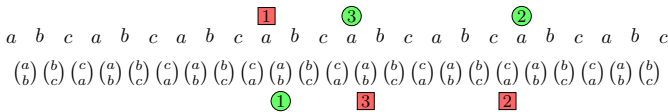


Fig. 2. An example of Duplicator's strategy for  $n = 3$ .

We have to show that this strategy of Duplicator allows him to play  $n$  rounds without losing the game. This proof is

similar to the classical one for  $(aa)^*$ , see e.g. [Lib04]. The main intuition is that the length of the non-matching intervals between  $u$  and  $v$  is at worst divided by 2 at each round, and it starts with a length of  $2^n$ , so Duplicator can survive  $n$  rounds. A detailed proof can be found in Appendix E.  $\square$

### B. Lyndon's Theorem

In this section we consider first-order logic on arbitrary signatures and unconstrained structures. Our goal is to see how Theorem 14 can be lifted to this general framework.

**Definition 18.** A formula  $\varphi$  is *monotone* in a predicate  $P$  if whenever a structure  $S$  is a model of  $\varphi$ , any structure  $S'$  obtained from  $S$  by adding tuples to  $P$  is also a model of  $\varphi$ .

**Example 19.** On graphs, where the only predicate is the edge predicate, the formula asking for the existence of a triangle is monotone, but the formula stating that the graph is not a clique is not monotone.

**Definition 20.** A formula  $\varphi$  is *positive* in  $P$  if it never uses  $P$  under a negation.

Let us recall the statement of Lyndon's Theorem, which holds on general (possibly infinite) structures:

**Theorem 21** ([Lyn59, Cor 2.1]). *If  $\psi$  is an FO formula monotone in predicates  $P_1, \dots, P_n$ , then it is equivalent to a formula positive in predicates  $P_1, \dots, P_n$ .*

We will now see explicitly how the language  $K$  from Section IV-A can be used to show that Lyndon's Theorem fails on finite structures.

The failure of this theorem on finite structures was first shown in [AG87] with a very difficult proof, then reproved in [Sto95] with a simpler one, using the Ehrenfeucht-Fraïssé technique. Still, the proof from [Sto95] is quite involved compared to the one we present here.

Since Lyndon's theorem can be found in the literature under different formulations, and since it is not clear at first sight that they are equivalent, we make it clear here how the construction of this paper applies to all of them. This also serves the purpose of making explicit the exact signature needed in each formulation to show the failure on finite structures with our method.

### Several monotone predicates

This is the formulation of Theorem 21. Let us show that our language  $K$  shows its failure on finite structures.

We will use here the fact that if  $P = \{a, b, c\}$  is a set of monadic predicates, then a finite model over the signature  $(\leq, a, b, c)$  where the order  $\leq$  is total is simply a finite word on the powerset alphabet  $A = \mathcal{P}(P)$ . Therefore, in order to view our words as general finite structures, it suffices to axiomatize the fact that  $\leq$  a total order. This can be done with a formula  $\psi_{tot} = (\forall x, y. x \leq y \vee y \leq x) \wedge (\forall x, y, z. x \leq y \wedge y \leq z \Rightarrow x \leq z) \wedge (\forall x, y. x \leq y \wedge y \leq x \Rightarrow x = y) \wedge (\forall x. x \leq x)$ . Notice that  $\psi_{tot}$  is not monotone in the predicate  $\leq$ .

Let  $\varphi$  be the FO-formula defining  $K$ , obtained in Lemma 15, and let  $\psi = \varphi \wedge \psi_{tot}$ . Then,  $\psi$  is monotone in predicates  $a, b, c$ , and finite structures on signature  $(\leq, a, b, c)$  satisfying  $\psi$  are exactly words of  $K$ . However, as we proved in Theorem 14, no first-order formula that is positive in predicates  $a, b, c$  can define the same class of structures, since the same formula interpreted on words would be an  $\text{FO}^+$ -formula for  $K$ .

### Single monotone predicate

Other formulations of Lyndon's Theorem use a single monotone predicate, as in [Sto95]. We can encode the language  $K$  in this framework, by using one binary predicate  $A$  to represent all letter predicates. Let  $K_3$  be  $K$  restricted to words of length at least 3. By Theorem 14 it is clear that  $K_3$  is FO-definable but not  $\text{FO}^+$ -definable.

Let  $\psi_3$  be an FO-formula stating that there are at least 3 elements 0, 1, 2, and that for all  $y \notin \{0, 1, 2\}$  and for all  $x$ ,  $A(x, y)$  holds. We build the FO formula  $\varphi'$  from the FO formula  $\varphi$  recognizing the language  $K_3$  by replacing every occurrence of  $a(x)$  (resp.  $b(x), c(x)$ ) by  $A(x, 0)$  (resp.  $A(x, 1), A(x, 2)$ ).

Finally, we define the FO formula  $\psi' = \psi_{tot} \wedge \psi_3 \wedge \varphi'$ . Finite structures on signature  $(\leq, A)$  accepted by  $\psi'$  are exactly those who encode words of  $K_3$ . No formula positive in  $A$  can recognize this class of structures, otherwise we could obtain from it an  $\text{FO}^+$ -formula for  $K_3$ , by replacing every occurrence of  $A(x, y)$  by  $(a(x) \wedge y = 0) \vee (b(x) \wedge y = 1) \vee (c(x) \wedge y = 2) \vee y \geq 3$ .

### Closure under surjective homomorphisms

Lyndon's theorem is also often stated in the following way: if an FO formula defines a class of structures closed under surjective homomorphisms, then it is equivalent to a positive formula. This formulation is equivalent to saying that the formula is monotone in all predicates. We can deal with this framework as well, by incorporating a predicate  $\not\leq$  to the signature. Let  $\psi_{\not\leq}$  be the formula obtained from  $\psi$  by pushing negations to the leaves and replacing all subformulas of the shape  $\neg(x \leq y)$  with  $x \not\leq y$ . Let

$$\psi'' = (\exists x, y. x \leq y \wedge x \not\leq y) \vee (\psi_{\not\leq} \wedge \forall x, y. (x \leq y \vee x \not\leq y)).$$

Finite structures on the signature  $(\leq, \not\leq, a, b, c)$  can be classified into three categories:

- if there are  $x, y$  such that  $x \leq y \wedge x \not\leq y$ , then the structure satisfies  $\psi''$  (because of the first disjunct)
- otherwise, if there are  $x, y$  such that  $\neg(x \leq y \vee x \not\leq y)$ , then the structure does not satisfy  $\psi''$  (because of the last conjunct)
- otherwise,  $\not\leq$  is the complement of  $\leq$ , and the structure satisfies  $\psi''$  if and only if it satisfies  $\psi_{\not\leq}$ .

Therefore, in  $\psi_{\not\leq}$  we can use  $\leq$  and  $\not\leq$  freely, assuming that  $\not\leq$  is actually the complement of  $\leq$ . In particular the  $\psi_{tot}$  subformula of  $\psi_{\not\leq}$  axiomatizes the fact that  $\leq$  is a total order, provided that  $\not\leq$  is its complement. So the structures of the third item are exactly the words of  $K$ , with an additional

predicate  $\not\leq$  which is the complement of  $\leq$ . The two first items guarantee that the class of finite structures accepted by  $\psi''$  is monotone with respect to  $\leq$  and  $\not\leq$  as well. As before, it is impossible to have a formula positive in all predicates accepting the same class of finite structures as  $\psi''$ , since replacing  $x \not\leq y$  with  $y < x$  in this formula would directly yield an  $\text{FO}^+$ -formula for  $K$ .

## V. UNDECIDABILITY OF $\text{FO}^+$ -DEFINABILITY

This section is dedicated to the proof of the following Theorem:

**Theorem 22.** *The following problem is undecidable: given  $L$  a regular language on a powerset alphabet, is  $L$   $\text{FO}^+$ -definable?*

We will start with an informal proof sketch to convey the main ideas of the proof, before going to the technical details.

### A. Proof sketch

The proof proceeds by reduction from the Turing Machine Mortality problem, known to be undecidable [Hoo66]. A deterministic Turing Machine (TM) is *mortal* if there is a uniform bound  $n \in \mathbb{N}$  on the length of its runs, starting from any arbitrary configuration.

Given a machine  $M$ , we want to build a regular language  $L$  such that  $L$  is  $\text{FO}^+$ -definable if and only if  $M$  is mortal.

### Configuration words

The intuitive idea is that  $L$  will mimic the language  $(a^\dagger b^\dagger c^\dagger)^*$  from earlier, but the letters will be replaced by words encoding configurations of  $M$ . We therefore design an ordered alphabet  $A$  and a language  $C$  of configurations words such that words from  $C$  encode configurations of  $M$ . These words will be of three possible types 1, 2, 3, playing the role of the letters  $a, b, c$  of the language  $K$ . This partitions  $C$  into  $C_1 \cup C_2 \cup C_3$ . We guarantee that the transitions of  $M$  will always change the types in the following way:  $1 \rightarrow 2$ ,  $2 \rightarrow 3$  or  $3 \rightarrow 1$ .

