Localities and modular Ehrenfeucht–Fraïssé games

Achim Blumensath

TU Darmstadt, Mathematik, AG Logik, Schloßgartenstraße 7, 64289 Darmstadt, Germany

ARTICLE INFO

Article history:
Accepted 7 September 2010
Available online 14 September 2011

Keywords:
Ehrenfeucht–Fraïssé games
Locality
First-order logic

ABSTRACT

We study ways to simplify Ehrenfeucht–Fraïssé games. In particular, we consider decompositions of a structure and their effect on Ehrenfeucht–Fraïssé games. We investigate notions of locality and we present a generalisation of the theorem of Gaifman to linearly ordered structures.

© 2011 Published by Elsevier B.V.

1. Introduction

Ehrenfeucht–Fraïssé games provide a versatile tool for the investigation of logics and their expressive power. When compared to other methods these games have the advantage that they can easily be adapted to many different logics and that they also work well in the context of finite structures. Unfortunately, in nontrivial applications the complexity of playing these games quickly becomes unmanageable. Therefore, it was suggested by Fagin, Stockmeyer, and Vardi [8] to create a library of tools that can be used to simplify Ehrenfeucht–Fraïssé games.

For first-order logic, progress in this direction has been made by the theorems of Hanf [11] and Gaifman [10], and by the more recent results of Libkin et al. [12,1]. Games for existential monadic second-order logic were investigated by Fagin et al. and Schwentick [8,14,7,2].

A shortcoming of most of these results is that they can only be used on sparse structures, i.e., structures where the relations contain few tuples (a notable exception being [14]). The reason for this is the notion of locality the statements are based on. In non-sparse structures all elements are in the vicinity of each other. Hence, if we are interested in results for non-sparse structures we have to adopt a different notion of locality.

In this paper we look at notions of locality and their effect on Ehrenfeucht–Fraïssé games. We try to isolate special cases where games can be simplified. Besides sparse structures such cases turn out to be structures with a hierarchical decomposition and linearly ordered structures. In the first part we present several simple ideas to simplify games on non-sparse structures. The second part consists of a generalisation of the theorem of Gaifman that also gives meaningful results for certain structures that are non-sparse.

2. Preliminaries

Let us recall some basic definitions and fix our notation. Let \([n] := \{0, \ldots, n-1\}\). We tacitly identify tuples \(\bar{a} = a_0 \ldots a_{n-1} \in A^n\) with functions \([n] \rightarrow A\) and frequently we write \(\bar{a}\) for the set \(\{a_0, \ldots, a_{n-1}\}\). This allows us to write \(\bar{a} \subseteq \bar{b}\) or \(\bar{a} = \bar{b}\) for \(I \subseteq [n]\). We denote the empty tuple by \(\langle \rangle\).
We assume that the reader is familiar with basic concepts of first-order logic (see e.g. [6] for definitions and notation). We will only consider purely relational structures and we work with infinitary first-order logic throughout. The relation defined by a formula $\varphi$ in a structure $\mathfrak{A}$ is denoted by $\varphi^\mathfrak{A}$.

Two $\Sigma$-structures $\mathfrak{A}$ and $\mathfrak{B}$ are $m$-equivalent, in symbols $\mathfrak{A} \equiv_m \mathfrak{B}$, if they satisfy the same infinitary first-order sentences of quantifier rank at most $m$. We denote the quantifier rank of $\varphi$ by $\text{qr}(\varphi)$. For a tuple $\bar{a} \in A^n$, we write $(\mathfrak{A}, \bar{a})$ for the expansion of $\mathfrak{A}$ by constant symbols for each component $a_i$. In particular, $(\mathfrak{A}, \bar{a}) \equiv_m (\mathfrak{B}, \bar{b})$ means that the tuples $\bar{a}$ and $\bar{b}$ satisfy in their respective structure the same formulae of quantifier rank up to $m$.

The $m$ round Ehrenfeucht–Fraïssé game $\text{EF}_m(\mathfrak{A}, \mathfrak{B})$ between two $\Sigma$-structures $\mathfrak{A}$ and $\mathfrak{B}$ is played as follows. There are two players Spoiler and Duplicator who make moves in turn. In every round Spoiler selects either some element $a \in A$ or an element $b \in B$. Duplicator replies with an element of the other structure. Let $\bar{a} = a_0 \ldots a_m \in A^m$ and $\bar{b} = b_0 \ldots b_m \in B^m$ be the elements selected during the $m$ rounds. Duplicator wins the play if and only if the mapping $p = [(a_0, b_0), \ldots, (a_m, b_m)]$ is a partial isomorphism, that is, an isomorphism between the substructures induced by $\bar{a}$ and $\bar{b}$, respectively. To simplify notation we will denote such mappings $p$ by $\bar{a} \rightsquigarrow \bar{b}$.

A more algebraic way to look at Ehrenfeucht–Fraïssé games is via back-and-forth systems. Such a system consists of a sequence $(J_k)_{k \leq m}$ of partial isomorphisms from $\mathfrak{A}$ to $\mathfrak{B}$ with the following properties:

- **Forth property.** For every $\bar{a} \mapsto \bar{b} \in J_{k+1}$ and all $c \in A$, there exists an element $d \in B$ such that $\bar{a}c \mapsto \bar{b}d \in J_k$.
- **Back property.** For every $\bar{a} \mapsto \bar{b} \in J_{k+1}$ and all $d \in B$, there exists an element $c \in A$ such that $\bar{a}c \mapsto \bar{b}d \in J_k$.

**Theorem 1 (Ehrenfeucht–Fraïssé).** Let $\mathfrak{A}$ and $\mathfrak{B}$ be $\Sigma$-structures. The following statements are equivalent:

(a) $\mathfrak{A} \equiv_m \mathfrak{B}$,

(b) Duplicator has a winning strategy for $\text{EF}_m(\mathfrak{A}, \mathfrak{B})$.

(c) There exists a back-and-forth system $(J_k)_{k \leq m}$ between $\mathfrak{A}$ and $\mathfrak{B}$ such that $J_m \neq \emptyset$.

If the signature $\Sigma$ is finite then we can replace (a) by $m$-equivalence with respect to finitary first-order logic. The reason why we consider infinitary logic is that, for some constructions below, we need to introduce infinite signatures.

3. Decomposing structures

The first thing that comes to mind when tasked with simplifying an Ehrenfeucht–Fraïssé game is trying to break it down into simpler games. One way of doing so consists in decomposing the structures $\mathfrak{A}$ and $\mathfrak{B}$ in question into several parts $\mathfrak{A}_0, \ldots, \mathfrak{A}_{n-1}$ and $\mathfrak{B}_0, \ldots, \mathfrak{B}_{n-1}$ on which one can play separately. Therefore, we will study operations $f$ on structures for which there exists a function $g : \omega \to \omega$ such that

$$\mathfrak{A}_i \equiv_{g(m)} \mathfrak{B}_i,$$

for all $i$, implies $f(\mathfrak{A}_0, \ldots, \mathfrak{A}_{n-1}) \equiv_m f(\mathfrak{B}_0, \ldots, \mathfrak{B}_{n-1})$.

