Arch. Math. Logic (2008) 47:625–651
DOI 10.1007/s00153-008-0101-7

Mathematical Logic

Topological complexity of locally finite ω-languages

Olivier Finkel

Received: 9 September 2002 / Published online: 29 July 2008

Abstract Locally finite omega languages were introduced by Ressayre [Formal languages defined by the underlying structure of their words. J Symb Log 53(4):1009–1026, 1988]. These languages are defined by local sentences and extend ω-languages accepted by Büchi automata or defined by monadic second order sentences. We investigate their topological complexity. All locally finite ω-languages are analytic sets, the class LOC_ω of locally finite ω-languages meets all finite levels of the Borel hierarchy and there exist some locally finite ω-languages which are Borel sets of infinite rank or even analytic but non-Borel sets. This gives partial answers to questions of Simonnet (Automates et Théorie Descriptive, Ph.D. Thesis. Université Paris 7, March 1992) and of Duparc et al. [Computer science and the fine structure of Borel sets. Theor Comput Sci 257(1–2):85–105, 2001].

Keywords Local sentences · Locally finite ω-languages · Topological complexity · Borel hierarchy · Analytic sets

Mathematics Subject Classification (2000) 03D05 · 03E15 · 68Q45 · 03B15 · 03C85

Contents

O. Finkel (✉)
Equipe Modèles de Calcul et Complexité, Laboratoire de l'Informatique du Parallélisme,
UMR 5668, CNRS, ENS Lyon, UCB Lyon, INRIA, CNRS et Ecole Normale Supérieure de Lyon,
46, Allée d'Italie, 69364 Lyon Cedex 07, France
e-mail: Olivier.Finkel@ens-lyon.fr

1 Introduction

Local sentences were introduced by Ressayre in [32]. He proved there some remarkable stretching theorems which established some links between the finite and the infinite model theory of these sentences. Some of these theorems can only be proved assuming the existence (or the consistency of the existence) of large cardinals like inaccessible or Mahlo cardinals. These theorems show that the existence of some well ordered models of a local sentence φ (a binary relation symbol is here assumed to belong to the signature of φ and to be interpreted by a linear order in every model of φ) is equivalent to the existence of some finite model of φ, generated by some particular kind of indiscernibles, like special, remarkable or monotonic ones. In particular, a local sentence φ has a model of order type ω if and only if it has a finite model generated by N_φ special indiscernibles (where N_φ is a positive integer depending on φ), and a similar result establishes a connection between the existence of a model of order type α (where α is an ordinal $<\omega^\omega$) and the existence of a finite model (of another local sentence φ_α) generated by semi-monotonic indiscernibles [14].

These theorems provide some decision algorithms which show the decidability of the following problem:

(P_1) "For a given local sentence φ and an ordinal $\alpha < \omega^\omega$, has φ a model of order type α ?"

These results look like Büchi's one about the decidability of the monadic second order theory of one successor over the integers [4], and even more like its extension: the decidability of the monadic second order theory of the structure $(\alpha, <)$ for a countable ordinal α.

In order to prove this result, Büchi studied in the sixties the class of ω-languages accepted by finite automata with what is now called Büchi acceptance condition. He showed that an ω-language, i.e. a set of words of length ω over a finite alphabet, is accepted by a finite automaton with the Büchi acceptance condition if and only if it is defined by a monadic second order sentence and he found algorithms to give such an automaton from the monadic second order sentence. Hence the decision problem cited above was reduced to the decidability of the emptiness problem for Büchi automata which is easily shown to be decidable [4,39]. The equivalence between definability by monadic second order sentences and acceptance by finite automata, which is also true for languages of finite words [3], has then been extended to α-languages, i.e. languages of words of length α, where α is a countable ordinal $\geq \omega$ [5]. This led to similar decision algorithms showing that the monadic second order theory of the structure $(\alpha, <)$ is decidable.

In order to compare the power of the above decidability results concerning local or monadic sentences, it is now interesting to compare the expressive power of monadic sentences and of local sentences, and then to consider languages defined by these sentences.

Ressayre introduced locally finite languages which are defined by local sentences. Local sentences are first order, but they define locally finite languages via existential quantifications over relations and functions which appear in the local sentence. These second order quantifications are more general than the monadic ones:

– When finite words are considered, each regular language is locally finite [32], each quasirational language is locally finite, and many context-free as well as non-context-free languages are locally finite [15].
– Each regular ω-language is a locally finite ω-language, and there exist locally finite ω-languages which are not regular [12,15].
– This is extended to languages of transfinite length words: when α is an ordinal $<\omega^\omega$, an α-language accepted by a Büchi automaton is also defined by a local sentence [15].

Thus the class LOC_α of locally finite α-languages, for $\omega \leq \alpha < \omega^\omega$, is a strict extension of the class REG_α of regular α-languages (defined by monadic second order sentences). Then the following question naturally arises:

How large is the extension of $\mathbf{REG}_\alpha$ by $\mathbf{LOC}_\alpha$?

A way to attack this problem is to study the topological complexity of α-languages in each of these classes, and firstly to locate them with regard to the Borel and projective hierarchies. We restrict here our study to ω-languages and then it is well known that all regular ω-languages are boolean combinations of $\mathbf{\Sigma}^0_2$-sets hence $\mathbf{\Delta}^0_3$-sets [30,39].

We shall see in this paper that locally finite ω-languages extend far beyond regular ω-languages: the class LOC_ω meets all finite levels of the Borel hierarchy, contains some Borel sets of infinite rank and even some analytic but non-Borel sets.

This will show that the decision algorithm for the sentences in the form $\exists R_1, \ldots,$ $\exists R_k\varphi$, where φ is local in the signature $S(\varphi) = \{<, R_1, \ldots, R_k\}$ and $R_1, \ldots, R_k$ are relations or n-ary function symbols with $n \geq 1$, provides a very large extension, for $\alpha < \omega^\omega$, of Büchi's result about the decidability of the monadic second order theory of $(\alpha, <)$. Moreover, at least for $\alpha = \omega$, the algorithm for local sentences (given by Theorem 2.7 below) is of much lower complexity than the corresponding algorithm for monadic second order sentences.

The question of the topological complexity of locally finite ω-languages is also motivated by the general project of studying the logical definability of classes of formal languages of (finite or) infinite words, (or of relational structures like graphs). This research area is now called "descriptive complexity", see [31,40] for a survey about this field of research.

The study of topological complexity of locally finite ω-languages was also asked by Simonnet [36] and also by Duparc et al. in [10] where they asked for extensions of the Wagner hierarchy of regular ω-languages.

The paper is organized as follows. In Sect. 2 we review the definitions and some properties of local sentences and locally finite (omega) languages. Then we give some examples of locally finite ω-languages. In Sect. 3 we study topological properties of locally finite ω-languages. Firstly we show that LOC_ω is included in the class of analytic sets. Duparc studied recently the Wadge hierarchy which is a great refinement of the Borel hierarchy. He gave a normal form for Borel sets of finite rank in each Wadge degree, using operations over sets of finite and infinite words [9]. Using Duparc's operation of exponentiation of sets, we prove that the class LOC_ω meets all finite levels of the Borel hierarchy. Then we show that there exist some locally finite ω-languages which are Borel sets of infinite rank, and some others which are analytic but non-Borel sets.

2 Review of local sentences and languages

2.1 Definitions and properties of local sentences

In this paper the (first order) signatures are finite, always contain one binary predicate symbol $=$ for equality, and can contain both functional and relational symbols. The terms, open formulas and formulas are built in the usual way.

When M is a structure in a signature Λ and $X \subseteq |M|$, we define:
$cl^1(X, M) = X \cup \bigcup_{\{f \text{ n-ary function of } \Lambda\}} f^M(X^n) \cup \bigcup_{\{a \text{ constant of } \Lambda\}} a^M$
$cl^{n+1}(X, M) = cl^1(cl^n(X, M), M)$ 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 local sentences. We shall denote S(φ) the signature of a first order sentence φ, i.e. the set of non-logical symbols appearing in φ.

Definition 2.1 A first order sentence φ is local 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 φ takes at most n steps).

Notation For a local sentence φ, let n_φ be the smallest integer $n \geq 1$ verifying b of the above definition.

Remark 2.2 Because of a of Definition 2.1, a local sentence φ is always equivalent to a universal sentence, so we may assume that φ is universal.

Let us now state first properties of local sentences.

Theorem 2.3 (a) *The set of local sentences is recursively enumerable.*
(b) *It is undecidable whether an arbitrary sentence φ is a local one.*
(c) *It is undecidable whether an arbitrary universal sentence φ is a local one.*
(d) *It is undecidable whether an arbitrary universal sentence φ, such that S(φ) contains only two unary function symbols, is a local one.*
(e) *It is undecidable whether an arbitrary universal sentence φ, such that S(φ) contains only one binary function symbol, is a local one.*

Items (a) and (b) are results of Ressayre, see [15]. The proof of item (b) relies on Church's Theorem: it is undecidable to determine whether an arbitrary first order sentence φ is consistent. But one can prove in the same way items (c), (d), and (e) because it is undecidable to determine whether an arbitrary universal first order sentence φ is consistent, even if we assume that the signature of φ contains only two unary function symbols or one binary function symbol [1].

Per contra to these negative results, there exists a "recursive presentation" up to logical equivalence of all local sentences.

Theorem 2.4 (Ressayre, see [15]) *There exist a recursive set* $\mathbf{L}$ *of local sentences and a recursive function* $\mathbf{F}$ *such that*:

(1) ψ *local* $\longleftrightarrow \exists \psi' \in \mathbf{L}$ *such that* $\psi \equiv \psi'$.
(2) $\psi' \in \mathbf{L} \longrightarrow n_{\psi'} = \mathbf{F}(\psi')$.

The elements of **L** are the $\psi \wedge C_n$, where ψ run over the universal formulas and C_n run over the universal formulas in the signature $S(\psi)$ which express that closure in a model takes at most n steps. $\psi \wedge C_n$ is local and $n_{\psi \wedge C_n} \leq n$. Then we can compute $n_{\psi \wedge C_n}$, considering only finite models of cardinal $\leq m$, where m is an integer depending on n. And each local sentence ψ is equivalent to a universal formula θ, hence $\psi \equiv \theta \wedge C_{n_\psi}$.

We shall restrict now our attention to local sentences with a binary predicate $<$ in their signature which is interpreted by a linear ordering in all of their models.

Let us now recall a fundamental result, the stretching theorem for local sentences, which shows the existence of remarkable connections between the finite and the infinite model theory of local sentences. Below, semi-monotonic, special, and monotonic indiscernibles are particular kinds of indiscernibles which satisfy some extra properties; they are precisely defined in [14].

Theorem 2.5 ([14]) *For each local sentence φ there exists a positive integer N_φ, which can be effectively computed, such that*

(A) *φ has arbitrarily large finite models if and only if φ has an infinite model if and only if φ has a finite model generated by N_φ indiscernibles.*
(B) *φ has an infinite well ordered model if and only if φ has a finite model generated by N_φ semi-monotonic indiscernibles.*
(C) *φ has a model of order type ω if and only if φ has a finite model generated by N_φ special indiscernibles.*
(D) *φ has well ordered models of unbounded order types in the ordinals if and only if φ has a finite model generated by N_φ monotonic indiscernibles.*

Remark 2.6 In the above theorem the integer N_φ can be effectively computed from n_φ and q where $\varphi = \forall x_1 \ldots \forall x_q \theta(x_1, \ldots, x_q)$ and θ is an open formula. Let $v(\varphi)$ be the maximum number of variables of terms of complexity $\leq n_\varphi + 1$ and $v'(\varphi)$ be the maximum number of variables of an atomic formula involving terms of complexity $\leq n_\varphi + 1$ then

$$N_\varphi = \max\{3v(\varphi); v'(\varphi) + v(\varphi); q.v'(\varphi)\}$$

Thus the stretching theorem implies the existence of decision procedures for several problems. Let us remark that the set of local sentences is not recursive but we can consider that the algorithms given by the following theorem are applied to local sentences in the recursive set **L** given by Proposition 2.4. In particular φ is given with the integer n_φ.