Moreover, we design the order  $\leq_A$  of the alphabet  $A$  so that given two words  $u_1, u_2$  from  $C$ , there is a word  $v$  that is bigger (for the order  $\leq_A$ ) than both  $u_1$  and  $u_2$  if and only if  $u_1$  and  $u_2$  are consecutive configurations of  $M$ . Such a word  $v$  will be written  $\binom{u_1}{u_2}$  in this proof sketch.

### Language $L$

Finally, the language  $L$  will be roughly defined as the upward-closure of  $(C_1 \# C_2 \# C_3 \#)^*$ , where  $\#$  is a separator symbol. The only requirement for configuration words appearing in a word of  $L$  is on their types. Apart from this, the configuration words can be arbitrary, they do not have to form a run of  $M$ .

We will then use the  $\text{EF}^+$ -game technique to show that  $L$  is  $\text{FO}^+$ -definable if and only if  $M$  is mortal.

### If $M$ not mortal

The easier direction is proving that if  $M$  is not mortal, then  $L$  is not  $\text{FO}^+$ -definable. Indeed, if  $M$  is not mortal, we can choose an arbitrarily long run  $u = u_1 \# u_2 \# \dots \# u_N$  of  $M$ .

We build the word  $v = \binom{u_1}{u_2} \# \binom{u_2}{u_3} \# \dots \# \binom{u_{N-1}}{u_N}$ , and we verify that  $u \in L$  and  $v \notin L$ . Then, using the same technique as in the proof of Lemma 17, with  $C_1, C_2, C_3$  playing the role of  $a, b, c$ , we show that Duplicator wins the  $\text{EF}^+$  game on  $u, v$  with  $\log(N)$  rounds. Therefore  $L$  is not  $\text{FO}^+$ -definable, by Corollary 13.

### If $M$ mortal

The converse direction is more difficult: we have to show that if there is a bound  $n$  on the length of runs of  $M$  from any configuration, then  $L$  is  $\text{FO}^+$ -definable. We will again use the  $\text{EF}^+$ -game and Corollary 13: we give an integer  $m$  (depending only on  $n$ ) such that for any  $u \in L$  and  $v \notin L$ , Spoiler wins  $\text{EF}_m^+(u, v)$ .

Without loss of generality, consider  $u = u_0 \# u_1 \# \dots \# u_N$  a word of  $L$ , where each  $u_i$  is in  $C$ , and  $v$  a word not in  $L$ . To describe the winning strategy of Spoiler, we will first rule out the problems in “local behaviours”: if  $v$  contains a factor containing at most 2 symbols  $\#$  preventing it from belonging to  $L$ , then Spoiler wins easily in a bounded number of rounds by pointing this local inconsistency in  $v$ , that cannot be mirrored in  $u$ . The only remaining problem is the “long-term inconsistency” occurring in the previous  $\text{EF}^+$  games: a long factor with two conflicting possible interpretations, each being forced by one of the endpoints. For instance, when dealing with the language  $K$ , such long-term inconsistencies were exhibited by words of the form  $\binom{a}{b} \binom{b}{c} \binom{c}{a} \binom{a}{b} \dots \binom{b}{c}$ , with the first letter being constrained to  $a$  and the last one to  $c$ . We have to show that contrarily to what happens with the language  $K$ , or in the case where  $M$  is not mortal, Spoiler can now point out such long-term inconsistencies in a bounded number of rounds.

To do that, let us abstract a configuration word  $w \in C$  by its *height*  $h(w)$ : the length of the run of  $M$  starting in the configuration  $w$ . Our mortality hypothesis can be rewritten as: the height of any configuration word is at most  $n$ . A word  $w \in C$  will be abstracted by a single letter  $h(w) \in [0, n]$ . We saw that if a word  $w'$  is of the form  $\binom{w_1}{w_2}$ , then  $w_1$  and  $w_2$  encode consecutive configurations of  $M$ , so their heights must be consecutive integers  $i+1$  and  $i$ . We will abstract such a word  $w'$  by the letter  $\binom{i+1}{i}$ . This allows us to design an abstracted version of the  $\text{EF}$ -game, called the *integer game*, where letters are integers or pairs of integers, and with special rules designed to reflect the constraints of the original  $\text{EF}^+$  game on  $(u, v)$ . The integer game makes explicit the core combinatorial argument making use of the mortality hypothesis. We show that Spoiler wins this integer game in  $2n$  rounds. We finally conclude by lifting this strategy to the original  $\text{EF}^+$ -game.

This ends the proof sketch, and we now go to the more detailed proof.

### B. The Turing Machine Mortality problem

We will start by describing the problem we will reduce from, called Turing Machine (TM) Mortality.

The TM Mortality problem asks, given a deterministic TM  $M$ , whether there exists a bound  $n \in \mathbb{N}$  such that from any finite configuration (state of the machine, position on the tape, and content of the tape), the machine halts in at most  $n$  steps. We say that  $M$  is *mortal* if such an  $n$  exists.

**Theorem 23** ([Hoo66]). *The TM Mortality problem is undecidable.*

**Remark 24.** The standard mortality problem as formulated in [Hoo66] does not ask for a uniform bound on the halting time, and allows for infinite configurations, but it is well-known that the two formulations are equivalent using a compactness argument. Indeed, if for all  $n \in \mathbb{N}$ , the TM has a run of length at least  $n$  from some configuration  $C_n$ , then we can find a configuration  $C$  that is a limit of a subsequence of  $(C_n)_{n \in \mathbb{N}}$ , so that  $M$  has an infinite run from  $C$ .

Notice that the initial and final states of  $M$  play no role here, so we will omit them in the description of  $M$ . Indeed, we can assume that  $M$  halts whenever there is no transition from the current configuration.

Let  $M = (\Gamma, Q, \Delta)$  be a deterministic TM, where  $\Gamma$  is the alphabet of  $M$ ,  $Q$  its set of states, and  $\Delta \subseteq Q \times \Gamma \times Q \times \Gamma \times \{\leftarrow, \rightarrow\}$  its (deterministic) transition table.

We will also assume without loss of generality that  $Q$  is partitioned into  $Q_1, Q_2, Q_3$ , and that all possible successors of a state in  $Q_1$  (resp.  $Q_2, Q_3$ ) are in  $Q_2$  (resp.  $Q_3, Q_1$ ). Remark that if  $M$  is not of this shape, it suffices to make three copies  $Q_1, Q_2, Q_3$  of its state space, and have each transition change copy according to the 1-2-3 order given above. This transformation does not change the mortality of  $M$ .

We will say that  $p$  has *type*  $i$  if  $p \in Q_i$ . The *successor type* of 1 (resp. 2, 3) is 2 (resp. 3, 1).

Our goal is now to start from an instance  $M$  of TM Mortality, and define a regular language  $L$  such that  $L$  is  $\text{FO}^+$ -definable if and only if  $M$  is mortal.

### C. The base language $L_{base}$

#### The base alphabet

We define first a *base alphabet*  $A_{base}$ . Words over this alphabet will be used to encode configurations of the TM  $M$ .

$$A_{base} = \Gamma \cup (\Delta \times \Gamma) \cup (\Gamma \times \Delta) \cup (\Delta \times \Gamma \times \Delta) \cup (Q \times \Gamma) \cup \{\#\}.$$

We will note  $a_\delta$  (resp.  $a^{\delta'}$ ,  $a_\delta^{\delta'}$ ) the letters from  $\Delta \times \Gamma$  (resp.  $\Gamma \times \Delta, \Delta \times \Gamma \times \Delta$ ), and  $[q.a]$  letters of  $Q \times \Gamma$ .

The letter  $[q.a]$  is used to encode the position of the reading head,  $q \in Q$  being the current state of the machine, and  $a \in \Gamma$  the letter it is reading.

A letter  $a_\delta$  will be used to encode a position of the tape that the reading head just left, via a transition  $\delta$  writing an  $a$  on this position. A letter  $a^{\delta'}$  will be used for a position of the tape containing  $a$ , and that the reading head is about to enter via a transition  $\delta'$ . We use  $a_\delta^{\delta'}$  if both are simultaneously true, i.e. the reading head is coming back to the position it just visited. Finally, the letter  $\#$  is used as separator between different configurations.



## Configuration words

The encoding of a configuration of  $M$  is therefore a word of the form (for example):

$$a_1 a_2 \dots (a_{i-1})^{\delta'} [q.a_i] (a_{i+1})_{\delta} \dots a_n.$$

The letter  $(a_{i+1})_{\delta}$  indicates that the reading head came from the right via a transition  $\delta = (\_, \_, q, a_{i+1}, \leftarrow)$  (where  $\_$  is a placeholder for an unknown element). The letter  $(a_{i-1})^{\delta'}$  indicates that it will go in the next step to the left via a transition  $\delta' = (q, a_i, \_, \leftarrow)$ .

