Non-Maximal Decidable Structures

Alexis Bès and Patrick Cégielski
LACL, Université Paris 12
E-mail:{bes,cegielski}@univ-paris12.fr

June 17, 2007

Abstract

Given any infinite structure $\mathcal{M}$ with a decidable first-order theory, we give a sufficient condition in terms of the Gaifman graph of $\mathcal{M}$, which ensures that $\mathcal{M}$ can be expanded with some non-definable predicate in such a way that the first-order theory of the expansion is still decidable.

LACL Technical Report 2007-06

1 Introduction

Elgot and Rabin ask in [2] whether there exist maximal decidable structures, i.e. structures $\mathcal{M}$ with a decidable elementary theory and such that the elementary theory of any expansion of $\mathcal{M}$ by a non-definable predicate is undecidable.

Soprunov proved in [8] (using a forcing argument) that every structure in which a regular ordering is interpretable is not maximal. A partial order $(B, <)$ is said to be regular if for every $a \in B$ there exist distinct elements $b_1, b_2 \in B$ such that $b_1 < a$, $b_2 < a$, and no element $c \in B$ satisfies both $c < b_1$ and $c < b_2$. As a corollary he also proved that there is no maximal decidable structure if we replace "elementary theory" by "weak monadic second-order theory".

In [1] we considered a weakening of the Elgot-Rabin question, namely the question of whether all structures $\mathcal{M}$ whose first-order theory is decidable can be expanded by some constant in such a way that the resulting structure still has a decidable theory. We answered this question negatively by proving that there exists a structure $\mathcal{M}$ whose monadic second-order theory is decidable and such that any expansion of $\mathcal{M}$ by a constant has an undecidable elementary theory.

In this paper we address the initial Elgot-Rabin question, and provide a criterion for non-maximality. More precisely, given any structure $\mathcal{M}$ with a decidable first-order theory, we give in Section 3 a sufficient condition in terms of the Gaifman graph of $\mathcal{M}$, which ensures that $\mathcal{M}$ can be expanded with some non-definable predicate in such a way that the first-order theory of the expansion is still decidable. The condition is the following: for every natural number r and every finite set X of elements of the base set $|\mathcal{M}|$ of $\mathcal{M}$ there exists an element $x \in |\mathcal{M}|$ such that the Gaifman distance between x and every element of X is greater than r. This condition holds e.g. for the structure $(\mathbb{N}, S)$, where S denotes the graph of the successor function, and more generally for any labelled infinite graph with finite degree and whose elementary theory is decidable, i.e. any structure $\mathcal{M} = (V, E, P_1, \ldots, P_n)$ where V is infinite, E is a binary relation of finite degree, the P_i's are unary relations, and the elementary theory of $\mathcal{M}$ is decidable. Unlike Soprunov's condition, our condition expresses some limitation on the expressive power of the structure $\mathcal{M}$.

In Section 2 we recall some important definitions and results. Section 3 deals with the main theorem. We conclude the paper with related questions.

2 Preliminaries

In the sequel we consider first-order logic with equality. We deal only with relational structures. Given a language $\mathcal{L}$ and a $\mathcal{L}-$structure $\mathcal{M}$, we denote by $|\mathcal{M}|$ the base set of $\mathcal{M}$. For every symbol R of $\mathcal{L}$ we denote by $R^{\mathcal{M}}$ the interpretation of R in $\mathcal{M}$. As usual we shall sometimes confuse symbols and their interpretation. We denote by $FO(\mathcal{M})$ the first-order (complete) theory of $\mathcal{M}$, i.e. the set of first-order $\mathcal{L}-$sentences true in $\mathcal{M}$. By "definable in $\mathcal{M}$ " we mean "first-order definable in $\mathcal{M}$ without parameters".

We denote by $qr(\phi)$ the quantifier rank of the formula ϕ, defined inductively by $qr(\phi) = 0$ if ϕ is atomic, $qr(\neg F) = qr(F)$, $qr(F \alpha G) = \max(qr(F), qr(G))$ for $\alpha \in \{\wedge, \vee, \rightarrow\}$, and $qr(\exists x F) = qr(\forall x F) = qr(F) + 1$. We define $FO_n(\mathcal{M})$ as the set of $\mathcal{L}-$sentences F such that $qr(F) \leq n$ and $\mathcal{M} \models F$.

We say that the elementary diagram of a structure $\mathcal{M}$ is computable if there exists an injective map $f : |\mathcal{M}| \rightarrow \mathbb{N}$ such that the range of f, as well as the relations $\{(f(a_1), \ldots, f(a_n)) \mid a_1, \ldots, a_n \in |\mathcal{M}| \text{ and } \mathcal{M} \models R(a_1, \ldots, a_n)\}$ for every relation R of $\mathcal{L}$, are recursive (see e.g. [7]).