Theorem 2.7 ([14]) *It is decidable, for a given local sentence φ, whether*

(1) *φ has arbitrarily large finite models.*
(2) *φ has an infinite model.*
(3) *φ has an infinite well ordered model.*
(4) *φ has a model of order type ω.*
(5) *φ has well ordered models of unbounded order types in the ordinals.*

Remark 2.8 As indicated by the referee of this paper, " the above theorem is still true even the local sentences were not assumed to be in the recursive set **L**. Given an arbitrary local sentence, the algorithm could begin by searching for an equivalent sentence in **L** (together with a formal proof of the equivalence) and then, when it finds one, apply the algorithm to this sentence in **L**. Of course this would be only a partial recursive function, defined on the class of local sentences, and its complexity would be much worse than the complexity given below, but it is still an algorithm."

Theorem 2.7 follows directly from the stretching Theorem 2.5. For instance Theorem 2.5 (C) states that a local sentence φ has a model of order type ω iff it has a *finite* model generated by N_φ special indiscernibles, where N_φ is a positive integer effectively computable from φ and n_φ. Thus the existence of a model of order type ω of φ can be checked by considering only models whose cardinals are bounded by an integer depending on n_φ and N_φ (because closure in models of φ takes at most n_φ steps). A similar argument is used to prove other items of Theorem 2.7.

The question of the complexity of these decidable problems naturally arises. It is easy to see that the problems (1)–(5) which are shown to be decidable by Theorem 2.7 are in the class

$$\mathbf{NTIME}(\mathbf{2}^{\mathbf{O(n.\log(n))}})$$

when the algorithms work with input (φ, N_φ). Using non-determinism a Turing machine may guess a finite structure M of signature $\mathrm{S}(\varphi)$ generated in at most n_φ steps by N_φ elements $y_1, \ldots y_{N_\varphi}$. Then, assuming $\varphi = \forall x_1 \ldots \forall x_q \theta(x_1, \ldots, x_q)$ where θ is an open formula, the Turing machine checks that $\theta(x_1, \ldots, x_q)$ holds for all $x_1 \ldots x_q$ in M, and that the elements $y_1, \ldots y_{N_\varphi}$ are indiscernibles (respectively, semi-monotonic, special, monotonic, indiscernibles) in M.

On the other side Büchi showed that one can decide whether a monadic second order formula of $S1S$ is true in the structure $(\omega, <)$. But for a formula of size n his procedure might run in time

$$\underbrace{2^{2^{\cdot^{\cdot^{2^n}}}}}_{\mathbf{O(n)}}$$

see [4,34] for more details. Moreover it has been proved by Meyer that one cannot essentially improve this result: the monadic second order theory of the structure $(\omega, <)$ is not elementary recursive [27].

Notice that the complexity of Büchi's algorithm for monadic sentences is in terms of the length of the formula and the complexity of the algorithms for local sentences is in terms of the length of a local sentence φ *and* the corresponding integer N_φ. But a sentence in **L** is of the form $\varphi = \psi \wedge C_n$, where ψ is a universal sentence and C_n is a universal sentence in the signature $S(\psi)$ which expresses that closure in a model takes at most n steps. The length of C_n is greater than n and $n_\varphi = n_{\psi \wedge C_n} \leq n$. So $n_\varphi \leq |\varphi|$ where $|\varphi|$ is the length of φ and we can easily get from the equality given in Remark 2.6 that $N_\varphi = \mathbf{O}(|\varphi|^3)$.

Thus the algorithms for local sentences given by Theorem 2.7 are of much lower complexity than the algorithm for decidability of $S1S$. This is remarkable because the expressive power of local sentences is also greater than the expressive power of monadic second order sentences.

Recall also that there is an extension of item (C) of the stretching Theorem 2.5 for ordinals $\alpha < \omega^\omega$ from which we can infer other decidability results.

Theorem 2.9 ([14]) *To every local sentence φ and every ordinal α such that $\omega \leq \alpha < \omega^\omega$ one can associate by an effective procedure a local sentence φ_α, a unary predicate symbol P being in the signature $S(\varphi_\alpha)$, such that the following equivalence holds*:

(C_α) *φ has a well ordered model of order type α if and only if φ_α has a finite model M generated by N_{φ_α} semi-monotonic indiscernibles in P^M.*

Theorem 2.10 ([14]) *It is decidable, for a given local sentence φ and a given ordinal $\alpha < \omega^\omega$, whether φ has a model of order type α.*

There are also other variations of the stretching theorem involving large cardinal axioms, see [14].

2.2 Definitions and first properties of local languages

Let us now introduce notations for words. Let Σ be a finite alphabet whose elements are called letters. A finite non-empty word over Σ is a finite sequence of letters: $x = a_1a_2 \ldots a_n$ where $\forall i \in [1; n]\ a_i \in \Sigma$. We shall denote $x(i) = a_i$ the i^{th} letter of x and $x[i] = x(1) \ldots x(i)$ for $i \leq n$. The length of x is $|x| = n$. The empty word will be denoted by λ and has 0 letters. Its length is 0. The set of finite words over Σ is denoted $\Sigma^\star$. $\Sigma^+ = \Sigma^\star - \{\lambda\}$ is the set of non-empty words over Σ. A (finitary) language L over Σ is a subset of $\Sigma^\star$. Its complement (in $\Sigma^\star$) is $L^- = \Sigma^\star - L$. The usual concatenation product of u and v will be denoted by $u.v$ or just uv. For $V \subseteq \Sigma^\star$, we denote $V^\star = \{v_1 \ldots v_n \mid n \in \mathbb{N} \ \ and \ \ \forall i \in [1; n] \ \ v_i \in V\}$.

The first infinite ordinal is ω. An ω-word over Σ is an ω -sequence $a_1a_2 \ldots a_n \ldots$, where $\forall i \geq 1\ a_i \in \Sigma$. When σ is an ω-word over σ, we write $\sigma = \sigma(1)\sigma(2) \ldots \sigma(n) \ldots$ and $\sigma[n] = \sigma(1)\sigma(2) \ldots \sigma(n)$ the finite word of length n, prefix of σ. The set of ω-words over the alphabet Σ is denoted by Σ^ω. An ω-language over an alphabet Σ is a subset of Σ^ω. For $V \subseteq \Sigma^\star$, $V^\omega = \{\sigma = u_1 \ldots u_n \ldots \in \Sigma^\omega \mid \forall i \geq 1\ u_i \in V\}$ is the ω-power of V. For a subset $A \subseteq \Sigma^\omega$, the complement of A (in Σ^ω) is $\Sigma^\omega - A$ denoted A^-. The concatenation product is extended to the product of a finite word u and an ω-word v: the infinite word $u.v$ is then the ω-word such that: $(u.v)(k) = u(k)$ if $k \leq |u|$, and $(u.v)(k) = v(k - |u|)$ if $k > |u|$.

The prefix relation is denoted $\sqsubseteq$: the finite word u is a prefix of the finite word v (respectively, the infinite word v), denoted $u \sqsubseteq v$, if and only if there exists a finite word w (respectively, an infinite word w), such that $v = u.w$.

A word over Σ may be considered as a structure in the following usual manner: Let Σ be a finite alphabet. We denote P_a a unary predicate for each letter $a \in \Sigma$ and Λ_Σ the signature $\{<, (P_a)_{a\in\Sigma}\}$. Let σ be a finite word over the alphabet Σ, $|\sigma|$ is the length of the word σ. We may write that $|\sigma| = \{1, 2, \ldots, |\sigma|\}$. σ is identified with the structure $(|\sigma|, <^\sigma, (P_a^\sigma)_{a\in\Sigma})$ of signature Λ_Σ where $P_a^\sigma = \{1 \leq i \leq |\sigma| \mid$ the i^{th}letter of σ is an $a\}$. In a similar manner if σ is an ω-word over the alphabet

Σ, then ω is the length of the word σ and we may write $|\sigma| = \{1, 2, 3, \ldots\}$. σ is identified to the structure $(|\sigma|, <^\sigma, (P_a^\sigma)_{a\in\sigma})$ of signature Λ_Σ where $P_a^\sigma = \{1 \leq i < \omega \mid \text{ the } i^{th} \text{ letter of } \sigma \text{ is an } a\}$.

Definition 2.11 Let Σ be a finite alphabet and $L \subseteq \Sigma^\star$ be a language of finite words (respectively, $L \subseteq \Sigma^\omega$ be a language of infinite words) over the alphabet Σ. Then L is a locally finite language (respectively, ω-language) $\longleftrightarrow$ there exists a local sentence φ in a signature $\Lambda \supseteq \Lambda_\Sigma$ such that $\sigma \in L$ iff $\exists$ finite M, (respectively, $\exists M$ of order type ω) $M \models \varphi$ and $M|\Lambda_\Sigma = \sigma$ (where $M|\Lambda_\Sigma$ is the reduction of M to the signature Λ_Σ).

We then denote $L = L^\Sigma(\varphi)$ (respectively, $L = L_\omega^\Sigma(\varphi)$), and to simplify, when there is no ambiguity, $L = L(\varphi)$ (respectively, $L = L_\omega(\varphi)$) the locally finite language (respectively, ω-language) defined by φ. The class of locally finite languages will be denoted LOC. The class of locally finite ω-languages will be denoted LOC_ω.

The empty word λ 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 [15].

Remark 2.12 The notion of locally finite language is very different from the usual notion of local language which represents a subclass of the class of rational languages. But from now on, as in [15], things being well defined and made precise, we shall call simply local languages the locally finite languages.

Let us state the following decidability results.

Theorem 2.13 *It is decidable, for a local sentence φ, given with the integer n_φ, and an alphabet Σ, whether*

(1) *The local language $L^\Sigma(\varphi)$ is empty.*
(2) *The local language $L^\Sigma(\varphi)$ is infinite.*
(3) *The local ω-language $L_\omega^\Sigma(\varphi)$ is empty.*

(1) follows directly from the fact that if a local sentence φ has a finite model then it has a model whose cardinal is bounded by a positive integer depending only on arities of the function symbols of the signature of φ and on n_φ.
(2) and (3) follows items (1) and (4) of Theorem 2.7.
(3) states that the emptiness problem for local ω-languages is decidable. It relies on a remarkable analogue to the property: “a Büchi language is non-empty iff it contains an ultimately periodic word, i.e. an ω-word in the form $u.v^\omega$ where u and v are finite words.”

When local ω-languages are considered, this property becomes: “a local ω-language is non-empty iff it contains an ω-word which is the reduction, to the signature of words, of an ω-model generated by special indiscernibles.”

2.3 Examples of local ω-languages

Example 2.14 ([18]) The ω-language which contains only the word $\sigma = abab^2ab^3ab^4\ldots$ is a local ω-language over the alphabet $\{a, b\}$.

Recall that for any family L of finitary languages, the ω-Kleene closure of L, is:

$$\omega-KC(L) = \left\{ \bigcup_{1 \le i \le n} U_i.V_i^{\omega} \mid \forall i \in [1, n]\ U_i, V_i \in L \right\}.$$

It is well known that the class REG_ω of regular ω-languages (respectively, the class CF_ω of context free ω-languages) is the ω-Kleene closure of the family REG of regular finitary languages (respectively, of the family CF of context free finitary languages) [38,39].

We showed that a similar characterization does not hold for local languages.

Theorem 2.15 ([18]) *The ω-Kleene closure of the class LOC of finitary local languages is strictly included in the class LOC_ω of local ω-languages.*

Then we easily derive the following example because every regular finitary language is local [32].

