



HAL
open science

Adding modular predicates

Luc Dartois, Charles Paperman

► **To cite this version:**

| Luc Dartois, Charles Paperman. Adding modular predicates. 2014. hal-00934622v2

HAL Id: hal-00934622

<https://hal.science/hal-00934622v2>

Preprint submitted on 19 May 2014 (v2), last revised 13 Nov 2015 (v3)

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.

Adding modular predicates^{*}

Luc Dartois¹ and Charles Paperman¹

LIAFA, Université Paris-Diderot and CNRS,
Case 7014, 75205 Paris Cedex 13, France

`luc.dartois, charles.paperman@liafa.univ-paris-diderot.fr`

Abstract. This paper is a contribution to the study of regular languages defined by fragments of first order or even monadic second order logic. More specifically, we consider the operation of enriching a given fragment by adding modular predicates. Our first result gives a simple algebraic counterpart to this operation in terms of semidirect products of varieties together with a combinatorial description based on elementary operations on languages. Now, a difficult question is to know whether the decidability of a given fragment is preserved under this enrichment. We first prove that this is always the case for so-called local varieties. The problem remains open in the nonlocal case but our main results also gives several sufficient conditions to preserve decidability. We use these latter results to establish the decidability of three fragments of the first order logic with two variables.

1 Introduction

The *decision problem* for a given class of regular languages consists in deciding, given a regular language, whether or not it belongs to this class. Solving the decision problem for various fragments of monadic second order is a well-studied problem on regular languages [5, 6, 8–11, 13, 15, 19–23]. Fragments of logic are usually defined in terms of their quantifier complexity (Σ_n -classes) or number of variables allowed in the formulae. Another possible parameter is to impose restrictions on the numerical predicates in the signature. A complete classification of the numerical predicates defining only regular languages was given by Péladeau [14] and Straubing [20]. There are essentially three basic groups of such predicates: the linear order, the local predicates LOC and the modular predicates MOD. Given a fragment $\mathcal{F}[\sigma]$ on the signature σ , the enrichment $\mathcal{F}[\sigma] \rightarrow \mathcal{F}[\sigma, \text{LOC}]$ has been widely studied [16, 19, 24]. For instance, Straubing [19] gave a nice algebraic interpretation of the enrichment $\mathcal{F}[\langle \rangle] \rightarrow \mathcal{F}[\langle \rangle, \text{LOC}]$ when \mathcal{F} is the fragment $\mathcal{B}\Sigma_n$ of Boolean combinations of Σ_n -formulae. The natural framework to state this kind of result is Eilenberg's theory of varieties and can be roughly summarized as follows:

^{*} The authors are supported by the project ANR 2010 BLAN 0202 02 FREC, the second author is supported by Fondation CFM.

1. In good cases (but not always) the enrichment by LOC corresponds to the operation $\mathbf{V} \rightarrow \mathbf{V} * \mathbf{LI}$ (the semidirect product by the variety \mathbf{LI} of locally trivial semigroups) on varieties.
2. If \mathbf{V} is a local variety, then \mathbf{V} is decidable if and only if $\mathbf{V} * \mathbf{LI}$ is decidable.
3. The nonlocal case requires advanced algebraic tools (notably derived categories) and is still the topic of intense research. Several important cases have been solved positively, but Auinger [2] exhibited an example of a decidable variety \mathbf{V} such that $\mathbf{V} * \mathbf{LI}$ is undecidable.

The aim of this paper is to establish similar results for the enrichment $\mathcal{F}[\sigma] \rightarrow \mathcal{F}[\sigma, \mathbf{MOD}]$. Our first result (Theorem 6) states that the algebraic counterpart to this enrichment is another semidirect product, the operation $\mathbf{V} \rightarrow \mathbf{V} * \mathbf{MOD}$, where \mathbf{MOD} is the variety of cyclic stamps. Our second result (Theorem 15) shows that when \mathbf{V} is local, then \mathbf{V} is decidable if and only if $\mathbf{V} * \mathbf{MOD}$ is decidable. Finally, our main result (Theorem 19) deals with the nonlocal case. Its proof relies on the equational theory for variety of finite categories. Figure 6, which can be found at the end of section 6, summarizes the consequences of our results for deciding various fragments of first-order logic.

2 Preliminaries

2.1 Words and logic

Let A be a finite alphabet and σ a relational signature. Given a word $u = a_0 \cdots a_{n-1}$ of length n , we associate to u the *relational structure* $M_u = \{[0, n-1], (\mathbf{a})_{a \in A}, (P^u)_{P \in \sigma}\}$, where P^u is the interpretation of the symbol P over the interval $[0, n-1]$ and $(\mathbf{a})_{a \in A}$ are disjoint monadic predicates given by the positions of the letters over the structure. For instance, if $u = aabbab$, then $\mathbf{a} = \{0, 1, 4\}$ and $\mathbf{b} = \{2, 3, 5\}$. Basic examples of predicates include the binary predicate $<$, interpreted as the usual order on integers. For each $k \geq 0$, we define the LOC_k predicates to be the unary predicates $x = \min + k$, which is true at the position k , the dual predicate $x = \max - k$ and the binary predicate $x = y + k$. The class LOC of local predicates is the union of all LOC_k . We also consider the modular predicate MOD_i^d , which holds at all positions equal to i modulo d , and the 0-ary modular predicate D_i^d which is true if the word length is equal to i modulo d . For $u = aabbab$, we have $\text{MOD}_0^2 = \{0, 2, 4\}$, and D_1^3 is *false* whereas D_0^3 is *true*. We denote by MOD_d the set of *modular* predicates modulo d . We define \mathbf{MOD} as the union of all MOD_d .

Formulae are interpreted on words in the usual way (see [20]). For instance the formula $\exists x \exists y \exists z \mathbf{a}x \wedge \mathbf{b}y \wedge \mathbf{a}z \wedge (x < y) \wedge (y < z)$ defines the language $A^*aA^*bA^*aA^*$. Since a sentence defines a language, one can naturally associate a class of languages to a class of sentences.

In [9], Kufleitner and Lauser defined *fragments of logic* as sets of formulae closed under some syntactical substitutions. Here, we only require substitutions on an atomic level. Thus in this paper, a *fragment of logic* is a set of formulae closed under atomic substitutions. Classical examples include $\mathcal{BS}\Sigma_n$, and its two

variable restriction $\mathcal{B}\Sigma_n^2$. If needed the alphabet will be specified. For instance $\mathcal{F}[\sigma](B^*)$ will denote the set of languages of B^* definable by a formula of the fragment \mathcal{F} on the signature σ .

2.2 Enriched words

We now fix a positive integer d and an alphabet A . Let \mathbb{Z}_d be the cyclic group of order d .

Definition 1 (Enriched alphabet). We call the set $A_d = A \times \mathbb{Z}_d$ the enriched alphabet of A , and we denote by $\pi_d : A_d^* \rightarrow A^*$ the projection defined by $\pi_d(a, i) = a$ for each $(a, i) \in A_d$. For example, the word $(a, 2)(b, 1)(b, 2)(a, 0)$ is an enriched word of $abba$ for $d = 3$. We say that $abba$ is the underlying word of $(a, 2)(b, 1)(b, 2)(a, 0)$.