Let us recall useful definitions and results related to the Gaifman graph of a structure [3] (see also [5]). Let $\mathcal{L}$ be a relational language, and $\mathcal{M}$ be a $\mathcal{L}-$structure. The *Gaifman graph of* $\mathcal{M}$, which we denote by $G(\mathcal{M})$, is the undirected graph whose set of vertices is $|\mathcal{M}|$, and such that for all $x, y \in |\mathcal{M}|$, there is an edge between x and y if and only if $x = y$ or if there exist some $n-$ary relational symbol $R \in \mathcal{L}$ and some $n-$tuple $\vec{t}$ of elements of $|\mathcal{M}|$ which contains both x and y and satisfies $\vec{t} \in R^{\mathcal{M}}$.

The distance $d(x, y)$ between two elements $x, y \in |\mathcal{M}|$ is defined as the usual distance in the sense of the graph $G(\mathcal{M})$. We denote by $B_r(x)$ the $r-$sphere with center x, i.e. the set of elements y of $|\mathcal{M}|$ such that $d(x, y) \leq r$. It should be noted that for every fixed r the binary relation "$y \in B_r(x)$" is definable in $\mathcal{M}$. For every $X \subseteq |\mathcal{M}|$ we define $B_r(X)$ as $B_r(X) = \bigcup_{x \in X} B_r(x)$.

A $r-$local formula $\varphi(x_1, \ldots, x_n)$ is a formula whose quantifiers are all relativized to $B_r(\{x_1, \ldots, x_n\})$. We shall use the notation $\varphi^{(r)}$ to indicate that φ is $r-$local.

Let us state Gaifman's theorem about local formulas.

Theorem 1 *([3]) Let $\vec{x} = (x_1, \ldots, x_n)$ and $\varphi(\vec{x})$ be a $\mathcal{L}-$formula. From φ one can compute effectively a formula which is equivalent to φ and is a boolean combination of formulas of the form:*

- $\psi^{(r)}(\vec{x})$
- $\exists x_1 \ldots \exists x_s \ (\bigwedge_{1\leq i\leq s} \alpha^{(r)}(x_i) \wedge \bigwedge_{1\leq i<j\leq s} d(x_i, x_j) > 2r)$

where $s \leq qr(\varphi) + n$ and $r \leq 7^k$.

Moreover if φ is a sentence then only sentences of the second kind occur in the resulting formula.

3 A sufficient condition for non-maximality

The aim of this section is to prove the following theorem.

Theorem 2 *Let $\mathcal{L}$ be a finite relational language, and $\mathcal{M}$ be an infinite countable $\mathcal{L}-$structure which satisfies the following conditions:*

1. *$FO(\mathcal{M})$ is decidable*
2. *every element of $|\mathcal{M}|$ is definable in $\mathcal{M}$*
3. *for every finite set $X \subseteq |\mathcal{M}|$ and every $r \in \mathbb{N}$, there exists $a \in |\mathcal{M}|$ such that $d(a, X) > r$.*

Then there exists a unary predicate symbol $R \notin \mathcal{L}$ and a $(\mathcal{L} \cup \{R\})-$expansion $\mathcal{M}'$ of $\mathcal{M}$ such that :

- *$FO(\mathcal{M}')$ is decidable*
- *the set $R^{\mathcal{M}'}$ is not definable in $\mathcal{M}$.*
- *the elementary diagram of $\mathcal{M}'$ is computable.*

Let us consider a few examples.

- The structure $\mathcal{M} = (\mathbb{N}; S)$, where S denotes the graph of the function $x \mapsto x + 1$, satisfies all conditions of Theorem 2. Indeed Langford [4] proved that $FO(\mathcal{M})$ is decidable. Moreover condition 2 is easy to prove, and condition 3 is a straightforward consequence of the fact that $d(x, y) = |x - y|$ for all natural numbers x, y.
- The same holds for any structure of the form $\mathcal{M} = (\mathbb{N}; S, P_1, \ldots, P_n)$ where the P_i's denote unary predicates and $FO(\mathcal{M})$ is decidable. Note that expanding a structure by unary predicates does not change its Gaifman graph.

- More generally Theorem 2 applies to any infinite labelled graph with finite degree, more precisely to any structure of the form $\mathcal{M} = (V; E, P_1, \dots, P_n)$ where V is infinite, E is a binary relation with finite degree, the P_i's denote unary predicates, $FO(\mathcal{M})$ is decidable, and every element of V is definable in $\mathcal{M}$. In this case the Gaifman graph of $\mathcal{M}$ has finite degree, which implies condition 3. Note that Theorem 2 also applies to some structures for which the degree of the Gaifman graph is infinite – see the last example.

