



**HAL**  
open science

# Closure Properties of Locally Finite Omega Languages

Olivier Finkel

► **To cite this version:**

Olivier Finkel. Closure Properties of Locally Finite Omega Languages. Theoretical Computer Science, 2004, 322 (1), pp.69-84. hal-00104203v1

**HAL Id: hal-00104203**

**<https://hal.science/hal-00104203v1>**

Submitted on 6 Oct 2006 (v1), last revised 12 Mar 2008 (v2)

**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.

# CLOSURE PROPERTIES OF LOCALLY FINITE $\omega$ -LANGUAGES<sup>\*</sup>

Olivier Finkel<sup>\*</sup>

*Equipe de Logique Mathématique  
CNRS et Université Paris 7, U.F.R. de Mathématiques  
2 Place Jussieu 75251 Paris cedex 05, France.*

**Dedicated to Denis Richard for his 60 th Birthday**

---

## Abstract

Locally finite  $\omega$ -languages, defined via second order quantifications followed by a first order locally finite sentence, were introduced by Ressayre in [Res88]. They enjoy very nice properties and extend  $\omega$ -languages accepted by finite automata or defined by monadic second order sentences. We study here closure properties of the family  $LOC_\omega$  of locally finite omega languages. In particular we show that the class  $LOC_\omega$  is neither closed under intersection nor under complementation, giving an answer to a question of Ressayre [Res89].

*Key words:* Formal languages; logical definability; infinite words; locally finite languages; closure properties.

---

## 1 Introduction

In the sixties J.R. Büchi was the first to study  $\omega$ -languages recognized by finite automata in order to prove the decidability of the monadic second order theory of one successor over the integers [Büc62]. In the course of his proof he showed that an  $\omega$ -language, i.e. a set of infinite words over a finite alphabet, is accepted by a finite automaton with the now called Büchi acceptance condition if and only if it is defined by an (existential) monadic second order sentence.

---

<sup>\*</sup> Partially supported by Intas 00-447.

<sup>\*</sup> Corresponding author

*Email address:* `finkel@logique.jussieu.fr` (Olivier Finkel).

Algorithms have been found to give such an automaton from the monadic second order sentence and conversely. Thus the above cited decision problem is reduced to the decidability of the emptiness problem for Büchi automata which is easily shown to be decidable. The equivalence between definability by monadic second order sentences and acceptance by finite automata holds also for languages of finite words [Büc60], and has been extended to languages of words of length  $\alpha$ , where  $\alpha$  is a countable ordinal  $\geq \omega$  [BS73].

The research area, now called “descriptive complexity”, found its origin in the above cited work of Büchi as well as in the fundamental result of Fagin who proved that the class **NP** is characterized by existential second order formulas, [Fag74]. Since then, a lot of work has been achieved about the logical definability of classes of formal languages of finite or infinite words, or of relational structures like graphs, see [Fag93] [Pin96] [Tho96] [Imm99] for a survey about this field of research.

Several extensions of existential monadic second order logic over words have been studied.

Lauteman, Schwentick and Therien proved that context free languages are characterized by existential second order formulas in which the second order quantifiers bear only on matchings, i.e. pairing relations without crossover, [LST94].

Parigot and Pelz, and more recently Yamasaki, extended monadic second order logic with two second order relation symbols and characterized classes of Petri net ( $\omega$ )-languages [PP85] [Pel87] [Yam99].

Eiter, Gottlob and Gurevich studied the relationship between monadic second order logic and syntactic fragments of existential second order logic over (finite) words [EGG00]. Distinguishing prefix classes, they determined which of them define only regular languages and which of them have the same expressive power as monadic second order logic.

Another extension has been introduced by Ressayre, in order to apply some powerful tools of model theory to the study of formal languages, [Res88]. He defined locally finite sentences (firstly called local). A locally finite sentence  $\varphi$  is a first order sentence which is equivalent to a universal one and whose models satisfy simple structural properties: closure under functions takes a finite number  $n_\varphi$  of steps.

These syntactic and semantic restrictions allow a meaningful use of the notion of indiscernables and lead to beautiful stretching theorems connecting the existence of some well ordered infinite models of  $\varphi$  with the existence of some finite models generated by indiscernables [FR96].

Locally finite languages are defined by second order formulas in the form

$\exists \bar{R} \exists \bar{f} \varphi$  where  $\varphi$  is a locally finite sentence and  $\bar{R}$  (respectively,  $\bar{f}$ ) represent the relation (respectively, function) symbols in the signature of  $\varphi$ .

These second order quantifications are much more general than the monadic ones as the following results show:

- Each regular language is locally finite, [Res88], and many context free as well as non context free languages are locally finite [Fin01].
- Each regular  $\omega$ -language is a locally finite  $\omega$ -language, [Fin01] [Fin89], and there exist many more locally finite  $\omega$ -languages as we shall see below.
- This result is extended to languages of transfinite length words: if  $\alpha$  is an ordinal  $< \omega^\omega$ , each regular  $\alpha$ -language is also locally finite [Fin01].

But a pumping lemma, following from a stretching theorem, makes locally finite  $\omega$ -languages keep important properties of regular  $\omega$ -languages, [Res88] [FR96]. It is an analogue for each locally finite  $\omega$ -language of the property:

“ A regular  $\omega$ -language is non empty if and only if it contains an ultimately periodic word ”.

This lemma implies in a similar manner the decidability of the emptiness problem for locally finite  $\omega$ -languages. Moreover for each countable ordinal  $\alpha < \omega^\omega$ , the decidability of the emptiness problem for locally finite  $\alpha$ -languages follows from similar arguments, [FR96].

Other decidability results, as the decidability of the problem: “is a given finitary locally finite language infinite?” follow from stretching theorems of [Res88][FR96].

These interesting properties of locally finite languages naturally lead to the question of the richness of the class of locally finite languages: how large is this class? What are its closure properties?

The study of locally finite languages of finite words was begun by Ressayre in [Res88] and continued in [Fin01]. We focus in this paper on the class  $LOC_\omega$  of locally finite  $\omega$ -languages and study classical closure properties for this class. In particular, we show that  $LOC_\omega$  is neither closed under intersection, nor under complementation. The proof uses the notion of rational cone of finitary languages which is important in formal language theory and the notion of indiscernables in a structure, often used in model theory.

This gives an answer to a question of Ressayre, [Res89]. Of course we would have preferred a positive answer to this question which would have provided a useful class of sentences for specification and verification of properties of non-terminating systems. But this leaves still open, for further study, the possibility to find such a useful class of sentences as a subclass of the class of locally finite sentences.

In section 2, we give the first definitions and some examples of locally finite

$\omega$ -languages. In section 3, closure properties for  $\omega$ -languages are investigated. We show that the class  $LOC_\omega$  is not closed under intersection with regular  $\omega$ -languages thus  $LOC_\omega$  is neither closed under intersection, nor under complementation (because  $LOC_\omega$  is closed under union). Then we prove that  $LOC_\omega$  is closed under  $\lambda$ -free morphism and  $\lambda$ -free substitution of locally finite (finitary) languages.

## 2 Definitions and examples

### 2.1 Definitions

We briefly indicate now some basic facts about first order logic and model theory. See for example [CK73] for more background on this subject.

We consider here formulas of first order logic. The language of first order logic contains (first order) variables  $x, y, z, \dots$  ranging over elements of a structure, logical symbols: the connectives  $\wedge$  (and),  $\vee$  (or),  $\rightarrow$  (implication),  $\neg$  (negation), and the quantifiers  $\forall$  (for all), and  $\exists$  (there exists), and also the binary predicate symbol of identity  $=$ .

A signature is a set of constant, relation ( different from  $=$  ) and function symbols. we shall consider here only finite signatures.

Let  $Sig$  be a finite signature. We define firstly the set of terms in the signature  $Sig$  which is built inductively as follows:

- (1) A variable is a term.
- (2) A constant symbol is a term.
- (3) If  $F$  is a  $m$ -ary function symbol and  $t_1, t_2, \dots, t_m$  are terms, then  $F(t_1, \dots, t_m)$  is a term.