Definition 2 (Well-formed words). A word $(a_0, i_0)(a_1, i_1) \cdots (a_n, i_n)$ of A_d^* is well-formed if for $0 \leq j \leq n$, $i_j = j \pmod{d}$. We denote by K_d the set of all well-formed words of A_d^* .

Let $\alpha_d : A^* \rightarrow A_d^*$ be the function defined for any word $u = a_0 a_1 \cdots a_n \in A^*$ by $\alpha_d(u) = (a_0, 0)(a_1, 1) \cdots (a_n, n \pmod{d})$. The word $\alpha_d(u)$ is called the well-formed word attached to u .

Note that the restriction of π_d to the set of well-formed words is one-to-one. For instance, the enriched word $(a, 0)(b, 1)(b, 2)(a, 0)$ is a well-formed word for $d = 3$. It is the unique well-formed word having the word $abba$ as underlying word. The following lemma is an easy consequence of this observation.

2.3 Algebraic notions

We refer to [16] for the standard definitions of semigroup theory.

A *stamp* is a surjective monoid morphism from A^* onto a finite monoid. For a stamp $\varphi : A^* \rightarrow M$, the set $\varphi(A)$ is an element of the powerset monoid of M . As such it has an idempotent power. The *stability index* of a stamp is the least positive integer s such that $\varphi(A^s) = \varphi(A^{2s})$. This set forms a subsemigroup called the *stable semigroup* of φ . The set $\varphi((A^s)^*)$ is called the *stable monoid* of φ and the morphism from $(A^s)^*$ onto the stable monoid induced by φ is called the *stable stamp*. The stable monoid of a regular language is the stable monoid of its syntactic stamp.

A (pseudo) *variety of finite monoids* is a class of finite monoids closed under division and finite products.

We say that a morphism between finitely generated monoids is *length-preserving* if the image of each letter is a letter. Let \mathcal{C} be a class of morphisms between finitely generated free monoids closed under composition and containing the length-preserving morphisms. Examples include the morphisms between finitely generated free monoids (*all*), the *non-erasing* (*ne*) morphisms (morphisms for which the image of letters are non empty words) and the *length-multiplying* (*lm*)

morphisms (morphisms for which there is an integer k such that the image of each letter is a word of size k).

Given such a class of morphism \mathcal{C} , we recall the notion of \mathcal{C} -varieties of stamps. A stamp $\varphi : A^* \rightarrow M$ \mathcal{C} -divides another stamp $\psi : B^* \rightarrow N$ if and only if there exists a pair (α, β) such that α is a \mathcal{C} -morphism from A^* to B^* , $\beta : N \rightarrow M$ is a partial surjective monoid morphism and $\varphi = \beta \circ \psi \circ \alpha$. The pair (α, β) is called an \mathcal{C} -division. Then a \mathcal{C} -variety of stamps is a class of stamps closed under \mathcal{C} -division and finite product.

$$\begin{array}{ccc} A^* & \xrightarrow{\alpha} & B^* \\ \varphi \downarrow & & \downarrow \psi \\ M & \xleftarrow{\beta} & N \end{array}$$

Example 3. Given $d > 0$, let \mathbf{MOD}_d be the class of all stamps of the form $\pi_d : A^* \rightarrow \mathbb{Z}_d$ with $\pi_d(a) = \pi_d(b)$ for all letters a and b . Then \mathbf{MOD}_d is a lm -variety of stamps and the corresponding lm -variety of languages \mathbf{MOD}_d is the lm -variety generated by the languages $\{(A^d)^* A^i \mid 0 \leq i < d\}$. The class $\mathbf{MOD} = \bigcup_{d>0} \mathbf{MOD}_d$ is also a lm -variety of stamps.

Example 4. Let \mathbf{DA} be the variety of monoids satisfying the equation $(xy)^\omega = (xy)^\omega x (xy)^\omega$ where $x \mapsto x^\omega$ is the application that maps an element to its idempotent power. Alternatively \mathbf{DA} is the variety of monoids whose regular \mathcal{D} -classes are aperiodic semigroups. The corresponding variety of languages is the class of languages definable in $\mathbf{FO}^2[<]$, the two-variable first order logic [22]. When adding the local predicates we obtain a ne -variety of languages. The variety of stamps corresponding to $\mathbf{FO}^2[<, \text{LOC}]$ is \mathbf{LDA} , the class of stamps $\eta : A^* \rightarrow M$ such that for every idempotent e of the semigroup $\eta(A^+)$, the submonoid eMe is in \mathbf{DA} . For instance, the syntactic stamp of the language $(ab)^*$ is in \mathbf{LDA} but the syntactic stamp of the language $c^*(ce^*bc^*)^*$ is not.

3 Wreath product

3.1 Wreath Product Principle for MOD

The wreath product is an algebraic operation on monoids that specializes the semidirect product. This operation has been studied intensively in semigroup theory. The reader is referred to [18] for applications to languages. In logic, this operation often encodes the addition of some new predicates. In particular, for many cases, the $- * \mathbf{LI}$ operation corresponds to adding local predicates to a given signature. The rather technical definition of the wreath product is omitted. We will only use it through the following theorem, a consequence of the Wreath Product Principle for stamps presented in [5].

Theorem 5 (Wreath Product Principle for MOD [5]). *Let \mathbf{V} be a (ne) -variety, let \mathcal{V} be the corresponding (ne) -variety of languages, L a regular language of A^* and d a positive integer. Then the following properties are equivalent:*

- (1) The language L is recognized by a stamp in $\mathbf{V} * \mathbf{MOD}_d$,
- (2) The language L belongs to the lattice of languages generated by the languages of the form $(A^d)^* A^i$ for $i < d$ and of the form $\pi_d(L' \cap K_d)$ where $L' \in \mathcal{V}(A_d^*)$.

Furthermore, a language L is recognized by a stamp in $\mathbf{V} * \mathbf{MOD}$ if and only if there exists $d > 0$ such that L is recognized by a stamp in $\mathbf{V} * \mathbf{MOD}_d$.

The next theorem is the main result of this section.

Theorem 6. Let $\mathcal{F}[\sigma]$ be a fragment equivalent to variety $(ne)\text{-}\mathbf{V}$, L a regular language and d a positive integer. Then the following properties are equivalent:

- (1) L is definable by a formula of $\mathcal{F}[\sigma, \mathbf{MOD}_d]$,
- (2) η_L belongs to $\mathbf{V} * \mathbf{MOD}_d$,
- (3) there exist some languages L_0, \dots, L_{d-1} of $\mathcal{V}(A_d^*)$ such that:

$$L = \bigcup_{i=0}^{d-1} ((A^d)^* A^i \cap \pi_d(L_i \cap K_d)) \quad (\text{a})$$

Furthermore, a language L is definable in $\mathcal{F}[\sigma, \mathbf{MOD}]$ if and only if L is recognized by a stamp in $\mathbf{V} * \mathbf{MOD}$.