A word  $u \in (A_{base})^*$  is a *configuration word* if it encodes a configuration of  $M$  with no incoherences. More formally,  $u$  is a configuration word if  $u$  contains no  $\#$ , exactly one letter from  $Q \times \Gamma$  (the reading head), and either one  $a_{\delta}$  and one  $b^{\delta'}$  located on each side of the head, or just one letter  $a_{\delta}^{\delta'}$  adjacent to the head. Moreover, the labels  $\delta$  and  $\delta'$  both have to be coherent with the current content of the tape.

**Remark 25.** Because we ask these  $\delta$  and  $\delta'$  labellings to be present, configuration words only encode TM configurations that have a predecessor and a successor configuration.

The *type* of a configuration word is simply the type in  $\{1, 2, 3\}$  of the unique state it contains.

Let us call  $C \subseteq (A_{base})^*$  the language of configuration words. This language  $C$  is partitioned into  $C_1, C_2, C_3$  according to the type of the configuration word. It is straightforward to verify that each  $C_i$  is an FO-definable language.

We can now define the language  $L_{base}$ . The basic idea is that we want  $L_{base}$  to be  $(C_1 \# C_2 \# C_3 \#)^*$ , but in order to avoid unnecessary bookkeeping later in the proof, we do not want to care about the endpoints being  $C_1$  and  $C_3$ . Let us also drop the last  $\#$  which is useless as a separator, and assume that  $C_1$  appears at least once, just for the sake of simplifying the final expression. This gives for  $L_{base}$  the expression:

$$(\varepsilon + C_3 \# + C_2 \# + C_3 \#)(C_1 \# C_2 \# C_3 \#)^*(C_1 + C_1 \# C_2 + C_1 \# C_2 \# C_3).$$

Notice that  $L_{base}$  cannot verify that the sequence is an actual run of  $M$ , since it just controls that the immediate neighbourhood of the reading head is valid, and that the types succeed each other according to the 1-2-3 cycle. The tape can be arbitrarily changed from one configuration word to the next.

### D. The alphabet $A$

We now define another alphabet  $A_{amb}$  (*amb* for ambiguous), consisting of some unordered pairs of letters from  $A_{base}$ . An unordered pair  $\{a, b\}$  is in  $A_{amb}$  if  $a$  can be replaced by  $b$  in the encodings of two successive configurations of  $M$  of the same length. Thus, let  $A_{amb}$  be the following set of unordered pairs (we note the “predecessor” element first to facilitate the reading):

- $\{a_{\delta}, a\}$ ,  $a \in \Gamma$ ,  $\delta \in \Delta$
- $\{a, a^{\delta'}\}$ ,  $a \in \Gamma$ ,  $\delta' \in \Delta$
- $\{a^{\delta'}, [q.a]\}$ ,  $\delta' = (\_, \_, q, \_, \_) \in \Delta$
- $\{a_{\delta}^{\delta'}, [q.a]\}$ ,  $\delta = (\_, \_, p, a, d) \in \Delta$ ,  $\delta' = (p, \_, q, \_, -d) \in \Delta$
- $\{[p.a], b_{\delta}\}$ ,  $\delta = (p, a, \_, b, \_) \in \Delta$
- $\{[p.a], b_{\delta}^{\delta'}\}$ ,  $\delta = (p, a, q, b, d) \in \Delta$ ,  $\delta' = (q, \_, \_, \_, -d) \in \Delta$

where  $\_$  stands for an arbitrary element, and  $-d$  is the direction opposite to  $d$ .

Notice that all letters of  $A_{amb}$  have a clear “predecessor” element: even the possible ambiguity regarding letters  $a_{\delta}^{\delta'}$  are resolved thanks to the type constraint on transitions of  $M$ . For readability, we will use the notation  $\binom{a}{b}$  instead of  $\{a, b\}$ , where the upper letter is the predecessor element.

We can now define the alphabet  $A = A_{base} \cup A_{amb}$ , partially ordered by  $a <_A b$  if  $a \in A_{base}$ ,  $b \in A_{amb}$ ,  $a \in b$ . For now we use the general formalism of ordered alphabet for simplicity. We will later describe in Remark 45 how the construction is easily modified to fit in the powerset alphabet framework.

### E. Superposing configuration words

We will see that thanks to the definition of the alphabet  $A_{amb}$ , two distinct configurations can be “superposed”, i.e. can be written simultaneously with letters of  $A$  including letters of  $A_{amb}$ , if and only if one follows from the other by a valid transition of  $M$ .

**Lemma 26.** *If  $u_1, u_2 \in C$  encode two successive configurations of the same length, then there exists  $v \in A^*$  such that  $u_1 \leq_A v$  and  $u_2 \leq_A v$ .*

*Proof.* It suffices to take the letters in  $v$  to be the union of letters in  $u_1, u_2$  when these letters differ. For instance if  $u_1 = aab^{\delta'}[p.a]c_{\delta}c$  and  $u_2 = aa^{\delta''}[q.b]d_{\delta'}cc$ , then  $v = a \binom{a}{a^{\delta''}} \binom{b^{\delta'}}{[q.b]} \binom{[p.a]}{d_{\delta'}} \binom{c_{\delta}}{c}$ .  $\square$

**Lemma 27.** *Let  $u_1, u_2 \in C$ , and assume that there exists  $v \in A^*$  satisfying  $u_1 \leq_A v$  and  $u_2 \leq_A v$ . Then either  $u_1 = u_2$ , or one is the successor configuration of the other.*

*Proof.* Let  $[p.a]$  and  $[q.b]$  be the reading heads in  $u_1$  and  $u_2$ , in positions  $i$  and  $j$  respectively.

If  $i = j$ , let us consider the letter  $v[i]$ , we know that  $[p.a] \leq_A v[i]$  and  $[q.b] \leq_A v[i]$ . Since no letter of the form  $\{[p.a], [q.b]\}$  exists in  $A_{amb}$ , we must have  $[p.a] = [q.b]$ . Let  $\delta' = (p, a, p', a', d)$  be the transition of  $M$  from  $[p.a]$ . Let  $\lambda = -1$  if  $d = \leftarrow$  and  $\lambda = 1$  if  $d = \rightarrow$ . By definition of  $C$ , we must have letters  $b_1, b_2$  such that  $u_1[i + \lambda] = b_1^{\delta'}$  and  $u_2[i + \lambda] = b_2^{\delta'}$ , with possible additional  $\delta$  subscripts. As before, there is no letter  $\{b_1^{\delta'}, b_2^{\delta'}\}$  in  $A_{amb}$ , even with optional additional  $\delta$  subscripts, so  $u_1[i + \lambda] = u_2[i + \lambda]$ . Finally, either both  $u_1[i - \lambda]$  and  $u_2[i - \lambda]$  are letters from  $\Gamma$  (if the  $\delta$  subscript is present  $u_1[i + \lambda] = u_2[i + \lambda]$ ), or are letters with a  $\delta$  subscript. In both cases, again by definition of  $A_{amb}$ , we must have  $u_1[i - \lambda] = u_2[i - \lambda]$ . The rest of the words  $u_1, u_2$  outside of these three positions  $\{i_1, i, i + 1\}$  are forced to be letters of  $\Gamma$  by definition of  $C$ , so again they cannot vary between  $u_1$  and  $u_2$ , since  $A_{amb}$  does not contain letters  $\{a, b\}$  with  $a, b \in \Gamma$ . We can conclude that  $u_1 = u_2$ .

Consider now the case where  $i \neq j$ . Assume without loss of generality that  $p$  is of type 1 (no type plays a particular role). Let  $\delta_1$  be the transition from  $[p.a]$ , of type  $1 \rightarrow 2$ . Let us call *enriched letter* a letter of the form  $a_{\delta}$ ,  $a^{\delta'}$ , or  $a_{\delta}^{\delta'}$ . By definition of  $A_{amb}$ , both  $u_2[i]$  and  $u_1[j]$  must be enriched letters. By definition of  $C$ , it means  $u_1[j]$  is just next to  $u_1[i]$  where the reading head is, say without loss of generality  $j = i + 1$ .