We then define the set of atomic formulas which are in the form given below:

- (1) If  $t_1$  and  $t_2$  are terms, then  $t_1 = t_2$  is an atomic formula.
- (2) If  $t_1, t_2, \dots, t_m$  are terms and  $R$  is a  $m$ -ary relation symbol, then  $R(t_1, \dots, t_m)$  is an atomic formula.

Finally the set of formulas is built inductively from atomic formulas as follows:

- (1) An atomic formula is a formula.
- (2) If  $\varphi$  and  $\psi$  are formulas, then  $\varphi \wedge \psi$ ,  $\varphi \vee \psi$ ,  $\varphi \rightarrow \psi$  and  $\neg\varphi$  are formulas.
- (3) If  $x$  is a variable and  $\varphi$  is a formula, then  $\forall x\varphi$  and  $\exists x\varphi$  are formulas.

An open formula is a formula with no quantifier.

We assume the reader to know the notion of free and bound occurrences of a variable in a formula. Then a sentence is a formula with no free variable. A sentence in prenex normal form is in the form  $\varphi = Q_1x_1 \dots Q_nx_n\varphi_0(x_1, \dots, x_n)$ , where each  $Q_i$  is either the quantifier  $\forall$  or the quantifier  $\exists$  and the formula  $\varphi_0$  is an open formula.

It is well known that every sentence is equivalent to a sentence written in prenex normal form.

A sentence is said to be universal if it is in prenex normal form and each quantifier is the universal quantifier  $\forall$ .

We then recall the notion of a structure in a signature  $Sig$ : A structure is in the form:

$$M = (|M|, (a^M)_{a \in Sig})$$

Where  $|M|$  is a set called the universe of the structure, and for  $a \in Sig$ ,  $a^M$  is the interpretation of  $a$  in  $M$ :

If  $f$  is a  $m$ -ary function symbol in  $Sig$ , then  $f^M$  is a function:  $M^m \rightarrow M$ .

If  $R$  is a  $m$ -ary relation symbol in  $Sig$ , then  $R^M$  is a relation:  $R^M \subseteq M^m$ .

If  $a$  is a constant symbol in  $Sig$ , then  $a^M$  is a distinguished element in  $M$ .

In order to simplify the notations we shall sometimes write  $a$  instead of  $a^M$  when the meaning is clear from the context.

When  $M$  is a structure and  $\varphi$  is a sentence in the same signature  $Sig$ , we write  $M \models \varphi$  for “ $M$  is a model of  $\varphi$ ”, which means that  $\varphi$  is satisfied in the structure  $M$ . A detailed exposition of these notions may be found in [CK73].

When  $M$  is a structure in the signature  $Sig$  and  $Sig_1$  is another signature such that  $Sig_1 \subseteq Sig$ , then the reduction of  $M$  to the signature  $Sig_1$  is denoted  $M|_{Sig_1}$ . It is a structure in the signature  $Sig_1$  which has same universe  $|M|$  as  $M$ , and the same interpretations for symbols in  $Sig_1$ . Conversely an expansion of a structure  $M$  in the signature  $Sig_1$  to a structure in the signature  $Sig$  has same universe as  $M$  and same interpretations for symbols in  $Sig_1$ .

When  $M$  is a structure in a signature  $Sig$  and  $X \subseteq |M|$ , we define:

$$cl^1(X, M) = X \cup \bigcup_{\{f \text{ n-ary function of } Sig\}} f^M(X^n) \cup \bigcup_{\{a \text{ constant of } Sig\}} a^M$$

$$cl^{n+1}(X, M) = cl^1(cl^n(X, M), M) \quad \text{for an integer } n \geq 1$$

and  $cl(X, M) = \bigcup_{n \geq 1} cl^n(X, M)$  is the closure of  $X$  in  $M$ .

Let us now define locally finite sentences. We shall denote  $S(\varphi)$  the signature of a first order sentence  $\varphi$ , i.e. the set of non logical symbols appearing in  $\varphi$ .

**Definition 2.1** *A first order sentence  $\varphi$  is locally finite if and only if:*

- a)  $M \models \varphi$  and  $X \subseteq |M|$  imply  $cl(X, M) \models \varphi$
- b)  $\exists n \in \mathbb{N}$  such that  $\forall M$ , if  $M \models \varphi$  and  $X \subseteq |M|$ , then  $cl(X, M) = cl^n(X, M)$  (closure in models of  $\varphi$  takes less than  $n$  steps).

**Notation.** For a locally finite sentence  $\varphi$ , we shall denote by  $n_\varphi$  the smallest integer  $n \geq 1$  satisfying b) of the above definition.

**Remark 2.2** *Because of a) of Definition 2.1, a locally finite sentence  $\varphi$  is always equivalent to a universal sentence.*

We now introduce some notations for finite or infinite words.

Let  $\Sigma$  be a finite alphabet whose elements are called letters. A finite word over  $\Sigma$  is a finite sequence of letters:  $x = a_0 \dots a_n$  where  $\forall i \in [0; n] a_i \in \Sigma$ . We shall denote  $x(i) = a_i$  the  $i + 1^{th}$  letter of  $x$  and  $x[i] = x(0) \dots x(i)$  for  $i \leq n$ . The length of  $x$  is  $|x| = n + 1$ . The empty word will be denoted by  $\lambda$  and has 0 letter. Its length is 0. The set of finite words over  $\Sigma$  is denoted  $\Sigma^*$ .  $\Sigma^+ = \Sigma^* - \{\lambda\}$  is the set of non-empty words over  $\Sigma$ . A (finitary) language  $L$  over  $\Sigma$  is a subset of  $\Sigma^*$ . Its complement (in  $\Sigma^*$ ) is  $L^- = \Sigma^* - L$ . The usual concatenation product of  $u$  and  $v$  will be denoted by  $u.v$  or just  $uv$ . The set of non negative integers is denoted by  $\mathbb{N}$ . For  $V \subseteq \Sigma^*$ , we denote  $V^* = \{v_1 \dots v_n \mid \forall i \in [1; n] v_i \in V\} \cup \{\lambda\}$ .

The first infinite ordinal is  $\omega$ .

An  $\omega$ -word over  $\Sigma$  is an  $\omega$ -sequence  $a_0 a_1 \dots a_n \dots$ , where  $\forall i \geq 0 a_i \in \Sigma$ .

For  $\sigma \in \Sigma^\omega$ ,  $\sigma(n)$  is the  $n + 1^{th}$  letter of  $\sigma$  and  $\sigma[n] = \sigma(0)\sigma(1) \dots \sigma(n)$ .

The set of  $\omega$ -words over the alphabet  $\Sigma$  is denoted by  $\Sigma^\omega$ . An  $\omega$ -language over  $\Sigma$  is a subset of  $\Sigma^\omega$ . The  $\omega$ -power of a finitary language  $V \subseteq \Sigma^*$  is the  $\omega$ -language  $V^\omega = \{\sigma = u_1 \dots u_n \dots \in \Sigma^\omega \mid \forall i \geq 1 u_i \in V\}$ .

For a subset  $A \subseteq \Sigma^\omega$ , the complement of  $A$  (in  $\Sigma^\omega$ ) is  $\Sigma^\omega - A$  denoted  $A^-$ .

The concatenation product is extended to the product of a finite word  $u$  and an  $\omega$ -word  $v$ : the infinite word  $u.v$  is then the  $\omega$ -word satisfying:

$(u.v)(k) = u(k)$  if  $k < |u|$ , and  $(u.v)(k) = v(k - |u|)$  if  $k \geq |u|$ .

A word over  $\Sigma$  may be considered as a structure in the usual manner:

Let  $\Sigma$  be a finite alphabet. For each letter  $a \in \Sigma$  we denote  $P_a$  a unary predicate and  $\Lambda_\Sigma$  the signature  $\{<, (P_a)_{a \in \Sigma}\}$ . The length  $|\sigma|$  of a non-empty finite word  $\sigma \in \Sigma^*$  may be written  $|\sigma| = \{0, 1, \dots, |\sigma| - 1\}$ .  $\sigma$  is identified to the structure  $(|\sigma|, <^\sigma, (P_a^\sigma)_{a \in \Sigma})$  of signature  $\Lambda_\Sigma$  where  $P_a^\sigma = \{i < |\sigma| \mid \text{the } i + 1^{th} \text{ letter of } \sigma \text{ is an } a\}$ .