Example 2.16 ([15]) Every regular ω-language is a local ω-language, i.e. $\mathrm{REG}_\omega \subseteq \mathrm{LOC}_\omega$.

Since numerous context free languages are local [15], $CF_\omega = \omega-KC(CF)$ implies that many context free ω-languages are local. The problem whether every context free ω-language is local is still open but by Theorem 2.15, $CF \subseteq \mathrm{LOC}$ would imply that $CF_\omega \subseteq \mathrm{LOC}_\omega$.

Example 2.17 The ω-languages U^ω and $U.a^\omega$, where $U = \{a^{n^2} b^{n^2} c^{n^2} \mid n \geq 1\}$ is a local finitary language over the alphabet $\{a, b, c\}$ [15], are examples of local but non context free ω-languages.

Example 2.18 ([18]) The ω-language $L = \{0^n 1^p 2^\omega \mid p \leq 2^n\}$ over the alphabet $\Sigma = \{0, 1, 2\}$ is local because the finitary language $\{0^n 1^p \mid p \leq 2^n\}$ is local [15]. But the ω-language $A = \{0^n 1^p 2^\omega \mid p > 2^n\}$ over the same alphabet Σ is not local [18]. From this we can easily deduce that the complement of L is not a local ω-language.

We shall construct some other local ω-languages in the sequel, see for example the construction of local ω-languages which are Borel of infinite rank in Sect. 3.3, or analytic but non Borel in Sect. 3.4.

Now we recall some closure properties of the class LOC_ω which allow us to generate many other local ω-languages from the known ones. The class LOC_ω is closed under union, left concatenation with local finitary languages, λ-free substitution of local (finitary) languages, λ-free morphism [18].

3 Topological complexity of local ω-languages

3.1 Borel and projective hierarchies

We assume the reader to be familiar with basic notions of topology which may be found in [22,23,28].

Topology is an important tool for the study of subsets of a set Σ^ω, where Σ is a finite or infinite set. We study here local ω-languages which are defined over a finite alphabet. Thus we shall restrict our study to subsets of spaces in the form Σ^ω, where Σ is a finite set, called here an alphabet, having at least two elements (because the case of an alphabet having a single letter is trivial). We shall consider Σ^ω as a topological space with the Cantor topology. The open sets of Σ^ω are the sets in the form $W.\Sigma^\omega$, where $W \subseteq \Sigma^\star$.

Define now the following classes of the Borel Hierarchy:

Definition 3.1 The classes $\mathbf{\Sigma}^0_n$ and $\mathbf{\Pi}^0_n$ of the Borel Hierarchy on the topological space Σ^ω are defined as follows:

$\mathbf{\Sigma}^0_1$ is the class of open subsets of Σ^ω.

$\mathbf{\Pi}^0_1$ is the class of closed subsets of Σ^ω.

And for any integer $n \geq 1$:

$\mathbf{\Sigma}^0_{n+1}$ is the class of countable unions of $\mathbf{\Pi}^0_n$-subsets of Σ^ω.

$\mathbf{\Pi}^0_{n+1}$ is the class of countable intersections of $\mathbf{\Sigma}^0_n$-subsets of Σ^ω.

The Borel Hierarchy is also defined for transfinite levels. The classes $\mathbf{\Sigma}^0_\alpha$ and $\mathbf{\Pi}^0_\alpha$, for a countable ordinal α, are defined in the following way:

$\mathbf{\Sigma}^0_\alpha$ is the class of countable unions of subsets of Σ^ω in $\bigcup_{\gamma<\alpha} \mathbf{\Pi}^0_\gamma$.

$\mathbf{\Pi}^0_\alpha$ is the class of countable intersections of subsets of Σ^ω in $\bigcup_{\gamma<\alpha} \mathbf{\Sigma}^0_\gamma$.

Recall some basic results about these classes [28]:

Theorem 3.2

(a) $\mathbf{\Sigma}^0_\alpha \cup \mathbf{\Pi}^0_\alpha \subsetneq \mathbf{\Sigma}^0_{\alpha+1} \cap \mathbf{\Pi}^0_{\alpha+1}$, *for each countable ordinal* $\alpha \geq 1$.
(b) $\bigcup_{\gamma<\alpha} \mathbf{\Sigma}^0_\gamma = \bigcup_{\gamma<\alpha} \mathbf{\Pi}^0_\gamma \subsetneq \mathbf{\Sigma}^0_\alpha \cap \mathbf{\Pi}^0_\alpha$, *for each countable limit ordinal* α.
(c) *A set* $W \subseteq \Sigma^\omega$ *is in the class* $\mathbf{\Sigma}^0_\alpha$ *iff its complement is in the class* $\mathbf{\Pi}^0_\alpha$.
(d) $\mathbf{\Sigma}^0_\alpha - \mathbf{\Pi}^0_\alpha \neq \emptyset$ *and* $\mathbf{\Pi}^0_\alpha - \mathbf{\Sigma}^0_\alpha \neq \emptyset$ *for every countable ordinal* $\alpha \geq 1$.

We shall say that a subset of Σ^ω is a Borel set of rank α, for a countable ordinal α, iff it is in $\mathbf{\Sigma}^0_\alpha \cup \mathbf{\Pi}^0_\alpha$ but not in $\bigcup_{\gamma<\alpha}(\mathbf{\Sigma}^0_\gamma \cup \mathbf{\Pi}^0_\gamma)$.

The class of Borel subsets of Σ^ω is strictly included in the class of analytic subsets of Σ^ω which we now define.

Definition 3.3 A subset A of Σ^ω is in the class $\mathbf{\Sigma}^1_1$ of **analytic** sets iff there exists another finite set Y and a Borel subset B of $(\Sigma \times Y)^\omega$ such that $x \in A \leftrightarrow \exists y \in Y^\omega$ such that $(x, y) \in B$, where (x, y) is the infinite word over the alphabet $\Sigma \times Y$ such that $(x, y)(i) = (x(i), y(i))$ for each integer $i \geq 1$.

Remark 3.4 In the above definition we could take B in the class $\mathbf{\Pi}^0_2$. Moreover analytic subsets of Σ^ω are the projections of $\mathbf{\Pi}^0_1$-subsets of $\Sigma^\omega \times \omega^\omega$, where ω^ω is the Baire space [28].

Recall that a set $F \subseteq \Sigma^\omega$ is said to be a $\mathbf{\Sigma}^0_\alpha$ (respectively, $\mathbf{\Pi}^0_\alpha$, $\mathbf{\Sigma}^1_1$)-complete set iff for any set $E \subseteq Y^\omega$, E is in $\mathbf{\Sigma}^0_\alpha$ (respectively, $\mathbf{\Pi}^0_\alpha$, $\mathbf{\Sigma}^1_1$) iff there exists a continuous function $f : Y^\omega \to \Sigma^\omega$, such that $E = f^{-1}(F)$.

Let us now recall the definition of the arithmetical hierarchy of ω-languages, see for example [38] or [28]. Let Σ be a finite alphabet. An ω-language $L \subseteq \Sigma^\omega$ belongs to the class Σ_n if and only if there exists a recursive relation $R_L \subseteq (\mathbb{N})^{n-1} \times \Sigma^\star$ such that

$$L = \{\sigma \in \Sigma^\omega \mid \exists a_1 \ldots Q_n a_n \quad (a_1, \ldots, a_{n-1}, \sigma[a_n + 1]) \in R_L\}$$

where Q_i is one of the quantifiers $\forall$ or $\exists$ (not necessarily in an alternating order). An ω-language $L \subseteq \Sigma^\omega$ belongs to the class Π_n if and only if its complement $\Sigma^\omega - L$ belongs to the class Σ_n. The inclusion relations that hold between the classes Σ_n and Π_n are the same as for the corresponding classes of the Borel hierarchy and the classes Σ_n and Π_n are strictly included in the respective classes $\mathbf{\Sigma}^0_n$ and $\mathbf{\Pi}^0_n$ of the Borel hierarchy.

As in the case of the Borel hierarchy, projections of arithmetical sets (of the second Π-class) lead beyond the arithmetical hierarchy, to the analytical hierarchy of ω-languages. The first class of the analytical hierarchy of ω-languages is the class Σ^1_1 (lightface). An ω-language $L \subseteq \Sigma^\omega$ belongs to the class Σ^1_1 if and only if there exists a recursive relation $R_L \subseteq (\mathbb{N}) \times \{0, 1\}^\star \times \Sigma^\star$ such that:

$$L = \{\sigma \in \Sigma^\omega \mid \exists \tau (\tau \in \{0, 1\}^\omega \wedge \forall n \exists m ((n, \tau[m], \sigma[m]) \in R_L))\}.$$

Thus an ω-language $L \subseteq \Sigma^\omega$ is in the class Σ^1_1 iff it is the projection of an ω-language over the alphabet $\{0, 1\} \times \Sigma$ which is in the class Π_2 of the arithmetical hierarchy.

Remark 3.5 Σ^1_1-subsets of Σ^ω are also projections of Π_1-subsets of $\Sigma^\omega \times \omega^\omega$, where ω^ω is the Baire space [28].

It turns out that an ω-language $L \subseteq \Sigma^\omega$ is in the class Σ^1_1 iff it is accepted by a non deterministic Turing machine reading ω-words with a Muller acceptance condition. (A Turing machine $\mathcal{T}$ is given with a set $\mathcal{F}$ of designated state sets which are particular subsets of its finite set K of states; then an ω-word σ is accepted by $\mathcal{T}$ iff there exists a run of $\mathcal{T}$ reading σ for which the set of states entered infinitely often by $\mathcal{T}$ during this run is in $\mathcal{F}$). This class is denoted $NT(\inf, =)$ (where $(\inf, =)$ indicates the Muller condition) in [38] and also called the class of recursive ω-languages REK_ω.[1]

With the above definitions, we can state the following:

Theorem 3.6 *The class LOC_ω is strictly included in the class Σ^1_1.*

Proof Let $L^\Sigma_\omega(\varphi)$ be a local ω-language defined by the local sentence φ. We may replace the constant and function symbols of $S(\varphi)$ by relation symbols in a usual

[1] In another presentation, as in [33], the recursive ω-languages are those which are in the intersection $\Sigma_1 \cap \Pi_1$.

manner. For example we replace an n-ary function f by a (n+1)-ary relation R_f and we express by a Π_2^0 formula that the relation R_f is functional:

$$\forall x_1 \dots x_n z z' \exists y [R_f(x_1, \dots, x_n, y) \wedge (R_f(x_1, \dots, x_n, z) \\ \wedge R_f(x_1, \dots, x_n, z') \rightarrow z = z')].$$

Then from φ we obtain another first order sentence which is not universal and not local but which defines the same ω-language when reductions of models to the signature Λ_Σ of words are considered. Let us call $\psi(R_1, \dots, R_k)$ the resulting first order sentence in the signature $\Lambda_\Sigma \cup \{R_1, \dots, R_k\}$ where $R_1, \dots, R_k$ are relation symbols of arities $n_1, \dots, n_k$.

An ω-model of $\psi(R_1, \dots, R_k)$ may be viewed as an element of:

$$\Sigma^\omega \times 2^{\omega^{n_1}} \times 2^{\omega^{n_2}} \times \cdots \times 2^{\omega^{n_k}}$$

because any n-ary relation R over ω can be identified with its characteristic function, i.e. a function $\omega^n \rightarrow 2 = \{0, 1\}$ which associates 1 to an n-tuple $(x_1, \dots, x_n)$ iff $R(x_1, \dots, x_n)$.

But $\Sigma^\omega \times 2^{\omega^{n_1}} \times 2^{\omega^{n_2}} \times \dots \times 2^{\omega^{n_k}}$ is a classical recursively presented Polish space (generalizing Σ^ω) and it is well known [28] that a subset of this space which is defined by a first order sentence where the quantifiers run only over the integers of ω is an arithmetical subset of $\Sigma^\omega \times 2^{\omega^{n_1}} \times 2^{\omega^{n_2}} \times \cdots \times 2^{\omega^{n_k}}$.