Proof. We only treat the case of a variety of monoids, since the proof for a variety is the same. (3) implies (2) follows from Theorem 5.

(2) implies (3). Assume that η_L belongs to $\mathbf{V} * \mathbf{MOD}$. By Theorem 5, we can suppose that L belongs to the lattice generated by languages of the form $(A^d)^* A^i$ for $i < d$ and $\pi_d(L' \cap K_d)$ with $L' \in \mathcal{V}(A_d^*)$. Recall that the lattice is distributive and the languages $(A^d)^* A^i$ form a partition of A^* . Therefore, there are languages H_0, \dots, H_{d-1} in the lattice of languages generated by $\pi_d(L' \cap K_d)$ with $L' \in \mathcal{V}(A_d^*)$ such that $L = \bigcup_{i=0}^{d-1} ((A^d)^* A^i \cap H_i)$. Thus, for $0 \leq i < d$, there exists a language $L_i \in \mathcal{V}(A_d^*)$ such that $H_i = \pi_d(L_i \cap K_d)$.

For the equivalence between (1) and (3) we need an auxiliary result which gives a decomposition of the language defined by a formula into smaller pieces.

Lemma 7. Let $\mathcal{F}[\sigma, \mathbf{MOD}]$ be a fragment of logic and φ a formula of $\mathcal{F}[\sigma, \mathbf{MOD}_d]$. Then there exists d formulae ψ_i of $\mathcal{F}[\sigma, \mathbf{MOD}_d]$ that do not contain any predicate D_j^d and such that $\varphi \equiv \bigvee_{i=0}^{d-1} (\psi_i \wedge D_i^d)$. Moreover, we have:

$$L(\varphi) = \bigcup_{i=0}^{d-1} ((A^d)^* A^i \cap L(\psi_i)).$$

The proof is omitted here. It relies on some elementary manipulations of formulae. We now conclude the proof of Theorem 6. Let φ be a formula of $\mathcal{F}[\sigma, \mathbf{MOD}]$. Then φ belongs to $\mathcal{F}[\sigma, \mathbf{MOD}_d]$ for some $d > 0$. Using Lemma 7, we know it is sufficient to consider a formula φ without any length predicate. We transform it into a formula ψ by replacing every predicate $\mathbf{MOD}_i^d(x)$ by $\bigvee_{a \in A} (\mathbf{a}, \mathbf{i})x$

and every predicate $\mathbf{a}x$ by $\bigvee_{0 \leq i < d} (\mathbf{a}, \mathbf{i})x$. The resulting formula ψ is in $\mathcal{F}[\sigma](A_d^*)$ and $L(\varphi) = \pi_d(L(\psi) \cap K_d)$. Conversely, we transform a formula ψ of $\mathcal{F}[\sigma](A_d^*)$ into a formula φ of $\mathcal{F}[\sigma, \text{MOD}_d]$ by replacing every predicate $(\mathbf{a}, \mathbf{i})x$ in ψ by $\mathbf{a}x \wedge \text{MOD}_i^d(x)$. We also get $L(\varphi) = \pi_d(L(\psi) \cap K_d)$. \square

The semidirect product does not necessarily preserve decidability (see e.g. [2]). The remainder of the article focuses on varieties of categories to obtain decidability results for this semidirect product.

4 Categories

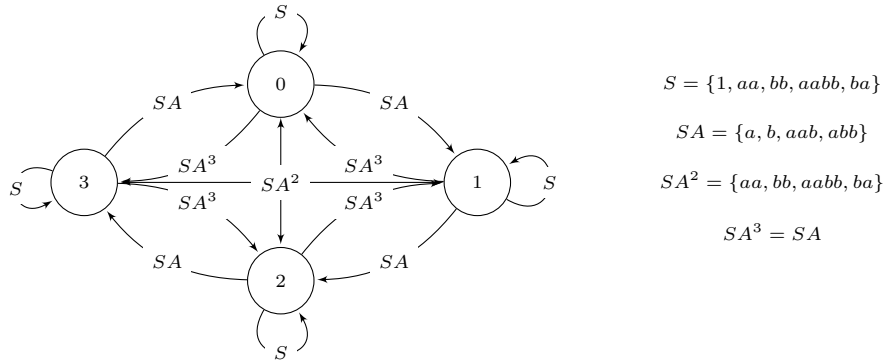
4.1 The Derived Category Theorem

A (small) *category* C is a set of objects denoted $\text{Ob}(C)$ equipped with a set of arrows between any pair of objects, with a composition law for consecutive arrows. A *loop* is an arrow whose initial object is the same as its final object. A category contains, for each object u , an identity denoted 1_u . Thus the set of loops around a given object, equipped with the composition law, forms a monoid, called the *local monoid* of that object. We refer to Tilson [24] for complete definitions. Here we only consider finite categories, seen as a generalization of finite monoids, since a monoid can be viewed as a one-object category. Here we give the definition of the derived category for **MOD** which is an adaptation of the one introduced by Tilson [24] and specialized for **MOD** in [5].

Definition 8. Let $\varphi : A^* \rightarrow M$ be a stamp and d an integer. The d -derived category of φ , denoted $C_d(\varphi)$, is the category with \mathbb{Z}_d as set of objects, and the arrows from i to j are the elements m of M such that there exists a word u satisfying $\varphi(u) = m$ and $i + |u| \equiv j \pmod{d}$. The d -derived category of a regular language L , denoted $C_d(L)$, is the category $C_d(\eta_L)$.

Lemma 9. Let d be a positive integer, and L be a regular language of stability index s . Then the local monoids of $C_d(L)$ are isomorphic to $\eta_L((A^d)^*)$. In particular, the local monoids of $C_s(L)$ are isomorphic to the stable monoid of L .

Example 10. The 4-derived category of the language $(aa)^*ab(bb)^*$ is given below. Let η be its syntactic morphism and S its stable monoid. Its stability index is 4.



To pursue the parallel with monoids, we recall the notion of division of categories, which extends the notion of division on monoids. First for a category C we denote by $\text{Ob}(C)$ the set of objects of C and by $C(u, v)$ the set of arrow between the objects u and v of C .

Let C, D be two categories. A *division* of categories $\tau : C \rightarrow D$ is given by a mapping $\tau : \text{Ob}(C) \rightarrow \text{Ob}(D)$, and for each pair of objects u and v , by a relation $\tau : C(u, v) \rightarrow D(\tau(u), \tau(v))$ such that

1. $\tau(x)\tau(y) \subseteq \tau(xy)$ for consecutive arrows x, y ,
2. $\tau(x) \neq \emptyset$ for any arrow x ,
3. $\tau(x) \cap \tau(y) \neq \emptyset$ implies $x = y$ if x and y are coterminial,
4. $1_{\tau(u)} \in \tau(1_u)$ for any object u of C .

One can see that this definition is exactly a generalization of a notion of division if we take, for instance, two categories C and D with only one object (ie monoids). Then C divides D in a sense of category if and only if C divide D in a sense of monoid. The *global* of a variety \mathbf{V} , denoted by \mathbf{gV} , is the variety of all categories that divide a monoid in \mathbf{V} , seen as a one-object category.