In a similar manner if  $\sigma$  is an  $\omega$ -word over the alphabet  $\Sigma$ , then  $\omega$  is the length of the word  $\sigma$  and we may write  $|\sigma| = \omega = \{0, 1, 2, 3, \dots\}$ .  $\sigma$  is identified to the structure  $(\omega, <^\sigma, (P_a^\sigma)_{a \in \Sigma})$  of signature  $\Lambda_\Sigma$  where  $P_a^\sigma = \{i < \omega \mid \text{the } i + 1^{th} \text{ letter of } \sigma \text{ is an } a\}$ .

**Definition 2.3** *Let  $\Sigma$  be a finite alphabet and  $L \subseteq \Sigma^*$ .*

*Then  $[L \text{ is a locally finite language}] \iff$*

*there exists a locally finite sentence  $\varphi$  in a signature  $\Lambda \supseteq \Lambda_\Sigma$  such that  $\sigma \in L$*

iff  $\exists M, M \models \varphi$  and  $M|_{\Lambda_\Sigma} = \sigma$ ].

We then denote  $L = L^\Sigma(\varphi)$  and, if there is no ambiguity,  $L = L(\varphi)$  the locally finite language defined by  $\varphi$ .

The class of locally finite languages will be denoted  $LOC$ .

The empty word  $\lambda$  has 0 letters. It is represented by the empty structure. Recall that if  $L(\varphi)$  is a locally finite language then  $L(\varphi) - \{\lambda\}$  and  $L(\varphi) \cup \{\lambda\}$  are also locally finite [Fin01].

**Definition 2.4** Let  $\Sigma$  be a finite alphabet and  $L \subseteq \Sigma^\omega$ .

Then  $[L \text{ is a locally finite } \omega\text{-language}] \iff [ \text{there exists a locally finite sentence } \varphi \text{ in a signature } \Lambda \supseteq \Lambda_\Sigma \text{ such that } \forall \sigma \in \Sigma^\omega \ \sigma \in L \text{ iff } \exists M, M \models \varphi \text{ and } M|_{\Lambda_\Sigma} = \sigma ]$ .

We then denote  $L = L_\omega^\Sigma(\varphi)$ , and, if there is no ambiguity,  $L = L_\omega(\varphi)$  the locally finite  $\omega$ -language defined by  $\varphi$ .

The class of locally finite  $\omega$ -languages will be denoted  $LOC_\omega$ .

**Remark 2.5** The notion of locally finite ( $\omega$ )-language is very different of the usual notion of local ( $\omega$ )-language which represents a subclass of the class of rational ( $\omega$ )-languages. But from now on, as in [Fin01], things being well defined and precised, we shall call simply local ( $\omega$ )-languages (respectively, local sentences) the locally finite ( $\omega$ )-languages (respectively, locally finite sentences).

## 2.2 Examples of local $\omega$ -languages

The following example should not be skipped because it is crucial to Theorem 2.9 below.

**Example 2.6** The  $\omega$ -language which contains only the word  $\sigma = abab^2ab^3ab^4 \dots$  (where the  $i$ -th occurrence of  $a$  is followed by the factor  $b^i a$ ) is a local  $\omega$ -language over the alphabet  $\{a, b\}$ .

**Proof.** Let the signature  $S(\varphi) = \{P_a, P_b, <, p, p', f\}$ , where  $p$  and  $p'$  are unary function symbols,  $f$  is a binary function symbol. And let  $\varphi$  be the following sentence, conjunction of:

- (1)  $\forall xyz[(x \leq y \vee y \leq x) \wedge (x \leq y \wedge y \leq x \leftrightarrow x = y) \wedge (x \leq y \wedge y \leq z \rightarrow x \leq z)]$ ,
- (2)  $\forall x[P_a(x) \leftrightarrow \neg P_b(x)]$ ,
- (3)  $\forall xy[(x < y \wedge P_a(x) \wedge P_a(y)) \rightarrow P_b(f(xy))]$ ,
- (4)  $\forall xy[x \geq y \rightarrow f(xy) = x]$ ,
- (5)  $\forall xy[\neg P_a(x) \vee \neg P_a(y) \rightarrow f(xy) = x]$ ,
- (6)  $\forall x[P_a(p(x)) \wedge P_a(p'(x))]$ ,
- (7)  $\forall x[\neg P_a(x) \rightarrow p(x) < p'(x)]$ ,
- (8)  $\forall x[P_a(x) \rightarrow p(x) = p'(x) = x]$ ,



- (9)  $\forall x[\neg P_a(x) \rightarrow x = f(p(x)p'(x))]$ ,  
(10)  $\forall xx'y \in P_a[x < x' < y \rightarrow f(x'y) < f(xy) < y]$ ,  
(11)  $\forall xyy' \in P_a[x \leq y' < y \rightarrow y' < f(xy) < y]$ .

Above sentence (1) means that “ $<$  is a linear order ”. The sentence (2) expresses that  $P_a, P_b$  form a partition in any model  $M$  of  $\varphi$ . The sentence (3) says that  $f$  is a function from  $\{(x, y) \mid x < y \wedge P_a(x) \wedge P_a(y)\}$  into  $P_b$  while sentences (4)-(5) express that the function  $f$  is trivially defined elsewhere. Sentences (6)-(7) say that  $p$  and  $p'$  are functions defined from  $\neg P_a = P_b$  into  $P_a$  and (8) states that  $p$  and  $p'$  are trivially defined on  $P_a$ .

The projections  $p$  and  $p'$  are used to say that the function  $f$  is surjective from  $\{(x, y) \mid x < y \wedge P_a(x) \wedge P_a(y)\}$  onto  $P_b$ ; this is implied by sentence (9).

The 10<sup>th</sup> and 11<sup>th</sup> conjunctions are used to order the elements of  $f(P_a \times P_a)$  in order to obtain the word  $\sigma = abab^2ab^3ab^4 \dots$  when the reduction to the signature of words is considered.

Notice that (10)-(11) imply also that  $f$  is injective hence is in fact a bijection from  $\{(x, y) \mid x < y \wedge P_a(x) \wedge P_a(y)\}$  into  $P_b$ .  $\square$

**Remark 2.7** *We have defined the functions  $f$  and  $p, p'$ , in a trivial manner (like  $f(xy) = x$  or  $p(x) = x$ ) where they were not useful for defining the local  $\omega$ -language  $\{\sigma\}$ , (see the conjunctions (4), (5) and (8)). This will imply that closure in models of  $\varphi$  takes at most a finite number of steps. This method will be applied in the construction of other local sentences in the sequel of this paper.*

We can easily check that  $\varphi$  is equivalent to a universal formula and that closure in its models takes at most  $n_\varphi = 2$  steps: one takes closure under the functions  $p, p'$  then by  $f$ .

Hence  $\varphi$  is a local sentence and by construction:  $L_\omega^{\{a,b\}}(\varphi) = \{abab^2ab^3 \dots\}$ .

Let us give some examples of closure in a model  $M$  of  $\varphi$  such that  $M \models \Lambda_{\{a,b\}} = \sigma = abab^2ab^3ab^4 \dots$ .

Let  $X_n \subseteq |M|$  be the segment of  $M$  corresponding to the segment  $ab^n a$  of  $\sigma$ . Then the closure of  $X_n$  under the functions  $p, p'$  is the set  $X_n \cup Z_n$  where  $Z_n$  corresponds to the set of the  $(n - 1)$  first letters  $a$  of  $\sigma$ .  $cl(X_n, M)$  is the closure of  $X_n \cup Z_n$  under  $f$  and it is the initial segment of  $M$  corresponding to the initial segment  $abab^2ab^3ab^4 \dots ab^n a$  of  $\sigma$ .