Let us recall several well-known instances of such operations. The canonical example consists of disjoint unions.

**Lemma 2.** If $\mathfrak{A}_0 \equiv_m \mathfrak{B}_0$ and $\mathfrak{A}_1 \equiv_m \mathfrak{B}_1$ then $\mathfrak{A}_0 \uplus \mathfrak{A}_1 \equiv_m \mathfrak{B}_0 \uplus \mathfrak{B}_1$.

For a proof, note that if Duplicator has strategies to win the games $\text{EF}_m(\mathfrak{A}_0, \mathfrak{B}_0)$ and $\text{EF}_m(\mathfrak{A}_1, \mathfrak{B}_1)$ then she can compose them to win the game $\text{EF}_m(\mathfrak{A}, \mathfrak{B})$. The key reason why this is possible is that one can select elements of one component without knowledge of which elements of the other component have been chosen. For unions that are not disjoint the situation is more complex since the component games are not independent. We will return to this more general case below.

An analogous result holds for direct products, although it will not be used in this article.

**Lemma 3.** If $\mathfrak{A}_0 \equiv_m \mathfrak{B}_0$ and $\mathfrak{A}_1 \equiv_m \mathfrak{B}_1$ then $\mathfrak{A}_0 \times \mathfrak{A}_1 \equiv_m \mathfrak{B}_0 \times \mathfrak{B}_1$.

In fact, this result and its version for disjoint unions can be generalised to infinitely many operands. (There are even stronger generalisations possible.)

**Theorem 4 (Feferman-Vaught).** If $\mathfrak{A}_i \equiv_m \mathfrak{B}_i$, for all $i \in I$, then

$$\bigcup_{i \in I} \mathfrak{A}_i \equiv_m \bigcup_{i \in I} \mathfrak{B}_i \quad \text{and} \quad \prod_{i \in I} \mathfrak{A}_i \equiv_m \prod_{i \in I} \mathfrak{B}_i.$$
Definition 5. Let $\Sigma$ and $\Gamma = \{R_0, \ldots, R_s\}$ be signatures. A $k$-dimensional first-order interpretation (from $\Sigma$ to $\Gamma$) is a list

$$I = (\delta(x), \epsilon(x, y), \varphi_{R_0}(\bar{x}_0, \ldots, \bar{x}_{n_0-1}), \ldots, \varphi_{R_s}(\bar{x}_0, \ldots, \bar{x}_{n_s-1}))$$

of infinitary first-order formulae over the signature $\Sigma$ where each of the tuples $\bar{x}, \bar{y}, \bar{z}$ above has length $k$. Such an interpretation defines the following operation on structures. It maps a $\Sigma$-structure $A$ to the $\Gamma$-structure

$$I(A) := (\delta^A, \epsilon^A, \varphi_{R_0}^A, \ldots, \varphi_{R_s}^A)/\epsilon^A.$$

That is, to construct $I(A)$ we first build the structure where the universe consists of all tuples $\bar{a} \in A^k$ satisfying $\delta$ and where the relations are defined by the formulae $\varphi_{R_i}$. Then we factorise by the relation defined by $\epsilon$. Of course, this only works if $\epsilon^A$ is a congruence of the structure $(\delta^A, \varphi_{R_0}^A, \ldots, \varphi_{R_s}^A)$. If this is not the case then $I(A)$ remains undefined.

Lemma 6. Let $I$ be a $k$-dimensional first-order interpretation where each formula has quantifier rank at most $r$. For every formula $\varphi$ over $\Gamma$, there exists a formula $I(\varphi)$ over $\Sigma$ of quantifier rank $qr(I(\varphi)) \leq k \cdot qr(\varphi) + r$ such that, for all structures $A$ where $I(A)$ is defined, we have

$$I(A) \models \varphi \iff A \models I(\varphi).$$

Corollary 7. Let $I$ be a $k$-dimensional first-order interpretation where each formula has quantifier rank at most $r$. If $A \equiv_{km+r} B$ then $I(A) \equiv_{m} I(B)$, provided these are defined.

One way to simplify proofs based on Ehrenfeucht–Fraïssé games with the help of interpretations consists in replacing the structures $A$ and $B$ by more convenient structures $A'$ and $B'$ such that $A = I(A')$ and $B = I(B')$, for some interpretation $I$. Then $A' \equiv_{km+r} B'$ implies $A \equiv_{m} B$. Of course, these new structures cannot be really simpler than the original ones since we can recover the latter from the former. But they might be more convenient to play on.

As an example of a more convenient game, suppose that we want to prove that spoiler wins the $m$ round game between $A$ and $B$. We might simplify his task by replacing these structures by expansions $A'$ and $B'$ with some additional, definable relations that make certain information directly available which, in the original game, spoiler would need several steps to check. For instance, one could add the immediate successor relation to a partial order. Then the player can check immediately whether two elements are immediate successors. Otherwise, he would need an additional move to select an element in between. If the definitions of the new relations have quantifier rank $r$ then, by Corollary 7, a proof that spoiler can win EF$_{m-r}$($A'$, $B'$) implies that he can also win the original game EF$_m(A, B)$.

Example. Consider finite linear orders $A = (A, <, \perp, \top)$ and $B = (B, <, \perp, \top)$ with constants for the least and greatest element. (Formally, we regard $\perp$ and $\top$ as unary predicates to remain in our purely relational framework.) If $|A|, |B| > 2^m$ then $A \equiv_{m} B$ (see, e.g., [6] Example 1.3.5). There exists an interpretation $I$ of quantifier rank 1 that defines the relation

$$E := \{(\top, \perp) \cup \{ (a, b) \mid b \text{ is the immediate successor of } a \}.$$

Thus, $I(A)$ is a cycle of length $|A|$. By Corollary 7, we have $I(A) \equiv_{m-1} I(B)$. It follows that, if $C$ and $D$ are cycles of length greater than $2^m$ then $C \equiv_{m-1} D$.

To obtain more substantial simplifications we can combine interpretations with other operations like disjoint unions. In the remainder of this section we will consider partitions of a structure that do not correspond to a disjoint union. We would like to apply the above techniques to this case.

Suppose that we have a partition $A_0 \cup A_1$ of $A$ and let $A_0$ and $A_1$ be the corresponding substructures of $A$. We would like to find an operation $f$ such that $f(A_0, A_1) = A$. It turns out that using the substructures $A_0$ and $A_1$ directly is not sufficient. We will use certain expansions $A_0^+$ and $A_1^+$ instead. The operations $f$ we will consider consist of a disjoint union followed by a one-dimensional quantifier-free interpretation. By the lemmas above it follows that, if $A$ and $B$ are structures that can be written as $A = f(A_0, A_1)$ and $B = f(B_0, B_1)$, for the same operation $f$, then

$$A_0^+ \equiv_m B_0^+ \text{ and } A_1^+ \equiv_m B_1^+ \text{ implies } A \equiv_m B.$$

In order to recover the structure $A$ from its substructures $A_0$ and $A_1$ we have to know which tuples $\bar{a}_0 \subseteq A_0$ and $\bar{a}_1 \subseteq A_1$ are connected by a relation. In the expansion $A_1^+$ we therefore colour all tuples by information about those tuples in the other component it is connected with.