- The structure $\mathcal{M} = (\mathbb{N}; <)$ does not satisfy condition 3 of Theorem 2 since $d(x, y) \leq 1$ for all $x, y \in \mathbb{N}$. Observe that $FO(\mathcal{M})$ is decidable [4], and moreover $\mathcal{M}$ is not maximal: consider e.g. the structure $\mathcal{M}' = (\mathbb{N}; <, +)$ where $+$ denotes the graph of addition; $FO(\mathcal{M}')$ is decidable [6], and $+$ is not definable in $\mathcal{M}$ since in $\mathcal{M}$ one can only define finite or co-finite subsets of $\mathbb{N}$.

  One can prove actually that for every infinite structure $\mathcal{M}$ in which some linear ordering of elements of $|\mathcal{M}|$, condition 3 does not hold. However the next example shows that Theorem 2 can be applied to some structures in which an infinite linear ordering is *interpretable*.

- Consider the disjoint union of ω copies of $(\mathbb{N}; <)$ equipped with a successor relation between copies, i.e. the structure $\mathcal{M} = (\mathbb{N} \times \mathbb{N}; <, Suc)$ where

  – $(x, y) < (x', y')$ if and only if $(x = x'$ and $y < y')$;
  – $Suc((x, y), (x', y'))$ if and only if $x' = x + 1$

  then $\mathcal{M}$ satisfies the conditions of Theorem 2: the first condition comes from the fact that $FO(\mathcal{M})$ reduces to $FO(\mathbb{N}; <)$ and the two other conditions are easy to check.

Let us explain informally the structure of the proof of Theorem 2. Given $\mathcal{M}$ which satisfies the conditions of Theorem 2, we define $R^{\mathcal{M}'}$ by marking gradually elements of $|\mathcal{M}|$, some in $R^{\mathcal{M}'}$ and some in its complement. More precisely we define by induction on n the sequence $(X_n)_{n \in \mathbb{N}}$ with $X_n = (R_n, S_n, T_n, F_n)$ where R_n corresponds to a set of elements of $R^{\mathcal{M}'}$(we will say "marked positively"), S_n corresponds to a set of elements marked in the complement of $R^{\mathcal{M}'}$ (we will say "marked negatively"), T_n roughly corresponds to a set of spheres whose elements are marked in the complement of $R^{\mathcal{M}'}$, and F_n denotes the set of formulas of quantifier rank $\leq n$ which will be true in $\mathcal{M}'$. At each step n, the partial marking X_n ensures that any subsequent marking will lead to a set $R^{\mathcal{M}'}$ not definable by any formula of quantifier rank n. Moreover X_n also *fixes* $FO_n(\mathcal{M}')$. Finally $R^{\mathcal{M}'}$ will be defined as the union of the sets R_n. In the construction we impose some sparsity condition on $R^{\mathcal{M}'}$; this condition ensures that there are few elements of $R^{\mathcal{M}'}$ in each $r-$sphere, which allows to express with $\mathcal{L}$-sentences whether a $r-$sphere of $\mathcal{M}$ can be marked conveniently, and then use the condition that $FO(\mathcal{M})$ is decidable in order to extend the marking in an effective way.

Proof of Theorem 2.

Assume that $\mathcal{M}$ is a $\mathcal{L}-$structure which satisfies the conditions of the theorem. Let $R \notin \mathcal{L}$ be a unary predicate symbol. For every $X \subseteq |\mathcal{M}|$ we shall denote by $\mathcal{M}(X)$ the $(\mathcal{L} \cup \{R\})-$expansion of $\mathcal{M}$ defined by interpreting R by X.

Throughout the proof we shall use the following interesting consequences of conditions 1 and 2:

- the elementary diagram of $\mathcal{M}$ is computable. Indeed since $\mathcal{L}$ is finite we can enumerate all formulas $\varphi(x)$ with one free variable. Let us denote by $(\varphi_i(x))_{i\geq 0}$ such an enumeration. Then the application $f : |\mathcal{M}| \to \mathbb{N}$ which maps every element e of $|\mathcal{M}|$ to the least integer i such that φ_i defines e is injective; moreover the range of f, and the relations $\{(f(a_1), \ldots, f(a_n)) : \mathcal{M} \models Q(a_1, \ldots, a_n)\}$ for every symbol Q of $\mathcal{L}$, are recursive.
- if $\psi(x)$ is a formula with one free variable and $\mathcal{M} \models \exists x \psi(x)$ then one can find in an effective way the first integer i who belongs to the range of f and such that $\mathcal{M} \models \exists x(\varphi_i(x) \wedge \psi(x))$. That is, one can find effectively some element $x \in |\mathcal{M}|$ for which $\psi(x)$ holds in $\mathcal{M}$.
- every finite or co-finite subset $A \subseteq |\mathcal{M}|$ is definable in $\mathcal{M}$. This will allow to use shortcuts such as "$x \in A$" when we write formulas in the language $\mathcal{L}$.