Let now  $Y \subseteq |M|$  be the segment of  $M$  corresponding to the three last letters  $b^2 a$  of the segment  $ab^n a$  of  $\sigma$ , for some integer  $n \geq 3$ . Then the closure of  $Y$  under the functions  $p, p'$  is  $Y \cup Z$  where  $Z$  corresponds to the set of the two first letters  $a$  of  $\sigma$ . The closure of  $Y \cup Z$  under  $f$  is the set  $cl(Y, M)$  which induces the word  $abab^2 a$  but which is not a segment of  $M$  because it contains the two first letters  $a$  and the  $(n + 1)$ -th letter  $a$  of  $\sigma$  but not any other letter  $a$  of  $\sigma$ .

We are going now to get more examples of local  $\omega$ -languages. Recall first the following:

**Definition 2.8** *The  $\omega$ -Kleene closure of a family  $\mathcal{L}$  of finitary languages is:*

$$\omega\text{-}KC(\mathcal{L}) = \{\cup_{i=1}^n U_i.V_i^\omega \mid \forall i \in [1, n] \ U_i, V_i \in \mathcal{L}\}$$

This notion of  $\omega$ -Kleene closure appears in the characterization of the class  $REG_\omega$  of regular  $\omega$ -languages (respectively, of the class  $CF_\omega$  of context free  $\omega$ -languages) which is the  $\omega$ -Kleene closure of the family  $REG$  of regular finitary languages (respectively, of the family  $CF$  of context free finitary languages), [Tho90] [PP02] [Sta97].

A natural question arises: does a similar characterization hold for local languages? The answer is given by the following:

**Theorem 2.9** *The  $\omega$ -Kleene closure of the class  $LOC$  of finitary local languages is strictly included into the class  $LOC_\omega$  of local  $\omega$ -languages.*

**Proof.** We have already proved that  $\omega\text{-}KC(LOC) \subseteq LOC_\omega$  in [Fin01]. In order to show that the inclusion is strict, remark that if an  $\omega$ -language  $L$  belongs to  $\omega\text{-}KC(LOC)$ , then  $L$  contains at least an ultimately periodic word, i.e. a word in the form  $u.v^\omega$  where  $u$  and  $v$  are finite words. Now we can easily check that the local  $\omega$ -language given in example 2.6 does not contain any ultimately periodic word because its single word is not ultimately periodic.  $\square$

A first consequence of Theorem 2.9 is that every regular  $\omega$ -language is a local  $\omega$ -language, i.e.  $REG_\omega \subseteq LOC_\omega$ , because every finitary regular language is local [Res88].

We had shown in [Fin01] that many context free languages are local thus  $CF_\omega = \omega\text{-}KC(CF)$  implies that many context free  $\omega$ -languages are local. The problem to know whether each context free language is local is still open but by Theorem 2.9,  $CF \subseteq LOC$  would imply that  $CF_\omega \subseteq LOC_\omega$ .

The  $\omega$ -language given in example 2.6 is local but non context free because every context free  $\omega$ -language contains at least one ultimately periodic word. We proved in [Fin01] that the finitary language  $U = \{a^n b^{n^2} \mid n \geq 1\} \subseteq \{a, b\}^*$  is local. Thus the  $\omega$ -language  $U.c^\omega \subseteq \{a, b, c\}^\omega$  is local but  $U.c^\omega$  is not context free because  $U \notin CF$ , [CG77].

These two examples show that the inclusion  $LOC_\omega \subseteq CF_\omega$  does not hold.

### 3 Closure properties of locally finite omega languages

**Theorem 3.1** *The class of locally finite omega languages is not closed under intersection with a regular  $\omega$ -language in the form  $L.a^\omega$  where  $L$  is a rational language,  $L \subseteq \Sigma^*$  and  $a \notin \Sigma$ . Hence  $LOC_\omega$  is neither closed under intersection nor under complementation.*

To prove this theorem, we shall proceed by successive lemmas. We shall assume that every language considered here is constituted of words over a finite alphabet included in a given countable set  $\Sigma_D$ .

We firstly define the family  $I$  of finitary languages by: for a finitary language  $L \subseteq \Sigma^*$ , where  $\Sigma \subseteq \Sigma_D$  is a finite alphabet,  $L \in I$  if and only if  $L.a^\omega \in LOC_\omega$  whenever  $a$  is a letter of  $\Sigma_D - \Sigma$ .

It is easy to see that if  $L.a^\omega$  is a local  $\omega$ -language and  $a \in \Sigma_D - \Sigma$ , then for all  $b \in \Sigma_D - \Sigma$ , it holds that  $L.b^\omega \in LOC_\omega$ . It suffices to replace the predicate  $P_a$  by  $P_b$  in the sentence defining  $L.a^\omega$ .

**Lemma 3.2**  *$I$  is closed under inverse alphabetic morphism.*

**Proof.** Let  $L \in I$ , i.e.  $L \subseteq \Gamma^*$  for some finite alphabet  $\Gamma$ ,  $a \notin \Gamma$  and  $L.a^\omega = L_\omega^{\Gamma \cup \{a\}}(\varphi)$  for a local sentence  $\varphi$ .

Let  $h$  be the alphabetic morphism:  $\Sigma^* \rightarrow \Gamma^*$ , defined by  $h(c) \in \Gamma \cup \{\lambda\}$  for  $c \in \Sigma$ , where  $\lambda$  is the empty word. And let  $\Sigma' = \{c \in \Sigma \mid h(c) = \lambda\}$ . We assume, possibly changing  $a$ , that  $a \notin \Sigma$ .

We first replace in  $\varphi$  the letter predicates  $(P_c)_{c \in \Gamma}$  by  $(Q_c)_{c \in \Gamma}$ .

The language  $h^{-1}(L).a^\omega$  is then defined by the following sentence  $\psi$ , in the signature  $S(\psi) = \{P, A, (P_c)_{c \in \Sigma}, P_a\} \cup S(\varphi)$ , where  $S(\varphi)$  contains the letter predicates  $Q_c$  for  $c \in \Gamma \cup \{a\}$ ,  $P$  is a unary predicate symbol and  $A$  is a constant symbol.  $\psi$  is the conjunction of:

- ( $<$  is a linear order ),
- $((P_c)_{c \in \Sigma}, P_a)$  form a partition,
- $\forall x_1 \dots x_n \in P[\varphi_0(x_1 \dots x_n) \wedge \bigwedge_{c \in \Gamma} (Q_c(x_1) \leftrightarrow \bigvee_{d \in h^{-1}(c)} P_d(x_1)) \wedge (Q_a(x_1) \leftrightarrow P_a(x_1))]$ , where  $\varphi = \forall x_1 \dots x_n \varphi_0(x_1 \dots x_n)$  with  $\varphi_0$  an open formula,
- $\forall x_1 \dots x_k [(\bigvee_{1 \leq j \leq k} \neg P(x_j)) \rightarrow g(x_1 \dots x_k) = x_1]$ , for each  $k$ -ary function  $g$  of  $S(\varphi)$ ,
- $(P_c)_{c \in \Sigma'}$  form a partition of  $\neg P$ ,
- $P(B)$ , for each constant  $B$  of  $S(\varphi)$ ,
- $\forall xy [\neg P_a(y) \wedge P_a(x) \rightarrow y < x]$ ,
- $P_a(A)$ .

$\psi$  is equivalent to a universal sentence and closure in its models takes at most  $n_\varphi + 1$  steps. Hence  $\psi$  is local and by construction  $L_\omega^{\Sigma \cup \{a\}}(\psi) = h^{-1}(L).a^\omega$ .  $\square$

**Lemma 3.3** *I is closed under non erasing alphabetic morphism.*

Recall that an alphabetic morphism  $h : \Sigma^* \rightarrow \Gamma^*$ , defined by  $h(c) \in \Gamma \cup \{\lambda\}$ , for  $c \in \Sigma$  is said to be non erasing if  $\forall c \in \Sigma, h(c) \in \Gamma$ .