Proposition 11. *Let L be a regular language. For any $0 < d \leq d'$, if d divide d' , then $C_{d'}(L)$ divides $C_d(L)$.*

Proof. Let L be regular language and $0 \leq d < d'$ be integers such that d divides d' . We define the relation $\tau : C_{d'}(L) \rightarrow C_d(L)$. The object application $\text{Ob}(\tau) : \mathbb{Z}_{d'} \rightarrow \mathbb{Z}_d$ is defined by $\text{Ob}(\tau)(x) = x \bmod d$ for any $x \in \mathbb{Z}_{d'}$. Let (x, m, y) be an arrow of $C_{d'}(L)$. By definition, there exists $u \in A^*$ such that $\eta_L(u) = m$ and $|u| \equiv y - x \bmod d'$. Let $a = x \bmod d$ and $b = y \bmod d$. Then, since d divides d' , $|u| \equiv b - a \bmod d$. Thus, the arrow (a, m, b) is in $C_d(L)$. We define $\tau(x, m, y) = (a, m, b)$. The application τ is a morphism and for any $(x, m, y) \neq (x', m', y')$, we have $\tau(x, m, y) \neq \tau(x', m', y')$. Therefore, τ defines a division from $C_{d'}$ to C_d . \square

The derived category theorem was originally proved by Tilson [24] for varieties of monoids and semigroups. In [4], Chaubard extended this theorem to \mathcal{C} -varieties. Here we give the specialization to **MOD** of this latter generalization.

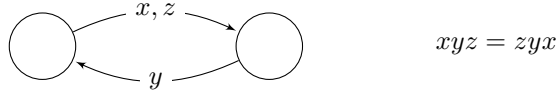
Theorem 12 (Derived Category's Theorem for MOD [4]). *Let \mathbf{V} be a variety and L a regular language. A language L has its syntactic stamp in $\mathbf{V} * \mathbf{MOD}$ if and only if there exists $d > 0$ such that $C_d(L)$ is in \mathbf{gV} .*

4.2 Rank of a variety

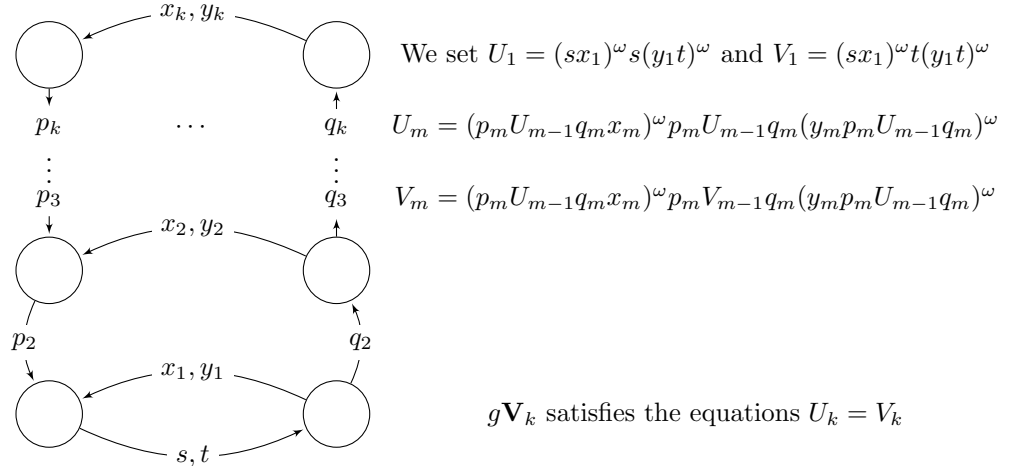
In the same way that varieties of monoids are characterized by profinite equations over alphabets, varieties of categories can be described by profinite equations on labelled graphs. Then a profinite path equation is given by a domain graph and the equality of two coterminial paths. As in the case of monoids, it is known [24] that varieties of categories are exactly the classes of categories satisfying sets of path equations.

Then given a variety of monoids \mathbf{V} , we say that \mathbf{V} is of *finite rank* if the variety of categories \mathbf{gV} can be described by paths equations over a finite graph. Furthermore, the *rank of \mathbf{V}* is defined as the minimal size of graphs describing \mathbf{gV} . We refer to Tilson [24] for complete definitions and proofs and we give here examples of varieties of finite rank below.

- Example 13.* 1. Varieties of rank 1 are called local. This is the case for several varieties. For instance, the variety of semilattice monoids $\mathbf{J}^1 = \llbracket xy = yx, x^2 = x \rrbracket$, the variety $\mathbf{DA} = \llbracket (xy)^\omega x(xy)^\omega = (xy)^\omega \rrbracket$ see Example 4, the variety of aperiodic monoids $\mathbf{A} = \llbracket x^\omega = x^{\omega+1} \rrbracket$.
2. The variety of commutative monoids $\mathbf{Com} = \llbracket xy = yx \rrbracket$. The variety of categories \mathbf{gCom} is defined by



3. A recent algebraic description of the languages definable by formulae of $\mathcal{BS}_{k+1}^2[<]$ was established in [8, 11]. In the subsequent we will denote by \mathbf{V}_k the equivalent variety of monoids. This result was extended to $\mathcal{BS}_{k+1}^2[<, +1]$ in [9]. From this latter result we derive the following description of \mathbf{gV}_k .



5 The local case

For any variety \mathbf{V} , we define \mathbf{QV} to be the lm -variety of stamps whose stable monoid is in \mathbf{V} . Following Tilson [24], we denote by $\ell\mathbf{V}$ the variety of categories whose local monoids are all in \mathbf{V} . The next theorem makes explicit the link between \mathbf{QV} and $\ell\mathbf{V}$.

Theorem 14. *Let \mathbf{V} be a variety and L a regular language of A^* of stability index s . The following properties are equivalent:*

- (1) L is recognized by a stamp in \mathbf{QV} ,
- (2) there exists an integer d such that $C_d(L)$ is in $\ell\mathbf{V}$,
- (3) $C_s(L)$ is in $\ell\mathbf{V}$.

Proof.

(1) \rightarrow (3). If L is recognized by a stamp in \mathbf{QV} , then its syntactic stamp is also in \mathbf{QV} and its stable monoid is in \mathbf{V} . But, thanks to Lemma 9, the local monoids of $C_s(L)$ belong to \mathbf{V} , and thus $C_s(L)$ is in $\ell\mathbf{V}$.

(3) \rightarrow (2). Is obvious.