And $L_\omega^\Sigma(\varphi)$ is the projection of this arithmetical set onto Σ^ω and it is well known that such a projection of an arithmetical set is a Σ_1^1-subset of Σ^ω.

Remark 3.7 Another way to show this result is to consider a non deterministic Turing machine $\mathcal{T}$ which accepts $L_\omega(\varphi)$. Let then σ be an ω-word over Σ. The non determinism of $\mathcal{T}$ is used to guess an expansion of the word σ (considered as a structure of signature Λ_Σ) to a structure in the signature $S(\varphi)$ which is coded by an ω-word. Then the Turing machine checks whether this expansion is a model of φ. This can be checked with a Muller acceptance condition. If such a model exists, the word σ is in $L_\omega(\varphi)$. And if no such model exists, the word σ is not in $L_\omega(\varphi)$. Then an ω-word σ over Σ is in $L_\omega(\varphi)$ iff there exists an accepting run of $\mathcal{T}$ on σ.

The strictness of the inclusion is easy to prove. The ω-language $A = \{0^n 1^p 2^\omega \mid p > 2^n\}$ over the alphabet $\Sigma = \{0, 1, 2\}$, given in Example 2.18, is not local but it is easily shown to be in the class Σ_1^1 and even in the class Σ_2^0. □

The inclusion $\Sigma_1^1 \subset \mathbf{\Sigma}_1^1$ is trivial and well known. Thus, when studying local ω-languages, we shall not have to consider non $\mathbf{\Sigma}_1^1$-sets.

Corollary 3.8 *Every local ω-language over a finite alphabet Σ is an analytic subset of Σ^ω.*

By Suslin's Theorem [22, p. 226], an analytic subset of Σ^ω is either countable or has the continuum power. Then we can infer the following:

Corollary 3.9 *Let Σ be a finite alphabet. Every local ω-language $L_\omega^\Sigma(\varphi)$ over the alphabet Σ is either countable or has the continuum power.*

3.2 Borel sets of finite rank and local ω-languages

We shall prove that the class LOC_ω meets all finite levels of the Borel hierarchy. The proof is very similar to our corresponding proof for the class of context free ω-languages in [16]. We shall use recent results of Duparc who studied the Wadge hierarchy which is a great refinement of the Borel hierarchy. He gave an inductive construction of a Borel set of every given degree of this hierarchy, introducing operations over sets of finite *or* infinite words over an alphabet Σ, called conciliating sets in [7,9]. So we shall sometimes consider subsets of $\Sigma^\star \cup \Sigma^\omega = \Sigma^{\leq\omega}$, for an alphabet Σ, and the correspondence $A \to A^d$ where for $A \subseteq \Sigma^{\leq\omega}$ and d a letter not in Σ:

$$A^d = \{x \in (\Sigma \cup \{d\})^\omega \mid x(/d) \in A\}$$

where $x(/d)$ is the sequence obtained from x when removing every occurrence of the letter d.

We shall only use in this paper Duparc's operation of exponentiation:

$$A \to A^\sim$$

which produces some sets of higher complexity, in the following sense:

Theorem 3.10 (Duparc [9]) *Let n be an integer ≥ 1 and $A \subseteq \Sigma^{\leq\omega}$. If $A^d \subseteq (\Sigma \cup \{d\})^\omega$ is a $\mathbf{\Sigma}^0_n$-complete (respectively, $\mathbf{\Pi}^0_n$-complete) set then $(A^\sim)^d$ is a $\mathbf{\Sigma}^0_{n+1}$-complete (respectively, $\mathbf{\Pi}^0_{n+1}$-complete) set.*

Let us now introduce Duparc's operation of exponentiation on sets.

Definition 3.11 Let Σ be a finite alphabet and $\twoheadleftarrow \notin \Sigma$, let $X = \Sigma \cup \{\twoheadleftarrow\}$. Let x be a finite or infinite word over the alphabet $X = \Sigma \cup \{\twoheadleftarrow\}$. Then $x^\twoheadleftarrow$ is inductively defined by:

$\lambda^\twoheadleftarrow = \lambda$,
and for a finite word $u \in (\Sigma \cup \{\twoheadleftarrow\})^\star$:
$(u.a)^\twoheadleftarrow = u^\twoheadleftarrow.a$, if $a \in \Sigma$,
$(u.\twoheadleftarrow)^\twoheadleftarrow = u^\twoheadleftarrow$ with its last letter removed, if $|u^\twoheadleftarrow| > 0$,
$(u.\twoheadleftarrow)^\twoheadleftarrow = \lambda$, if $|u^\twoheadleftarrow| = 0$,
and for u infinite:
$(u)^\twoheadleftarrow = \lim_{n\in\omega}(u[n])^\twoheadleftarrow$, where, given β_n and v in $\Sigma^\star$,
$v \sqsubseteq \lim_{n\in\omega} \beta_n \leftrightarrow \exists n \forall p \geq n \quad \beta_p[|v|] = v$.

Remark 3.12 For $x \in X^{\leq\omega}$, $x^\twoheadleftarrow$ denotes the string x, once every $\twoheadleftarrow$ occurring in x has been "evaluated" to the back space operation (the one familiar to your computer!), proceeding from left to right inside x. In other words $x^\twoheadleftarrow = x$ from which every interval of the form "$a \twoheadleftarrow$" ($a \in \Sigma$) is removed.

For example if $u = (a \twoheadleftarrow)^n$, for n an integer ≥ 1, or $u = (a \twoheadleftarrow)^\omega$, or $u = (a \twoheadleftarrow\twoheadleftarrow)^\omega$, then $(u)^\twoheadleftarrow = \lambda$. If $u = (ab \twoheadleftarrow)^\omega$ then $(u)^\twoheadleftarrow = a^\omega$ and if $u = bb(\twoheadleftarrow a)^\omega$ then $(u)^\twoheadleftarrow = b$.

We define now the operation $A \to A^{\sim}$ of exponentiation of conciliating sets:

Definition 3.13 For $A \subseteq \Sigma^{\leq\omega}$ and $\twoheadleftarrow \notin \Sigma$, let

$$A^{\sim} = \{x \in (\Sigma \cup \{\twoheadleftarrow\})^{\leq\omega} \mid x^{\twoheadleftarrow} \in A\}.$$

We now prove that the class LOC_ω is closed under this operation $\sim$.

Proposition 3.14 *If $A \subseteq \Sigma^\omega$ is in LOC_ω, then $A^{\sim} \subseteq (\Sigma \cup \{\twoheadleftarrow\})^\omega$ is also in LOC_ω.*

Proof We remark that an ω-word $\sigma \in A^{\sim}$ may be considered as an ω-word $\sigma^{\twoheadleftarrow} \in A$ to which we possibly add, before the first letter $\sigma^{\twoheadleftarrow}(1)$ of $\sigma^{\twoheadleftarrow}$ (respectively, between two consecutive letters $\sigma^{\twoheadleftarrow}(n)$ and $\sigma^{\twoheadleftarrow}(n+1)$ of $\sigma^{\twoheadleftarrow}$), a finite word w_1 (respectively, w_{n+1}) where:
w_{n+1} belongs to the context free (finitary) language $\mathcal{C}_1$ generated by the context free grammar with the following production rules:
$S \to aS \twoheadleftarrow S$ with $a \in \Sigma$ and $S \to \lambda$ where λ is the empty word.
This language $\mathcal{C}_1$ corresponds to words where every letter of Σ has been erased after using the back space operation. And w_1 belongs to the finitary language $\mathcal{C}_2 = (\mathcal{C}_1.(\twoheadleftarrow)^\star)^\star$. This language corresponds to words where every letter of Σ has been removed after using the back space operation and this operation may be has been used also when there was not any letter to erase. Then for $A \subseteq \Sigma^\omega$, the ω-language $A^{\sim} \subseteq (\Sigma \cup \{\twoheadleftarrow\})^\omega$ is obtained by substituting in A the language $a.\mathcal{C}_1$ for each letter $a \in \Sigma$, and then making a left concatenation by the language $\mathcal{C}_2$.

Now we easily show that the language $\mathcal{C}_1$ is local, defined by the following sentence φ in the signature $S(\varphi) = \{<, (P_a)_{a\in(\Sigma\cup\{\twoheadleftarrow\})}, s\}$, where s is a unary function symbol. φ is the conjunction of:

- $\forall xyz[(x \leq y \vee y \leq x) \wedge ((x \leq y \wedge y \leq x) \leftrightarrow x = y) \wedge ((x \leq y \wedge y \leq z) \to x \leq z)]$ (this means: " $<$ is a linear order "),
- $\forall x[(\bigvee_{a\in(\Sigma\cup\{\twoheadleftarrow\})} P_a(x)) \wedge (\bigwedge_{(a,a')\in(\Sigma\cup\{\twoheadleftarrow\})^2, a\neq a'} \neg(P_a(x) \wedge P_{a'}(x)))]$ (this means: "$(P_a)_{a\in(\Sigma\cup\{\twoheadleftarrow\})}$ form a partition "),
- $\forall x[P_a(x) \to (x < s(x) \wedge P_{\twoheadleftarrow}(s(x))]$, for each $a \in \Sigma$,
- $\forall xz[(P_a(x) \wedge P_b(z) \wedge x < z) \to (s(x) < z \vee s(z) < s(x))]$, for all $a, b \in \Sigma$,
- $\forall x[(\bigvee_{a\in\Sigma} P_a(x)) \leftrightarrow P_{\twoheadleftarrow}(s(x))]$,
- $\forall x[s(s(x)) = x]$.

φ is equivalent to a universal formula and closure in its models takes only one step because $\varphi \to \forall x[s(s(x)) = x]$. Then φ is a local sentence and we easily check that $L(\varphi) = \mathcal{C}_1$ (the function s is used to associate a letter $a \in \Sigma$ with the eraser $\twoheadleftarrow$ which erases a). Hence $\mathcal{C}_1$ is a local language and so is $a.\mathcal{C}_1$ for $a \in \Sigma$. But $\mathcal{C}_2 = (\mathcal{C}_1.(\twoheadleftarrow)^\star)^\star$ and the class LOC is closed under concatenation product and star operation [15]. Thus the language $\mathcal{C}_2$ is also local.

LOC_ω is closed under substitution of local finitary languages and left concatenation by local finitary languages [18], therefore if $A \subseteq \Sigma^\omega$ is a local ω-language then the ω-language $A^{\sim}$ is a local ω-language. □

Consider now subsets of $\Sigma^{\leq\omega}$ in the form $A \cup B$, where $A = L^\Sigma(\varphi)$ is a local finitary language and $B = L^\Sigma_\omega(\psi)$ is a local ω-language. Remark that A and B might not be defined by the same sentence. Let us prove the following:

Proposition 3.15 *If $C = A \cup B$, where $A \subseteq \Sigma^\star$ is in LOC and $B \subseteq \Sigma^\omega$ is in LOC_ω, then $C^\sim$ is also the union of a local finitary language and a local ω-language over the alphabet $\Sigma \cup \{\twoheadleftarrow\}$.*

Proof Let $A \subseteq \Sigma^\star$ be a local finitary language and let $B \subseteq \Sigma^\omega$ be a local ω-language. It follows from the definition of the operation $A \rightarrow A^\sim$ that if $C = A \cup B$ then $C^\sim$=$A^\sim \cup B^\sim$. But if $B = L_\omega^\Sigma(\psi)$, where ψ is a local sentence, then, by Proposition 3.14, there exists a local sentence ψ_1 such that $B^\sim = L_\omega^{\Sigma\cup\{\twoheadleftarrow\}}(\psi_1)$.