**Proof.** Let  $L \in I, L \subseteq \Sigma^*$ . Let  $a \notin \Sigma$  and  $L.a^\omega = L_\omega^{\Sigma \cup \{a\}}(\varphi)$  for a local sentence  $\varphi$ . Let  $h$  be a non erasing alphabetic morphism given by  $h : \Sigma \rightarrow \Gamma$ . Moreover we assume that  $a \notin \Gamma$  (possibly changing  $a$ ). Then the language  $h(L).a^\omega$  is defined by the following formula  $\psi$ , in the signature  $S(\psi) = S(\varphi) \cup \{(Q_c)_{c \in \Gamma}\}$ . The sentence  $\psi$  is the conjunction of:

- $\varphi$ ,
- $\forall x [\bigwedge_{c \in \Sigma} (P_c(x) \rightarrow Q_{h(c)}(x))]$ ,
- $[(Q_c)_{c \in \Gamma}, (P_a)]$  form a partition.

$\psi$  is local and if the predicates  $(Q_c)_{c \in \Gamma}, P_a$ , are the letter predicates,  $\psi$  defines the  $\omega$ -language  $L_\omega^{\Gamma \cup \{a\}}(\psi) = h(L).a^\omega$ .  $\square$

**Lemma 3.4** *I contains the finitary local languages.*

**Proof.**  $LOC_\omega$  contains the omega Kleene closure of the class  $LOC$  of finitary local languages, and for each letter  $a$  the language  $\{a\}$  is local.  $\square$

Recall now the definitions of the Antidyck language and of a rational cone of languages.

**Definition 3.5** *The Antidyck language over two sorts of parentheses is the language  $Q_2'^* = \{v \in (Y \cup \bar{Y})^* \mid v \rightarrow^* \lambda\}$ , where  $Y = \{y_1, y_2\}$ ,  $\bar{Y} = \{\bar{y}_1, \bar{y}_2\}$  and  $\rightarrow^*$  is the transitive closure of  $\rightarrow$  defined in  $(Y \cup \bar{Y})^*$  by:*

$$\forall y \in Y \quad yv_1\bar{y}v_2 \rightarrow v_1v_2 \quad \text{if and only if} \quad v_1 \in Y^*.$$

The Antidyck language  $Q_2'^*$  may be seen as the language containing words with two sorts of parentheses, such that: “the first parenthesis to be opened is the first to be closed”

**Definition 3.6** ([Ber79]) *A rational cone is a class of languages which is closed under morphism, inverse morphism, and intersection with a rational language. (Or, equivalently to these three properties, closed under rational transduction).*

The notion of rational cone has been much studied. In particular the Antidyck language  $Q_2'^*$  is a generator of the rational cone of the recursively enumerable languages, [FZV80].

On the other hand Nivat's Theorem states that a class of languages which is closed under alphabetic morphism, inverse alphabetic morphism, and intersection with a rational language, is a rational cone, [Ber79]. Moreover every rational transduction  $t$  is in the form  $t(u) = g[h^{-1}(u) \cap R]$ , where  $g$  et  $h$  are alphabetic morphisms and  $R$  is a rational language. Thus every recursively enumerable language may be written in the form  $g[h^{-1}(Q_2^*) \cap R]$ , where  $R$  is a rational language,  $g$  and  $h$  are alphabetic morphisms.

This result will be used here because the language  $Q_2^*$  is local, [Fin01].

Return now to the proof of Theorem 3.1 and suppose that  $LOC_\omega$  were closed under intersection with the languages  $R.a^\omega$ , where  $R \subseteq \Sigma^*$  is a rational language and  $a \notin \Sigma$ .

**Claim 3.7** *I would be closed under intersection with a rational language.*

**Proof.** Let  $L \in I$ ,  $L \subseteq \Sigma^*$ ,  $R \subseteq \Sigma^*$  be a rational language and  $a \notin \Sigma$ .  $L.a^\omega$  is a local  $\omega$ -language because  $L \in I$  and  $(L.a^\omega) \cap (R.a^\omega) = (L \cap R).a^\omega$  would be a local  $\omega$ -language, hence by definition of  $I$ ,  $L \cap R$  would belong to  $I$ .  $\square$

**Claim 3.8** *I would contain every language in the form  $g[h^{-1}(Q_2^*) \cap R]$ , where  $h$  is an alphabetic morphism,  $g$  is a non erasing alphabetic morphism and  $R$  is a rational language.*

**Proof.** It follows from the lemmas 3.2, 3.3, 3.4, the fact that  $Q_2^*$  is local and Claim 3.7.  $\square$

**Claim 3.9** *There would exist an erasing alphabetic morphism  $h$  and  $L \in I$  such that  $\{0^n 1^p \mid p > 2^n\} = h(L)$ , where an alphabetic morphism is said to be erasing if it is in the form  $h : \Sigma \rightarrow \Sigma \cup \{\lambda\}$ , with  $h(c) = c$  if  $c \in A$  and  $h(c) = \lambda$  if  $c \in \Sigma - A$ , for some subset  $A \subseteq \Sigma$ .*

**Proof.** We know that every recursively enumerable language may be written in the form  $g[h^{-1}(Q_2^*) \cap R]$ , where  $R$  is a rational language,  $g$  and  $h$  are alphabetic morphisms.

But every alphabetic morphism may be obtained as composed firstly by a non erasing alphabetic morphism followed by an erasing alphabetic morphism.

Thus it follows from Claim 3.8 that each recursively enumerable language, and in particular the language  $\{0^n 1^p \mid p > 2^n\}$ , would be the image by an erasing alphabetic morphism of a language of  $I$ .  $\square$

Let then  $h$  be an erasing morphism  $\Sigma^* \rightarrow \Sigma^*$  where  $\{0, 1\} \subseteq \Sigma$ ,  $h(0) = 0$  and  $h(1) = 1$  and  $h(c) = \lambda$  if  $c \in \Sigma - \{0, 1\}$  and let  $L \subseteq \Sigma^*$  be a language such that  $h(L) = \{0^n 1^p \mid p > 2^n\}$ .

Assume that  $L$  belongs to  $I$ , so if  $a \notin \Sigma$ ,  $L.a^\omega$  is a local  $\omega$ -language and there exists a local sentence  $\varphi$  such that  $L.a^\omega = L_\omega^{\Sigma \cup \{a\}}(\varphi)$ .

For all  $n \geq 1$  let  $M_n$  be a model of  $\varphi$  of order type  $\omega$  such that  $M_n \mid \Lambda_{\Sigma \cup \{a\}} =$

$\sigma_n \cdot a^\omega$ , where  $\sigma_n \in L$  and the number of occurrences of 0 in  $\sigma_n$  is  $n$  and the number of occurrences of 1 in  $\sigma_n$  is  $p_n > 2^n$ .

Let us now set the following definition in view of next lemma.

**Definition 3.10** *Let  $X$  be a set included in a structure  $M$  and  $P \subseteq |M|$ .  $X$  is a set of indiscernables above  $P$  for the atomic formulas of complexity  $\leq k$ , i.e. whose terms result by at most  $k$  applications of function symbols, for  $k \in \mathbb{N}$ , if and only if:*

- i)  $X$  is linearly ordered by  $<$ .
- ii) Whenever  $\bar{x}$  and  $\bar{y}$  are some  $n$ -tuples of elements of  $X$  which are isomorphic for the order of  $(X, <)$ ,  $\bar{x}$  and  $\bar{y}$  satisfy the same atomic formulas of complexity  $\leq k$ , with parameters in  $P$ .

**Lemma 3.11** *In the above conditions where  $M_n$  is defined for every integer  $n \geq 1$ , there exists in  $M_n$  an infinite set  $X_n$  of indiscernables above  $P_0^{M_n}$  for the atomic formulas of complexity  $\leq n_\varphi$ , with  $X_n \subseteq P_a^{M_n}$*

**Proof.** Let  $m(\varphi)$  be the maximum number of variables of the atomic formulas of complexity  $\leq n_\varphi$  i.e. whose terms result by at most  $n_\varphi$  applications of function symbols. These terms form a finite set  $T_\varphi$ .