So we have  $v[i] = \{[p.a], u_2[i]\} \in A_{amb}$ , and as we saw,  $u_2[i]$  is either predecessor or successor to  $[p.a]$  in this case. Let us assume for now that  $u_2[i]$  is successor to  $[p.a]$ . This means it has a  $\delta_1$  subscript. Since  $u_2 \in C$ , this forces the state  $q$  to be of type 2, the target type of  $\delta_1$ . Consider now the enriched letter  $u_1[j]$ , which is such that  $\{[q.b], u_1[j]\} \in A_{amb}$ . Either  $u_1[j]$  has a superscript  $\delta'$  with target  $q$ , or a subscript  $\delta_2$  with source  $p$ . This latter case is not possible, as from the fact that  $u_1 \in C$ , it would force  $p$  to be of type 3, the target type of  $\delta_2$ . So we have indeed that both  $u_1[i]$  and  $u_1[j]$  are the predecessors of  $u_2[i]$  and  $u_2[j]$  in the letters  $v[i], v[j]$  of  $A_{amb}$  respectively. It is straightforward to verify that this implies that  $u_1$  is the predecessor configuration of  $u_2$ : the only discrepancies allowed between  $u_1$  and  $u_2$  outside of positions  $i, j$  are of the form  $\{a_\delta, a\}$  and  $\{a, a^{\delta'}\}$  and do not influence the underlying letter of  $\Gamma$ . Similarly, in the other case, where  $u_2[i]$  is predecessor to  $[p.a]$  in the letter  $v[i]$ , we obtain that  $u_1$  is the successor configuration to  $u_2$ .  $\square$

**Lemma 28.** *It is impossible to have three distinct words  $u_1, u_2, u_3 \in C$  and  $v \in A^*$  such that for all  $i \in \{1, 2, 3\}$ ,  $u_i \leq_A v$ .*

*Proof.* By Lemma 27, any pair from  $\{u_1, u_2, u_3\}$  must encode two consecutive configurations of  $M$ . However, since the reading head must move at each step, from  $u_1$  to  $u_2$  and from  $u_2$  to  $u_3$ , this means the reading head moves either 0 or 2 positions between  $u_1$  and  $u_3$ , which yields a contradiction.  $\square$

#### F. The language $L$

We finally define  $L$  to be the monotone closure of  $L_{base}$  on alphabet  $A$ , so that  $L$  can contain letters from  $A_{amb}$ .

By Lemma 4, since  $L$  is the monotone closure of a regular language, it is regular (and monotone).

As a side remark, we can observe the following:

**Remark 29.**  $L$  is FO-definable. Since it is not crucial to the following, we only give here a rough intuition on why  $L$  is FO-definable. The language  $K$  from Section IV-A can be seen as an abstraction of  $L$ , with  $a, b, c$  playing the role of  $C_1, C_2, C_3$  respectively. In this light, and since  $C_1, C_2, C_3$  are all FO-definable, we can use the fact that the language  $K$  is FO-definable as well, by Lemma 15, to obtain an FO-formula for  $L$ . We also need Lemma 28 to guarantee that the equivalent of the letter  $\top$  from  $K$  never appears.

We will now prove in the next sections that  $L$  is  $\text{FO}^+$ -definable if and only if  $M$  is mortal, using Corollary 13.

#### G. $M$ not mortal $\implies L$ not $\text{FO}^+$ -definable

Let  $n \in \mathbb{N}$ , we aim to build  $(u, v) \in L \times \bar{L}$  such that  $u \leq_n v$ .

There is a configuration from which  $M$  has a run of length  $N + 3$ , with  $N = 2^{n+1} + 1$ . Let  $u = u_0 \# u_1 \# \dots \# u_N$  be an encoding of this run where each  $u_i \in C$ , and where we omitted the first and last configurations of the run, which may not be representable in  $C$  by Remark 25. Here all the  $u_i$ 's are of the same length, which is the size of the tape needed for this run.

By Lemma 26, for each  $i \in [0, N - 1]$ , there exists  $v_i \in A^*$  such that  $u_i \leq_A v_i$  and  $u_{i+1} \leq_A v_i$ .

We build  $v = u_0 \# v_1 \# \dots \# v_{N-2} \# u_N$ . Notice that  $v \notin L$ , because the types of  $u_0$  and  $u_N$  forces them to be separated by  $N - 1 \pmod 3$  configurations as in  $u$ , but in  $v$  they are separated by  $N - 2 \pmod 3$  configurations.

We describe a strategy for Duplicator witnessing  $u \leq_n v$ . It is a simple adaptation from the proof of Lemma 17, so we will just sketch the idea.

Let us consider that initially, there is a pair of initial (resp. final) tokens at the beginning (resp. end) of  $u, v$ . We will consider that the initial tokens are “blue”, and the final ones are “yellow”. In the following, a pair of corresponding tokens in  $u, v$  will be blue (resp. yellow) if they are at the same distance to the beginning (resp. end) of the word.

When Spoiler plays a token in  $u_i$  (resp.  $v_i$ ), Duplicator will look at the color of the closest token in  $u_i$ , (resp.  $v_i$ ), and answer with a token of the same color, i.e. by playing in  $v_i$  (resp.  $u_i$ ) for blue, and in  $v_{i-1}$  (resp.  $u_{i+1}$ ) for yellow. Of course, the same strategy applies to tokens played on  $\#$  positions.

This strategy preserves the following invariant: after  $k$  rounds, the number of  $\#$  between the last blue token and the first yellow token on the same word ( $u$  or  $v$ ) is at least  $2^{n-k}$ . This invariant guarantees that Duplicator wins the  $n$ -round game, since this gap will never be empty.

#### H. $M$ mortal $\implies L$ $\text{FO}^+$ -definable

Let  $M$  be a mortal  $TM$ , and  $n$  be the length of a maximal run of  $M$ , starting from any configuration.

We will show that  $L$  is  $\text{FO}^+$ -definable, by giving a strategy for Spoiler in  $EF_{f(n)}^+(u, v)$  for any  $(u, v) \in L \times \bar{L}$ , where the number of rounds  $f(n)$  depends only on  $n$ , and not on  $u, v$ .

Let  $(u, v) \in L \times \bar{L}$ . Without loss of generality we can assume that  $u \in L_{base}$ . This is because there exists  $u' \in L_{base}$  with  $u' \leq_A u$ , and we can consider the pair  $(u', v)$  instead of  $(u, v)$ . Indeed, if Spoiler wins on  $(u', v)$ , then the same strategy is winning on  $(u, v)$ , where his winning condition only gets easier.

Thus we can write  $u = u_0 \# u_1 \# \dots \# u_N$ , where each  $u_i$  is in  $C$ . Let us also write  $v = v_0 \# v_1 \# \dots \# v_T$ , where each  $v_i$  does not contain  $\#$ . Let us emphasize that no assumption is made on  $N$  and  $T$ , they can be any integers.

We will now describe a strategy for Spoiler in  $EF^+(u, v)$ , that is winning in a number  $f(n)$  of rounds only depending on  $n$ .

#### Ruling out local inconsistencies

As explained in the proof scheme of Section V-A, we will first show that if  $v$  presents *local inconsistencies* (that we define here formally via the notion of forbidden local factor), Spoiler can point them out in a bounded number of moves.

**Definition 30.** Let us call *local factor* a factor containing at most two symbols  $\#$ . A local factor is *forbidden* if it is not a factor of any word in  $L$ .

**Lemma 31.** *If  $v$  contains a forbidden local factor, Spoiler can win in a constant number of moves (at most 5).*

*Proof.* This can be seen by verifying that the language of words containing no forbidden local factors is  $\text{FO}^+$ -definable, with a formula  $\varphi_{loc}$  using at most 5 nested quantifiers. We sketch here how such a formula  $\varphi_{loc}$  can be built.

Notice that formulas of  $\text{FO}^+$  can use the letter  $\#$  either positively or negatively, since it is not comparable with any other letter. If  $(p, a) \in Q \times \Gamma$ , we define  $S(p, a) := \{(\sigma, \tau) \in (A_{base})^2 \mid \sigma[p.a]\tau \in C\}$  as the possible neighbourhoods of  $[p.a]$  in the base alphabet. Let us define an  $\text{FO}^+$ -formula  $\varphi_C(x)$  with free variable  $x$ , to be a formula that verifies that the maximal  $\#$ -free factor containing position  $x$  is in  $C^\uparrow$ , and that this is witnessed by the reading head at position  $x$ . This formula will verify that  $x$  contains some  $[p.a]$ , that the immediate neighbourhood of  $x$  is compatible with  $[p.a]$ , via a formula  $\bigvee_{(\sigma, \tau) \in S(p, a)} \sigma^\uparrow(x-1) \wedge \tau^\uparrow(x+1)$ , and that all other letters (not on positions  $\{x-1, x, x+1\}$ ) between the neighbouring  $\#$  symbols are in  $\Gamma^\uparrow$ .

We now give a description of the formula  $\varphi_{loc}$ . The formula will state that for all positions  $x < y$  of successive  $\#$  symbols (i.e. with no  $\#$  between them), there must be positions  $i_1 < x < i_2 < y < i_3$ , with only two  $\#$  symbols in  $[i_1, i_3]$ , such that  $\varphi_C(i_1) \wedge \varphi_C(i_2) \wedge \varphi_C(i_3)$ . Additionally, the types of the states in  $i_1, i_2, i_3$ , must be respectively either 1-2-3, 2-3-1, or 3-1-2.  $\square$

From now on, we will therefore assume that  $v$  does not contain forbidden local factors.

### Finding long-term inconsistencies

We will see how the only remaining cause for  $v$  not belonging to  $L$  is what we called *long-term inconsistencies* in the proof scheme of Section V-A. We will formalize this with the notion of non-coherent maximal ambiguous factor.

Let us start with an auxiliary definition.

**Definition 32.** A factor  $v_i$  of  $v$  is *compatible* with type  $j \in \{1, 2, 3\}$  if there exists  $u' \in C_j$  with  $u' \leq_A v_i$ . The *set-type* of  $v_i$  is  $\{j \mid v_i \text{ is compatible with } j\}$ .

