Записки научных
семинаров ПОМИ
Том 358, 2008 г.

A. Bès and P. Cégielski

NON–MAXIMAL DECIDABLE STRUCTURES

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

Dedicated to Yu. Matiyasevich on the occasion of his 60th birthday

1. Introduction

Elgot and Rabin ask in [3] whether there exist maximal decidable structures, i.e., structures $\mathfrak{M}$ with a decidable elementary theory and such that the elementary theory of any expansion of $\mathfrak{M}$ by a non-definable predicate is undecidable.

Soprunov proved in [10] (using a forcing argument) that every structure in which a regular ordering is interpretable is not maximal. A partial ordering $(B, <)$ 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"[1].

In [1] we considered a weakening of the Elgot-Rabin question, namely the question of whether all structures $\mathfrak{M}$ 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 $\mathfrak{M}$ whose monadic second-order theory is decidable and such that any expansion of $\mathfrak{M}$ 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 $\mathfrak{M}$ with

[1]These results, and the Elgot-Rabin question itself, were brought to our attention by Semenov's paper [8].

a decidable first-order theory, we give in Section 3 a sufficient condition in terms of the Gaifman graph of $\mathfrak{M}$, which ensures that $\mathfrak{M}$ 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 $|\mathfrak{M}|$ of $\mathfrak{M}$ there exists an element $x \in |\mathfrak{M}|$ 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., for any structure $\mathfrak{M} = (V, E, P_1, \ldots, P_n)$ where V is infinite, E is a binary relation of finite degree, the P_i's are unary relations, and the elementary theory of $\mathfrak{M}$ is decidable. Unlike Soprunov's condition, our condition expresses some limitation on the expressive power of the structure $\mathfrak{M}$.

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 $\mathfrak{L}$ and an $\mathfrak{L}$-structure $\mathfrak{M}$, we denote by $|\mathfrak{M}|$ the base set of $\mathfrak{M}$. For every symbol $R \in \mathfrak{L}$ we denote by $R^{\mathfrak{M}}$ the interpretation of R in $\mathfrak{M}$. As usual we shall often confuse symbols and their interpretation. We denote by $FO(\mathfrak{M})$ the first-order (complete) theory of $\mathfrak{M}$, i.e., the set of first-order $\mathfrak{L}$-sentences φ such that $\mathfrak{M} \models \varphi$.

We say that an n-ary relation R over $|\mathfrak{M}|$ is elementary definable (shortly: *definable*) in $\mathfrak{M}$ if there exists an $\mathfrak{L}$-formula φ with n free variables such that $R = \{(a_1, \ldots, a_n) : \mathfrak{M} \models \varphi(a_1, \ldots, a_n)\}$.

We denote by $qr(F)$ the quantifier rank of the formula F, defined inductively by $qr(F) = 0$ if F 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 $FO_n(\mathfrak{M})$ as the set of $\mathfrak{L}-$sentences F such that $qr(F) \leq n$ and $\mathfrak{M} \models F$.

We say that the elementary diagram of a structure $\mathfrak{M}$ is computable if there exists an injective map $f : |\mathfrak{M}| \rightarrow \mathbb{N}$ such that the range of f, as well as the relations

$$\{(f(a_1), \ldots, f(a_n)) \mid a_1, \ldots, a_n \in |\mathfrak{M}| \quad \text{and} \quad \mathfrak{M} \models R(a_1, \ldots, a_n)\}$$

for every relation R of $\mathfrak{L}$, are recursive (see e.g. [9]).

Let us recall useful definitions and results related to the Gaifman graph of a structure [4] (see also [6]). Let $\mathfrak{L}$ be a relational language, and $\mathfrak{M}$ be an $\mathfrak{L}$-structure. The *Gaifman graph of* $\mathfrak{M}$, which we denote by $G(\mathfrak{M})$, is the undirected graph whose set of vertices is $|\mathfrak{M}|$, and such that for all $x, y \in |\mathfrak{M}|$, there is an edge between x and y if and only if either $x = y$ or there exist some n-ary relational symbol $R \in \mathfrak{L}$ and some n-tuple $\vec{t}$ of elements of $|\mathfrak{M}|$ which contains both x and y and satisfies $\vec{t} \in R^{\mathfrak{M}}$.