Consider now the set $A^\sim \subseteq (\Sigma \cup \{\twoheadleftarrow\})^{\leq\omega}$: it is constituted of finite *and* infinite words. Let h be the substitution: $\Sigma \rightarrow P((\Sigma \cup \{\twoheadleftarrow\})^\star)$ defined by $a \rightarrow a.\mathcal{C}_1$ where $\mathcal{C}_1$ is the local language defined above. Then it is easy to see that the finite words of $A^\sim$ are obtained by substituting in A the language $a.\mathcal{C}_1$ for each letter $a \in \Sigma$ and concatenating on the left by the language $\mathcal{C}_2$. But LOC is closed under substitution and concatenation [15], so this language is a local language $L(\varphi_1)$ defined by a local sentence φ_1.

The infinite words in $A^\sim$ constitutes the ω-language

$L(\varphi_1).(\mathcal{C}_1 - \{\lambda\})^\omega$ if $\lambda \notin A$, and
$L(\varphi_1).(\mathcal{C}_1 - \{\lambda\})^\omega \cup (\mathcal{C}_2 - \{\lambda\})^\omega$ if $\lambda \in A$,

The languages $\mathcal{C}_2 - \{\lambda\}$ and $\mathcal{C}_1 - \{\lambda\}$ are local. Thus the set of infinite words in $A^\sim$ is a local ω-language $L_\omega(\varphi_2)$ because $\omega- KC(\text{LOC}) \subseteq \text{LOC}_\omega$ by Theorem 2.15. Finally we have got

$$C^\sim = L_\omega(\psi_1) \cup L_\omega(\varphi_2) \cup L(\varphi_1)$$

But LOC_ω is closed under union [18] hence $L_\omega(\psi_1) \cup L_\omega(\varphi_2)$ is a local ω-language. This ends the proof. □

We have seen above that the correspondence $A \rightarrow A^d$ is involved in Theorem 3.10. Hence we shall need the following proposition.

Proposition 3.16 (a) *if $A \subseteq \Sigma^\star$ is a local language, then A^d is a local ω-language.*
(b) *if $A \subseteq \Sigma^\omega$ is a local ω-language, then A^d is a local ω-language.*
(c) *If $A = L^\Sigma(\varphi) \cup L_\omega^\Sigma(\psi)$ is the union of a finitary local language and of a local ω-language over the same alphabet Σ, then A^d is a local ω-language over the alphabet $\Sigma \cup \{d\}$.*

Proof of (a) Let $A = L^\Sigma(\varphi)$ be a local finitary language over the alphabet Σ. Let P_d be a new letter unary predicate symbol and a be a new constant symbol. Let φ' be the following sentence in the signature $S(\varphi') = S(\varphi) \cup \{P_d, a\}$, which is the conjunction of the following formulas:

1. ($<$ is a linear order),
2. ($(P_e)_{e\in\ (\Sigma\cup\{d\})}$ form a partition),
3. $\forall x_1 \ldots x_j \in \neg P_d[\varphi_0(x_1, \ldots, x_j)]$, where $\varphi = \forall x_1 \ldots x_j\varphi_0(x_1, \ldots, x_j)$ with φ_0 an open formula,
4. $\forall x_1 \ldots x_m \in \neg P_d[f(x_1, \ldots, x_m) \in \neg P_d]$, for each m-ary function f of $S(\varphi)$,

5. $\neg P_d(c)$, for each constant c of $S(\varphi)$,
6. $\forall x_1 \ldots x_m [\bigvee_{1 \leq i \leq m} P_d(x_i) \rightarrow f(x_1, \ldots, x_m) = min(x_1, \ldots, x_m)]$, for each m-ary function f of $S(\varphi)$.
7. $\forall x [x \geq a \rightarrow P_d(x)]$.

This sentence is equivalent to a universal one and closure in its models takes at most $n_\varphi + 1$ steps. By construction $L(\varphi') = A^d$ holds. □

Remark 3.17 We have defined the function f by $f(x_1, \ldots, x_m) = min(x_1, \ldots, x_m)$ when at least one of the x_i was in P_d (see the conjunct numbered 6). In that case the function f is not useful for defining the local ω-language A^d, but this will imply that closure in models of φ' takes at most a finite number of steps, because $f(x_1, \ldots, x_m)$ is then equal to one of the x_i. This method will be applied in the construction of most local sentences in the sequel of this paper, where some functions are somewhere trivially defined (like $f(x, y) = x$ or $p(x) = x$ for a binary function f or a unary function p) in order to make the sentence local.

Proof of (b) Assume that $A = L_\omega(\varphi)$ where φ is a local sentence and $d \notin \Sigma$. A^d is defined by the following sentence ψ of signature $S(\psi) = S(\varphi) \cup \{P_d, s\}$, where P_d is a new unary predicate symbol and s is a new unary function symbol. ψ is the conjunction of:

- The same formulas (1) to (6) as in the proof of a),
- $\forall x[\neg P_d(x) \rightarrow s(x) = x]$,
- $\forall x[P_d(x) \rightarrow \neg P_d(s(x))]$,
- $\forall xy[(P_d(x) \wedge P_d(y) \wedge x \neq y) \rightarrow s(x) \neq s(y)]$.

This sentence is equivalent to a universal one and closure in its models takes at most $n_\varphi + 1$ steps (one applies first the function s and then the functions of $S(\varphi)$). In a model M of ψ, it is easy to see that s^M is an injective function from P_d^M into $\neg P_d^M$ and then, if M has order type ω, $\neg P_d^M$ is infinite and induces an ω-word which is a word of $L_\omega(\varphi)$. So $L_\omega(\psi) = (L_\omega(\varphi))^d$. □

Proof of (c) Let A and B be subsets of $\Sigma^{\leq\omega}$ for a finite alphabet Σ. Then we easily see that if $C = A \cup B$, $C^d = A^d \cup B^d$ holds. (c) is now an easy consequence of (a) and (b) because LOC_ω is closed under finite union [18]. □

We can now state the following result:

Theorem 3.18 *For each integer $n \geq 1$, there exist $\boldsymbol{\Sigma}_n^0$-complete and $\boldsymbol{\Pi}_n^0$-complete local ω-languages.*

Proof Consider first S_1 (respectively, P_1) being the following subsets of $\{0, 1\}^{\leq\omega}$: $S_1 = \{x \in \{0, 1\}^{\leq\omega} \mid \exists\, i\ x(i) = 1\}$ and $P_1 = \{\lambda\}$. Then $(S_1)^d$ (respectively, $(P_1)^d$) are $\boldsymbol{\Sigma}_1^0$-complete (respectively, $\boldsymbol{\Pi}_1^0$-complete).

We can now apply $n \geq 1$ times the operation of exponentiation of sets. More precisely, we define, for a set $A \subseteq \Sigma^{\leq\omega}$:

$A^{\sim.0} = A$
$A^{\sim.1} = A^{\sim}$ and
$A^{\sim.(n+1)} = (A^{\sim.n})^{\sim}$.

Now apply n times (for an integer $n \geq 1$) the operation $\sim$ (with different new letters $\twoheadleftarrow_1, \twoheadleftarrow_2, \twoheadleftarrow_3, \ldots, \twoheadleftarrow_n$) to S_1 and P_1.

By Theorem 3.10, it holds that for an integer $n \geq 1$:

$(S_1^{\sim.n})^d$ is a $\mathbf{\Sigma}^0_{n+1}$-complete subset of $\{0, 1, \twoheadleftarrow_1, \ldots, \twoheadleftarrow_n, d\}$.
$(P_1^{\sim.n})^d$ is a $\mathbf{\Pi}^0_{n+1}$-complete subset of $\{0, 1, \twoheadleftarrow_1, \ldots, \twoheadleftarrow_n, d\}$.

It is easy to see that S_1 and P_1 are in the form $L^{\{0,1\}}(\varphi) \cup L_\omega^{\{0,1\}}(\psi)$ where φ and ψ are local sentences (they are in fact unions of a finitary regular language and of a regular ω-language). Then it follows from Propositions 3.15 and 3.16 that the ω-languages $(S_1^{\sim.n})^d$ and $(P_1^{\sim.n})^d$ are local. Hence the class LOC_ω meets all finite levels of the Borel hierarchy. □

Remark 3.19 For $n = 1$ and $n = 2$, we could get some $\mathbf{\Sigma}^0_n$-complete and $\mathbf{\Pi}^0_n$-complete sets by considering well known examples of regular ω-languages [25,26,30], because $\mathrm{REG}_\omega \subseteq \mathrm{LOC}_\omega$:

$A_1 = \{\alpha \in \{0, 1\}^\omega \mid \exists i \quad \alpha(i) = 1\}$ is $\mathbf{\Sigma}^0_1$-complete,

$B_1 = \{\alpha \in \{0, 1\}^\omega \mid \forall i \quad \alpha(i) = 0\}$ is $\mathbf{\Pi}^0_1$-complete,

$A_2 = \{\alpha \in \{0, 1\}^\omega \mid \exists^{<\omega} i \quad \alpha(i) = 1\}$ is $\mathbf{\Sigma}^0_2$-complete,

$B_2 = \{\alpha \in \{0, 1\}^\omega \mid \exists^{\omega} i \quad \alpha(i) = 0\}$ is $\mathbf{\Pi}^0_2$-complete,

where $\exists^{<\omega} i$ means: “ there exist only finitely many i such that…”, and
$\exists^\omega i$ means: “ there exist infinitely many i such that…”.

Remark 3.20 Reasoning as in [16] for ω-powers of finitary context free languages, we can prove a similar result for local languages: for each integer $n \geq 1$, there exists a local language L_n such that $(L_n)^\omega$ is a $\mathbf{\Pi}^0_n$-complete set.

3.3 Borel sets of infinite rank and local ω-languages

We are going now to prove that there exist some local ω-languages which are Borel sets of infinite rank. More precisely:

Theorem 3.21 *There exists a local ω-language which is a $\mathbf{\Delta}^0_\omega$-set but not a Borel set of finite rank.*

Proof Recall that we can define the following operation on ω-languages: Let $(A_i)_{i\in\mathbb{N}}$ be a countable infinite family of subsets of X^ω for X a finite alphabet containing at least two letters a and b. Then [9]:

$$\sup_{i\in\mathbb{N}} A_i = \bigcup_{i\in\mathbb{N}} a^i.b.A_i.$$

Assume now that each set A_i is a Borel set of finite rank and that for every integer $j \geq 1$ there exists an integer i_j such that A_{i_j} is of Borel rank greater than j. Then the set $\sup_{i\in\mathbb{N}} A_i$ is a Borel set which is in $\mathbf{\Delta}^0_\omega = \mathbf{\Sigma}^0_\omega \cap \mathbf{\Delta}^0_\omega$. Firstly, it is easy to see that the Borel rank of the set $a^i.b.A_i$ is the same as the Borel rank of the set A_i. Thus the set $\sup_{i\in\mathbb{N}} A_i = \bigcup_{i\in\mathbb{N}} a^i.b.A_i$ is a $\mathbf{\Sigma}^0_\omega$-set because it is a countable union of Borel sets of finite ranks. Secondly $\bigcup_{i\in\mathbb{N}} a^i.b.A_i$ is the intersection of the sets $B_i = \bigcup_{j\neq i} a^j.b.X^\omega \cup a^i.b.A_i$. But for each integer i the set B_i is the union of two Borel sets of finite rank (the set $\bigcup_{j\neq i} a^j.b.X^\omega = (\bigcup_{j\neq i} a^j.b).X^\omega$ is an open set). Thus $\sup_{i\in\mathbb{N}} A_i = \bigcup_{i\in\mathbb{N}} a^i.b.A_i$ is a countable intersection of Borel sets of finite rank hence it is a $\mathbf{\Pi}^0_\omega$ set. Moreover the set $\sup_{i\in\mathbb{N}} A_i$ is not a Borel set of finite rank because otherwise assume that it is in the Borel class $\mathbf{\Sigma}^0_J$ for an integer $J \geq 1$. Then for each i, the language $a^i.b.A_i$ would be the intersection of the open set $a^i.b.X^\omega$ and of $\sup_{i\in\mathbb{N}} A_i$. But each class $\mathbf{\Sigma}^0_J$ is closed under finite intersection and then for each $i \in \mathbb{N}$, $a^i.b.A_i$ would be in the class $\mathbf{\Sigma}^0_J$. This would imply that, for all i, $A_i \in \mathbf{\Sigma}^0_J$ also holds which is in contradiction with the hypothesis.