(2) \rightarrow (1). Suppose that $C_d(L)$ is in $\ell\mathbf{V}$. Then the local monoids of $C_d(L)$, which are isomorphic to $\eta_L((A^d)^*)$ by Lemma 9, belong to \mathbf{V} . Thus $\eta_L((A^{ds})^*)$, which is a submonoid of $\eta_L((A^d)^*)$, also belongs to \mathbf{V} . Finally, by definition of the stability index, the monoid $\eta_L((A^s)^*) = \eta_L((A^{ds})^*)$ is in \mathbf{V} and thus η_L is in \mathbf{QV} . \square

Observe that any monoid of \mathbf{V} , viewed as a one-object category, belongs to $\ell\mathbf{V}$. Therefore by definition of \mathbf{gV} , any category of \mathbf{gV} divides a category of $\ell\mathbf{V}$, and thus $\mathbf{gV} \subseteq \ell\mathbf{V}$. The varieties satisfying $\mathbf{gV} = \ell\mathbf{V}$ are exactly the local varieties. Combining Theorem 12 and Theorem 14 yields the following theorem.

Theorem 15. *Let \mathbf{V} be a variety. Then $\mathbf{V} * \mathbf{MOD} \subseteq \mathbf{QV}$. If furthermore \mathbf{V} is local, then $\mathbf{V} * \mathbf{MOD} = \mathbf{QV}$.*

Since the stability index and the stable monoid of a given regular language are computable, one gets the following corollary.

Corollary 16. *Let $\mathcal{F}[\sigma]$ be a fragment equivalent to a local variety. Then $\mathcal{F}[\sigma]$ is decidable if and only if $\mathcal{F}[\sigma, \mathbf{MOD}]$ is decidable.*

Given a variety of monoids \mathbf{V} , we define \mathbf{LV} to be the variety of semigroups S such that for any idempotent e of S , the set eSe belongs to \mathbf{V} . It is known [24] that for local varieties, this operator is equal to the semidirect product by the variety \mathbf{LI} . The next corollary extends the previous result to \mathbf{LV} when only \mathbf{V} is known to be local. Recall that \mathbf{J}_1 denotes the variety of idempotent and commutative monoids.

Corollary 17. *Let \mathbf{V} be a local variety containing \mathbf{J}_1 . Then, $\mathbf{V} * \mathbf{LI} * \mathbf{MOD} = \mathbf{QLV}$.*

Remark 18. The equality $\mathbf{V} * \mathbf{MOD} = \mathbf{QV}$ does not always hold. A counterexample is the variety \mathbf{J} , which is known to be nonlocal. Chaubard, Pin and Straubing proved the decidability of $\mathbf{J} * \mathbf{MOD}$ [5], using the characterization of $g\mathbf{J}$ given by Knast in [7]. Using this characterization, we can prove that the language $(aa)^*ab(bb)^*$, whose stable monoid is in \mathbf{J} does not satisfy Knast's equation, proving that $\mathbf{J} * \mathbf{MOD} \subsetneq \mathbf{QJ}$ (see Example 10).

6 Main results

Theorem 6 gives a description of the languages definable in $\mathcal{F}[\sigma, \text{MOD}]$ together with a non effective algebraic characterization. What this characterization tells us is that the decidability problem can be divided into two parts. The first one consists in computing effectively this integer d , given the language L . We call it the *Delay problem* for MOD in reference to the Delay Theorem [19, 24] which solves a similar problem for the operation $\mathbf{V} \rightarrow \mathbf{V} * \mathbf{LI}$. The second problem is to find effectively some languages L_0, \dots, L_{d-1} occurring in Theorem 6 (a). Finding these languages can be reduced to the membership problem for \mathbf{gV} . We now state our main result, which gives a delay index assuming that the decidability and the rank of the variety \mathbf{gV} are known.

Theorem 19 (A Delay Theorem for MOD). *Let \mathbf{V} be a variety of rank k . Then a stamp with stability index s belongs to $\mathbf{V} * \text{MOD}$ if and only if it belongs to $\mathbf{V} * \text{MOD}_{ks}$.*

A proof of this Theorem can be found in the Appendix. The main idea of the proof is that when you consider the ks -derived category, you can ensure that the vertices from any projection from graphs of size k are far enough to use the main property of the stability index, that is words of length s can be pumped to any length multiple of s with the same behaviour regarding L . Thus projections from the graph description of \mathbf{gV} can be extended to any multiple of s greater than ks , allowing us to conclude.

Let us now deduce several corollaries from this result. First, since it is known that \mathbf{gV} is decidable if and only if $\mathbf{V} * \mathbf{LI}$ is decidable [2, 24], the next result follows from Theorems 12 and 19.

Corollary 20. *Let \mathbf{V} be a variety of finite rank. If $\mathbf{V} * \mathbf{LI}$ is decidable, then $\mathbf{V} * \text{MOD}$ is also decidable.*

Since the global of any decidable variety containing the syntactic stamp of $(ab)^*$ is known to be decidable [24], the following corollary is straightforward.

Corollary 21. *Let \mathbf{V} be a variety of finite rank that contains the syntactic stamp of the language $(ab)^*$. Then \mathbf{V} is decidable if and only if $\mathbf{V} * \text{MOD}$ is decidable.*

Another result with a very similar approach applies to variety of semigroups of the form $\mathbf{V} * \mathbf{LI}$.

Theorem 22. *Let \mathbf{V} be a variety of rank k containing the variety \mathbf{J}_1 . Then a stamp with stability index s belongs to $\mathbf{V} * \mathbf{LI} * \text{MOD}$ if and only if it belongs to $\mathbf{V} * \mathbf{LI} * \text{MOD}_{ks}$. Consequently, if \mathbf{V} is decidable then $\mathbf{V} * \mathbf{LI} * \text{MOD}$ is decidable*

The following table summarizes some effective algebraic characterizations obtained as consequence of our results.

	$\mathcal{B}\Sigma_1 = \mathcal{B}\Sigma_1^2$	$\mathcal{B}\Sigma_{k+1}^2$	\mathbf{FO}^2	\mathbf{FO}
[<]	J [15, 23]	\mathbf{V}_k [11, 8]	DA [22]	A [13, 17]
[<, LOC]	J * LI [7]	$\mathbf{V}_k * \mathbf{LI}$ [10]	LDA [22]	A [13, 17]
[<, MOD]	J * MOD Corollary 20 or [5]	$\mathbf{V}_k * \mathbf{MOD}$ Corollary 20 New	QDA Theorem 15 or [6]	QA Theorem 15 or [20, 3]
[<, LOC, MOD]	J * LI * MOD Theorem 22 or [12]	$\mathbf{V}_k * \mathbf{LI} * \mathbf{MOD}$ Theorem 22 New	QLDA Corollary 17 New	QA Theorem 15 or [20, 3]

Conclusion

We presented a study of the enrichment operation on logical fragments: $\mathcal{F}[\sigma] \rightarrow \mathcal{F}[\sigma, \mathbf{MOD}]$. For fragments equivalent to a variety, this operation exactly corresponds to the algebraic operation $\mathbf{V} \rightarrow \mathbf{V} * \mathbf{MOD}$. Our main result states that decidability can be obtained from the decidability through equations of a known variety of categories. This work subsumes several known results and leads to the decidability of new fragments. The main ingredients are the partial Delay Theorem for \mathbf{MOD} and a decision process for the global of \mathbf{V} . Both of them might be improved. Indeed, in the case of \mathbf{MOD} , the decidability of a weaker version of the global might be sufficient for the wreath product by \mathbf{MOD} . On the other hand our Delay Theorem require some more complex notions than the traditional variety of monoids. We know that our bounds are not always tight, as it is the case for the variety \mathbf{J} whose rank is 2. Moreover, the decidability remains open for varieties where little is known about \mathbf{gV} . An interesting case of study would be the variety generated by the syntactic monoid of the language $(ab)^*$, sometimes referred to as the *universal counterexample*. Indeed, this variety does not fall into the scope of any of our theorems.