The distance $d(x, y)$ between two elements $x, y \in |\mathfrak{M}|$ is defined as the usual distance in the sense of the graph $G(\mathfrak{M})$. We denote by $B_r(x)$ the r-ball with center x, i.e., the set of elements y of $|\mathfrak{M}|$ 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 $\mathfrak{M}$. For every $X \subseteq |\mathfrak{M}|$ we define $B_r(X)$ as $B_r(X) = \bigcup_{x \in X} B_r(x)$.

An 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 now state Gaifman's theorem about local formulas.

Theorem 1 ([4])**.** *Let $\vec{x} = (x_1, \ldots, x_n)$, and let $\varphi(\vec{x})$ be an $\mathfrak{L}$-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 $\mathfrak{L}$ be a finite relational language, and $\mathfrak{M}$ be an infinite countable $\mathfrak{L}$-structure which satisfies the following conditions:*

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

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

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

Note that in the above Theorem, the construction of $\mathfrak{M}'$ from $\mathfrak{M}$ can be repeated starting from $\mathfrak{M}'$. Indeed $\mathfrak{M}'$ clearly satisfies conditions 1 and 2. Moreover expanding a structure by unary predicates does not modify its Gaifman graph, therefore we have $G(\mathfrak{M}') = G(\mathfrak{M})$, which implies that condition 3 also holds for $\mathfrak{M}'$.

Let us illustrate Theorem 2 with a few examples.

- The structure $\mathfrak{M} = (\mathbb{N}; S)$, where S denotes the graph of the function $x \mapsto x + 1$, satisfies all conditions of Theorem 2. Indeed Langford [5] proved that $FO(\mathfrak{M})$ 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 $\mathfrak{M} = (\mathbb{N}; S, P_1, \ldots, P_n)$ where the P_i's denote unary predicates and $FO(\mathfrak{M})$ is decidable (the Gaifman graph of any such structure is equal to the one of $(\mathbb{N}; S)$, see the remark above).
- More generally Theorem 2 applies to any infinite labelled graph with finite degree, more precisely to any structure of the form $\mathfrak{M} = (V; E, P_1, \ldots, P_n)$ where V is infinite, E is a binary relation with finite degree, the P_i's denote unary predicates, $FO(\mathfrak{M})$ is decidable, and every element of V is definable in $\mathfrak{M}$. In this case the Gaifman graph of $\mathfrak{M}$ 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 $\mathfrak{M} = (\mathbb{N}; <)$ does not satisfy condition 3 of Theorem 2 since $d(x, y) \leq 1$ for all $x, y \in \mathbb{N}$. Observe that $FO(\mathfrak{M})$ is decidable [5], and moreover $\mathfrak{M}$ is not maximal: consider e.g. the structure $\mathfrak{M}' = (\mathbb{N}; <, +)$ where $+$ denotes the graph of addition; $FO(\mathfrak{M}')$ is decidable [7], and $+$ is not definable in $\mathfrak{M}$ since in $\mathfrak{M}$ one can only define finite or co-finite subsets of $\mathbb{N}$.

One can prove actually that for every infinite structure $\mathfrak{M}$ in which some linear ordering of elements of $|\mathfrak{M}|$ is definable, 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

$$\mathfrak{M} = (\mathbb{N} \times \mathbb{N}; <, Suc),$$

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 $\mathfrak{M}$ satisfies the conditions of Theorem 2: the first condition comes from the fact that $FO(\mathfrak{M})$ 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 $\mathfrak{M}$ which fulfills all conditions of Theorem 2, we define $R^{\mathfrak{M}'}$ by marking gradually elements of $|\mathfrak{M}|$, some in $R^{\mathfrak{M}'}$ 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 finite set of elements of $R^{\mathfrak{M}'}$ (we will say "marked positively");
- S_n corresponds to a finite set of elements of the complement of $R^{\mathfrak{M}'}$ (we will say "marked negatively");
- T_n corresponds to a finite set of centers of balls whose elements (apart from elements of R_n) are marked in the complement of $R^{\mathfrak{M}'}$;
- F_n denotes the set of formulas of quantifier rank $\leq n$ which will be true in $\mathfrak{M}'$.

The set $R^{\mathfrak{M}'}$ will be defined as the union of the sets R_n. At each step n, the partial marking X_n ensures that $R^{\mathfrak{M}'}$ is not definable by any formula of quantifier rank n, and also fixes $FO_n(\mathfrak{M}')$. The possibility to fix $FO_n(\mathfrak{M}')$ whereas $R^{\mathfrak{M}'}$ is only partially defined, comes from Gaifman's Theorem 1 which reduces the satisfaction of sentences in $\mathfrak{M}'$ to the one of sentences which only speak about a finite number of r-balls in $|\mathfrak{M}'|$ (these are sentences of the second kind in Theorem 1), and thus can be evaluated as soon as $R^{\mathfrak{M}'}$ is completely defined in these r-balls.