In order to simplify the following proof, we now introduce a variant of $A^\sim$ which was already defined in [16]:

Definition 3.22 For $A \subseteq \Sigma^{\leq\omega}$ and $\twoheadleftarrow \notin \Sigma$, let $X = \Sigma \cup \{\twoheadleftarrow\}$ and
$A^\approx = \{x \in (\Sigma \cup \{\twoheadleftarrow\})^{\leq\omega} \mid x^\twoheadleftarrow \in A\}$,
where $x^\twoheadleftarrow$ is inductively defined by:

$\lambda^\twoheadleftarrow = \lambda$,

and for a finite word $u \in (\Sigma \cup \{\twoheadleftarrow\})^\star$:

$(u.a)^\twoheadleftarrow = u^\twoheadleftarrow.a$, if $a \in \Sigma$,
$(u.\twoheadleftarrow)^\twoheadleftarrow = u^\twoheadleftarrow$ with its last letter removed if $|u^\twoheadleftarrow| > 0$,
$(u.\twoheadleftarrow)^\twoheadleftarrow$ is **undefined** if $|u^\twoheadleftarrow| = 0$,

and for u infinite:

$(u)^\twoheadleftarrow = \lim_{n\in\omega}(u[n])^\twoheadleftarrow$, where, given β_n and v in $\Sigma^\star$,
$v \sqsubseteq \lim_{n\in\omega} \beta_n \leftrightarrow \exists n \forall p \geq n \quad \beta_p[|v|] = v$.

The only difference between the previous definition and this one is that here $(u.\twoheadleftarrow)^\twoheadleftarrow$ is **undefined** if $|u^\twoheadleftarrow| = 0$. Recall that if A is a $\mathbf{\Pi}^0_2$-complete subset of Σ^ω, then for each integer $n \geq 1$ the set $A^{\approx.n}$ is a $\mathbf{\Pi}^0_{n+2}$-complete subset of $(\Sigma \cup \{\twoheadleftarrow_1, \ldots, \twoheadleftarrow_n\})^\omega$ [16]. Then the set $\sup_{i\in\mathbb{N}} A^{\approx.i}$ is a Borel set of rank ω.

In fact this latter result is true only when countable infinite alphabets are allowed because we see from the definition of $A^{\approx.n}$ that this is a set over the alphabet $\Sigma \cup \{\twoheadleftarrow_1$

$, \ldots, \twoheadleftarrow_n\}$. So if we want to find such a set in LOC_ω we have to modify this set by coding the infinite number of erasers $\twoheadleftarrow_1, \ldots, \twoheadleftarrow_n, \ldots$ by finite words over a finite alphabet. We shall then code the eraser $\twoheadleftarrow_n$ by the word $a.b^n$ where a and b are two letters which are not in Σ.

It is easy to see that the resulting set $A^{\approx.n}$ will still be a $\mathbf{\Pi}^0_{n+2}$-complete subset (of $(\Sigma \cup \{a, b\})^\omega$). The proof is left to the reader.

Let then $A = L_\omega(\varphi)$ be a local ω-language over the alphabet Σ. We are going to show that $\sup_{i\in\mathbb{N}} A^{\approx.i}$ is a local ω-language.

An ω-word of $\sup_{i\in\mathbb{N}} A^{\approx.i}$ is in the form $a^n.b.u$ where $u \in A^{\approx.n}$.

Remark that in such an ω-word, there are only finitely many (codes of) erasers and that the number of erasers is fixed by the initial segment $a^n.b$.

We have now to find a local sentence which defines this ω-language. As in the proof of closure of the class LOC [15], (respectively, LOC_ω [18]) under substitution by finitary local languages, we use a unary function I which marks the first letters of the subwords, in order to divide an ω-word into omega (finite) subwords (the function I will be constant on each such "subword" and $I(x)$ will indicate the first letter of the subword containing x). This is expressed by the following sentence θ_1 conjunction of:

- " $<$ is a linear order ",
- $\forall xy[(I(y) \leq y) \wedge (y \leq x \rightarrow I(y) \leq I(x)) \wedge (I(y) \leq x \leq y \rightarrow I(x) = I(y))]$.

Every subword will have a last letter (and then it will be finite). We use another unary function e to designate this last letter. This is expressed by the following sentence θ_2 conjunction of:

- $\forall x[I(e(x)) = I(x)]$,
- $\forall x[x \leq e(x)]$,
- $\forall xy[I(x) = I(y) \rightarrow (e(x) = e(y))]$.

The initial segment of the word in the form $a^n.b$ will be indicated by a unary predicate P_0 and a constant B. Notice that we can assume, without loss of generality, that 0 is not a letter of the alphabet Σ, so the predicate P_0 cannot be a letter predicate. This is expressed by the following sentence θ_3 conjunction of:

- $\forall xy[P_0(x) \wedge \neg P_0(y) \rightarrow x < y]$,
- $P_0(B)$,
- $\forall x[P_0(x) \rightarrow x \leq B]$,
- $P_b(B)$,
- $\forall x[P_0(x) \wedge x < B \rightarrow P_a(x)]$.

We shall say that if a subword on which the function I is constant has length 1 it designates a letter in P_0 or a letter of the alphabet Σ, and otherwise (if such a subword has length >1) it designates an eraser $a.b^n$ where n is an integer ≥ 1. We use a unary predicate P to indicate the letters in Σ. This is expressed by the following sentence θ_4 conjunction of:

- $\forall x[P_0(x) \rightarrow I(x) = x = e(x)]$,
- $\forall x[P(x) \leftrightarrow (I(x) = x = e(x) \wedge \neg P_0(x))]$,
- $\forall x[P(x) \leftrightarrow \bigvee_{c\in\Sigma} P_c(x)]$,
- $\forall x[I(x) \neq e(x) \rightarrow P_a(I(x))]$,

– $\forall x[(I(x) \neq e(x) \wedge x \neq I(x)) \rightarrow P_b(x)]$.

We have now to say that if the ω-word begins with $a^n.b$ the erasers are in the finite set $\{a.b^1, \ldots, a.b^n\}$. We shall use a unary function i which will be injective from each subword into the initial segment designated by P_0; and we add that i is strictly increasing on each subword, this will be useful in the sequel. This is expressed by the following sentence θ_5 conjunction of:

– $\forall x[\neg P_0(x) \rightarrow P_0(i(x))]$,
– $\forall xy[(I(x) = I(y) \wedge x < y) \rightarrow i(x) < i(y)]$,
– $\forall x[P_0(x) \rightarrow i(x) = x]$.
(this third conjunct expresses that i is trivially defined on P_0).

Now we want to be able to compare the erasers because an eraser $\twoheadleftarrow_k = a.b^k$ is allowed to erase another eraser $\twoheadleftarrow_j = a.b^j$ if and only if $k > j$, because of the inductive definition of the sets $A^{\approx.n}$. Then we will compare each eraser to an initial segment of P_0. We use for that purpose another binary function f such that, for $I(x) \notin P_0$, $f(I(x), .)$ will be a function from P_0 into $\{y \mid I(y) = I(x)\}$. This is expressed by the following sentence θ_6 conjunction of:

– $\forall xy[(\neg P_0(x) \wedge P_0(y)) \rightarrow I(f(I(x), y)) = I(x)]$,
– $\forall xy[P_0(x) \rightarrow f(x, y) = x]$,
– $\forall xy[I(x) \neq x \rightarrow f(x, y) = x]$,
– $\forall xy[\neg P_0(y) \rightarrow f(x, y) = x]$.
(these three latest conjuncts are used to trivially define the function f when it is not useful for our purpose, see Remark 3.17)

Now we are going to say that $f(I(x), .)$ is strictly increasing, hence also injective, from $\{z \in P_0 \mid z \leq i(e(x))\}$ into $\{y \mid I(y) = I(x)\}$. This ensures that i is an injection from $\{y \mid I(y) = I(x)\}$ into $\{z \in P_0 \mid z \leq i(e(x))\}$ (because i is increasing) and **conversely** $f(I(x), .)$ is an injection from $\{z \in P_0 \mid z \leq i(e(x))\}$ into $\{y \mid I(y) = I(x)\}$. Therefore these sets have the same cardinal because they are finite and, for $x \notin P_0$, i is a strictly increasing bijection from $\{y \mid I(y) = I(x)\}$ onto an initial segment of P_0. Hence we shall be able to compare two erasers by comparing the images by the function i of the last elements $e(x)$ and $e(y)$ of the segments which code these erasers. This is expressed by the following sentence θ_7:

– $\forall xyz[(\neg P_0(x) \wedge P_0(y) \wedge P_0(z) \wedge y < z \leq i(e(x)) \rightarrow f(I(x), y) < f(I(x), z)]$.

Now we are able to associate an eraser $a.b^j$ which **really erases** with the letter of Σ or the other eraser of type $a.b^k$, with $k < j$, which is **erased** by $a.b^j$. Indeed we shall use a unary function s which associates the first element of the eraser with the letter of Σ or the first element of the eraser which is erased. Let P_1 and P_2 be two new unary predicate symbols, the first one will indicate the first elements of the erasers which **really erase** and the second one will indicate the letters of Σ or the first elements of the erasers which are **erased**. This is expressed by the following sentence θ_8, conjunction of:

– $\forall x[(P_1(x) \vee P_2(x)) \rightarrow (\neg P_0(x) \wedge I(x) = x)]$,
– $\forall x[I(x) \neq e(x) \rightarrow (P_1(I(x)) \vee P_2(I(x)))]$,

– $\forall x[\neg(P_1(x) \vee P_2(x)) \rightarrow s(x) = x]$,
– $\forall x[P_2(x) \leftrightarrow P_1(s(x))]$,
– $\forall x[s(s(x)) = x]$,
– $\forall x[P_2(x) \rightarrow x < s(x)]$.

Remark that some letters of Σ will not be erased by any eraser, hence we have not added the conjunct $\forall x[P(x) \rightarrow P_2(x)]$.

Now we have to ensure, as already mentioned above, that an eraser erase a letter of Σ or an another eraser it is **allowed to erase**. This is expressed by the following sentence θ_9:

– $\forall x[P_2(x) \rightarrow i(e(x)) < i(e(s(x)))]$.

More, the operations of erasing have to be done in a good order, i.e. in an ω-word which contains only the erasers $\twoheadleftarrow_1, \ldots, \twoheadleftarrow_n$, the first operation of erasing uses the last eraser $\twoheadleftarrow_n$, then the second one uses the eraser $\twoheadleftarrow_{n-1}$, and so on. Moreover there is not any letter of Σ which is not erased between an eraser and the segment it erases. This is expressed by the following sentence θ_{10} conjunction of:

– $\forall xy[(P_1(x) \wedge P_1(y) \wedge x < y) \rightarrow ((s(x) < x < s(y) < y) \vee (s(y) < s(x) < x < y \wedge i(e(x)) \geq i(e(y))))]$
– $\forall xy[(P_1(x) \wedge s(x) < y < x \wedge I(y) = e(y)) \rightarrow P_2(y)]$.