Acknowledgements We would like to thank Olivier Carton for his helpful advices and Jean-Éric Pin for his time, commitment and tenacity during the genesis of this article.

References

1. J. ALMEIDA AND P. WEIL, Profinite categories and semidirect products, *J. Pure Appl. Algebra* **123**,1-3 (1998), 1–50.
2. K. AUINGER, On the decidability of membership in the global of a monoid pseudovariety, *Internat. J. Algebra Comput.* **20**,2 (2010), 181–188.
3. D. A. M. BARRINGTON, K. COMPTON, H. STRAUBING AND D. THÉRIEN, Regular languages in \mathbf{NC}^1 , *J. Comput. System Sci.* **44**,3 (1992), 478–499.
4. L. CHAUBARD, *Méthodes algébriques pour les langages formels*, PhD thesis, Université Paris Diderot.

5. L. CHAUBARD, J.-É. PIN AND H. STRAUBING, First order formulas with modular predicates, in *21st Annual IEEE Symposium on Logic in Computer Science (LICS 2006)*, pp. 211–220, IEEE, 2006.
6. L. DARTOIS AND C. PAPERMAN, Two-variable first order logic with modular predicates over words, in *International Symposium on Theoretical Aspects of Computer Science (STACS)*, N. Portier and T. Wilke (eds.), Dagstuhl, Germany, 2013, pp. 329–340, *LIPICs*, Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik.
7. R. KNAST, A semigroup characterization of dot-depth one languages, *RAIRO Inform. Théor.* **17**,4 (1983), 321–330.
8. A. KREBS AND H. STRAUBING, An effective characterization of the alternation hierarchy in two-variable logic, in *FSTTCS*, pp. 86–98, 2012.
9. M. KUFLEITNER AND A. LAUSER, Lattices of Logical Fragments over Words, in *ICALP (2)*, pp. 275–286, 2012.
10. M. KUFLEITNER AND A. LAUSER, Quantifier Alternation in Two-Variable First-Order Logic with Successor Is Decidable, in *International Symposium on Theoretical Aspects of Computer Science (STACS)*, pp. 305–316, 2013.
11. M. KUFLEITNER AND P. WEIL, The FO^2 alternation hierarchy is decidable, in *Computer science logic 2012*, pp. 426–439, *LIPICs. Leibniz Int. Proc. Inform.* vol. 16, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2012.
12. A. MACIEL, P. PÉLADEAU AND D. THÉRIEN, Programs over semigroups of dot-depth one, *Theoret. Comput. Sci.* **245**,1 (2000), 135–148. Semigroups and algebraic engineering (Fukushima, 1997).
13. R. MCNAUGHTON AND S. PAPERT, *Counter-free automata*, The M.I.T. Press, Cambridge, Mass.-London, 1971.
14. P. PÉLADEAU, Logically defined subsets of \mathbf{N}^k , *Theoret. Comput. Sci.* **93**,2 (1992), 169–183.
15. D. PERRIN AND J.-É. PIN, First-order logic and star-free sets, *J. Comput. System Sci.* **32**,3 (1986), 393–406.
16. J.-É. PIN, Syntactic semigroups, in *Handbook of formal languages, Vol. 1*, pp. 679–746, Springer, Berlin, 1997.
17. M. P. SCHÜTZENBERGER, On finite monoids having only trivial subgroups, *Information and Control* **8** (1965), 190–194.
18. H. STRAUBING, Families of recognizable sets corresponding to certain varieties of finite monoids, *J. Pure Appl. Algebra* **15**,3 (1979), 305–318.
19. H. STRAUBING, Finite semigroup varieties of the form $V * D$, *J. Pure Appl. Algebra* **36**,1 (1985), 53–94.
20. H. STRAUBING, *Finite automata, formal logic, and circuit complexity*, Birkhäuser Boston Inc., Boston, MA, 1994.
21. H. STRAUBING, On logical descriptions of regular languages, in *LATIN 2002: Theoretical informatics*, pp. 528–538, *Lect. Notes Comp. Sci.* vol. 2286, Springer, Berlin, 2002.
22. D. THÉRIEN AND T. WILKE, Over words, two variables are as powerful as one quantifier alternation, in *STOC '98 (Dallas, TX)*, pp. 234–240, ACM, New York, 1999.
23. W. THOMAS, Classifying regular events in symbolic logic, *J. Comput. System Sci.* **25**,3 (1982), 360–376.
24. B. TILSON, Categories as algebra: an essential ingredient in the theory of monoids, *J. Pure Appl. Algebra* **48**,1-2 (1987), 83–198.

7 Appendix

7.1 Wreath Product

Here we give the definition of the wreath product by **MOD**. The wreath product of a monoid M and the cyclic group of order d , \mathbb{Z}_d , is denoted by $M \circ \mathbb{Z}_d$ and is defined on the set $M^{\mathbb{Z}_d} \times \mathbb{Z}_d$ equipped with the following product:

$$(f, i)(g, j) = (f \cdot jg, i + j)$$

with \cdot being the pointwise product on $M^{\mathbb{Z}_d}$ and $jg : \mathbb{Z}_d \rightarrow M$ is defined by $jg(t) = g(t + j)$. Having define the wreath product of two monoids, we can now define the wreath product of (*ne*-)variety by **MOD**. Let \mathbf{V} be a (*ne*-)variety. A stamp $\eta : A^* \rightarrow M$ belongs to $\mathbf{V} * \mathbf{MOD}$ if and only if there exists $d > 0$ and $\mu : (B \times \mathbb{Z}_d)^* \rightarrow N$, with $\mu \in \mathbf{V}$ such that η *lm*-divides the stamp $\mu' : B^* \rightarrow N \circ \mathbb{Z}_d$, defined by $\mu'(b) = (f, 1)$ with $f(i) = \mu(b, i)$.