By Lemma 28, each  $v_i$  is compatible with at most 2 distinct types in  $\{1, 2, 3\}$ . If  $v_i$  is compatible with 2 types, then one is the predecessor (resp. successor) of the other in the 1-2-3 cycle order, and we call it the *first type* (resp. *second type*) of  $v_i$ . We will consider that  $v_0$  (resp.  $v_T$ ) is only compatible with  $\text{type}(u_0)$  (resp.  $\text{type}(u_N)$ ). Indeed, if Duplicator matches  $v_0$  to a word  $u_i$  with  $i \neq 0$ , Spoiler can win the game in the next round, by choosing a  $\#$  position before  $u_i$  (and same argument for  $v_T$ ).

**Definition 33.** A factor of the form  $v_i\#v_{i+1}\#\dots\#v_j$  of  $v$  is called *ambiguous* if each  $v_k$  is compatible with two types, and the set-types succeed each other in the cycle order  $\{1, 2\} \rightarrow \{2, 3\} \rightarrow \{3, 1\}$ . For instance if the set-type of  $v_i$  is  $\{3, 1\}$ , then  $v_{i+1}$  must have set-type  $\{1, 2\}$ . An ambiguous factor is *maximal* if it is not contained in a strictly larger ambiguous factor.

**Definition 34.** A factor  $v_i$  of  $v$  is called an *anchor* if either  $i = 0, i = T$  or if  $v_{i-1}\#v_i\#v_{i+1}$  is not ambiguous.

If  $v_i$  is an anchor, we can uniquely define its *anchor type*. It is simply its type if  $i = 0$  or  $T$ , and otherwise since  $v_{i-1}\#v_i\#v_{i+1}$  is not ambiguous, we define the anchor type of  $v_i$  to be the only possible type for  $v_i$  that does not create an incoherence with its two neighbours. Notice that such a type exists, since we assumed  $v$  does not contain forbidden local factors.

**Example 35.** Assume  $v_5$  has set-type  $\{2, 3\}$ ,  $v_6$  has set-type  $\{3, 1\}$ , and  $v_7$  has set-type  $\{2, 3\}$ . Then  $v_6$  is an anchor, and its anchor type is 1. The type 3 is indeed impossible for  $v_6$ , since its successor type 1 is not in the set-type of  $v_7$ .

Notice that if Duplicator maps an anchor  $v_i$  to a word  $u_j$  such that  $\text{type}(u_j)$  is not the anchor type of  $v_i$ , then Spoiler can win in at most 5 moves, by pointing to a contradiction with the immediate neighbourhood of  $v_i$ .

**Definition 36.** Let  $v_i\#v_{i+1}\#\dots\#v_j$  be a maximal ambiguous factor. It is called *coherent* if  $v_{i-1}\#v_i\#\dots\#v_{j+1} \in L$ , and this is witnessed by the anchor types of  $v_{i-1}$  and  $v_{j+1}$ .

In other words,  $v_i\#v_{i+1}\#\dots\#v_j$  is coherent if the anchor types at  $v_{i-1}, v_{j+1}$  are either both concatenable with the first type of both  $v_i, v_j$ , or are both concatenable with their second type. Here by ‘‘concatenable’’, we mean to respect the 1-2-3 order, for instance type 3 must be followed by type 1.

**Example 37.** Let  $w = v_i\#v_{i+1}\#\dots\#v_j$  be a maximal ambiguous factor, where  $v_i$  has set-type  $\{1, 2\}$  and  $v_j$  has set-type  $\{2, 3\}$ . Assume  $v_{i-1}$  has anchor type 1, so it is concatenable with the second type of  $v_i$ . This means that for  $w$  to be coherent, we need  $v_{j+1}$  to have anchor type 1, in order to be concatenable with the second type of  $v_j$  as well.

**Lemma 38.**  *$v$  contains a maximal ambiguous factor  $w$  that is not coherent.*

*Proof.* Assume that all maximal ambiguous factors of  $v$  are coherent. Since  $v$  does not contain forbidden local factors, we have that the anchor types of two consecutive anchors follow the 1-2-3 order. This means that the anchor types, together with the coherence of maximal ambiguous factors, give us a witness that  $v \in L$ . This witness is a word of  $L_{base}$ , obtained by choosing the anchor types on all anchors, and either the first type or the second type uniformly in maximal ambiguous factors, as fixed by the anchors at the extremities. Since we know that  $v \notin L$ , this is a contradiction.  $\square$

We are now ready to describe Spoiler’s strategy. Spoiler starts by placing two tokens delimiting a maximal ambiguous factor  $w$  that is not coherent, as obtained in Lemma 38. Because  $w$  is not coherent, Duplicator is forced to answer with the first type for one of these tokens, and with the second type for the other: otherwise Spoiler immediately wins by exposing a local inconsistency with the anchors delimiting  $w$ .

We will now show how Spoiler can win on such a factor  $w$ , starting with these two tokens. For this, we will introduce the notion of height of a configuration word, and the integer game that will abstract the  $EF^+$ -game on  $w$ . This is where we finally make use of the hypothesis that  $M$  is mortal.

### Abstracting words by integers

If  $u \in C$  is a configuration word, let us define its *height*  $h(u)$  to be the length of the run starting in  $u$ , and not going outside of the tape specified in  $u$ .

If  $u \in C$ , let us also define its  $n$ -approximation  $\alpha_n(u)$  as the maximal word in  $(A_{base})^{\leq n} \cdot (Q \times \Gamma) \cdot (A_{base})^{\leq n}$  that is an infix on  $u$ . That is, we remove letters whose distance to the reading head is bigger than  $n$ .

Here are a few properties of the height:

**Lemma 39.** *For all  $u \in C$ , the following hold:*

- $0 < h(u) < n$ .
- For all  $x, y \in \Gamma^*$ , we have  $h(xuy) \geq h(u)$ .
- $h(u) = h(\alpha_n(u))$ .
- If  $v \in C$  is the successor configuration of  $u \in C$ , then  $h(v) = h(u) - 1$ .

*Proof.* The first item is a consequence of the fact that  $M$  is mortal with bound  $n$ . Notice that the inequalities are strict because of Remark 25: words from  $C$  must have a predecessor and a successor configuration. The second item comes from the fact that the run of length  $h(u)$  starting in  $u$  is still possible when adding a context  $x, y$ , which is not affected by this run. The third item uses the fact that a run can only visit the  $n$ -approximation of  $u$ , so the context outside of  $\alpha_n(u)$  does not affect the height  $h(u)$ . The fourth item is a basic consequence of the definition of the height.  $\square$

**Corollary 40.** *The height of a configuration word  $u$  is an  $FO^+$ -definable property, i.e for all  $k \in \mathbb{N}$  there exists an  $FO^+$  formula  $h_k$  such that  $h_k$  accepts a configuration word  $u \in C$  if and only if  $h(u) = k$ .*

*Proof.* By Lemma 39, the formula  $h_k$  can simply use a lookup table to verify that  $\alpha_n(u)$  is of height  $k$ , using a finite disjunction listing possibilities for  $\alpha_n(u)$  being of height  $k$ . When evaluated on  $A^*$ , the formula  $h_k$  will accept the monotone closure of configuration words of height  $k$ .  $\square$

**Remark 41.** We use here the fact that computation is done locally around the reading head to obtain Corollary 40. This seems to make Turing Machines more suited to this reduction than e.g. cellular automata, where computation is done in parallel on the whole tape.

Thanks to the height abstraction, we will show that we can focus on playing a special kind of abstracted EF-game.

### The integer game

The idea is to abstract a configuration word  $u_i \in C$  by its height  $h(u_i)$ . If  $u'$  is the predecessor configuration of  $u''$ , and  $u', u'' \leq_A v$ , we will abstract the word  $v$  by  $\binom{h(u')}{h(u'')}$ .

Let  $\Sigma_{base} = [0, n]$  and  $\Sigma_{amb} = \{\binom{i}{i-1} \mid 1 \leq i \leq n\}$ . Let  $\Sigma = \Sigma_{base} \cup \Sigma_{amb}$ , ordered by  $i \leq_\Sigma \binom{i}{i-1}$  and  $i-1 \leq_\Sigma \binom{i}{i-1}$  for all  $\binom{i}{i-1} \in \Sigma_{amb}$ .

We define the  $n$ -integer game as follows: It is played on an arena  $(u, v)$  with  $u \in (\Sigma_{base})^*$  and  $v \in (\Sigma_{amb})^*$ . If we note  $i$  (resp.  $j$ ) the first (resp. last) letter of  $u$ , then the first (resp. last) letter of  $v$  is  $\binom{i}{i-1}$  (resp.  $\binom{j+1}{j}$ ).