We now define by induction on n the sequence $(X_n)_{n\in\mathbb{N}}$ such that for every n, $X_n = (R_n, S_n, T_n, F_n)$ where

1. R_n, S_n, T_n are finite subsets of $|\mathcal{M}|$;
2. $R_n \cap S_n = \varnothing$;
3. F_n is a set of $(\mathcal{L} \cup \{R\})-$sentences with quantifier rank $\leq n$;
4. $d(R_n, R_{n+1} \setminus R_n) \geq 7^{n+1}$;
5. $d(x, y) \geq 7^{n+1}$ for every pair of distinct elements of $R_{n+1} \setminus R_n$;
6. for every $R' \subseteq |\mathcal{M}|$ such that $R_n \subseteq R'$ and
$$R' \cap ((S_n \cup \bigcup_{i\leq n} B_{7^i}(T_i)) \setminus R_n) = \varnothing,$$
R' is not definable by any $\mathcal{L}-$formula of quantifier rank $\leq n$;
7. For every $R' \subseteq |\mathcal{M}|$ such that $R_n \subseteq R'$,
$$R' \cap ((S_n \cup \bigcup_{i\leq n} B_{7^i}(T_i)) \setminus R_n) = \varnothing,$$

$$d(R', R' \setminus R_n) \geq 7^{n+1},$$

and $d(x, y) \geq 7^{n+1}$ whenever x, y are distinct elements of $R' \setminus R_n$, we have

$$FO_n(\mathcal{M}(R')) = F_n.$$

Induction hypothesis: assume that $(X_i)_{i<n}$ is defined and satisfies the required conditions.

Let us define X_n. The definition consists in two main steps: during the first step we extend the marking in order to ensure that $R^{\mathcal{M}'}$ will not be definable by any formula with quantifier rank n; this is the easiest step, and it uses condition (3) of the theorem. During the second step, we extend again the marking in order to fix $FO_n(\mathcal{M}')$.

We set $r = 7^n$.

First step: during this step we mark a finite number of elements in order to ensure that $R^{\mathcal{M}'}$ will not be definable by any $\mathcal{L}-$formula with quantifier rank n.

Since we deal with a finite relational language, there exist up to equivalence finitely many formulas with quantifier rank n. From $\mathcal{L}$ one can compute an integer k_n and a finite set of $\mathcal{L}-$formulas $\{\alpha_{n,i}(x) : 1 \leq i \leq k_n\}$ such that every $\mathcal{L}-$formula with quantifier rank n is equivalent to a disjunction of some of the $\alpha_{n,i}$'s, and moreover such that the formulas $\alpha_{n,i}$ are incompatible. For $i = 1, \ldots, k_n$, let us denote by $E_{n,i}$ the subset of $|\mathcal{M}|$ defined by $\alpha_{n,i}(x)$. By construction the sequence $(E_{n,1}, \ldots, E_{n,k_n})$ is a partition of $|\mathcal{M}|$, and every subset of $|\mathcal{M}|$ definable by a formula of quantifier rank n is a finite union of some of the subsets $E_{n,i}$.

We shall mark elements in order that for some i, the subset $E_{n,i}$ contains at least an element marked positively and another element marked negatively. This will ensure that condition 6 is satisfied. More precisely, for $i = 1, \ldots, k_n$, we mark positively (respectively negatively) at most one new element of $E_{n,i}$. We define the sets $R'_{n,i}$ (resp. $S'_{n,i}$) such that $R'_{n,i}$ contains the set of new elements to mark positively (resp. negatively) in $E_{n,i}$ (each of the sets $R'_{n,i}$ and $S'_{n,i}$ is either empty or reduced to a singleton). We proceed as follows:

- if there exists some element of $E_{n,i}$ which is not marked yet, and moreover all marked elements of $E_{n,i}$ are marked positively, then we mark negatively the first unmarked element of $E_{n,i}$.

  Formally, assume that the sets $R'_{n,j}$ and $S'_{n,j}$ have been defined for every $j < i$, and let

$$Z_{n,i} = R_{n-1} \cup \bigcup_{j<i} R'_{n,j} \cup S_{n-1} \cup \bigcup_{j<i} S'_{n,j} \cup \bigcup_{i<n} B_{(7^i)}(T_i)$$

If

$$\mathcal{M} \models \exists x(\alpha_{n,i}(x) \wedge x \notin Z_{n,i})$$

and moreover

$$\mathcal{M} \models (E_{n,i} \cap Z_{n,i}) \subseteq (R_{n-1} \cup \bigcup_{j<i} R'_{n,j})$$

(this set-theoretic property is expressible as a $\mathcal{L}-$sentence) then we set $S'_{n,i}$ as the singleton set consisting in the first x such that

$$\mathcal{M} \models \exists x(\alpha_{n,i}(x) \wedge x \notin Z_{n,i}).$$

Otherwise we set $S'_{n,i} = \varnothing$.