In the construction we impose some sparsity condition on $R^{\mathfrak{M}'}$; this condition implies that there are few elements of $R^{\mathfrak{M}'}$ in each r-ball, which in turn allows to express with $\mathfrak{L}$-sentences that an r-ball of $|\mathfrak{M}|$ can be marked conveniently, and then use the hypothesis that $FO(\mathfrak{M})$ is decidable in order to extend the marking in an effective way.

Proof of Theorem 2.

Assume that $\mathfrak{M}$ is an $\mathfrak{L}$-structure which satisfies all conditions of the theorem. Let $R \notin \mathfrak{L}$ be a unary predicate symbol. For every $X \subseteq |\mathfrak{M}|$ we shall denote by $\mathfrak{M}(X)$ the $(\mathfrak{L}\cup\{R\})$-expansion of $\mathfrak{M}$ 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 $\mathfrak{M}$ is computable. Indeed since $\mathfrak{L}$ 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 $f : |\mathfrak{M}| \to \mathbb{N}$ which maps every element e of $|\mathfrak{M}|$ to the least integer i such that φ_i defines e is injective; moreover the range of f, and the relations $\{(f(a_1),\dots,f(a_n)) : \mathfrak{M} \models Q(a_1,\dots,a_n)\}$ for every symbol Q of $\mathfrak{L}$, are recursive.
- if $\psi(x)$ is a formula with one free variable and $\mathfrak{M} \models \exists x\psi(x)$ then one can find in an effective way the first integer i which belongs to the range of f and such that $\mathfrak{M} \models \exists x(\varphi_i(x) \wedge \psi(x))$. That is, one can find effectively some element $x \in |\mathfrak{M}|$ for which $\psi(x)$ holds in $\mathfrak{M}$.
- every finite or co-finite subset $A \subseteq |\mathfrak{M}|$ is definable in $\mathfrak{M}$. This will allow to use shortcuts such as "$x \in A$" when we write formulas in the language $\mathfrak{L}$.

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

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

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

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

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

Conditions 4, 5 and 6 express that the marking associated with X_n extends the one associated with X_{n-1}, and 7 and 8 specify that elements of $R_n \setminus R_{n-1}$ (i.e., new elements marked positively) are far away from each other and also from elements of R_{n-1}. Conditions 9 and 10 ensure that for any set $R' \subseteq |\mathfrak{M}|$ which extends R_n "sparsely" (this will hold in particular for the sets $R_{n+1}, R_{n+2}, \ldots$ and eventually for $R^{\mathfrak{M}'}$), R' is not definable in $\mathfrak{M}$ by any $\mathfrak{L}$-formula of quantifier rank $\leq n$, and moreover $FO_n(\mathfrak{M}(R')) = F_n$, i.e., the partial marking X_n fixes $FO_n(\mathfrak{M}(R'))$.

We now define the sequence $(X_n)_{n \in \mathbb{N}}$.

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 get condition 9, i.e., to ensure that $R^{\mathfrak{M}'}$ will not be definable in $\mathfrak{M}$ with any $\mathfrak{L}$-formula with quantifier rank n; this is the easiest step, and it involves condition 3 of the Theorem. During the second step, we extend again the marking in order to get condition 9, i.e., to fix $FO_n(\mathfrak{M}')$.

We set $r = 7^n$.

First step: during this step we mark a finite number of elements in order to ensure that $R^{\mathfrak{M}'}$ will not be definable by any $\mathfrak{L}$-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 $\mathfrak{L}$ one can compute an integer k_n and a finite set of $\mathfrak{L}$-formulas $\{\alpha_{n,i}(x) : 1 \leq i \leq k_n\}$ such that every $\mathfrak{L}$-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 $|\mathfrak{M}|$ defined by $\alpha_{n,i}(x)$. By construction the sequence $(E_{n,1}, \ldots, E_{n,k_n})$ is a partition of $|\mathfrak{M}|$, and every subset of $|\mathfrak{M}|$ 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 9 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

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

and moreover

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

(this property is expressible with an $\mathfrak{L}$-sentence), then we set $S'_{n,i}$ as the singleton set consisting of the first element x such that

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

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

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

and moreover

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

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

Note that the above construction is effective (see the remarks at the beginning of the proof).