For all strictly increasing sequences  $\bar{x}$  and  $\bar{y}$  of length  $m(\varphi)$  of  $P_a^{M_n}$ , let us set  $\bar{x} \sim \bar{y}$  if and only if  $\bar{x}$  and  $\bar{y}$  satisfy in  $M_n$  the same atomic formulas with parameters in  $P_0^{M_n}$  and of complexity  $\leq n_\varphi$ .

$P_0^{M_n}$  is a finite set of cardinal  $n$ , hence the set of atomic formulas with parameters in  $P_0^{M_n}$  and of complexity  $\leq n_\varphi$  is also finite.

Then applying the infinite Ramsey Theorem, we can find  $X_n \subseteq P_a^{M_n}$  homogeneous for  $\sim$  and infinite. This is the set we are looking for.  $\square$

We return now to the proof of Theorem 3.1 and consider in  $|M_n|$  the subset  $X_n \cup P_0^{M_n} = Y_n$ . This subset is infinite hence it is of order type  $\omega$  in  $M_n$  and it generates in  $M_n$  a model of order type  $\omega$  too, which will be denoted by  $M_n(Y_n) = A_n$ .

This model of  $\varphi$  induces a word  $u_n \cdot a^\omega$  of  $L \cdot a^\omega$ , such that there are in  $u_n$ :  $n$  occurrences of the letter 0 and  $q_n \leq p_n$  occurrences of the letter 1. But  $u_n \cdot a^\omega \in L \cdot a^\omega$  implies that  $2^n < q_n$  also holds.

$A_n$  is generated from  $Y_n$  by the use of only a finite set  $T_\varphi$  of terms of less than  $k_\varphi$  variables. If  $n$  is big enough with regard to  $k_\varphi$  and  $\text{card}(T_\varphi)$ , because  $q_n > 2^n$ , there exist parameters  $a_1, \dots, a_k$ , elements of  $P_0^{M_n}$ , and some indiscernables  $x_1, \dots, x_j$ , and  $y_1, \dots, y_j$  of  $X_n$ , such that  $x_1 < \dots < x_j$  and  $y_1 < \dots < y_j$  and  $\bar{x} \neq \bar{y}$  and a term  $t \in T_\varphi$  such that  $t(a_1, \dots, a_k, x_1, \dots, x_j) <$

$t(a_1, \dots, a_k, y_1, \dots, y_j)$  and these two elements being in  $P_1^{M_n}$ .

But then we could find in  $X_n$  a sequence  $(\bar{x}_i)_{1 \leq i \leq N}$ , with  $N$  arbitrarily large, such that for each  $i$ ,  $1 \leq i \leq N$ ,  $\bar{x}_i \bar{x}_{i+1}$  is of the order type of  $\bar{x} \bar{y}$ .

Then for all integers  $i$  such that  $1 \leq i \leq N$ ,  $P_1^{M_n}(t(a_1, \dots, a_k, \bar{x}_i))$  and the terms  $t(a_1, \dots, a_k, \bar{x}_i)$  are distinct two by two. This would imply that, for all integers  $N \geq 1$ ,  $\text{card}(P_1^{M_n}) \geq N$ . So there would be a contradiction with  $\text{card}(P_1^{M_n}) = p_n$  and we have proved that  $L$  does not belong to  $I$ .

Thus we can infer Theorem 3.1 from Claim 3.9. The non closure under complementation of the class  $LOC_\omega$  can be deduced from the non closure under intersection and the fact that  $LOC_\omega$  is closed under union (see next Theorem) or from an example, deduced from preceding proof, of a local  $\omega$ -language which complement is not local (see next remark).  $\square$

**Remark 3.12** The above proof shows in particular that the  $\omega$ -language  $A = \{0^n 1^p 2^\omega \mid p > 2^n\}$  is not local. From which we can easily deduce that the local  $\omega$ -language  $\{0^n 1^p 2^\omega \mid p \leq 2^n\} = L$  has a complement which is not a local  $\omega$ -language. (This  $\omega$ -language  $L$  is local because  $\{0^n 1^p \mid p \leq 2^n\}$  is a local finitary language [Fin01]).

Indeed if its complement was  $\omega$ -local, we would deduce, from a local sentence  $\varphi$  such that  $L_\omega(\varphi) = L^-$ , another local sentence  $\psi$  such that  $L_\omega(\psi) = A$ .

For example the sentence  $\psi$ , conjunction of:

- $\varphi$ ,
- $\forall xy[(P_0(x) \wedge P_1(y)) \rightarrow x < y]$ ,
- $\forall xy[(P_1(x) \wedge P_2(y)) \rightarrow x < y]$ ,
- $\forall xy[(P_0(x) \wedge P_2(y)) \rightarrow x < y]$ ,
- $P_2(c)$ , where  $c$  is a new constant symbol.

Now we establish that  $LOC_\omega$  is closed under several operations.

**Theorem 3.13** *The class  $LOC_\omega$  is closed under union, left concatenation with local (finitary) languages,  $\lambda$ -free substitution of local (finitary) languages,  $\lambda$ -free morphism.*

**Proof.**

**Closure under union.**

Let  $\varphi_1$  and  $\varphi_2$  be two local sentences defining local  $\omega$ -languages  $L_\omega(\varphi_1)$  and  $L_\omega(\varphi_2)$  over a finite alphabet  $\Sigma$ . Let us define a new local sentence  $\varphi_1 \cup \varphi_2$  which defines the local  $\omega$ -language  $L_\omega(\varphi_1 \cup \varphi_2) = L_\omega(\varphi_1) \cup L_\omega(\varphi_2)$ :

We may assume that  $S(\varphi_1) \cap S(\varphi_2) = \Lambda_\Sigma$ . Then  $S(\varphi_1 \cup \varphi_2)$  will be  $S(\varphi_1) \cup S(\varphi_2)$ . And the sentence  $\varphi_1 \cup \varphi_2$  is the following sentence:

$$\begin{aligned} & [\varphi_1 \wedge_{n\text{-ary function symbol } f \in S(\varphi_2)} (\forall x_1, \dots, x_n f(x_1, \dots, x_n) = \min(x_1, \dots, x_n))] \\ & \vee [\varphi_2 \wedge_{n\text{-ary function symbol } g \in S(\varphi_2)} (\forall x_1, \dots, x_n g(x_1, \dots, x_n) = \min(x_1, \dots, x_n))] \end{aligned}$$

### Closure under left concatenation by a local finitary language.

Consider a finitary local language  $L(\varphi)$  and a local  $\omega$ -language  $L_\omega(\psi)$  over the same alphabet  $\Gamma$ .

We may easily assume that  $L_\omega(\varphi)$  is empty, possibly adding a constant symbol  $c$  to the signature of  $\varphi$  and adding the conjunction  $\forall x[x \leq c]$  to the sentence  $\varphi$  (this means that every model of  $\varphi$  has a greatest element).

We may also assume that  $S(\varphi) \cap S(\psi) = \{<, (P_a)_{a \in \Gamma}\} = \Lambda_\Gamma$ .