Lemma 9. *Let $\mathcal{F}[\sigma, \mathbf{MOD}]$ be a fragment of logic and φ a formula of $\mathcal{F}[\sigma, \mathbf{MOD}_d]$. Then there exists d formulae ψ_i of $\mathcal{F}[\sigma, \mathbf{MOD}_d]$ that do not contain any predicate D_j^d and such that $\varphi \equiv \bigvee_{i=0}^{d-1} (\psi_i \wedge D_i^d)$. Moreover, we have:*

$$L(\varphi) = \bigcup_{i=0}^{d-1} ((A^d)^* A^i \cap L(\psi_i)).$$

Proof. For $i < d$, we define the formula ψ_i to be the formula φ where we replaced every predicate D_i^d by *true* and every D_j^d with $j \neq i$ by *false*. One should notice that, by definition of a fragment [9], the formulae ψ_i are in $\mathcal{F}[\sigma, \mathbf{MOD}_d]$. We can conclude the proof since the formula (D_i^d) recognizes the language $(A^d)^* A^i$. \square

7.2 The local case

Corollary 17. *Let \mathbf{V} be a local variety containing \mathbf{J}_1 . Then, $\mathbf{V} * \mathbf{LI} * \mathbf{MOD} = \mathbf{QLV}$.*

Proof. This corollary is proved by considering the notion of locality adapted to the semigroups framework. The *ne*-variety \mathbf{LV} can be seen as a variety of semigroups. Therefore, one can use the notion of locality obtained in [1] for varieties of semigroups. In particular, we claim that the variety of semigroups \mathbf{LV} is local. A semigroupoid S belongs to $g\mathbf{LV}$ if and only if the consolidated semigroup S_{cd} belongs to \mathbf{LV} . Note that an idempotent of S_{cd} is an idempotent loop of S . Therefore S_{cd} belongs to \mathbf{LV} if and only if the local semigroup of S belongs to \mathbf{LV} . Since this is exactly the definition of a local variety, this show that \mathbf{LV} is a local variety of semigroups. It is tempting to use the version of Theorem 15 for the *ne*-variety. Unfortunately, it is not clear that the locality of a variety of semigroups is equivalent to the locality of the corresponding *ne*-variety. However, the proof of Theorem 15 is easily adaptable to varieties of semigroups.

7.3 Main Results

Theorem 22. *Let \mathbf{V} be a variety of rank k containing the variety \mathbf{J}_1 . Then a stamp with stability index s belongs to $\mathbf{V} * \mathbf{LI} * \mathbf{MOD}$ if and only if it belongs to $\mathbf{V} * \mathbf{LI} * \mathbf{MOD}_{ks}$.*

The proof stems from results from [24]. We first introduce two definitions from [24].

Definition 23 (Consolidated semigroup). *Let C be a finite category and $\text{Arr}(C)$ the set of arrows of C . We denote by C_{cd} the semigroup defined on the set*

$$E = \text{Arr}(C) \cup \{0\}$$

with for any $x \in E$, $0x = x0 = 0$, and for $x, y \in \text{Arr}(C)$,

$$x.y = \begin{cases} xy & \text{if } x \text{ and } y \text{ are consecutives arrows,} \\ 0 & \text{otherwise.} \end{cases}$$

Definition 24 (Idempotent category). *Let S be a semigroup. We denote by S_E the finite category whom object are idempotents of S and arrows from e to f the set eSf .*

The next proposition is stated in [24, Proposition 16.1]

Proposition 25. *Let C be a finite category. Then $(C_{cd})_E$ and C divide each other.*

The next Theorem corresponds to [24, Theorem 17.1]

Theorem 26 (Delay Theorem). *Let S be a semigroup and \mathbf{V} a non trivial variety of monoids. The semigroup S belongs to $\mathbf{V} * \mathbf{LI}$ if and only if S_E belongs to \mathbf{gV} .*

We now conclude the proof of Theorem 22. Let \mathcal{W} be the ne -variety of languages equivalent to $\mathbf{V} * \mathbf{LI}$. First remark that for any alphabet A and any $a \in A$, the syntactic stamps of the languages A^*a and, for any d , K_d belong to $\mathbf{J}_1 * \mathbf{LI}$. Therefore, they belong to $\mathbf{V} * \mathbf{LI}$. In this case, item 3 of Theorem 6 simplifies as the union can be transferred inside the π_d morphism. Then the syntactic stamp of the language L belongs to $\mathbf{V} * \mathbf{LI} * \mathbf{MOD}$ if and only if there exists $d > 0$ such that

$$L_d = \pi_d^{-1}(L) \cap K_d \in \mathcal{W}(A_d^*).$$

We denote by $\eta_d : A_d^+ \rightarrow S_d(L)$ the syntactic stamp of L_d .

Lemma 27. *The semigroup $S_d(L)$ divides $C_d(L)_{cd}$ and $C_d(L)_{cd}$ divides a product of $S_d(L)$. Furthermore, $S_d(L)$ belongs to $\mathbf{V} * \mathbf{LI}$ if and only if $C_d(L)_{cd}$ belongs to $\mathbf{V} * \mathbf{LI}$.*

Proof. We set $\eta : A_d^* \rightarrow C_d(L)_{cd}^1$ with $\eta(a, i) = (i, \eta_L(a), i + 1)$. Then, $L_d = \eta^{-1}(P)$ where

$$P = \{(0, m, j) \mid m \in \eta_L(L) \text{ and } j \in \mathbb{Z}_d\}.$$

Therefore $C_d(L)_{cd}$ recognizes L_d and consequently $S_d(L)$ divides $C_d(L)_{cd}$. We denote by \sim the syntactic congruence of P on $C_d(L)_{cd}$. It appears that $C_d(L)_{cd}/\sim$ is isomorphic to $S_d(L)$. Furthermore, for any $m, m' \in M$ and $j \in \mathbb{Z}_d$, $(0, m, j) \sim (0, m', j)$ implies $m = m'$. Finally $C_d(L)_{cd}$ is isomorphic to the sub-semigroup of $(C_d(L)_{cd}/\sim)^d$ containing the elements

$$((i, m, j), (i + 1, m, j + 1), \dots, (i + d - 1, m, j + d - 1)).$$

□

The syntactic stamp of the language L belongs to $\mathbf{V} * \mathbf{LI} * \mathbf{MOD}$ if and only if there exists $d > 0$ such that $C_d(L)_{cd}$ is in $\mathbf{V} * \mathbf{LI}$. By Theorem 26, for all $d > 0$, $C_d(L)_{cd}$ is in $\mathbf{V} * \mathbf{LI}$ if and only if $(C_d(L)_{cd})_E \in \mathbf{gV}$. By Proposition 25, for all $d > 0$, $(C_d(L)_{cd})_E$ is in \mathbf{gV} if and only if $C_d(L)$ is in \mathbf{gV} . Therefore, the language L belongs to $\mathbf{V} * \mathbf{LI} * \mathbf{MOD}$ if and only if there exists $d > 0$ such that $C_d(L)$ is in \mathbf{gV} . By Theorem 12 and Theorem 19, since \mathbf{V} has rank k , there exists $d > 0$ such that $C_d(L)$ is in \mathbf{gV} if and only if $C_{ks}(L) \in \mathbf{gV}$, which concludes the proof.

7.4 Proof of the main result

First notice that since the if condition is trivial, we only need to prove the only if implication. Remark now that if $\text{rank}(\mathbf{V}) = 1$ then the variety is local and we know that we can restrict to congruence modulo the stability index. For the rest of the proof we assume that $\text{rank}(\mathbf{V}) = k > 1$.

Let now d be such that $C_d(L) \in \mathbf{gV}$. Without loss of generality, we assume that d is greater than k . Indeed if $d \leq k$, we consider $d' = dk$. Then by Proposition 11 $C_{d'}(L)$ divides $C_d(L)$ and thus also belongs to \mathbf{gV} .