Second step: during this step we extend the marking in order to fix $FO_n(\mathfrak{M}')$.

Up to equivalence, there exist finitely many $(\mathfrak{L}\cup\{R\})$-sentences F such that $qr(F)=n$. By Theorem 1, every such sentence F is equivalent to a boolean combination of sentences of the form

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

Consider an enumeration $G_{n,1},\ldots,G_{n,m_n}$ of all sentences 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 sentences $G_{n,j}$ will be true in $\mathfrak{M}'$, which will suffice (using again Theorem 1 to fix which sentences F with quantifier rank n will be true in $\mathfrak{M}'$).

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

We define by induction on $j \le m_n$ 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-balls 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

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

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

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

We want to check whether there exists $R' \subseteq |\mathfrak{M}|$ such that
1. $\mathfrak{M}(R') \models G_{n,j}$;
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 an $\mathfrak{L}$-sentence.

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

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

Conditions 3 and 4 of $(*)$ imply that each ball $B_r(x_i)$ 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 $y_1, \ldots, y_t \in |\mathfrak{M}|$ 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-balls $B_r(x_i)$ for $i \leq t$. By definition of x_i we have $\mathfrak{M}(R') \models \alpha^{(r)}_{n,j}(x_i)$. Now y_i is the unique element of $R' \cap B_r(x_i)$ thus we have $\mathfrak{M} \models \alpha'_{n,j}(x_i, y_i)$ 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-balls $B_r(x_i)$ for $i > t$. By definition we have $\mathfrak{M}(R') \models \alpha^{(r)}_{n,j}(x_i)$, and $B_r(x_i)$ contains no element of $R' \setminus R^+_{n,j}$. Thus we have $\mathfrak{M} \models \gamma^{(r)}_{n,j}(x_i)$ 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 $\mathfrak{M} \models G'_{n,j}$ where $G'_{n,j}$ is the $\mathfrak{L}$-sentence $G'_{n,j}$ defined as follows:

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

where

$$H_{n,j,t} : \exists x_1 \dots \exists x_s \exists y_1 \dots \exists y_t \Big(\bigwedge_{1 \leq i < j \leq s} d(x_i, x_j) > 2r \wedge$$
$$\wedge \bigwedge_{1 \leq i < j \leq t} d(y_i, y_j) > 7r \wedge$$
$$\wedge \bigwedge_{1 \leq i \leq t} d(y_i, R^+_{n,j}) > 7r \wedge \bigwedge_{1 \leq i \leq t} \beta^{(r)}_{n,j}(x_i, y_i) \wedge \bigwedge_{t < i \leq s} \gamma^{(r)}_{n,j}(x_i) \Big)$$

with

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

Conversely, assume that $\mathfrak{M} \models G'_{n,j}$. Let t, $x_1, \dots, x_s$, and $y_1, \dots, y_t$ be such that $H_{n,j,t}$ holds in $\mathfrak{M}$. Then if we set $R' = R^+_{n,j} \cup \{y_1, \dots, y_t\}$, 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 $\mathfrak{M} \models G'_{n,j}$ for some $\mathfrak{L}$-sentence which can be constructed effectively from $G_{n,j}$.

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

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

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

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

Note that the above definition of $T'_{n,j}$ means that all elements which were not marked yet and belong to some r-ball $B_r(x_i)$ are now marked negatively.

This completes the second step of the construction of X_n.

We can now define X_n as follows: for $n \geq 1$ 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

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

For $n = 0$, the definitions are the same but we omit the set R_{n-1} (respectively S_{n-1}) in the definition of R_n (respectively S_n).