Let then  $P$  be a new unary predicate symbol not in  $S(\varphi) \cup S(\psi)$ , and let  $\varphi.\psi$  be the following sentence in the signature  $S(\varphi) \cup S(\psi) \cup \{P\}$ , which is the conjunction of:

- ( $<$  is a linear order ),
- $((P_a)_{a \in \Gamma}$  form a partition),
- $\forall xy[P(x) \wedge \neg P(y) \rightarrow x < y]$ ,
- $\forall x_1, \dots, x_j \in P[\varphi_0(x_1, \dots, x_j)]$ ,  
where  $\varphi = \forall x_1, \dots, x_j \varphi_0(x_1, \dots, x_j)$  with  $\varphi_0$  an open formula,
- $\forall x_1, \dots, x_m \in P[f(x_1, \dots, x_m) \in P]$ ,  
for each  $m$ -ary function  $f$  of  $S(\varphi)$ ,
- $\forall x_1, \dots, x_m[\bigvee_{1 \leq i \leq m} \neg P(x_i) \rightarrow f(x_1, \dots, x_m) = \min(x_1, \dots, x_m)]$ ,  
for each  $m$ -ary function  $f$  of  $S(\varphi)$ ,
- $P(c)$ , for each constant  $c$  of  $S(\varphi)$ ,
- $\forall x_1, \dots, x_j \in \neg P[\psi_0(x_1, \dots, x_j)]$ ,  
where  $\psi = \forall x_1, \dots, x_j \psi_0(x_1, \dots, x_j)$  with  $\psi_0$  an open formula,
- $\forall x_1, \dots, x_m \in \neg P[f(x_1, \dots, x_m) \in \neg P]$ ,  
for each  $m$ -ary function  $f$  of  $S(\psi)$ ,
- $\forall x_1, \dots, x_m[\bigvee_{1 \leq i \leq m} P(x_i) \rightarrow f(x_1, \dots, x_m) = \min(x_1, \dots, x_m)]$ ,  
for each  $m$ -ary function  $f$  of  $S(\psi)$ ,
- $\neg P(c)$ , for each constant  $c$  of  $S(\psi)$ ,

This sentence  $\varphi.\psi$  is equivalent to a universal formula and closure in its models takes at most  $\max(n_\varphi, n_\psi)$  steps, hence it is a local sentence and by construction it holds that  $L(\varphi.\psi) = L(\varphi).L(\psi)$ . Moreover when  $\omega$ -words are considered  $L_\omega(\varphi.\psi) = L(\varphi).L_\omega(\psi)$  holds because by hypothesis  $L_\omega(\varphi)$  is empty.

### Closure under $\lambda$ -free substitution of local languages.

The proof is very similar to our proof of the closure of the class *LOC* under substitution by local finitary languages in [Fin01]. We recall it now.



Recall first the notion of substitution:

A substitution  $f$  is defined by a mapping  $\Sigma \rightarrow P(\Gamma^*)$ , where  $\Sigma = \{a_1, \dots, a_n\}$  and  $\Gamma$  are two finite alphabets,  $f : a_i \rightarrow L_i$  where  $\forall i \in [1; n]$ ,  $L_i$  is a finitary language over the alphabet  $\Gamma$ . The substitution is said to be  $\lambda$ -free if  $\forall i \in [1; n]$ ,  $L_i$  does not contain the empty word  $\lambda$ . It is a ( $\lambda$ -free) morphism when every language  $L_i$  contains only one (nonempty) word.

Now this mapping is extended in the usual manner to finite words and to finitary languages: for some letters  $x(0), \dots, x(n)$  in  $\Sigma$ ,  $f(x(0)x(1)\dots x(n)) = \{u_0u_1\dots u_n \mid \forall i \in [0; n] \ u_i \in f(x(i))\}$ , and for  $L \subseteq \Sigma^*$ ,  $f(L) = \cup_{x \in L} f(x)$ .

If the substitution  $f$  is  $\lambda$ -free, we can extend this to  $\omega$ -words and  $\omega$ -languages:  $f(x(0)x(1)\dots x(n)\dots) = \{u_0u_1\dots u_n\dots \mid \forall i \geq 0 \ u_i \in f(x(i))\}$  and for  $L \subseteq \Sigma^\omega$ ,  $f(L) = \cup_{x \in L} f(x)$ .

Let then  $\Sigma = \{a_1, \dots, a_n\}$  be a finite alphabet and let  $f$  be a  $\lambda$ -free substitution of local languages:  $\Sigma \rightarrow P(\Gamma^*)$ ,  $a_i \rightarrow L_i$  where  $\forall i \in [1; n]$ ,  $L_i$  is a local language defined by the sentence  $\varphi_i$ , over the alphabet  $\Gamma$ . We may assume that  $L_\omega(\varphi_i)$  is empty, possibly adding a constant symbol  $c_i$  to the signature of  $\varphi_i$  and adding the conjunction  $\forall x[x \leq c_i]$  to the sentence  $\varphi_i$  (this means that every model of  $\varphi_i$  has a greatest element). We also assume that the signatures of the sentences  $\varphi_i$  verify  $S(\varphi_i) \cap S(\varphi_j) = \{<, (P_a)_{a \in \Gamma}\}$  for  $i \neq j$ . Let now  $L \subseteq \Sigma^\omega$  be a local  $\omega$ -language defined by a local sentence  $\varphi$ . We shall denote  $Q_{a_i}$  the unary predicate of  $S(\varphi)$  which indicates the places of the letters  $a_i$  in a word of  $L$ , so if  $a_i \in \Gamma \cap \Sigma$  for some indice  $i$ , there will be two distinct predicates  $Q_{a_i}$  and  $P_{a_i}$ . We may also assume that  $\forall i \in [1, \dots, n]$ ,  $S(\varphi_i) \cap S(\varphi) = \{<\}$ . Then we can construct a local sentence  $\psi$  (already given in [Fin01]) such that  $L_\omega(\psi) = f(L)$ .

$\psi$  is the conjunction of the following sentences, which meaning is explained below:

- “  $<$  is a linear order ”,
- $\forall xy[(I(y) \leq y) \wedge (y \leq x \rightarrow I(y) \leq I(x)) \wedge (I(y) \leq x \leq y \rightarrow I(x) = I(y))]$ ,
- $\forall x[I(x) = x \leftrightarrow P(x)]$ ,
- $P(c)$ , for each constant  $c$  of  $S(\varphi)$ ,
- $\forall x_1, \dots, x_k[R(x_1, \dots, x_k) \rightarrow P(x_1) \wedge \dots \wedge P(x_k)]$ ,  
for each predicate  $R(x_1, \dots, x_k)$  of  $S(\varphi)$ ,
- $\forall x_1, \dots, x_j[(P(x_1) \wedge \dots \wedge P(x_j)) \rightarrow P(f(x_1, \dots, x_j))]$ ,  
for each  $j$ -ary function symbol  $f$  of  $S(\varphi)$ ,
- $\forall x_1, \dots, x_j[(\bigvee_{1 \leq i \leq j} \neg P(x_i)) \rightarrow f(x_1, \dots, x_j) = \min(x_1, \dots, x_j)]$ ,  
for each  $j$ -ary function symbol  $f$  of  $S(\varphi)$ ,
- $\forall x_1, \dots, x_m[(P(x_1) \wedge \dots \wedge P(x_m)) \rightarrow \varphi_0(x_1, \dots, x_m)]$ ,  
where  $\varphi = \forall x_1, \dots, x_m \varphi_0(x_1, \dots, x_m)$  with  $\varphi_0$  an open formula,
- $\forall x_1, \dots, x_j[\bigvee_{i, k \leq j} (I(x_i) \neq I(x_k)) \rightarrow f(x_1, \dots, x_j) = \min(x_1, \dots, x_j)]$ ,  
for every function  $f$  of  $S(\varphi_l)$  for an integer  $l \leq n$ ,
- $\forall xy_1, \dots, y_j[(\bigwedge_{1 \leq l \leq j} I(y_l) = I(x)) \rightarrow I(f(y_1, \dots, y_j)) = I(x)]$ ,  
for each  $j$ -ary function symbol  $f$  of  $S(\varphi_i)$  for an integer  $i \leq n$ ,

Finally, for each  $i \leq n$ :

- $\forall x y_1, \dots, y_p [(\bigwedge_{1 \leq l \leq p} I(y_l) = I(x) \wedge Q_{a_i}(I(x))) \rightarrow \varphi_i^0(y_1, \dots, y_p) \wedge \{(e_j(y_1) = I(x) \wedge f_j(y_1, \dots, y_p) = y_1 \wedge \neg R_j(y_1, \dots, y_p) \wedge \bigwedge_{e_i \in S(\varphi_i)} e_i(y_1) = e_i(x); \text{ where } n \geq j \neq i, \text{ and } e_j, f_j, R_j \text{ run over the constants, functions, and predicates of } S(\varphi_j)\})]$ ,

Above , 1) to each constant  $e_l$  of  $S(\varphi_l)$  is associated a new unary function  $e_l(y)$  and 2) whenever  $\varphi_i = \forall y_1, \dots, y_p \psi_i(y_1, \dots, y_p)$  with  $\psi_i$  an open formula,  $\varphi_i^0$  is  $\psi_i$  in which every constant  $e_i$  has been replaced by the function  $e_i(y)$ .