- Then, if all currently marked elements of $E_{n,i}$ are marked negatively, and moreover there exists some unmarked element x of $E_{n,i}$ at distance $\geq 7^{n+1}$ from already marked elements, then we mark positively the first such element x.

  Formally, let

$$Z'_{n,i} = Z_{n,i} \cup S'_{n,i}$$

  If

$$\mathcal{M} \models (E_{n,i} \cap (R_{n-1} \cup \bigcup_{j<i} R'_{n,j})) = \varnothing$$

  and moreover

$$\mathcal{M} \models \exists x(\alpha_{n,i}(x) \wedge d(x, Z'_{n,i}) \geq 7^{n+1})$$

  then let $R'_{n,i}$ be the singleton set consisting in the first such x. Otherwise we set $R'_{n,i} = \varnothing$.

Note that the previous procedure is effective (see the remarks at the beginning of the proof).

Second step: during this step we extend the marking in order to fix $FO_n(\mathcal{M}')$. Up to equivalence, there exist finitely many $(\mathcal{L} \cup \{R\})-$formulas F such that $qr(F) = n$. By Proposition 1 every such formula F is equivalent to a boolean combination of formulas of the form

$$\exists x_1 \dots \exists x_s \, (\bigwedge_{1 \leq i \leq s} \alpha^{(r)}(x_i) \wedge \bigwedge_{1 \leq i < j \leq s} d(x_i, x_j) > 2r).$$

Consider an enumeration $G_{n,1}, \dots, G_{n,m_n}$ of all formulas of the previous form which arise when we apply Theorem 1 to formulas F such that $qr(F) = n$.

During this step we shall fix which formulas $G_{n,j}$ will be true in $\mathcal{M}'$, which will suffice (using again Theorem 1) to fix which formulas F with quantifier rank n will be true in $\mathcal{M}'$.

The first idea is to check, for every j, whether there exists $R' \subseteq |\mathcal{M}|$ which extends in a convenient way the current marking and such that $\mathcal{M}(R') \models G_{n,j}$. If the answer is positive, then we shall extend our marking just enough to ensure that every subsequent extension of the marking will satisfy $\mathcal{M}' \models G_{n,j}$. If the answer is negative, then we do not extend the marking, and then every subsequent extension of the marking will satisfy $\mathcal{M}' \models \neg G_{n,j}$.

We define by induction on $j \leq m_n$ the sets $R''_{n,j}$ and $T'_{n,j}$, such that $R''_{n,j}$ contains new elements to mark positively, and $T'_{n,j}$ contains the centers of new $r-$spheres whose elements are marked negatively.

We proceed as follows. Fix j, and assume that the sets $R''_{n,i}$ and $T'_{n,i}$ have been defined for every $i < j$. We have

$$G_{n,j} : \exists x_1 \ldots \exists x_s \ (\bigwedge_{1\leq i\leq s} \alpha^{(r)}_{n,j}(x_i) \wedge \bigwedge_{1\leq i<j\leq s} d(x_i, x_j) > 2r)$$

for some $r-$local formula $\alpha^{(r)}_{n,j}$ (formally s depend on n and j, but we omit the subscripts for the sake of readability).

Let $R^+_{n,j}$ be the set of elements currently marked positively, i.e.

$$R^+_{n,j} = R_{n-1} \cup \bigcup_{i<k_n} R'_{n,i} \cup \bigcup_{i<j} R''_{n,i},$$

and let $R^-_{n,j}$ be the set of elements currently marked negatively, that is

$$R^-_{n,j} = (S_{n-1} \cup \bigcup_{i<k_n} S'_{n,i} \cup \bigcup_{i<n} B_{(7^i)}(T_i) \cup \bigcup_{i<j} B_{(7^n)}(T'_{n,i})) \setminus R^+_{n,j}.$$

Let $P_{n,j} = R^+_{n,j} \cup R^-_{n,j}$.

We want to check whether there exists $R' \subseteq |\mathcal{M}|$ such that

1. $\mathcal{M}(R') \models G_{n,j}$;
2. $R^+_{n,j} \subseteq R'$ and $R^-_{n,j} \cap R' = 0$ (i.e. R' extends the current marking);
3. $d(R^+_{n,j}, R' \setminus R^+_{n,j}) \geq 7^{n+1}$;
4. $d(x, y) \geq 7^{n+1}$ for every pair of distinct elements of $R' \setminus R^+_{n,j}$.