Definition 8. Let $A$ be a structure and let $r$ be the maximal arity of a relation of $A$. For $1 \leq n < r$, let $C_n$ be a set of colours. A $C$-colouring of $A$ is a function $\chi$ that maps every tuple $\bar{a} \in A^n$ with $1 \leq n < r$ to a colour $\chi(\bar{a}) \in C_n$. By $\langle A, \chi \rangle$ we denote the expansion of $A$ by relations $R_c := \chi^{-1}(c)$, for every $c \in \bigcup_n C_n$. 


The following theorem is an immediate consequence of Lemma 2 and Corollary 7.

**Theorem 9.** Let $\mathfrak{A}$ and $\mathfrak{B}$ be $\Sigma$-structures with partitions $A = A_0 \cup A_1$ and $B = B_0 \cup B_1$. Suppose that there exists a quantifier-free interpretation $\mathcal{I}$ and colourings $\chi_i$ of $A_i$ and $\eta_i$ of $B_i$ such that

\[ \mathfrak{A} = \mathcal{I}(\langle \mathfrak{A}_0, \chi_0 \rangle \cup \langle \mathfrak{A}_1, \chi_1 \rangle) \quad \text{and} \quad \mathfrak{B} = \mathcal{I}(\langle \mathfrak{B}_0, \eta_0 \rangle \cup \langle \mathfrak{B}_1, \eta_1 \rangle). \]

If $(\mathfrak{A}_0, \chi_0) \equiv_m (\mathfrak{B}_0, \eta_0)$ and $(\mathfrak{A}_1, \chi_1) \equiv_m (\mathfrak{B}_1, \eta_1)$ then we have $\mathfrak{A} \equiv_m \mathfrak{B}$.

In order to use this theorem we have to find suitable colourings and interpretations. Let $A = A_0 \cup A_1$ be a partition of $\mathfrak{A}$. We start by defining colourings $\chi_i$ of $\mathfrak{A}_i$ such that,

\[ \mathfrak{A} = \mathcal{I}(\langle \mathfrak{A}_0, \chi_0 \rangle \cup \langle \mathfrak{A}_1, \chi_1 \rangle), \]

for some quantifier-free interpretation $\mathcal{I}$. There is a canonical choice for such colours. We can colour a tuple $\bar{a}$ by its external type as defined below.

**Definition 10.** Let $\mathfrak{A}$ be a $\Sigma$-structure and $X, U \subseteq A$.

(a) A formula $\varphi(\bar{x})$ is a literal if it is either atomic or the negation of an atomic formula. If, in addition, every variable $x_i$ really appears in $\varphi$ then we call $\varphi(\bar{x})$ a strict literal.

(b) The atomic type of a tuple $\bar{a} \subseteq A$ over a set $U \subseteq A$ of parameters is the set

\[ \text{atp}(\bar{a}/U) := \{ \varphi(\bar{x}, \bar{c}) \mid \varphi(\bar{x}, \bar{y}) \text{ a literal, } \bar{c} \subseteq U, \mathfrak{A} \models \varphi(\bar{a}, \bar{c}) \}. \]

For $U = \emptyset$, we just write $\text{atp}(\bar{a})$. The external type of $\bar{a}$ is the set

\[ \text{etp}(\bar{a}/U) := \text{atp}(\bar{a}/U) \smallsetminus \text{atp}(\bar{a}). \]

(c) For $\bar{a}, \bar{b} \subseteq A$, we define the type equivalence relation

\[ \bar{a} \equiv_U \bar{b} \quad \text{iff} \quad \text{etp}(\bar{a}/U) = \text{etp}(\bar{b}/U). \]

In the following we will not distinguish between an $\equiv_U$-class and the corresponding external type. In particular, we will speak of external types when we are actually dealing with $\equiv_U$-classes.

(d) We denote the set of all external $n$-types over $U$ realised in $X$ by

\[ I_n(X/U) := X^n/\equiv_U. \]

The union over all $n$ is

\[ I(X/U) := I_1(X/U) \cup \cdots \cup I_{r-1}(X/U), \]

where $r$ is the maximal arity of relations of $\mathfrak{A}$. For $U = A \smallsetminus X$, we introduce the shorthands

\[ I_n(X) := I_n(X/A \smallsetminus X) \quad \text{and} \quad I(X) := I(X/A \smallsetminus X). \]

(e) Set $\Sigma_{\mathcal{L}}(X) := \Sigma \cup \{ R_\tau \mid \tau \in I(X) \}$. The localisation of $\mathfrak{A}$ to $X$ is the $\Sigma_{\mathcal{L}}(X)$-structure

\[ \mathfrak{L}(X) := (\mathfrak{A}|_{X, \chi}) \]

where $\chi$ is the $I(X)$-colouring with $\chi(\bar{a}) := \text{etp}(\bar{a}/A \smallsetminus X)$.

Intuitively, the external type $\text{etp}(\bar{a}/U)$ of a tuple $\bar{a}$ records how $\bar{a}$ is connected via relations to the parameters in $U$. For instance, in a graph $\text{etp}(\bar{a}/U)$ contains all edges between components $a_i$ and elements of $U$. External types were introduced in [3,4] generalising work of Courcelle [5]. See also [13] for similar techniques.

**Example.** Let $(A, \leq, \bar{P})$ be a linear order with unary predicates $\bar{P}$. For every convex subset $C \subseteq A$, we have

\[ \bar{a} \equiv_{A \smallsetminus C} \bar{b} \quad \text{for all } \bar{a}, \bar{b} \in C^n. \]

When labelling tuples by their external type we can recover the original structure from its substructures with the help of a disjoint union and a quantifier-free interpretation.
Lemma 11. Let $\mathfrak{A}$ be a $\Sigma$-structure and $X \subseteq A$. There exists a one-dimensional quantifier-free interpretation $\mathcal{I}$ such that

$$\mathfrak{A} \equiv \mathcal{I}(\mathfrak{A}(X) \cup \mathfrak{A}(A \setminus X)).$$

If we are given two structures $\mathfrak{A}$ and $\mathfrak{B}$ and partitions $A_0 \cup A_1 = A$ and $B_0 \cup B_1 = B$ then it follows that there are interpretations $\mathcal{I}$ and $\mathcal{J}$ that reconstruct $\mathfrak{A}$ and $\mathfrak{B}$ from the respective localisations. But, in order to apply Theorem 9 we furthermore require that $\mathcal{I} = \mathcal{J}$. Note that, with our current definitions, this is never the case for the trivial reason that the sets of colours used by $\mathcal{L}(A_i)$ and $\mathcal{L}(B_i)$ are disjoint. Hence, we have to unify them by finding a suitable bijection mapping colours of $\mathcal{L}(A_i)$ to those of $\mathcal{L}(B_i)$. In order to be able to use the same interpretation for both structures we cannot use an arbitrary bijection between $I(A_i)$ and $I(B_i)$. We need one that respects the relations between tuples of these types. The following definition formalises this idea.