The rest of the rules is very close to those of  $EF^+(u, v)$ : in each round, Spoiler plays a token in  $u$  or  $v$ , Duplicator has to answer with a token in the other word, while maintaining the order between tokens, and the constraint that the label of a token in  $u$  is  $\leq_\Sigma$ -smaller than the label of its counterpart in  $v$ . We add an additional *neighbouring constraint* for Duplicator: consecutive tokens in one word must be related to consecutive tokens in the other, and in this case, if two tokens of  $v$  are in consecutive positions labelled  $\binom{i}{i-1} \binom{j}{j-1}$ , the corresponding tokens in  $u$  must be either labelled  $i, j$  or  $i-1, j-1$ . A mix  $i, j-1$  or  $i-1, j$  is not allowed.

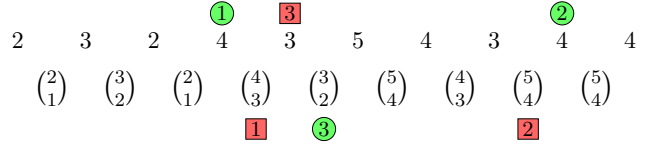


Fig. 3. A position of the integer game.

**Lemma 42.** *For all  $n \in \mathbb{N}$ , Spoiler can win any  $n$ -integer game in  $2n$  rounds.*

To keep the flow of the global proof, we defer the proof of Lemma 42 to Section V-I. We will first see how to use this Lemma to conclude the proof.

**Remark 43.** Lemma 42 still holds if the definition of  $n$ -integer game is generalized to include the symmetric case where, if we note  $i, j$  the first and last letters of  $u$  respectively,  $v$  starts with  $\binom{i+1}{i}$  and ends with  $\binom{j}{j-1}$ . Indeed, it suffices to consider the mirrored images of  $u$  and  $v$  to show that Spoiler wins in the same amount of rounds.

### Lifting the strategy to the original $EF^+$ -game

Let us go back to the  $EF^+$ -game on  $u, v$ , where Spoiler has placed two tokens delimiting a non-coherent maximal ambiguous factor  $w$  in  $v$ .

Spoiler can now play only between these existing tokens, and import the strategy from the integer game, by abstracting each word  $u_i$  by its height and each word  $v_j$  by  $\binom{h(u')}{h(u'')}$ , where  $u', u'' \in C$  are such that  $u \leq_A v$ ,  $u' \leq_A v$ , and  $h(u') = 1 + h(u'')$ . Each factor of  $u, v$  delimited by  $\#$  corresponds to a single position in the abstracted integer game. Spoiler can for instance mimic a move of the integer game by playing in the first position of the corresponding factor in  $u$  or  $v$ , i.e. just after a  $\#$ -labelled position.

**Lemma 44.** *If Duplicator does not comply with the rules of the integer game, then Spoiler can punish it in at most  $\log n$  rounds.*

*Proof.* Assume  $u_i$  is matched to  $v_j$ , but there is no  $u' \leq_A v_j$  such that  $h(u_i) = h(u')$ . It means that  $\alpha_n(u)$  is not compatible with  $v$ , and this can be punished by Spoiler using  $\log n$  rounds (with a dichotomy strategy, or  $n$  rounds with a naive strategy). Thus by Lemma 39, Spoiler can enforce the basic rule of the integer game, stating that if integer  $t$  is matched to  $\binom{s+1}{s}$ , then  $t = s + 1$  or  $t = s$ . Using the correspondence between  $\text{EF}^+$ -games and  $\text{FO}^+$ -definability, this property can also be seen via Corollary 40.

If neighbours are matched with non-neighbours, then it suffices for Spoiler to point the two  $\#$  positions between the non-neighbours, that cannot be matched in the other word, so he wins in 2 moves. We show that the rest of the neighbourhood rule is also enforced. Assume  $u_i \# u_{i+1}$  is matched to  $v_j \# v_{j+1}$ . Assume  $\text{type}(u_i)$  is the first (resp. second) type of  $v_j$  while  $\text{type}(u_{i+1})$  is the second (resp. first) type of  $v_{j+1}$ . By definition of  $L$ ,  $\text{type}(u_{i+1})$  must be the successor type of  $\text{type}(u_i)$ , for instance without loss of generality,  $\text{type}(u_i) = 1$  and  $\text{type}(u_{i+1}) = 2$ . Then, the set-type of  $v_i$  is  $\{1, 2\}$  (resp.  $\{3, 1\}$ ) and the set-type of  $v_{j+1}$  is  $\{1, 2\}$  (resp.  $\{2, 3\}$ ). This contradicts the fact that  $v_i \# v_{i+1}$  is part of an ambiguous factor, as set-types should follow each other in the order  $\{1, 2\}$ - $\{2, 3\}$ - $\{3, 1\}$ .  $\square$

Combining these arguments and by Lemma 42, we obtain that following this strategy, Spoiler will win in at most  $f(n) = 2 + 2n + \log n + 5$  rounds, by punishing Duplicator as soon as Duplicator loses the  $n$ -integer game.

Using Corollary 13, we obtain that  $L$  is  $\text{FO}^+$ -definable, with a formula of quantifier rank at most  $f(n)$ .

**Remark 45.** The alphabet  $A$  can be turned into a powerset alphabet, by adding all subsets of  $A_{\text{base}}$  absent from  $A_{\text{amb}}$ , rejecting any word containing  $\emptyset$  but no new non-empty subset, and accepting any word containing a new non-empty subset. This shows that this undecidability result still holds in the special case of powerset alphabets.

This concludes the proof of Theorem 22, up to the proof of Lemma 42 which is done in the next section.

### 1. Winning the integer game

Let us show Lemma 42: Spoiler can win any  $n$ -integer game in  $2n$  rounds.

*Proof.* Let  $(u, v)$  be an arena for an  $n$ -integer game. We proceed by induction on  $n$ .

For  $n = 1$ , the constraints on the game forces  $u \in 1(0+1)^*0$  and  $v \in \binom{1}{0}^*$ .

We can have Spoiler play on the last occurrence of 1 in  $u$ , and on the successor position labelled 0. Duplicator cannot respond to these two moves while respecting the neighbouring constraint, so Spoiler wins in 2 moves.

Assume now that for some  $n \geq 1$ , Spoiler wins any  $n$ -integer game in  $2n$  moves, and consider an  $(n + 1)$ -integer game arena  $(u, v)$ . If the letters  $n + 1$  and  $\binom{n+1}{n}$  do not appear in  $u, v$  respectively, then Spoiler can win in  $2n$  moves by induction hypothesis.

If the letter  $n + 1$  does not appear in  $u$ , then let  $y$  be the first position labelled  $\binom{n+1}{n}$  in  $v$ . By definition of the integer game  $y$  cannot be the first position of  $v$ , otherwise  $u$  should start with  $n + 1$ . We will choose position  $y$  in  $v$  for the first move of Spoiler, let  $x$  be the position in  $u$  answered by Duplicator, we have  $u[x] = n$ . We can assume that  $x$  is not the first position of  $u$ , otherwise Spoiler can win in the next move. If Spoiler were to play  $x - 1$  in  $u$ , with  $u[x - 1] = i$ , by the neighbouring constraint Duplicator would be forced to answer  $y - 1$  in  $v$ , with label  $\binom{i+1}{i}$ . This shows that the words  $u[.x-1]$  and  $v[.y-1]$  form a correct  $n$ -integer arena, as the integer  $n + 1$  is not present anymore, and all other constraints are respected. Therefore, Spoiler can win by playing  $2n$  moves in these prefixes. This gives a total of  $2n + 1$  moves in the original  $(n + 1)$ -integer game.

Finally, if the letter  $n + 1$  does appear in  $u$ , Spoiler starts by playing the position  $x$  in  $u$  corresponding to the last occurrence of  $n + 1$  in  $u$ . Duplicator must answer a position  $y$  labelled  $\binom{n+1}{n}$ . Notice that neither  $x$  nor  $y$  can be a last position, so  $u[x + 1]$  and  $v[y + 1]$  are well-defined. As before, using the neighbouring constraint, we know that if  $i = u[x + 1]$ , then  $v[y + 1] = \binom{i}{i-1}$ . Therefore, the words  $u[x+1..]$  and  $v[y+1..]$  form an  $(n + 1)$ -integer game arena, and moreover the letter  $n + 1$  does not appear in  $u[x+1..]$  (by choice of  $x$ ). Using the precedent case, we know that Spoiler can win from there in  $2n + 1$  moves, playing only on  $u[x+1..]$  and  $v[y+1..]$ . This gives a total of  $2n + 2$  moves in the original  $(n + 1)$ -integer game, thereby completing the induction proof.  $\square$

### Conclusion

We believe this paper gives an example of fruitful interaction between automata theory and model theory. Indeed, a classical result of model theory, the failure of Lyndon's theorem on finite structures, has been greatly simplified by using the toolbox of regular languages. Conversely, this question coming from model theory, when considered on regular languages, yields the first (to our knowledge) natural fragment of regular languages with undecidable membership problem, and opens new techniques for proving undecidability of expressibility in positive logics. We hope that the tools developed in this paper can be further used in both fields, and that this will encourage more interactions of this form in the future.