Let us denote by $(*)$ the conjunction of these four conditions. Let us prove that one can express $(*)$ with a $\mathcal{L}-$sentence.

Assume first that there exists R' which satisfies $(*)$. Let $x_1, \ldots, x_s \in |\mathcal{M}|$ be such that

$$\mathcal{M}(R') \models (\bigwedge_{1\leq i\leq s} \alpha^{(r)}_{n,j}(x_i) \wedge \bigwedge_{1\leq i<j\leq s} d(x_i, x_j) > 2r)$$

Conditions 3 and 4 of $(*)$ imply that each sphere $B_r(x_i)$ contains at most one element of $R' \setminus R^+_{n,j}$, and moreover that if such an element exists, it is the unique element of R' in $B_r(x_i)$. Thus we can assume without loss of generality that there exist $t \leq s$ and $y_1, \ldots, y_t \in |\mathcal{M}|$ such that

$$B_r(x_i) \cap (R' \setminus R^+_{n,j}) = \{y_i\}$$

for every $i \leq t$, and

$$B_r(x_i) \cap (R' \setminus R^+_{n,j}) = \varnothing$$

for every $i > t$. Condition (3) yields $d(R^+_{n,j}, y_i) \geq 7^{n+1}$ for every i, and condition (4) yields $d(y_i, y_j) \geq 7^{n+1}$ for all distinct integers i, j.

Let us consider first the $r-$spheres $B_r(x_i)$ for $i \leq t$. By definition of x_i we have $\mathcal{M}(R') \models \alpha^{(r)}_{n,j}(x_i)$. Now y_i is the unique element of $R' \cap B_r(x_i)$ thus we have $\mathcal{M} \models \alpha'_{n,j}(x_i, y_i)$ where $\alpha'_{n,j}(x_i, y_i)$ is obtained from $\alpha^{(r)}_{n,j}(x_i)$ by replacing every atomic formula of the form $R(z)$ by $(z = y_i)$.

Now consider the $r-$spheres $B_r(x_i)$ for $i > t$. By definition we have $\mathcal{M}(R') \models \alpha^{(r)}_{n,j}(x_i)$, and $B_r(x_i)$ contains no element of $R' \setminus R^+_{n,j}$. Thus we have $\mathcal{M} \models \gamma^{(r)}_{n,j}(x_i)$ where $\gamma^{(r)}_{n,j}(x_i)$ is obtained from $\alpha^{(r)}_{n,j}(x_i)$ by replacing every atomic formula of the form $R(z)$ by $(z \in B_r(x_i) \cap R^+_{n,j})$.

The previous arguments show that $\mathcal{M} \models G'_{n,j}$ where $G'_{n,j}$ is the $\mathcal{L}-$sentence $G'_{n,j}$ defined as follows:

$$G'_{n,j} : \bigvee_{t \leq s} H_{n,j,t}$$

where

$$H_{n,j,t} : \exists x_1 \ldots \exists x_s \exists y_1 \ldots \exists y_t (\bigwedge_{1 \leq i < j \leq s} d(x_i, x_j) > 2r \wedge \bigwedge_{1 \leq i < j \leq t} d(y_i, y_j) > 7r \wedge$$

$$\wedge \bigwedge_{1 \leq i \leq t} d(y_i, R^+_{n,j}) > 7r \wedge \bigwedge_{1 \leq i \leq t} \beta^{(r)}_{n,j}(x_i, y_i) \wedge \bigwedge_{t < i \leq s} \gamma^{(r)}_{n,j}(x_i))$$

with

$$\beta^{(r)}_{n,j}(x_i, y_i) : \; y_i \in B_r(x_i) \wedge y_i \notin P_{n,j} \wedge B_r(x_i) \cap R^+_{n,j} = \varnothing \; \wedge \alpha'^{(r)}_{n,j}(x_i, y_i).$$

Conversely, assume that $\mathcal{M} \models G'_{n,j}$. Let t, $x_1, \ldots, x_s$, and $y_1, \ldots, y_t$ be such that $H_{n,j,t}$ holds in $\mathcal{M}$. Then if we set $R' = R^+_{n,j} \cup \{y_1, \ldots, y_t\}$, one checks easily that R' satisfies $(*)$

Therefore we have shown that the question whether there exists R' which satisfies $(*)$ is equivalent to the question whether $\mathcal{M} \models G'_{n,j}$ for some $\mathcal{L}-$formula which can be constructed effectively from $G_{n,j}$.