### Construction of $\psi$ :

Using the function  $I$  which marks the first letters of the subwords, we divide an  $\omega$ -word into omega (finite) subwords (the function  $I$  is constant on each subword and  $I(x)$  is the first letter of the subword containing  $x$ ). In every model  $M$  of order type  $\omega$  of  $\psi$ , the set of the “first letters of subwords” ,  $P^M$ , grows richer in a model of order type  $\omega$  of  $\varphi$  (therefore will constitute an  $\omega$ -word of  $L$ ).

Then we “substitute”: for each letter  $a_i$  in  $P^M$ , we substitute a (finite) word of  $L_i$ , using for that the formula  $\varphi_i$ .

If closure takes at most  $n_\varphi$  (respectively  $n_{\varphi_i}$ ) steps in every model of  $\varphi$  (respectively of  $\varphi_i$ ), then closure takes at most  $[1 + n_\varphi + \sup_i(n_{\varphi_i})]$  steps in each model of  $\psi$  (one takes closure under the function  $I$ , then under the functions of  $S(\varphi)$ , and finally under the functions of  $S(\varphi_i)$ ,  $1 \leq i \leq n$ ).

Therefore  $\psi$  is a local sentence and by construction  $\psi$  defines the  $\omega$ -language  $f(L)$ .

### Closure under $\lambda$ -free morphism.

It is just a particular case of the preceding one, when every language  $L_i$  contains only one non-empty finite word.  $\square$

**Acknowledgments.** Thanks to the anonymous referees for useful comments on the preliminary version of this paper.

### References

- [Ber79] J. Berstel, Transductions and Context Free Languages, Teubner Studienbücher Informatik, 1979.
- [Büc60] J.R. Büchi, Weak Second Order Arithmetic and Finite Automata, Zeitschrift für Mathematische Logik und Grundlagen der Mathematik, 6 (1960), pp 66-92.

- [Büc62] J.R. Büchi, On a Decision Method in Restricted Second Order Arithmetic, Logic Methodology and Philosophy of Science, ( Proc. 1960 Int. Congr. ). Stanford University Press, 1962, 1-11.
- [BS73] J.R. Büchi and D. Siefkes, The Monadic Second Order Theory of All Countable ordinals, Decidable Theories 2, 1973, S.L.N.M. , number 328.
- [CG77] R. S. Cohen and A. Y. Gold, Theory of  $\omega$ -Languages, Parts one and two, Journal of Computer and System Science 15 (1977) 2, 169-184 and 185-208.
- [CK73] C.C. Chang and H.J. Keisler, Model Theory, American Elsevier Publishing Company, Inc, New York, 1973. (North Holland, Amsterdam, 3rd ed., 1990).
- [DLT98] A. Durand, C. Lautemann and T. Schwentick, Subclasses of Binary-NP, Journal of Logic and Computation, vol. 8(2), 1998, p. 189-207.
- [EF95] H. D. Ebbinghaus and J. Flum, Finite Model Theory, Springer Verlag, New York, 1995.
- [EGG00] T. Eiter, G. Gottlob and Y. Gurevich, Existential Second-Order Logic over Strings, IEEE Symposium on Logic in Computer Science (LICS'98), Journal of the ACM, vol. 47, no. 1, Jan. 2000, 77–131.
- [Fag74] R. Fagin, Generalized First-Order Spectra and Polynomial-Time Recognizable Sets, in: R.M. Karp, ed., Complexity of Computation, SIAM-AMS Proc. 7 (1974) p. 43-73.
- [Fag93] R. Fagin, Finite-Model Theory, A Personal Perspective, Theoretical Computer Science, Volume 116 (1993) p. 3-31.
- [Fin89] O.Finkel, Langages de Büchi et  $\omega$ -Langages Locaux, C.R.A.S. Paris, t.309, Série 1, p. 991-994, 1989.
- [Fin93] O.Finkel, Théorie des Modèles des Formules Locales et Etude des Langages Formels qu'elles Définissent, Ph. D. Thesis, University Paris 7, May 1993.
- [FR96] O. Finkel and J-P. Ressayre, Stretchings, Journal of Symbolic Logic, Volume 61, number 2, June 1996, p. 563-585.
- [Fin01] O. Finkel, Locally Finite Languages, Theoretical Computer Science, Volume 255 (1-2), March 2001, p.223-261.
- [FZV80] P. Franchi-Zanettacci and B. Vauquelin, Automates à Files, Theoretical Computer Science, Volume 11 (1980), p.221-225.
- [GKS00] G. Gottlob, P. G. Kolaitis, T. Schwentick, Existential Second-Order Logic over Graphs: Charting the Tractability Frontier, FOCS 2000, p. 664-674.
- [Hac01] Y. Hachaichi, Contributions à la Théorie des Modèles Finis et à la Complexité Descriptive, Ph. D. Thesis, university Paris 7, July 2001.
- [HU69] J.E. Hopcroft and J.D. Ullman, Formal Languages and their Relation to Automata, Addison-Wesley Publishing Company, Reading, Massachusetts, 1969.

- [Imm87] N. Immerman, Languages That Capture Complexity Classes, *SIAM J. of Computing* 16:4 (1987), p. 347-354.
- [Imm99] N. Immerman, *Descriptive Complexity*, Springer Graduate Texts in Computer Science, 1999.
- [LST94] C. Lauteman, T. Schwentick and D. Thérien, Logics for Context Free Languages, in the *Proceedings of Computer Science Logic 94*, p. 205-216.
- [LT94] H. Lescow and W. Thomas, Logical Specifications of Infinite Computations, In: "A Decade of Concurrency" (J. W. de Bakker et al., eds), Springer LNCS 803 (1994), 583-621.
- [PP85] M. Parigot and E. Pelz, A Logical Approach of Petri Net Languages, *Theoretical Computer Science*, Volume 39 (1985) p. 155-169.
- [Pel87] E. Pelz, Infinitary Languages of Petri Nets and Logical Sentences, *Advances in Petri Nets 1987* (Oxford, 1986), *Lecture Notes in Comput. Sci.*, 266, Springer, Berlin, 1987, p. 224-237.
- [PP02] D. Perrin and J.-E. Pin, *Infinite Words*, Book in preparation, available from <http://www.liafa.jussieu.fr/~jep/InfiniteWords.html>
- [Pin96] J-E. Pin, Logic, Semigroups and Automata on Words, *Annals of Mathematics and Artificial Intelligence* 16 (1996), p. 343-384.
- [Res88] J-P. Ressayre, Formal Languages Defined by the Underlying Structure of their Words, *Journal of Symbolic Logic*, 53, number 4, December 1988, p. 1009-1026.
- [Res89] J-P. Ressayre, Question posed during a Working Group on Automata and Logic at University Paris 7, 1989.
- [Sta97] L. Staiger,  $\omega$ -Languages, Chapter of the *Handbook of Formal languages*, Vol. 3, edited by G. Rozenberg and A. Salomaa, Springer-Verlag, Berlin.
- [Tho90] W. Thomas, Automata on Infinite Objects, in: J. Van Leeuwen, ed., *Handbook of Theoretical Computer Science*, Vol. B ( Elsevier, Amsterdam, 1990 ), p. 133-191.
- [Tho96] W. Thomas, Languages, Automata and Logic, in *Handbook of Formal Languages Theory*, Vol. 3, Edited by G. Rozenberg and A. Salomaa, Springer Verlag, 1996.
- [Yam99] H. Yamasaki, Logical Characterization of Petri Net  $\omega$ -Languages, Report 99-17, Department of Mathematics and Computer Science, Martin-Luther-Universität Halle-Wittenberg, 1999.