Consider now an ω-word of the form $a^n.b.u$ where $u \in A^{\approx.n}$. When the operations of erasing (with the erasers $\twoheadleftarrow_1, \ldots, \twoheadleftarrow_n$) have been completed in u, then the resulting word must be in $A = L_\omega(\varphi)$. Let P_3 be a new unary predicate, we shall say that P_3 induces this resulting word. This is expressed by the following sentence θ_{11} conjunction of:

– $\forall x[P_3(x) \leftrightarrow (P(x) \wedge \neg P_2(x))]$,
– $\forall x_1 \ldots x_j \in P_3[\varphi_0(x_1, \ldots, x_j)]$, where $\varphi = \forall x_1 \ldots x_j \varphi_0(x_1, \ldots, x_j)$ with φ_0 an open formula,
– $\forall x_1 \ldots x_m \in P_3[g(x_1, \ldots, x_m) \in P_3]$, for each m-ary function g of $S(\varphi)$,
– $\forall x_1 \ldots x_m[\bigvee_{1 \leq i \leq m} \neg P_3(x_i) \rightarrow g(x_1, \ldots, x_m) = min(x_1, \ldots, x_m)]$, for each m-ary function g of $S(\varphi)$,
– $P_3(c)$, for each constant c of $S(\varphi)$.

We add the following sentence θ_{12} which expresses that j is an injective function from P_2 into P_3, where j is a new unary function symbol. This will ensure that in an ω-model, P_3 is infinite and hence it induces an ω-word of $L_\omega(\varphi)$ (which remains when the operations of erasing have been made). θ_{12} is the conjunction of:

– $\forall x[P_2(x) \rightarrow P_3(j(x))]$,
– $\forall xy[(P_2(x) \wedge P_2(y) \wedge x \neq y) \rightarrow j(x) \neq j(y)]$,
– $\forall x[\neg P_2(x) \rightarrow j(x) = x]$.
(this latest conjunct is used to define trivially the function j on $\neg P_2$, see Remark 3.17).

Now the conjunction $\bigwedge_{1 \leq i \leq 12} \theta_i$ is a sentence which is equivalent to a universal sentence, because it is the conjunction of a finite number of universal sentences, and closure in its models takes at most $n_\varphi + 5$ steps: one takes first closure under the functions I and e, then under s, and again under I and e, then under i and j, then under f and the functions of $S(\varphi)$.

By construction we check that:

$$L_\omega\left(\bigwedge_{1\le i\le 12}\theta_i\right) = \sup_{i\in\mathbb{N}}(L_\omega(\varphi))^{\approx.i}.$$

$\square$

Remark 3.23 The above proof is the first step for the study of local ω-languages which are Borel sets of infinite rank. Using this first result and other methods, we have constructed some local ω-languages which are Borel sets of every Borel rank smaller than the Cantor ordinal ε_0 [17]. On the other side, Kechris et al. proved in [24] that the supremum of the set of Borel ranks of (lightface) Π^1_1, so also of (lightface) Σ^1_1, sets is the ordinal γ^1_2. This ordinal is strictly greater than the first non-Δ^1_2 ordinal [24]. Thus it holds that $\omega_1^{\mathrm{CK}} < \gamma^1_2$, where ω_1^{CK} is the first non-recursive ordinal. The question is left open to determine completely the set of *all* Borel ranks of local ω-languages and in particular to find its supremum which is of course smaller than or equal to γ^1_2.

3.4 Beyond Borel sets

The question naturally arises: are there local ω-languages which are analytic but not Borel sets?

Theorem 3.24 *There exist local ω-languages which are $\mathbf{\Sigma}^1_1$-complete hence non Borel sets.*

Proof We shall use here results about languages of infinite binary trees whose nodes are labeled in a finite alphabet Σ. A node of an infinite binary tree is represented by a finite word over the alphabet $\{l, r\}$ where r means "right" and l means "left". Then an infinite binary tree whose nodes are labeled in Σ is identified with a function $t : \{l, r\}^\star \to \Sigma$. The set of infinite binary trees labeled in Σ will be denoted T^ω_Σ.

There is a natural topology on this set T^ω_Σ [22,26,28]. It is defined by the following distance: Let t and s be two distinct infinite trees in T^ω_Σ. Then the distance between t and s is $\frac{1}{2^n}$ where n is the smallest integer such that $t(x) \neq s(x)$ for some word $x \in \{l, r\}^\star$ of length n. The open sets are then in the form $T_0.T^\omega_\Sigma$ where T_0 is a set of finite labeled trees. $T_0.T^\omega_\Sigma$ is the set of infinite binary trees which extend some finite labeled binary tree $t_0 \in T_0$, t_0 is here a sort of prefix, an "initial subtree" of a tree in $t_0.T^\omega_\Sigma$.

The Borel hierarchy and the projective hierarchy on T^ω_Σ are defined from open sets in the same manner as in the case of the topological space Σ^ω.

Let t be a tree. A branch B of t is a subset of the set of nodes of t which is linearly ordered by the tree partial order R ($R(xy) \leftrightarrow x \sqsubseteq y$) and which is closed under the prefix relation, i.e. if x and y are nodes of t such that $y \in B$ and $x \sqsubseteq y$ then $x \in B$. A branch B of a tree is said to be maximal iff there is not any other branch of t which strictly contains B.

Let t be an infinite binary tree in T_Σ^ω. If B is a maximal branch of t, then this branch is infinite. Let $(u_i)_{i\geq 0}$ be the enumeration of the nodes in B which is strictly increasing for the prefix order. The infinite sequence of labels of the nodes of such a maximal branch B, i.e. $t(u_0)t(u_1)\ldots t(u_n)\ldots$ is called a path. It is an ω-word over the alphabet Σ.

Let then $L \subseteq \Sigma^\omega$ be an ω-language over Σ. Then we denote Path(L) the set of infinite trees t in T_Σ^ω such that t has (at least) a path in L.

It is well known that if $L \subseteq \Sigma^\omega$ is an ω-language over Σ which is a $\mathbf{\Pi}_2^0$-complete subset of Σ^ω (or a set of higher complexity in the Borel hierarchy) then the set Path(L) is a $\mathbf{\Sigma}_1^1$-complete subset of T_Σ^ω. Hence Path(L) is not a Borel set [22,30,37].

For $L_\omega^\Sigma(\varphi)$ a local ω-language, we shall find another local ω-language $L_\omega^{(\Sigma\cup\{0,1\})}(\psi)$ and a continuous function

$$h : T_\Sigma^\omega \to (\Sigma \cup \{0, 1\})^\omega$$

such that Path$(L_\omega^\Sigma(\varphi)) = h^{-1}(L_\omega^{(\Sigma\cup\{0,1\})}(\psi))$. For that we shall code trees labeled in Σ by words over $\Sigma \cup \{0, 1\}$, where 0 and 1 are supposed to be two new letters not in Σ. We use two new unary predicate symbols, P and B. The first one will indicate the set of nodes of the tree and the second one will indicate a maximal branch of the tree which provides a word of $L_\omega^\Sigma(\varphi)$ when the labels are considered. We first express that R (a binary new relation) is a strict partial order over P by the following sentence ϕ_1, conjunction of:

- $\forall xy[R(xy) \to P(x) \wedge P(y)]$,
- $\forall xyz[R(xy) \wedge R(yz) \to R(xz)]$,
- $\forall xy[R(xy) \to \neg R(yx)]$.

We have to say that this order is the order of a tree, i.e. that the predecessors of an element $x \in P$ are linearly ordered by R. This is expressed by the following sentence ϕ_2:

- $\forall xyz[R(xz) \wedge R(yz) \to (R(xy) \vee R(yx) \vee x = y)]$.

Now we use a new constant symbol S and the following sentence ϕ_3 expresses that this constant is interpreted by the root node of the tree:

- $P(S) \wedge \forall x \in P[x \neq S \to R(Sx)]$.

The trees are labeled in Σ, and we use the two other letters to code the relation R in a word. So let ϕ_4 be the following sentence, conjunction of:

- $((P_a)_{a\in(\Sigma\cup\{0,1\})}$ form a partition),
- $\forall x[P(x) \leftrightarrow \bigvee_{a\in\Sigma} P_a(x)]$,
- $\forall x[\neg P(x) \leftrightarrow P_0(x) \vee P_1(x)]$.

We use a binary new function f and two unary new functions p and p' to say that a model M of ψ is the disjoint union of P^M and of $f^M(P^M \times P^M)$. f^M will be an injective function from $P^M \times P^M$ into $\neg P^M$, and the projections p^M and p'^M will ensure that $f^M(P^M \times P^M) = \neg P^M$. This is expressed by the following sentence ϕ_5, conjunction of:

- $\forall xy \in P[\neg P(f(xy))]$,
- $\forall x[\neg P(x) \to P(p(x)) \wedge P(p'(x))]$,
- $\forall xy[P(x) \wedge P(y) \to x = p(f(xy)) \wedge y = p'(f(xy))]$,
- $\forall x[\neg P(x) \to x = f(p(x)p'(x))]$,
  (these four latest conjuncts imply that the function f is a bijection from $P \times P$ onto $\neg P$),
- $\forall xy[\neg P(x) \vee \neg P(y) \to f(xy) = x]$,
- $\forall x[P(x) \to p(x) = p'(x) = x]$,
  (these two latest conjuncts trivially define somewhere the functions f, p and p' according to Remark 3.17).

The order of the elements of $f^M(P^M \times P^M)$ for $<^M$ in M will be also determined by the order $<^M$ on P^M. Let us remark that we choose such an order on $f^M(P^M \times P^M)$ but we could have made another choice. But we want this order to be determined by ψ. Then once the enumeration of order type ω of the nodes has been chosen, the code of a tree as an ω-word over the alphabet $\Sigma \cup \{0, 1\}$ is completely fixed. This is expressed by the following sentence ϕ_6, conjunction of:

- $\forall xyx'y' \in P[\max(xy) < \max(x'y') \to f(xy) < f(x'y')]$,
  (where $\max(xy) = y$ iff $x \leq y$ and $\max(xy) = x$ iff $y \leq x$),
- $\forall xyz \in P[y < z \leq x \to (f(xy) < f(xz) \wedge f(zx) > f(yx) \wedge f(xy) < f(zx))]$,
- $\forall xyz \in P[y \leq x < z \to (x < f(xy) < z \wedge x < f(yx) < z)]$.

This will fix the order of the letters 0 and 1 which code the tree order and in order to really code the tree order by the letters 0 and 1 of the word, we use the following sentence ϕ_7, conjunction of:

- $\forall xy[R(xy) \to P_0(f(xy))]$,
- $\forall xy[\neg R(xy) \to P_1(f(xy))]$.

In order to say that the branches of the tree have at most length ω when the word coding the tree is an ω-word we use the following sentence ϕ_8 which expresses that the order R is compatible with the order $<$ of the words:

- $\forall xy[R(xy) \to x < y]$.

The unary predicate B will indicate the nodes of a branch of the tree, this is expressed by using the following sentence ϕ_9, conjunction of:

- $\forall x[B(x) \to P(x)]$,
- $\forall xy[(B(x) \wedge B(y) \wedge x \neq y) \to (R(xy) \vee R(yx))]$,
- $\forall xy[B(x) \wedge R(yx) \to B(y)]$.

This branch will be a maximal branch (this will be useful for having an infinite branch when infinite trees are considered). We use a new unary function i which is trivial on B and which associates to each node x of $\neg B$ another node $i(x)$ of the branch B such that x and $i(x)$ are incomparable with regard to the relation R of the tree. This is expressed by the following sentence ϕ_{10}, conjunction of:

- $\forall x[(P(x) \wedge \neg B(x)) \to B(i(x))]$,
- $\forall x[(P(x) \wedge \neg B(x)) \to (\neg R(xi(x)) \wedge \neg R(i(x)x))]$,

- $\forall x[\neg P(x) \to i(x) = x]$,
- $\forall x[B(x) \to i(x) = x]$.
  (these two latest conjuncts trivially define the function i on B and on $\neg P$).