If $\mathcal{M} \models \neg G'_{n,j}$ (which can be checked effectively since by our hypotheses $FO(\mathcal{M})$ is decidable), then we set

$$R''_{n,j} = T'_{n,j} = F'_{n,j} = \varnothing.$$

Now if $\mathcal{M} \models G'_{n,j}$ one can find effectively the least value of t such that $\mathcal{M} \models H_{n,j,t}$, and then $x_1, \ldots, x_s$ and $y_1, \ldots, y_t$ for which the formula holds. We set

$$R''_{n,j} = \{y_1, \ldots, y_t\}, \; T'_{n,j} = \{x_1, \ldots, x_s\}, \text{ and } F'_{n,j} = \{G_{n,j}\}.$$

This completes the second step of the construction of X_n.

We can now define X_n as follows: we set

$$R_n = R_{n-1} \cup \bigcup_{i \leq k_n} R'_{n,i} \cup \bigcup_{j \leq m_n} R''_{n,j}$$

$$S_n = S_{n-1} \cup \bigcup_{i \leq k_n} S'_{n,i}$$

and

$$T_n = \bigcup_{j \leq m_n} T'_{n,j}.$$

In order to define F_n, consider a formula F with quantifier rank n. By Theorem 1, F is equivalent to a formula F' which is a boolean combination of formulas of the form $G_{n,j}$. Consider the truth value of F' determined by setting "true" all formulas $G_{n,j} \in F'_{n,j}$, and "false" formulas $G_{n,j} \notin F'_{n,j}$. Then we define F_n as the union of F_{n-1} and of all formulas F for which F' is true.

We have defined X_n. There remains to show that X_n satisfies all conditions required in the definition.

- Conditions (1) to (5) are easy consequences of the construction of X_n (and the induction hypotheses).
- Let us consider condition (6). Let $R' \subseteq |\mathcal{M}|$ be such that $R_n \subseteq R'$ and

  $$R' \cap ((S_n \cup \bigcup_{i \leq n} B_{7^i}(T_i)) - R_n) = \varnothing.$$

  Let us prove that R' is not definable by any $\mathcal{L}-$formula of quantifier rank $\leq n$. Since every subset of $|\mathcal{M}|$ definable by a $\mathcal{L}-$formula with quantifier rank n is the union of some of the sets $E_{n,i}$, it suffices to prove that R' and its complement intersect some $E_{n,i}$.

  By construction, the set $X = R_n \cup S_n \cup \bigcup_{i \leq n} T_i$ is finite. Now by hypothesis $\mathcal{M}$ satisfies condition 3 of Theorem 2, thus there exists $x \in |\mathcal{M}|$ such that $d(X, x) > 7^n$. The element x belongs to some set $E_{n,i}$. Let us prove that R' and its complement intersect $E_{n,i}$.

  Consider the step of the construction of X_n during which we marked elements of $E_{n,i}$. Recall that just before this step the set of marked elements was

  $$Z_{n,i} = R_{n-1} \cup \bigcup_{j<i} R'_{n,j} \cup S_{n-1} \cup \bigcup_{j<i} S'_{n,j} \cup \bigcup_{i<n} B_{(7^i)}(T_i)$$

  Since $x \in E_{n,i}$ and $d(X, x) > 7^n$, the set $E_{n,i} \setminus Z_{n,i}$ is non-empty. Thus either $E_{n,i}$ already contained an element marked negatively (and in this case $S'_{n,i} = \varnothing$), or we marked one (from $E_{n,i} \setminus Z_{n,i}$) and put it in $S'_{n,i}$. Therefore the complement of R' intersects $E_{n,i}$.

Just after this step, then either $E_{n,i}$ already contained some element marked positively, or by definition of x there existed an element y of $E_{n,i}$ at distance $\geq 7^n$ from currently marked elements, and thus we could mark positively the first such element y. In both cases this ensures that R' intersects $E_{n,i}$.

- Let us prove now that X_n satisfies condition (7). Let $R' \subseteq |\mathcal{M}|$ be such that $R_n \subseteq R'$,

$$R' \cap ((S_n \cup \bigcup_{i \leq n} B_{7^i}(T_i)) \setminus R_n) = \varnothing,$$