In the short term, we are interested in extending these techniques to the framework of cost functions, see [Kup14], [Kup], and to other extensions of regular languages.

**Acknowledgements.** I am grateful to Thomas Colcombet for bringing this topic to my attention, and in particular for asking the question  $\text{FO}^+ \stackrel{?}{=} \text{monotone FO}$ , as well as for many interesting exchanges. Thanks also to Amina Doumane and Sam Van Gool for helpful discussions, and to the anonymous reviewers, as well as Anupam Das and Natacha Portier, for their comments on earlier versions of this document.

### REFERENCES

- [AG87] Miklos Ajtai and Yuri Gurevich. Monotone versus positive. *J. ACM*, 34(4):1004–1015, October 1987.

- [AG97] Natasha Alechina and Yuri Gurevich. *Syntax vs. semantics on finite structures*, pages 14–33. Springer Berlin Heidelberg, Berlin, Heidelberg, 1997.
- [Boj04] Mikołaj Bojańczyk. A bounding quantifier. In Jerzy Marcinkowski and Andrzej Tarlecki, editors, *Computer Science Logic*, pages 41–55, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg.
- [Col11] Thomas Colcombet. Green’s relations and their use in automata theory. In *Language and Automata Theory and Applications - 5th International Conference, LATA 2011, Tarragona, Spain, May 26-31, 2011. Proceedings*, volume 6638 of *Lecture Notes in Computer Science*, pages 1–21. Springer, 2011.
- [Col12] Thomas Colcombet. Regular cost functions, part i: Logic and algebra over words. volume 9, 12 2012.
- [Col13] Thomas Colcombet. Magnitude monadic logic over words and the use of relative internal set theory. In *2013 28th Annual ACM/IEEE Symposium on Logic in Computer Science*, pages 123–123, 2013.
- [DG08] Volker Diekert and Paul Gastin. First-order definable languages. In *Logic and Automata: History and Perspectives, Texts in Logic and Games*, pages 261–306. Amsterdam University Press, 2008.
- [FSS81] M. Furst, J. B. Saxe, and M. Sipser. Parity, circuits, and the polynomial-time hierarchy. In *22nd Annual Symposium on Foundations of Computer Science (sfcs 1981)*, pages 260–270, 1981.
- [GS92] Michelangelo Grigni and Michael Sipser. Monotone complexity. In *Proceedings of the London Mathematical Society Symposium on Boolean Function Complexity*, page 57–75, USA, 1992. Cambridge University Press.
- [Hoo66] Philip K. Hooper. The undecidability of the turing machine immortality problem. *Journal of Symbolic Logic*, 31(2):219–234, 1966.
- [Kam68] Hans W. Kamp. *Tense Logic and the Theory of Linear Order*. Phd thesis, University of Warsaw, 1968.
- [Kup] Denis Kuperberg. Erratum for [Kup14]. <http://perso.ens-lyon.fr/denis.kuperberg/papers/Erratum.pdf>.
- [Kup14] Denis Kuperberg. Linear temporal logic for regular cost functions. *Logical Methods in Computer Science*, 10(1), 2014.
- [KVB12] Denis Kuperberg and Michael Vanden Boom. On the expressive power of cost logics over infinite words. In *Automata, Languages, and Programming*, pages 287–298, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.
- [Lib04] Leonid Libkin. *Elements of Finite Model Theory*. Springer, August 2004.
- [LSS96] C. Lautemann, T. Schwentick, and I. A. Stewart. On positive p. In *2012 IEEE 27th Conference on Computational Complexity*, page 162. IEEE Computer Society, 1996.
- [Lyn59] Roger C. Lyndon. Properties preserved under homomorphism. *Pacific J. Math.*, 9(1):143–154, 1959.
- [MP71] Robert McNaughton and Seymour A. Papert. *Counter-Free Automata (M.I.T. Research Monograph No. 65)*. The MIT Press, 1971.
- [PST89] Jean-Eric Pin, Howard Straubing, and Denis Therien. New results on the generalized star-height problem. volume 349, pages 458–467, 02 1989.
- [PZ19] Thomas Place and Marc Zeitoun. Going higher in first-order quantifier alternation hierarchies on words. *J. ACM*, 66(2), March 2019.
- [Ros08] Benjamin Rossman. Homomorphism preservation theorems. *J. ACM*, 55, 07 2008.
- [Sch65] M.P. Schützenberger. On finite monoids having only trivial subgroups. *Information and Control*, 8(2):190 – 194, 1965.
- [Ste94] Iain A. Stewart. Logical Description of Monotone NP Problems. *Journal of Logic and Computation*, 4(4):337–357, 1994.
- [Sto95] Alexei P. Stolboushkin. Finitely monotone properties. In *LICS, San Diego, California, USA, June 26-29, 1995*, pages 324–330. IEEE Computer Society, 1995.

## APPENDIX

### A. Proof of Lemma 10

We prove here that any language definable by  $\text{FO}^+$  is monotone.

This is done by induction on the  $\text{FO}^+$  formula  $\varphi$ , where the induction property is strengthened to include possible free

variables: for all  $(u, \alpha) \in \llbracket \varphi \rrbracket$  and  $v \geq_A u$ , we have  $(v, \alpha) \in \llbracket \varphi \rrbracket$ .

#### Base cases:

Let  $(u, \alpha) \in \llbracket a^\dagger(x) \rrbracket$  and  $v \geq_A u$ , we have  $v[\alpha(x)] \geq_A u[\alpha(x)] \geq_A a$ , so  $(v, \alpha) \in \llbracket a^\dagger(x) \rrbracket$ .

Let  $(u, \alpha) \in \llbracket x \leq y \rrbracket$  and  $v \geq_A u$ . We have  $\alpha(x) \leq \alpha(y)$  so  $(v, \alpha) \in \llbracket x \leq y \rrbracket$ . The argument for  $<$  instead of  $\leq$  is identical.

#### Induction cases:

Let  $(u, \alpha) \in \llbracket \varphi \vee \psi \rrbracket$  and  $v \geq_A u$ . We have  $(u, \alpha) \in \llbracket \varphi \rrbracket$  or  $(u, \alpha) \in \llbracket \psi \rrbracket$ . Therefore, by induction hypothesis,  $(v, \alpha) \in \llbracket \varphi \rrbracket$  or  $(v, \alpha) \in \llbracket \psi \rrbracket$ , hence  $(v, \alpha) \in \llbracket \varphi \vee \psi \rrbracket$ . The argument for  $\varphi \vee \psi$  is identical.

Let  $(u, \alpha) \in \llbracket \exists x.\varphi \rrbracket$  and  $v \geq_A u$ . There exists  $i \in \text{dom}(u)$  such that  $(u, \alpha[x \mapsto i]) \in \llbracket \varphi \rrbracket$ . By induction hypothesis,  $(v, \alpha[x \mapsto i]) \in \llbracket \varphi \rrbracket$ . Hence,  $(v, \alpha) \in \llbracket \exists x.\varphi \rrbracket$ . The argument for  $\forall$  is identical.

### B. Proof of Theorem 12

The proof is an adaptation of the classical proof for correctness of EF-games, see e.g. [Lib04].

Since  $\text{FO}^+$  is a fragment of FO, we can directly use the following Lemma:

**Lemma 46** ([Lib04, Lem 3.13]). *Let  $n, k \in \mathbb{N}$ . Up to logical equivalence, there are finitely many formulas of quantifier rank at most  $n$  using  $k$  free variables.*

We will now show a strengthening of Theorem 12, where free variables are incorporated in the statement:

**Theorem 47.** *Let  $n, k \in \mathbb{N}$ ,  $u, v \in A$ ,  $\alpha : [1, k] \rightarrow \text{dom}(u)$  and  $\beta : [1, k] \rightarrow \text{dom}(v)$  be valuations for  $k$  variables  $x_1, \dots, x_k$  in  $u, v$  respectively. Then Duplicator wins  $\text{EF}_n^+(u, \alpha, v, \beta)$  if and only if for any  $\text{FO}^+$  formula  $\varphi$  with  $\text{qr}(\varphi) \leq n$  using  $k$  free variables  $x_1 \dots x_k$ , we have  $u, \alpha \models \varphi \Rightarrow v, \beta \models \varphi$ .*

*Proof.* We prove this by induction on  $n$ .

#### Base case $n = 0$ :

Notice that quantifier-free formulas of  $\text{FO}^+$  are just positive boolean combinations of atomic formulas, that either compare the values of the free variables, or assert that the label of a free variable is  $\leq_A$ -greater than some letter  $a \in A$ . Consider that there is a quantifier-free formula  $\varphi$  with  $k$  free variables accepting  $u, \alpha$  but rejecting  $v, \beta$ . This happens if and only if there is a variable  $x_i$  such that  $u[\alpha(x_i)] \not\leq_A v[\beta(x_i)]$ , or if two variables  $x_i, x_j$  are not in the same order according to  $\alpha$  and  $\beta$ . That is, this happens if and only if  $(u, \alpha, v, \beta)$  is not a valid  $k$ -position, i.e. if and only if Spoiler wins the 0-round game  $\text{EF}_0^+(u, \alpha, v, \beta)$ .