So in the remainder of the proof we will assume that $ds > ks$. Since $C_d(L) \in \mathbf{gV}$ we know that $C_{ds}(L) \in \mathbf{gV}$. Then $C_{ds}(L)$ satisfies every path equation $(X, u = v)$ defining \mathbf{gV} . The goal of this proof is to show if $C_{ks}(L)$ does not satisfy a path equation defining \mathbf{gV} , then $C_{ds}(L)$ cannot satisfy it either.

So assume that there exists a path equation $(X, u = v)$ of rank k defining \mathbf{gV} that is not satisfied by $C_{ks}(L)$. Then, there exists a category morphism $\varphi : X^* \rightarrow C_{ks}(L)$ such that $\widehat{\varphi}(u) \neq \widehat{\varphi}(v)$. We define $V = \varphi(\text{Ob}(X))$ the set of objects of $C_{ks}(L)$ that has a preimage by φ , and $E = \{(i, m, j) \in C_{ks}(L) \mid \exists e \in X \varphi(e) = (i, m, j)\}$ the set of arrows that have a preimage by φ . Notice that $E \subseteq V \times M_L \times V$.

We will construct a category morphism $\psi : X^* \rightarrow C_{ds}(L)$ such that $\widehat{\psi}(u) \neq \widehat{\psi}(v)$. In order to do that, we define a map $\theta : V \rightarrow C_{ds}(L)$ such that for all (i, m, j) in E , $(\theta(i), m, \theta(j))$ is an arrow of $C_{ds}(L)$.

Lemma 28. *There exists a smallest integer $i_V < ks$ such that $\{i_V + 1, \dots, i_V + s - 1 \pmod{ks}\} \cap V = \emptyset$.*

Proof. As the size of X is k , the size of V is at most k . Then the maximal distance between two consecutive vertices of V is at least $ks/k = s$. \square

We define $\theta : V \rightarrow \text{Ob}(C_{ds}(L))$ as follow :

$$\theta : \begin{cases} i \mapsto i \bmod ds & \text{if } i \leq i_V \\ i \mapsto ds + i - ks & \text{otherwise.} \end{cases}$$

The idea behind this is that $i = \theta(i)$ if i appears before the gap and $ks - i = ds - \theta(i)$ if i appears after it. Then each arrow from E will either appear directly as it does for $C_{ks}(L)$ if it does not go over the gap, and since the gap is of size s , we will be able to pump the arrows that go over it.

Lemma 29. *For any arrow (i, m, j) of E , $(\theta(i), m, \theta(j))$ is an arrow of $C_{ds}(L)$.*

Proof. Let (i, m, j) be an arrow of E . Then there exists a word u such that $\eta_L(u) = m$ and $i + |u| = j \bmod ks$. We now distinguish the cases depending on the length of u .

- If $|u| \geq s$, then we know, by definition of the stability index, that for any positive integer ℓ , there exists a word u_ℓ such that $\ell s \leq |u_\ell| < (\ell + 1)s$, $|u| = |u_\ell| \bmod s$ and $u \equiv_L u_\ell$. Then as θ preserves the congruence modulo s , $(\theta(i), \eta_L(u_\ell), \theta(j)) = (\theta(i), m, \theta(j))$ is an arrow of $C_{ds}(L)$.
- If $|u| < s$, then we have to treat several subcases:
 - If $\theta(i) = i$ and $\theta(j) = j$, then $\theta(i) + |u| = \theta(j) \bmod ds$. Thus $(\theta(i), m, \theta(j))$ is an arrow of $C_{ds}(L)$.
 - If $\theta(i) = ds + i - ks$ and $\theta(j) = ds + j - ks$, then as u has a size smaller than s , we have $i < j$ and $\theta(j) - \theta(i) = j - i$. Consequently $\theta(i) + |u| = \theta(j) \bmod ds$ and $(\theta(i), m, \theta(j))$ is an arrow of $C_{ds}(L)$.
 - If $\theta(i) = ds + i - ks$ and $\theta(j) = j$, then $i + |u| = j + ks$. So $\theta(i) + |u| = ds + i - ks + |u| = j + ds$. The same word u labels an arrow from $\theta(i)$ to j and thus $(\theta(i), m, \theta(j))$ is an arrow of $C_{ds}(L)$.
 - Finally, the case where $\theta(i) = i$ and $\theta(j) = ds + j - ks$ cannot happen since it implies that $i \leq i_V$ and $j > i_V + s$, and that $|u| = j - i > s \bmod ks$ which contradicts the $|u| < s$ hypothesis.

\square

We now define a new morphism $\psi : X^* \rightarrow C_{ds}$. We proceed as follow:

- First we define $\text{Ob}(\psi)$ to be $\theta \circ \text{Ob}(\varphi)$.
- We now have to define ψ on arrows. Let e be an arrow of X and $\varphi(e) = (i, m, j)$. We set $\psi(x) = (\theta(i), m, \theta(j))$. This is well defined thanks to Lemma 29.

Lemma 30. *Let u be a path in X^* . If $\varphi(u) = (i, m, j)$, then $\psi(u) = (\theta(i), m, \theta(j))$.*

Proof. Let $u = u_1 \cdots u_n \in X^*$ such that $\varphi(u_\ell) = (i_\ell, m_\ell, j_\ell)$ and $\varphi(u) = (i, m, j)$. Therefore, $\psi(u_\ell) = (\theta(i_\ell), m_\ell, \theta(j_\ell))$. However, since for all $1 \leq \ell < n$ $j_\ell = i_{\ell+1}$, we have $\psi(u) = (i_1, m_1 \cdots m_n, j_n) = (i, m, j)$ and $\psi(u) = (\theta(i_1), m_1 \cdots m_n, \theta(j_n)) = (\theta(i), m, \theta(j))$. \square

Recall that $\widehat{\varphi}(u) \neq \widehat{\varphi}(v)$. Then we can find $u', v' \in X^*$ co-terminal paths of X^* such that $\varphi(u') = \widehat{\varphi}(u)$, $\varphi(v') = \widehat{\varphi}(v)$, $\psi(u') = \widehat{\psi}(u)$ and $\psi(v') = \widehat{\psi}(v)$. We set $u' = u_1 \cdots u_n$ with $u_i \in X$ for any i and $v' = v_1 \cdots v_p$ with $v_i \in X$ for any i . To conclude we show that $\psi(u') \neq \psi(v')$ which is absurd. Indeed, if $\varphi(u') = (i, m, j) \in C_{ks}(L)$ and $\varphi(v') = (i, m', j) \in C_{ks}(L)$, then $m' \neq m$ since $\widehat{\varphi}$ separates u and v . Furthermore, by Lemma 30, we also have $\psi(u') = (\theta(i), m, \theta(j))$ and $\psi(v') = (\theta(i), m', \theta(j))$ in $C_{sd}(L)$. Finally $\psi(u') \neq \psi(v')$ and thus C_{ds} does not satisfy $(X, u = v)$, holding a contradiction.