Definition 12. Let $\mathfrak{A}$ and $\mathfrak{B}$ be $\Sigma$-structures with partitions $A = X_0 \cup \cdots \cup X_{m-1}$ and $B = Y_0 \cup \cdots \cup Y_{m-1}$.

(a) A tuple $\bar{a}$ is scattered if there are at least two indices $i$ with $\bar{a} \cap X_i \neq \emptyset$.

(b) Two tuples $\bar{a} \in A^n$ and $\bar{b} \in B^n$ are congruent if we have

$$a_i \in X_k \iff b_i \in Y_k, \quad \text{for all } i < n \text{ and } k < m.$$

(c) For a family $\bar{g} = (g^n_i)_{n,i}$ of bijections $g^n_i : I_n(X_i) \rightarrow I_n(Y_i)$, for $i < m$ and $0 < n < r$,

we write $\bar{g} : \bar{X} \equiv \bar{Y}$ if, whenever $\bar{a}^i \subseteq X_i$ and $\bar{b}^i \subseteq Y_i$, $i < m$, are tuples such that

- $\bar{a}^0 \cdots \bar{a}^{m-1}$ and $\bar{b}^0 \cdots \bar{b}^{m-1}$ are congruent and scattered, and

- $g^n_i(\text{etp}(\bar{a}^i/A \setminus X_i))) = \text{etp}(\bar{b}^i/B \setminus Y_i)$, for all $i$,

then we have

$$\mathfrak{A} \models \varphi(\bar{a}^0, \ldots, \bar{a}^{m-1}) \iff \mathfrak{B} \models \varphi(\bar{b}^0, \ldots, \bar{b}^{m-1}),$$

for every strict literal $\varphi(\bar{x}^0, \ldots, \bar{x}^{m-1}).$

Example. Consider the following partitions of circles $\mathcal{C}_6$ and $\mathcal{C}_7$ where the labels represent the external 1-types:

![Diagram of partitions of circles]

Then we have

$$g_0^1 : X_0X_1 \approx Y_0Y_1,$$

where $g_0^1 : X_0 \rightarrow Y_0$ and $g_1^1 : X_1 \rightarrow Y_1$ are the functions $x \mapsto x'$.

Remark. (a) Let $\mathfrak{A}$ and $\mathfrak{B}$ be structures with partitions $\bar{X}$ and $\bar{Y}$. We denote the substructure of $\mathfrak{A}$ induced by $X_i$ by $\mathfrak{A}_i$ and the substructure of $\mathfrak{B}$ induced by $Y_i$ by $\mathfrak{B}_i$. There exist functions $\bar{g}$ with $\bar{g} : \bar{X} \equiv \bar{Y}$ if and only if, there exist colourings $\chi_i$ of $\mathfrak{A}_i$ and $\eta_i$ of $\mathfrak{B}_i$ and a one-dimensional quantifier-free interpretation $\mathcal{I}$ such that

$$\mathfrak{A} = \mathcal{I}((\mathfrak{A}_0, \chi_0) \cup \cdots \cup (\mathfrak{A}_{m-1}, \chi_{m-1}))$$

and

$$\mathfrak{B} = \mathcal{I}((\mathfrak{B}_0, \eta_0) \cup \cdots \cup (\mathfrak{B}_{m-1}, \eta_{m-1})).$$

(To decide whether $\mathcal{I}$ should add a tuple $\bar{c}$ to a relation $R$, it only has to consider which components of $\bar{c}$ belong to which set $X_i$ and what the colours $\chi_i(c_i)$ are, where $\bar{c}_i \subseteq \bar{c}$ is the subtuple in $X_i$.)

(b) Suppose that $\bar{g} : \bar{X} \approx \bar{Y}$. If $\bar{a}^i \subseteq X_i$ and $\bar{b}^i \subseteq Y_i$ are tuples such that, for all $i$,

$$\text{atp}(\bar{a}^i) = \text{atp}(\bar{b}^i) \quad \text{and} \quad g_i^{[\bar{a}^i]} = [\bar{b}^i],$$

then we have

$$\text{atp}(\bar{a}^0 \cdots \bar{a}^{m-1}) = \text{atp}(\bar{b}^0 \cdots \bar{b}^{m-1}).$$
Definition 13. Let $\Sigma$ and $\Gamma$ be signatures and $g : \Sigma \rightarrow \Gamma$ an arity preserving bijection. If $\mathfrak{A}$ is a $\Sigma$-structure and $\mathfrak{B}$ a $\Gamma$-structure then we write $\mathfrak{A} \equiv_{g}^{E} \mathfrak{B}$ if we have
\[ \mathfrak{A} \models \varphi \iff \mathfrak{B} \models \varphi^{g}, \] for all $\varphi$ of quantifier rank at most $m$.

where $\varphi^{g}$ is the formula obtained from $\varphi$ by replacing every relation symbol $R$ by $g(R)$.

Lemma 14. Let $\mathfrak{A}$ and $\mathfrak{B}$ be $\Sigma$-structures, $X_0, \ldots, X_{m-1} \subseteq A$ and $Y_0, \ldots, Y_{m-1} \subseteq B$ sequences of disjoint subsets, and $g : X \approx Y$. If $g_i : \Sigma_{\leq}(X_i) \rightarrow \Sigma_{\leq}(Y_i), i < m$, are the corresponding bijections between the signatures then
\[ \sigma(X_i) \equiv_{0}^{g_i} \sigma(Y_i), \] for all $i < m$.

implies $\text{atp}(\bar{a}^0 \ldots \bar{a}^{m-1}) = \text{atp}(\bar{b}^0 \ldots \bar{b}^{m-1})$.

Proof. Note that $(\ast)$ implies
\[ |\bar{a}^i| = |\bar{b}^i|, \quad \text{atp}(\bar{a}^i) = \text{atp}(\bar{b}^i), \quad \text{and} \quad g_i^{[\bar{a}^i]}[\bar{a}^i] = [\bar{b}^i]. \]

If there is at most one index $i$ with $|\bar{a}^i| > 0$ then we are done. Otherwise, the claim follows from $g : X \approx Y$ and the remark above. $\Box$.

With the help of Lemma 11 we can rewrite Theorem 9 in the following form.

Theorem 15. Let $\mathfrak{A}$ and $\mathfrak{B}$ be $\Sigma$-structures with partitions $X_0 \cup X_1 = A$ and $Y_0 \cup Y_1 = B$ of their universes. Suppose that $g : X_0 X_1 \approx Y_0 Y_1$ and let $g_i : \Sigma_{\leq}(X_i) \rightarrow \Sigma_{\leq}(Y_i), i < 2$, be the corresponding bijections between the signatures. If
\[ \mathcal{L}(X_0) \equiv_{m}^{g_0} \mathcal{L}(Y_0) \quad \text{and} \quad \mathcal{L}(X_1) \equiv_{m}^{g_1} \mathcal{L}(Y_1) \]
then we have $\mathfrak{A} \equiv_{m} \mathfrak{B}$.