In order to define F_n, consider a sentence F with quantifier rank n. By Theorem 1, F is equivalent to a formula F' which is a boolean combination of sentence of the form $G_{n,j}$. Consider the truth value of F' determined by setting "true" all sentences $G_{n,j} \in F'_{n,j}$, and "false" sentences $G_{n,j} \notin F'_{n,j}$. Then we define F_n as the union of F_{n-1} and of all sentences 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 8 are easy consequences of the construction of X_n (and the induction hypotheses).
- Let us consider condition 9. Let $R' \subseteq |\mathfrak{M}|$ be such that $R_n \subseteq R'$ and

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

  Let us prove that R' is not definable by any $\mathfrak{L}$-formula of quantifier rank $\leq n$. Since every subset of $|\mathfrak{M}|$ definable by a $\mathfrak{L}$-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 $\mathfrak{M}$ satisfies condition 3 of Theorem 2, thus there exists $x \in |\mathfrak{M}|$ 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 10. Let $R' \subseteq |\mathfrak{M}|$ be such that $R_n \subseteq R'$,

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

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

and $d(x,y) \geq 7^{n+1}$ whenever x, y are distinct elements of $R' \setminus R_n$. Let us prove that $FO_n(\mathfrak{M}(R')) = F_n$. 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 sentences $G_{n,j}$. Thus it is sufficient to prove that for every j we have $\mathfrak{M}(R') \models G_{n,j}$ 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 dealt with the sentence $G_{n,j}$. If $\mathfrak{M} \models G'_{n,j}$ then in this case $F'_{n,j} = \{G_{n,j}\}$, and the definition of $R''_{n,j}$ and $T'_{n,j}$ imply that the sentence $G_{n,j}$ holds for every R' which extends (in a convenient way) the marking (R_n, S_n, T_n), thus we have $\mathfrak{M}(R') \models G_{n,j}$. On the other hand if $\mathfrak{M} \not\models G'_{n,j}$, 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 $\mathfrak{M}(R') \not\models G_{n,j}$.

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

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

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

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

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

Let us prove that the elementary diagram of $\mathfrak{M}'$ is computable. Consider the function f used for the elementary diagram of $\mathfrak{M}$; it is sufficient to prove that $\{f(a) \mid \mathfrak{M}' \models R(a)\ ,\ a \in |\mathfrak{M}|\}$ is recursive. Since every element e of $|\mathfrak{M}|$ 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 $|\mathfrak{M}|$ is marked in $R^{\mathfrak{M}'}$ or in its complement. Moreover the whole construction is effective. This implies that both $\{f(a) \mid \mathfrak{M}' \models R(a)\ ,\ a \in |\mathfrak{M}|\}$ and $\{f(a) \mid \mathfrak{M}' \not\models R(a)\ ,\ a \in |\mathfrak{M}|\}$ 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 $\mathfrak{M}$ which ensures that $\mathfrak{M}$ is not maximal. A natural problem is to extend Theorem 2 to structures $\mathfrak{M}$ which do not satisfy condition 3. In particular one can consider 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; for some recent progress see [2].

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

5. Acknowledgement

We wish to thank the anonymous referee for useful suggestions.

References

1. A. Bès and P. Cégielski, *Weakly maximal decidable structures.* — RAIRO – Theor. Inf. Appl., **42**, No. 1 (2008), 137–135.
2. A. Blumensath, *Locality and Modular Ehrenfeucht-Fraisse Games.* — Preprint (2006).
3. Calvin C. Elgot and Michael O. Rabin, *Decidability and undecidability of extensions of second (first) order theory of (generalized) successor.* — J. Symb. Log., **31**, No. 2 (1966), 169–181.
4. H. Gaifman, *On local and non-local properties.* — In: Logic Colloquium 81, Proc. Herbrand Symp., Marseille 1981, Stud. Logic Found. Math. 107 (1982), pp. 105–135.
5. C. H. Langford, *Theorem on deducibility (second paper).* — *Annals of Math.*, **2** (1927), 459–471.
6. L. Libkin, *Elements of Finite Model Theory*, Springer (2004).
7. M. Presburger, *Über de vollständigkeit eines gewissen systems der arithmetic ganzer zahlen, in welchen, die addition als einzige operation hervortritt.* — In: Comptes Rendus du Premier Congrès des Mathématicienes des Pays Slaves, Warsaw (1927), pp. 92–101, 395.
8. Alexei L. Semenov, *Decidability of monadic theories.* — In: Michal Chytil and Václav Koubek, editors, *MFCS, Lecture Notes in Computer Science*, **176** (1984), pp. 162–175.
9. Valentina S. Harizanov, *Computably-theoretic complexity of countable structures.* — Bulletin of Symbolic Logic, **8** (2002), 457–477.
10. S. Soprunov, *Decidable expansions of structures.* — Vopr. Kibern., **134** (1988), 175–179.

LACL, Université Paris-Est, France

Поступило 22 мая 2007 г.

E-mail: E-mail:{bes,cegielski} @univ-paris12.fr