Now we have to say that the branch B induces a word of $L_\omega(\varphi)$ (when the branch is infinite of length ω).
This is expressed by the following sentence ϕ_{11}, conjunction of:

- $B(c)$, for each constant c of $S(\varphi)$,
- $\forall x_1 \ldots x_k[S(x_1 \ldots x_k) \to B(x_1) \wedge \ldots \wedge B(x_k)]$, for each predicate $S(x_1 \ldots x_k)$ of $S(\varphi)$,
- $\forall x_1 \ldots x_j[(B(x_1) \wedge \ldots \wedge B(x_j)) \to B(g(x_1 \ldots x_j))]$, for each j-ary function symbol g of $S(\varphi)$,
- $\forall x_1 \ldots x_j[(\bigvee_{1 \le i \le j} \neg B(x_i)) \to g(x_1 \ldots x_j) = min(x_1 \ldots x_j)]$, for each j-ary function symbol g of $S(\varphi)$,
- $\forall x_1 \ldots x_m[(B(x_1) \wedge \ldots \wedge B(x_m)) \to \varphi_0(x_1 \ldots x_m)]$, where $\varphi = \forall x_1 \ldots x_m \varphi_0 (x_1 \ldots x_m)$ with φ_0 an open formula,

Consider now the conjunction:

$$\psi = \bigwedge_{1 \le i \le 11} \phi_i.$$

This sentence is written in the signature:

$$S(\psi) = S(\varphi) \cup \{S, P, B, R, P_0, P_1, f, p, p', i\}$$

where S is a constant symbol, P, B, P_0, P_1 are unary predicate symbols, R is a binary predicate symbol, p, p', i are unary function symbols and f is a binary function symbol. ψ is equivalent to a universal sentence, because it is the conjunction of a finite number of universal sentences, and closure in its models takes at most $n_\varphi + 3$ steps (one takes closure under the functions p and p', then under S and i, then under the functions of $S(\varphi)$ and finally under f). Hence ψ is a local sentence and it defines a local ω-language over the alphabet $\Sigma \cup \{0, 1\}$.

Consider now the set $\{l, r\}^\star$ of nodes of the infinite binary tree, and the lexicographic order on this set (assuming that l is before r for this order). Then, in the enumeration of the nodes with regard to this order, the first nodes will be $\lambda, l, r, ll, lr, rl, rr, lll, llr, \ldots$ Let then h be the mapping from T_Σ^ω into $(\Sigma \cup \{0, 1\})^\omega$ such that for every labeled binary infinite tree t of T_Σ^ω, $h(t)$ is the code of the tree as defined above (by the sentences ϕ_1 to ϕ_8), where the enumeration of length ω of the nodes is in lexicographic order as explained above. Then for a tree $t \in T_\Sigma^\omega$, $h(t) \in L_\omega(\psi)$ if and only if t has a path in $L_\omega(\varphi)$ thus $\mathrm{Path}(L_\omega^\Sigma(\varphi)) = h^{-1}(L_\omega^{\Sigma \cup \{0,1\}}(\psi))$ holds.

Hence if $L_\omega(\varphi)$ is a Borel set which is at least a $\mathbf{\Pi}_2^0$-complete subset of Σ^ω, the language $\mathrm{Path}(L_\omega(\varphi)) = h^{-1}(L_\omega(\psi))$ is a $\mathbf{\Sigma}_1^1$-complete subset of T_Σ^ω. But it is easy to see from the definition of h and of the lexicographic order on $\{l, r\}^\star$ that h is a continuous function from T_Σ^ω into $(\Sigma \cup \{0, 1\})^\omega$. Then the ω-language $L_\omega(\psi)$ is at least $\mathbf{\Sigma}_1^1$-complete because $h^{-1}(L_\omega(\psi))$ is a $\mathbf{\Sigma}_1^1$-complete set and it is in fact a

$\mathbf{\Sigma}_1^1$-complete subset of $(\Sigma \cup \{0, 1\})^\omega$ because every local ω-language is an analytic set by Theorem 3.8. Then in that case $L_\omega(\psi)$ is not a Borel set because a $\mathbf{\Sigma}_1^1$-complete set is not a Borel set.

Indeed this gives infinitely many $\mathbf{\Sigma}_1^1$-complete local ω-languages, because there exist infinitely many local ω-languages which are $\mathbf{\Pi}_2^0$-complete (for example the regular ω-languages which are $\mathbf{\Pi}_2^0$-complete). □

A natural question arises about the recursive analogue to Theorem 3.24: are there local languages which are in the class Σ_1^1 but in not any class of the arithmetical hierarchy? The answer can be easily derived from the inclusions $\Sigma_n \subseteq \mathbf{\Sigma}_n^0$ and $\Pi_n \subseteq \mathbf{\Pi}_n^0$ and Theorem 3.24:

Corollary 3.25 *There exist some local ω-languages in $\Sigma_1^1 - \bigcup_{n \geq 1} \Sigma_n$.*

Remark 3.26 The method we have used in the above proof to code the tree order relation may be used more generally to code the ω-models of some local sentence φ. Then we can show that the set of codes of ω-models of φ is itself a local ω-language.

Acknowledgments We wish to thank the referee for useful comments on a preliminary version of this paper.

References

1. Börger, E., Grädel, E., Gurevich, Y.: The Classical Decision Problem. Springer, Heidelberg (1997)
2. Büchi, J.R., Landweber, L.H.: Solving sequential conditions by finite state strategies. Trans. Am. Math. Soc. **138**, 295–311 (1969)
3. Büchi, J.R.: Weak second order arithmetic and finite automata. Zeitschrift Fur Mathematische Logik Und Grundlagen der Mathematik **6**, 66–92 (1960)
4. Büchi, J.R.: On a decision method in restricted second order arithmetic, logic methodology and philosophy of science. In: Proceedings of the 1960 International Congress. Stanford University Press, pp. 1–11 (1962)
5. Büchi, J.R., Siefkes, D.: The monadic second order theory of all countable ordinals, decidable theories 2. Lect. Notes Math. **328**, 1–217 (1973)
6. Chang, C.C., Keisler, H.J.: Model Theory, American Elsevier Publishing Company, Inc., New York (1973) (3rd edn., North Holland, Amsterdam, 1990)
7. Duparc, J.: La Forme Normale des Boréliens de Rang Fini. Ph.D. Thesis, Université Paris 7 (1995)
8. Duparc, J.: The Normal Form of Borel Sets, Part 1: Borel Sets of Finite Rank. Comptes Rendus de l'Académie des Sciences, Paris, t.320, Série 1, pp. 651–656 (1995)
9. Duparc, J.: Wadge hierarchy and Veblen hierarchy: Part 1: Borel sets of finite rank. J. Symb. Log. **66**(1), 56–86 (2001)
10. Duparc, J., Finkel, O., Ressayre, J-P.: Computer science and the fine structure of Borel sets. Theor. Comput. Sci. **257**(1–2), 85–105 (2001)
11. Engelfriet, J., Hoogeboom, H.J.: X-automata on ω-words. Theor. Comput. Sci. **110**(1), 1–51 (1993)
12. Finkel, O.: Langages de Büchi et ω-Langages Locaux. Comptes Rendus de l'Académie des Sciences, Paris, t.309, Série 1, pp. 991–994 (1989)
13. Finkel, O.: Théorie des Modèles des Formules Locales et Etude des Langages Formels qu'elles Définissent. Ph.D. Thesis, Université Paris 7 (1993)
14. Finkel, O., Ressayre, J-P.: Stretchings. J. Symb. Log. **61**(2), 563–585 (1996)
15. Finkel, O.: Locally finite languages. Theor. Comput. Sci. **255**(1–2), 223–261 (2001)
16. Finkel, O.: Topological properties of omega context free languages. Theor. Comput. Sci. **262**(1–2), 669–697 (2001)
17. Finkel, O.: Locally Finite ω-Languages and Borel Sets of Infinite Rank (preprint)

18. Finkel, O.: Closure properties of locally finite omega languages. Theor. Comput. Sci. **322**(1), 69–84 (2004)
19. Finkel, O.: On decidability properties of local sentences. In: Proceedings of the 11th Workshop on Logic, Language, Information and Computation WoLLIC 2004. Electronic Notes in Theoretical Computer Science, vol. 123, pp. 75–92. Elsevier, Amsterdam (2005)
20. Finkel, O.: On decidability properties of local sentences. Theor. Comput. Sci. **364**(2), 196–211 (2006)
21. Hopcroft, J.E., Ullman, J.D.: Formal Languages and their Relation to Automata. Addison-Wesley Publishing Company, Reading (1969)
22. Kechris, A.S.: Classical Descriptive Set Theory. Springer, Heidelberg (1995)
23. Kuratowski, K.: Topology. Academic Press, New York (1966)
24. Kechris, A.S., Marker, D., Sami, R.L.: Π_1^1 Borel sets. J. Symb. Log. **54**(3), 915–920 (1989)
25. Landweber, L.H.: Decision problems for ω-automata. Math. Syst. Theory **3**(4), 376–384 (1969)
26. Lescow, H., Thomas, W.: Logical specifications of infinite computations. In: A Decade of Concurrency. de Bakker, J.W., et al. (eds.) Lecture Notes in Computer Science, vol. 803, pp. 583–621. Springer, Heidelberg (1994)
27. Meyer, A.R.: Weak monadic second order theory of successor is not elementary recursive, logic colloquium (Boston, Mass., 1972–1973). Lecture Notes in Mathematics, vol. 453, pp. 132–154. Springer, Berlin (1975)
28. Moschovakis, Y.N.: Descriptive Set Theory. North-Holland, Amsterdam (1980)
29. Parigot, M., Pelz, E.: A logical approach to Petri net languages. Theor. Comput. Sci. **39**, 155–169 (1985)
30. Perrin, D., Pin, J.-E.: Infinite words, automata, semigroups, logic and games. Pure and Applied Mathematics, vol. 141 . Elsevier, Amsterdam (2004)
31. Pin, J-E.: Logic, semigroups and automata on words. Ann. Math. Artif. Intell. **16**, 343–384 (1996)
32. Ressayre, J-P.: Formal languages defined by the underlying structure of their words. J. Symb. Log. **53**(4), 1009–1026 (1988)
33. Rogers, H.: Theory of Recursive Functions and Effective Computability. McGraw-Hill, New York (1967)
34. Safra, S.: Complexity of Automata on Infinite Objects. Ph.D. Thesis, Weizmann Institute of Science, Rehovot, Israel (1989)
35. Shoenfield, J.R.: Mathematical Logic. Addison-Wesley, Reading (1967)
36. Simonnet, P.: Automates et Théorie Descriptive. Ph.D. Thesis, Université Paris 7, March (1992)
37. Simonnet, P.: Automate d'Arbres Infinis et Choix Borélien. Comptes Rendus de l'Académie des Sciences, Paris, t.316, Série 1, pp. 97–100, (1993)
38. Staiger L. (1997) ω-Languages. In: Rozenberg, G., Salomaa, A. (eds.) Handbook of Formal Languages, vol. 3, pp. 339–388. Springer, Berlin (1997)
39. Thomas, W.: Automata on Infinite Objects. In: Van Leeuwen, J. (eds.) Handbook of Theoretical Computer Science, vol. B, pp. 133–191. Elsevier, Amsterdam (1990)
40. Thomas, W.: Languages, automata and logic, In: Rozenberg, G., Salomaa A. (eds.) Handbook of Formal Languages, vol. 3, pp. 389–456. Springer, Heidelberg (1997)
41. Wadge, W.W.: Reducibility and determinateness on the Baire space. Ph.D. Thesis, Berkeley (1983)
42. Wagner, K.: On omega regular sets. Inf. Control **43**, 123–177 (1979)