Of course, whether we achieve a simplification this way largely depends on the existence of suitable partitions of the given structures, preferably with few external types between the components.

Example. Let $\mathfrak{A} = (A, E, <)$ and $\mathfrak{B} = (B, E, <)$ be undirected graphs equipped with an additional linear order. Suppose that $A = X_0 \cup X_1$ and $B = Y_0 \cup Y_1$ where every element of $X_0$ and $Y_0$ is less than all elements of, respectively, $X_1$ and $Y_1$. Set $\mathfrak{A}_i = \mathfrak{A}|_{X_i}$ and $\mathfrak{B}_i = \mathfrak{B}|_{Y_i}$. Let $\bar{a}^i$ be an enumeration of all elements of $\mathfrak{A}_i$ that are adjacent to some element of $\mathfrak{A}_{i-1}$, and let $\bar{b}^i$ be the elements of $\mathfrak{B}_i$ adjacent to some element of $\mathfrak{B}_{i-1}$.

Note that all tuples (of a given arity) disjoint from $\bar{a}^0, \bar{a}^1, \bar{b}^0, \bar{b}^1$ have the same external type. Consequently, if we have
\[ \text{atp}(\bar{a}^0 \bar{a}^1) = \text{atp}(\bar{b}^0 \bar{b}^1), \]
then we can find functions $g : X_0 X_1 \approx Y_0 Y_1$. ($g^n$ maps the unique external $n$-type over $X_{1-i}$ of a tuple disjoint from $\bar{a}^i$ to the unique external $n$-type over $Y_{1-i}$ of a tuple disjoint from $\bar{b}^i$. It maps the external $n$-type over $X_{1-i}$ of a tuple of the form $\bar{c} \bar{a}^i$ with $\bar{a}' \subseteq \bar{a}^i$ and $\bar{c} \cap \bar{a}^i = \emptyset$ to the external $n$-type over $Y_{1-i}$ of a tuple of the form $\bar{d} \bar{b}^i$ where $\bar{b}' \subseteq \bar{b}^i$ is the corresponding subtuple and $\bar{d} \cap \bar{b}^i = \emptyset$. Hence, the conditions
\[ (\mathfrak{A}_0, \bar{a}^0) \equiv_{m} (\mathfrak{B}_0, \bar{b}^0), \]
\[ (\mathfrak{A}_1, \bar{a}^1) \equiv_{m} (\mathfrak{B}_1, \bar{b}^1), \]
\[ \text{atp}(\bar{a}^0 \bar{a}^1) = \text{atp}(\bar{b}^0 \bar{b}^1) \]
imply that $\mathfrak{A} \equiv_{m} \mathfrak{B}$.

4. Contracting structures

In many applications the systems under consideration have a hierarchical structure. For instance, when designing a circuit diagram one usually assembles it in a modular way by using several predefined units. Usually, these units in turn consist of subunits which, again, might be built up from even simpler parts. When playing Ehrenfeucht–Fraïssé games on such structures one would like to take this hierarchy into account, e.g., by playing on various levels of abstraction where all units of a lower level are considered as black-boxes without internal structure. To do so we introduce an operation on structures that contracts a unit to a single point. After contracting all subunits we can play the game on the remaining structure.
Definition 16. Let \( \mathfrak{A} \) be a \( \Sigma \)-structure and \( X_0, \ldots, X_{n-1} \subseteq A \) a sequence of disjoint subsets such that, for all \( i, k < n \), there are functions 
\[
\bar{g}_{ik}: X_i, (A \setminus X_i) \cong X_k, (A \setminus X_k).
\]
We denote the components of \( \bar{g}_{ik} \) by \( (\bar{g}_{ik})_j^j \) for \( j \in [2] \) and \( 0 < n < r \).

(a) The \( \bar{X} \)-contraction \( C(\mathfrak{A}, \bar{X}) \) of \( \mathfrak{A} \) is obtained by replacing each set \( X_i \) by a single element \( x_i \) and adding auxiliary relations \( P_\tau \) that encode how the remaining elements were connected to those in \( X_i \). Formally, we define the universe of the contraction as 
\[
C := \left( A \setminus \bigcup_i X_i \right) \cup \{x_0, \ldots, x_{n-1}\},
\]
and the relations are
\[
R^{C(\mathfrak{A}, \bar{X})} := R^{\mathfrak{A}}|_C, \quad \text{for } R \in \Sigma,
\]
\[
P^{C(\mathfrak{A}, \bar{X})}_\tau := \left\{ (\bar{a}, x_i) \mid \atp(\bar{a}/X_i) = (g_{0i})^{(a)}_i(\tau) \right\}, \quad \text{for } \tau \in I(A \setminus X_0/X_0),
\]
and
\[
Q^{C(\mathfrak{A}, \bar{X})} := \{x_0, \ldots, x_{n-1}\}.
\]
(b) For a tuple \( \bar{a} \in A^m \), we define the local type w.r.t. \( \bar{X} \) as 
\[
\ltp(\bar{a}) := \langle \sim, F, (\tau_p)_p \rangle,
\]
where
\[
s \sim t \quad \text{iff} \quad a_s \in X_i \iff a_t \in X_i, \quad \text{for all } i < n,
\]
\[
F := \left\{ s < m \mid a_s \in A \setminus (X_0 \cup \ldots \cup X_{n-1}) \right\},
\]
\[
\tau_p := \atp(\bar{a}|_{F \cup p}), \quad \text{for } p \in [m]/\sim.
\]
(c) We call \( \mathfrak{A} \) globally uniform w.r.t. \( \bar{X} \) if, for all \( m < \omega \) and all tuples \( \bar{a}, \bar{b} \in A^m \), 
\[
\ltp(\bar{a}) = \ltp(\bar{b}) \quad \text{implies} \quad \atp(\bar{a}) = \atp(\bar{b}).
\]
(d) Suppose that \( \mathfrak{A} \) is globally uniform w.r.t. \( \bar{X} \) and let \( m < \omega \). The global \( m \)-type of \( \mathfrak{A}, \bar{X} \) is the set 
\[
gtp_m(\mathfrak{A}, \bar{X}) := \left\{ \ltp(\bar{a}) \mid \atp(\bar{a}) \in A^m \right\}.
\]
By definition, a local type of a tuple \( \bar{a} \) only records the relations between the components \( a_i \) in the same set \( X_k \) and the relations between those components and the components outside of \( \bigcup_k X_k \). A structure is globally uniform, if the missing relations, i.e., the relations between components in different sets \( X_k \) and \( X_i \), are determined by this information.