$$d(R', R \setminus R_n) \geq 7^{n+1}$$

and $d(x,y) \geq 7^{n+1}$ whenever x, y are distinct elements of $R' \setminus R_n$. Let us prove that $FO_n(\mathcal{M}(R')) = F_n$. The case of formulas with quantifier rank $< n$ follows from our induction hypotheses. Consider now formulas with quantifier rank n. Their truth values are completely determined by the truth values of formulas $G_{n,j}$. Thus it is sufficient to prove that for every j we have $\mathcal{M}(R') \models G_{n,j}$ if and only if $F'_{n,j} = \{G_{n,j}\}$. Fix j, and consider the step of the construction of X_n during which we delt with the formula $G_{n,j}$. If $\mathcal{M} \models G'_{n,j}$ then in this case $F'_{n,j} = \{G_{n,j}\}$, and the definition of $R''_{n,j}$ and $T'_{n,j}$ imply that the formula $G_{n,j}$ holds for every R' which extends (in a convenient way) the marking (R_n, S_n, T_n), thus we have $\mathcal{M}(R') \models G_{n,j}$. On the other hand if $\mathcal{M} \not\models G'_{n,j}$, then the property $(*)$ cannot be satisfied, and we have set $F_{n,j} = \varnothing$. In particular R' does not satisfy $(*)$. Now the hypotheses on R' yield that R' satisfies the three last conditions of $(*)$, thus the first condition is not satisfied, that is $\mathcal{M}(R') \not\models G_{n,j}$.

This concludes the proof that there exists a sequence $(X_n)_{n\geq 0}$ which satisfies all conditions required in the definition.

Now let $\mathcal{M}'$ be the $(\mathcal{L} \cup \{R\})-$expansion of $\mathcal{M}$ defined by

$$R^{\mathcal{M}'} = \bigcup_{n \geq 0} R_n.$$

Let us prove that $\mathcal{M}'$ satisfies the properties required in Theorem 2.

The definition of $R^{\mathcal{M}'}$ implies that for every n, $R^{\mathcal{M}'}$ is not definable by any $\mathcal{L}-$sentence with quantifier rank n, and moreover that $FO_n(\mathcal{M}') = F_n$. Therefore $R^{\mathcal{M}'}$ is not definable in $\mathcal{M}$, and $FO(\mathcal{M}')$ is decidable.

Let us prove that the elementary diagram of $\mathcal{M}'$ is computable. Consider the function f used for the elementary diagram of $\mathcal{M}$; it is sufficient to prove that $\{f(a) \mid \mathcal{M}' \models R(a)\ ,\ a \in |\mathcal{M}|\}$ is recursive. Since every element e of $|\mathcal{M}|$ is definable, there exists n, i such that $E_{n,i} = \{e\}$. During the construction of X_n, and more precisely just before the marking of $E_{n,i}$, then either e had already been marked, or e is marked during this step. Thus eventually every element of $|\mathcal{M}|$ is marked in $R^{\mathcal{M}'}$ or in its complement. This implies that both $\{f(a) \mid \mathcal{M}' \models R(a)\ ,\ a \in |\mathcal{M}|\}$ and $\{f(a) \mid \mathcal{M}' \not\models R(a)\ ,\ a \in |\mathcal{M}|\}$ are recursively enumerable, from which the result follows.

This concludes the proof of Theorem 2.

4 Conclusion

We gave a sufficient condition in terms of the Gaifman graph of the structure $\mathcal{M}$ which ensures that $\mathcal{M}$ is not maximal. A natural problem is to extend Theorem 2 to structures $\mathcal{M}$ which do not satisfy condition (3). We currently investigate the case of labelled linear orderings, i.e. infinite structures $(A; <, P_1, \ldots, P_n)$ where $<$ is a linear ordering over A and the P_i's denote unary predicates; the Gaifman distance is trivial for these structures. Another related general problem is to find a way to refine the notion of Gaifman distance.

Finally, it would also be interesting to study the complexity gap between the decision procedure for the theory of $\mathcal{M}$ and the one for the structure $\mathcal{M}'$ constructed in the proof of Theorem 2.

References

[1] A. Bès and P. Cégielski. Weakly maximal decidable structures. 2007. submitted.

[2] Calvin C. Elgot and Michael O. Rabin. Decidability and undecidability of extensions of second (first) order theory of (generalized) successor. *J. Symb. Log.*, 31(2):169–181, 1966.

[3] H. Gaifman. On local and non-local properties. In *Logic colloquium '81, Proc. Herbrand Symp., Marseille 1981, Stud. Logic Found. Math. 107*, pages 105–135, 1982.

[4] C.H. Langford. Theorems on deducibility. (second paper.). *Annals of Math.*, 2:459–471, 1927.

[5] Leonid Libkin. *Elements of Finite Model Theory.* Springer, 2004.

[6] M. Presburger. Über de vollständigkeit eines gewissen systems der arithmetik ganzer zahlen, in welchen, die addition als einzige operation hervortritt. In *Comptes Rendus du Premier Congrès des Mathématicienes des Pays Slaves*, pages 92–101, 395, Warsaw, 1927.

[7] Valentina S.Harizanov. Computably-theoretic complexity of countable structures. *Bulletin of Symbolic Logic*, 8:457–477, 2002.

[8] S. Soprunov. Decidable expansions of structures. *Vopr. Kibern.*, 134:175–179, 1988.