**Induction case:** Assume there is an  $\text{FO}^+$  formula  $\varphi$  with  $\text{qr}(\varphi) \leq n$ , accepting  $u, \alpha$  but not  $v, \beta$ . The formula  $\varphi$  is a positive combination of atomic formulas, formulas of the form  $\exists x.\psi$ , and formulas of the form  $\forall x.\psi$ . Therefore, one of these formulas accepts  $u, \alpha$  but not  $v, \beta$ . If it is an atomic formula,

then Spoiler immediately wins  $\text{EF}_n^+(u, \alpha, v, \beta)$  as in the base case.

If it is a formula of the form  $\exists x.\psi$ , then Spoiler can use the following strategy: pick a position  $p$  witnessing that the formula is true for  $u, \alpha$ , and play the position  $p$  in  $u$ . Duplicator will answer a position  $p'$  in  $v$ , and the game will move to  $(u, \alpha', v, \beta')$ , where  $\alpha' = \alpha[x \mapsto p]$  and  $\beta' = \beta[x \mapsto p']$ . Since the formula  $\psi$  has quantifier rank at most  $n - 1$ , and accepts  $u, \alpha'$  but not  $v, \beta'$ , by induction hypothesis Spoiler can win in the remaining  $n - 1$  rounds of the game.

Now if it is a formula of the form  $\forall x.\psi$ , then Spoiler can do the following: pick a position  $p'$  witnessing that the formula is false for  $v, \beta$ , and play the position  $p'$  in  $v$ . Duplicator will answer a position  $p$  in  $u$ , and the game will move to  $(u, \alpha', v, \beta')$ , where  $\alpha' = \alpha[x \mapsto p]$  and  $\beta' = \beta[x \mapsto p']$ . Since the formula  $\psi$  has quantifier rank at most  $n - 1$ , and accepts  $u, \alpha'$  but not  $v, \beta'$ , by induction hypothesis Spoiler can win in the remaining  $n - 1$  rounds of the game.

Let us now show the converse implication. We assume any formula of quantifier rank at most  $n$  accepting  $u, \alpha$  must accept  $v, \beta$ , and we give a strategy for Duplicator in  $\text{EF}_n^+(u, \alpha, v, \beta)$ .

Suppose Spoiler places token  $x$  at position  $p$  in  $u$ . Let  $\alpha' = \alpha[x \mapsto p]$ . By Lemma 46, up to logical equivalence, there is only a finite set  $F$  of  $\text{FO}^+$  formulas of rank at most  $n - 1$  with  $k + 1$  free variables accepting  $u, \alpha'$ . Let  $\psi = \bigwedge_{\varphi \in F} \varphi$ . Then  $u, \alpha$  satisfies the formula  $\exists x.\psi$  of rank  $n$  (as witnessed by  $p$ ), so by assumption we also have  $v, \beta \models \exists x.\psi$ . This means there is a  $p' \in \text{dom}(v)$  such that  $v, \beta' \models \psi$ , where  $\beta' = \beta[x \mapsto p']$ . Duplicator can answer position  $p'$  in  $v$ , and by induction hypothesis he will win the remaining of the game, since every formula of  $F$  accepts  $v, \beta'$ .

Suppose now that Spoiler places token  $x$  at position  $p'$  in  $v$ . Let  $\beta' = \beta[x \mapsto p']$ . Let  $F$  be the finite set of formulas (up to equivalence) of quantifier rank at most  $n - 1$  and with  $k + 1$  free variables, that reject  $v, \beta'$ . Let  $\psi = \bigvee_{\varphi \in F} \varphi$ , and  $\psi' = \forall x.\psi$ . By construction,  $x = p'$  witnesses that  $\psi'$  does not accept  $v, \beta$ . Our assumption implies that it does not accept  $u, \alpha$  either. So there is  $p \in \text{dom}(u)$  such that  $u, \alpha' \not\models \forall x.\psi$ , where  $\alpha' = \alpha[x \mapsto p]$ . Duplicator can answer position  $p$  in  $u$ . If a formula  $\varphi$  of rank at most  $n - 1$  is true in  $u, \alpha'$ , then by construction it cannot appear in  $F$ , therefore it is also true in  $v, \beta'$ . By induction hypothesis, Duplicator wins the remaining  $(n - 1)$ -round game starting from  $(u, \alpha', v, \beta')$ .  $\square$

### C. Syntactic monoid for the language $K$

It is instructive to see what the syntactic monoid of  $K$  looks like, in particular to get a first intuition on how an FO formula can be defined for  $K$ .

We depict this monoid  $M$  in Figure 4, using the eggbox representation based on Green's relations: boxes are  $\mathcal{J}$ -classes, lines are  $\mathcal{R}$ -classes, columns are  $\mathcal{L}$ -classes, and cells are  $\mathcal{H}$ -classes. See [Col11] for an introduction to Green's relations and eggbox representation.

The syntactic morphism  $h : A^* \rightarrow M$  is easily inferred, as elements of the monoid in  $h(A)$  are directly named after

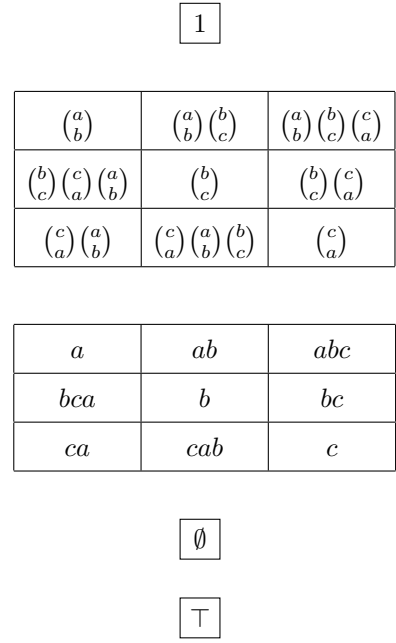


Fig. 4. The syntactic monoid  $M$  of  $K$

the letter mapping to them. The accepting part of  $M$  is  $F = \{1, \begin{pmatrix} a \\ b \\ c \\ a \end{pmatrix}, \begin{pmatrix} c \\ a \\ b \\ c \end{pmatrix}, abc, \top\}$ .

To show that  $K$  is FO-definable, it suffices to verify that  $M$  is aperiodic, which is directly visible on Figure 4, as all  $\mathcal{H}$ -classes are singletons (see [Col11]).

### D. An explicit FO formula for the language $K$

Recall that  $K = (a^\uparrow b^\uparrow c^\uparrow)^* + A^* \top A^*$ . We describe here the behaviour of a formula witnessing that  $K$  is FO-definable.

The  $A^* \top A^*$  part of  $K$  is just to rule out words containing  $\top$  by accepting them, which can be done by a formula  $\exists x.\top(x)$ . So we just need to design a formula  $\varphi$  for  $K' = (a^\uparrow b^\uparrow c^\uparrow)^* \setminus (A^* \top A^*)$ , assuming the letter  $\top$  does not appear, the final formula will then be  $\varphi \vee \exists x.\top(x)$ .

We will call *forbidden pattern* any word that is not an infix of a word in  $K'$ . Let us call *anchor* a position  $x$  such that either  $x$  is labelled by a singleton, or  $x$  is labelled by  $\begin{pmatrix} a \\ b \end{pmatrix}$  (resp.  $\begin{pmatrix} b \\ c \end{pmatrix}, \begin{pmatrix} c \\ a \end{pmatrix}$ ) with  $x + 1$  labelled by a letter different from  $\begin{pmatrix} b \\ c \end{pmatrix}$  (resp.  $\begin{pmatrix} c \\ a \end{pmatrix}, \begin{pmatrix} a \\ b \end{pmatrix}$ ). The idea is that if  $x$  is an anchor position of  $u \in K'$ , then there is only one possibility for the value of  $x \bmod 3$ . If the first position is labelled by a letter from  $a^\uparrow$ , we will consider that it is an anchor labelled  $a$ , otherwise we will reject the input word. Similarly, the last position is either a  $c$  anchor or causes immediate rejection of the word. If  $x, y$  are successive anchor positions (i.e. with no other anchor positions between them), the word  $u[x + 1..y - 1]$  is necessarily an infix of  $(\begin{pmatrix} a \\ b \\ c \\ a \end{pmatrix} \begin{pmatrix} b \\ c \\ a \end{pmatrix} \begin{pmatrix} c \\ a \\ b \end{pmatrix})^*$ . We say that an anchor  $x$  *goes right-up* (resp. *right-down*) if we can replace the letter  $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$  by  $\alpha$  (resp.  $\beta$ ) at position  $x + 1$  without having a forbidden pattern in the immediate neighbourhood of  $x$ . Notice that  $x$  can not go both right-up and right-down. We define in the same way the left-up and left-down property by replacing  $x + 1$  with  $x - 1$ . For instance