Theorem 17. Let \( \mathfrak{A} \) and \( \mathfrak{B} \) be structures and suppose that \( X_0, \ldots, X_{m-1} \subseteq A \) and \( Y_0, \ldots, Y_{n-1} \subseteq B \) are subsets such that \( \mathfrak{A} \) is globally uniform w.r.t. \( \bar{X} \) and \( \mathfrak{B} \) is globally uniform w.r.t. \( \bar{Y} \). Let 
\[
\bar{g}_{ik}: X_i, (A \setminus X_i) \cong X_k, (A \setminus X_k)
\]
and
\[
\bar{h}_{ik}: Y_i, (B \setminus Y_i) \cong Y_k, (B \setminus Y_k)
\]
be the corresponding functions and suppose that there exists functions \( \bar{f} \) such that 
\[
\bar{f} : X_0, (A \setminus X_0) \cong Y_0, (B \setminus Y_0).
\]
Let \( f \) be the bijection between signatures induced by \( \bar{f} \) and let \( f_{ij} \) be the corresponding bijection induced by the composition of \( \bar{g}_{i0}, \bar{f}, \) and \( \bar{h}_{0j} \). If 
\[
gtp_k(\mathfrak{A}, \bar{X}) = gtp_k((\mathfrak{B}, \bar{Y}),
\]
\[
\mathfrak{C}(\mathfrak{A}, \bar{X}) \equiv \mathfrak{C}(\mathfrak{B}, \bar{Y}),
\]

and
\[ \mathcal{L}(X_i) \equiv_{k} \mathcal{L}(Y_j), \quad \text{for all } i, j, \]
then \( \mathfrak{A} \equiv_{k} \mathfrak{B}. \)

**Proof.** Let \( \lambda : \mathfrak{A} \to \mathfrak{C}(\mathfrak{A}, \bar{X}) \) and \( \mu : \mathfrak{B} \to \mathfrak{C}(\mathfrak{B}, \bar{Y}) \) be the contraction maps with \( \lambda(X_i) = \{ x_i \} \) and \( \mu(Y_i) = \{ y_i \} \). We call a map \( \bar{a} \mapsto \bar{b} \) \( l \)-good, if \( \bar{a} \in \mathfrak{A}^{k-l} \) and \( \bar{b} \in \mathfrak{B}^{k-l} \),
\[ (\mathfrak{C}(\mathfrak{A}, \bar{X}), \lambda(\bar{a})) \equiv_{l} (\mathfrak{C}(\mathfrak{B}, \bar{Y}), \mu(\bar{b})), \]
and, for all indices \( i < m \) and \( j < n \) such that \( I := \{ s | a_s \in X_i \} \neq \emptyset \) and \( \bar{b}_{| I} \subseteq Y_j \), we have
\[ (\mathcal{L}(X_i), \bar{a}_{| I}) \equiv_{l} (\mathcal{L}(Y_j), \bar{b}_{| I}). \]
Let \( J_l \) be the set of all \( l \)-good maps. We claim that \( (J_l)_{l \in \mathbb{K}} : \mathfrak{A} \equiv_{k} \mathfrak{B}. \)

We have \( (\cdot) \mapsto (\cdot) \in J_k \neq \emptyset \). To check the forth property assume that \( \bar{a} \mapsto \bar{b} \in J_l \) and \( \bar{c} \in A \).

If \( \bar{c} \in \mathfrak{A} \setminus (X_0 \cup \cdots \cup X_{m-1}) \) then
\[ (\mathfrak{C}(\mathfrak{A}, \bar{X}), \lambda(\bar{a})) \equiv_{l} (\mathfrak{C}(\mathfrak{B}, \bar{Y}), \mu(\bar{b})). \]
implies that there is some element \( d \in B \) such that
\[ (\mathfrak{C}(\mathfrak{A}, \bar{X}), \lambda(\bar{a}c)) \equiv_{l-1} (\mathfrak{C}(\mathfrak{B}, \bar{Y}), \mu(\bar{b})d). \]
Hence, \( \bar{a}c \mapsto \bar{b}d \) is \( (l-1) \)-good.

Suppose that \( \bar{c} \in X_i \). If \( \bar{a} \cap X_i = \emptyset \) then, for suitable \( j \) and \( I \), we have
\[ (\mathcal{L}(X_i), \bar{a}_{| I}) \equiv_{l} (\mathcal{L}(Y_j), \bar{b}_{| I}), \]
which implies that we can find an element \( d \in Y_j \) with
\[ (\mathcal{L}(X_i), \bar{a}_{| I}c) \equiv_{l-1} (\mathcal{L}(Y_j), \bar{b}_{| I}d). \]
Hence, \( \bar{a}c \mapsto \bar{b}d \) is \( (l-1) \)-good. If \( \bar{a} \cap X_i = \emptyset \) then
\[ (\mathfrak{C}(\mathfrak{A}, \bar{X}), \lambda(\bar{a})) \equiv_{l} (\mathfrak{C}(\mathfrak{B}, \bar{Y}), \mu(\bar{b})). \]
implies that there is some index \( j < n \) such that
\[ (\mathfrak{C}(\mathfrak{A}, \bar{X}), \lambda(\bar{a})x_i) \equiv_{l-1} (\mathfrak{C}(\mathfrak{B}, \bar{Y}), \mu(\bar{b})y_j). \]
In particular, \( \bar{b} \cap Y_j = \emptyset \). Let \( d \) be an element such that
\[ (\mathcal{L}(X_i), c) \equiv_{l-1} (\mathcal{L}(Y_j), c). \]
Then \( \bar{a}c \mapsto \bar{b}d \) is \( (l-1) \)-good.

The back property follows by symmetry. It therefore remains to prove that every \( \bar{a} \mapsto \bar{b} \in J_0 \) is a partial isomorphism. Fix a subtuple \( \bar{a}^i \subseteq \bar{a} \) and let \( \bar{b}^i \subseteq \bar{b} \) be the corresponding subtuple of \( \bar{b} \). We prove by induction on \( |\bar{a}^i| = |\bar{b}^i| \) that \( \text{atp}(\bar{a}^i) = \text{atp}(\bar{b}^i) \).

First, suppose that \( \bar{a}^i \) is not scattered. Then \( \bar{a}^i = \bar{a}_0 \cup \bar{a}_1 \) and \( \bar{b}^i = \bar{b}_0 \cup \bar{b}_1 \) where \( \bar{a}_0 \subseteq X_i \) and \( \bar{b}_0 \subseteq Y_j \), for some \( i, j \), while \( \bar{a}_1 \subseteq A \setminus \bigcup_i X_i \) and \( \bar{b}_1 \subseteq B \setminus \bigcup_j Y_j \). Hence,
\[ (\mathcal{L}(X_i), \bar{a}_0) \equiv_{0} (\mathcal{L}(Y_j), \bar{b}_0) \]
implies
\[ \text{etp}(\bar{a}_0/A \setminus X_i) = \text{etp}(\bar{b}_0/B \setminus Y_j). \]
Therefore \( \bar{a}^i \) and \( \bar{b}^i \) satisfy the same strict literals. By induction hypothesis, it follows that \( \text{atp}(\bar{a}^i) = \text{atp}(\bar{b}^i) \).

For a scattered tuple \( \bar{a}^i \), we can use the induction hypothesis to show that \( \text{ltp}(\bar{a}^i) = \text{ltp}(\bar{b}^i) \). Since \( \mathfrak{A} \) and \( \mathfrak{B} \) are globally uniform with the same global type, it follows that \( \text{atp}(\bar{a}^i) = \text{atp}(\bar{b}^i) \). \( \square \)
Example. Let $\mathfrak{A}_n$ be the graph consisting of a cycle of length $n$ to every point of which is attached a path of length $n$ as follows:

Let $X_{n0}, \ldots, X_{n,n-1}$ be the sets of vertices of the attached paths. Then $\mathcal{E}(X_{ni})$ is a path of length $n$ and $\mathcal{C}(\mathfrak{A}_n, \bar{X}^n)$ is a cycle of length $n$ with additional edges attached at every vertex. Note that all elements of $X_{ni}$ have the same external type, while there are two different external types for the elements of $\mathfrak{A}_n \setminus X_{ni}$. It follows that there are functions $\bar{g}_{ik}: X_{ni}, (\mathfrak{A}_n \setminus X_{ni}) \approx X_{nk}, (\mathfrak{A}_n \setminus X_{nk})$.

Furthermore, the structure $\mathfrak{A}_n$ is globally uniform w.r.t. $\bar{X}^n$ and $\text{gtp}_k(\mathfrak{A}_m, \bar{X}^m) = \text{gtp}_k(\mathfrak{A}_n, \bar{X}^n)$, for all $k, m, n$.

For $m, n \geq 2^k$, we have

$$\mathcal{E}(\mathfrak{A}_m, \bar{X}^m) \equiv_k \mathcal{E}(\mathfrak{A}_n, \bar{X}^n) \quad \text{and} \quad \mathcal{L}(X^n) \equiv_k \mathcal{L}(X^n).$$

By the theorem it follows that $\mathfrak{A}_m \equiv_k \mathfrak{A}_n$.

5. Gluing structures

In the preceding sections we have considered decompositions of a structure into disjoint parts. Now we study decompositions into two parts that overlap. Suppose that we have a subset $Z \subseteq A$ and some notion of distance between elements of $\mathfrak{A}$. For $i < n$, let $X_i$ be the set of all elements whose distance from $Z$ is $i$, and let $X_n$ contain the remaining elements. We are interested in the decomposition of $\mathfrak{A}$ into the sets $X_0 \cup \cdots \cup X_{n-1}$ and $X_1 \cup \cdots \cup X_n$. This situation arose in [14]. We present a slightly generalised version of those results rephrased to fit our terminology.

**Definition 18.** Let $\mathfrak{A}$ be a $\Sigma$-structure and $X_0 \cup \cdots \cup X_n = A$ a partition of its universe. We set $X_{<k} := X_0 \cup \cdots \cup X_{k-1}$ and $X_{>k} := X_{k+1} \cup \cdots \cup X_n$.

(a) The **inner part** of $\mathfrak{A}$ is

$$\mathfrak{J}(\bar{X}) := (\mathfrak{A}|_{A \setminus X_n}, X_0, \ldots, X_{n-1}, (p^1_\tau), \ldots, (p^{n-1}_\tau))$$

where

$$p^i_\tau := \{ \bar{a} \subseteq X_{<i} \mid \text{etp}(\bar{a}/X_{>i}) = \tau \}.$$ 

Analogously, we define the **outer part** of $\mathfrak{A}$ by $\mathfrak{O}(X_0 \ldots X_n) := \mathfrak{J}(X_n \ldots X_0)$.

(b) Let $\mathfrak{B}$ be another $\Sigma$-structure and $Y_0 \cup \cdots \cup Y_n = B$ a partition of its universe. For bijections

$$g^k_i : I_k(X_{<i}/X_{>i}) \to I_k(Y_{<i}/Y_{>i}), \quad 0 \leq i < n, \quad 0 < k < r,$$

$$h^k_i : I_k(X_{>i}/X_{<i}) \to I_k(Y_{<i}/Y_{>i}), \quad 0 \leq i < n, \quad 0 < k < r,$$

we write...
\[ \bar{g}, \bar{h} : \bar{X} \approx \bar{Y} \]

if, whenever there is an index \(0 < l < n\) and tuples \(\bar{a} \subseteq X_{<l}, \bar{c} \subseteq X_{<l}, \bar{b} \subseteq Y_{<l}, \) and \(\bar{d} \subseteq Y_{<l}\) such that

- \(|\bar{a}| = |\bar{b}| > 0\) and \(|\bar{c}| = |\bar{d}| > 0\),
- \(g_i[\bar{a}] = [\bar{b}]\) and \(h_i[\bar{c}] = [\bar{d}]\)

then we have

\[ \mathfrak{A} \models \varphi(\bar{a}, \bar{c}) \iff \mathfrak{B} \models \varphi(\bar{b}, \bar{d}), \]

for every strict literal \(\varphi(\bar{x}, \bar{y})\).

If the distance between \(X_0\) and \(X_n\) is large enough then we can play the game separately on \(\mathcal{J}(\bar{X})\) and \(\mathcal{O}(\bar{X})\).

**Theorem 19.** Let \(\mathfrak{A}\) and \(\mathfrak{B}\) be structures and \(X_0 \cup \cdots \cup X_n = A\) and \(Y_0 \cup \cdots \cup Y_n = B\) partitions of their universes. Suppose that \(m\) is a number such that \(2^m \leq n\). If there are bijections \(\bar{g}\) and \(\bar{h}\) such that

\[ \bar{g}, \bar{h} : \bar{X} \approx \bar{Y}, \quad \mathcal{J}(\bar{X}) \equiv_m^h \mathcal{J}(\bar{Y}), \quad \text{and} \quad \mathcal{O}(\bar{X}) \equiv_m^h \mathcal{O}(\bar{Y}), \]

where \(g\) and \(h\) are the bijections of the signatures corresponding to, respectively, \(\bar{g}\) and \(\bar{h}\). Then \(\mathfrak{A} \equiv_m \mathfrak{B}\).

**Proof.** For a set \(Z \subseteq A\), we set

\[ \mu(Z) : = \max\{|-1| \cup \{k \mid Z \cap X_k \neq \emptyset\}\}, \]

and

\[ \nu(Z) : = \min\{|n + 1| \cup \{k \mid Z \cap X_k \neq \emptyset\}\}. \]

We partition every tuple \(\bar{a} \subseteq A\) as \(\bar{a} = \lambda(\bar{a}) \cup \varrho(\bar{a})\) with left part \(\lambda(\bar{a}) \subseteq A \setminus X_n\) and right part \(\varrho(\bar{a}) \subseteq A \setminus X_0\) by induction on \(|\bar{a}|\). We set \(\lambda(\emptyset) : = \emptyset\) and \(\varrho(\emptyset) : = \emptyset\). For nonempty tuples \(\bar{ac}\), we consider two cases. Suppose that \(c \in X_l\). If \(c\) is nearer to the left part of \(\bar{a}\) than to its right part, that is, if

\[ l \leq \frac{1}{2}(\nu(\varrho(\bar{a})) + \mu(\lambda(\bar{a}))), \]

then we add \(c\) to the left part, i.e., we set

\[ \lambda(\bar{ac}) : = \lambda(\bar{a})c \quad \text{and} \quad \varrho(\bar{ac}) : = \varrho(\bar{a}). \]

Otherwise, we define

\[ \lambda(\bar{ac}) : = \lambda(\bar{a}) \quad \text{and} \quad \varrho(\bar{ac}) : = \varrho(\bar{a})c. \]

For \(\bar{b} \subseteq B\), we define \(\lambda(\bar{b})\) and \(\varrho(\bar{b})\) analogously. Note that, if \(k : = |\bar{a}| \leq m\) then

\[ \nu(\varrho(\bar{a})) - \mu(\lambda(\bar{a})) \geq (n + 2)2^{-k} > n2^{-m} \geq 1. \]

Define

\[ J_k : = \{ \bar{a} \mapsto \bar{b} \mid (\mathcal{J}(\bar{X}), \lambda(\bar{a})) \equiv_k^g (\mathcal{J}(\bar{Y}), \lambda(\bar{b})) \} \]

We claim that \((J_k)_{k < m} : \mathfrak{A} \equiv_m \mathfrak{B}\). By definition, we have \(\emptyset \mapsto \emptyset \in J_m \neq \emptyset\).

To check the forth property, let \(\bar{a} \mapsto \bar{b} \in J_k\) and \(c \in A\). By symmetry, we may assume that \(\lambda(\bar{ac}) \neq \lambda(\bar{a})\) and \(\varrho(\bar{ac}) = \varrho(\bar{a})\).

Since \((\mathcal{J}(\bar{X}), \lambda(\bar{a})) \equiv_k^g (\mathcal{J}(\bar{Y}), \lambda(\bar{b}))\) there is some element \(d \in B \cup Y_n\) such that

\[ (\mathcal{J}(\bar{X}), \lambda(\bar{ac})c) \equiv_{k-1}^g (\mathcal{J}(\bar{Y}), \lambda(\bar{bd})). \]

Consequently, we have \(\bar{ac} \mapsto \bar{bd} \in J_{k-1}\).

The back property follows by symmetry. It remains to show that every \(\bar{a} \mapsto \bar{b} \in J_0\) is a partial isomorphism. Suppose that \(\mathfrak{A} \models \varphi(\bar{a})\), for some literal \(\varphi\). If \(\lambda(\bar{a}) = \emptyset\) then

\[ (\mathcal{O}(\bar{X}), \varrho(\bar{a})) \equiv_0^h (\mathcal{O}(\bar{Y}), \varrho(\bar{b})). \]

implies \(\mathfrak{B} \models \varphi(\bar{b})\). In a similar way it follows that \(\varrho(\bar{a}) = \emptyset\) implies \(\mathfrak{B} \models \varphi(\bar{b})\).

Therefore, we may assume that \(\lambda(\bar{a})\) and \(\varrho(\bar{a})\) are both nonempty. There exists some index \(l < n\) such that \(\lambda(\bar{a}) \subseteq X_{<l}\) and \(\varrho(\bar{a}) \subseteq X_{<l}\). Since

\[ (\mathcal{J}(\bar{X}), \lambda(\bar{a})) \equiv_0^g (\mathcal{J}(\bar{Y}), \lambda(\bar{b})) \]


we have
\[ g(\text{etp}(\lambda(a)/X_{\omega})) = \text{etp}(\lambda(b)/Y_{\omega}) \].
Again it follows that \( \mathfrak{B} \models \varphi(\bar{b}) \).

**Example.** Let \( \mathfrak{A} = (A, <, \bar{R}) \) be a finite linearly ordered structure. We define the distance between two elements \( a, b \in A \) by
\[ d(a, b) := |\{c \in A \mid a < c \leq b \text{ or } b < c \leq a\}|. \]
Suppose that the relations \( R_i \) are local in the sense that there is a number \( k \) such that, for every tuple \( \bar{a} \in R_i \), we have \( d(a_i, a_j) \leq k \), for all \( i, j \).
Let \( C \subseteq A \) be a subset that is convex with respect to \( < \). If we are given a second structure \( \mathfrak{B} \) with a convex subset \( D \subseteq B \) such that \( \mathfrak{B} \models \varphi \models \mathfrak{A} \), then we can apply the above machinery by defining
\[ X_i := \{a \in A \mid k(i-1) < d(a, c) \leq ki \text{ for some } c \in C\}, \]
\[ Y_i := \{b \in B \mid k(i-1) < d(b, c) \leq ki \text{ for some } c \in D\}. \]

Since \( O(\bar{X}) \models O(\bar{Y}) \) we only have to prove that \( \mathfrak{I}(\bar{X}) = \mathfrak{I}(\bar{Y}) \).

6. The theorem of Gaifman for globally uniform structures

The theorem of Gaifman provides a powerful method for proving expressibility results. Informally, the theorem states that, in order to determine whether a given first-order formula holds, we only have to count how many disjoint, \( m \)-equivalent neighbourhoods we can find in the structure under consideration. This theorem is mainly useful if the neighbourhoods are small or if the structure is sparsely sparse, i.e., if its relations contain few tuples.

Unfortunately, if the structures in question are non-sparse the statement of the theorem can become trivial since neighbourhoods might encompass the whole structure. Nevertheless there are examples of successful arguments using Ehrenfeucht–Fraïssé games on non-sparse structures like linear orderings or Presburger Arithmetic (see, e.g., [6,9]). Furthermore, these arguments seem also to be based on a notion of locality. Therefore, there is hope to generalise Gaifman’s theorem to cover these cases. For linear orderings, we will present such a generalisation in the next section.

In order to obtain a meaningful generalisation of the theorem of Gaifman we need to consider other metrics. Hence, we start in this section with defining a quite general notion of a metric. For every element \( a \) of our structure, we assume that we are given some set \( N_k(a) \) which we interpret as the set of all elements whose distance to \( a \) is at most \( k \). In order for these sets \( N_k(a) \) to induce a reasonable notion of distance we require them to satisfy some simple axioms.

**Definition 20.** Let \( \mathfrak{A} \) be a structure.

(a) Let \( N_k(a) \subseteq A \), for \( a \in A \) and \( k < \omega \), be a family of sets. We call \( N = (N_k(a))_{a,k} \) a system of neighbourhoods if, for all \( a \in A \) and every \( k < \omega \), the following conditions are satisfied:
- \( a \in N_0(a) \),
- \( N_k(a) \subseteq N_{k+1}(a) \),
- There is an increasing function \( \zeta : \omega \rightarrow \omega \) such that, for all \( a, b \in A \),
\[ b \in N_k(a) \text{ implies } N_k(a) \subseteq N_{\zeta(k)}(b). \]

For \( X \subseteq A \), we set \( N_k(X) := \bigcup_{a \in X} N_k(a) \). We write \( \zeta^n(x) \) for the \( n \)-th iteration of \( \zeta \), i.e., \( \zeta^0(x) := x \) and \( \zeta^{n+1}(x) := \zeta(\zeta^n(x)) \).
(b) Let \( N \) be a system of neighbourhoods. A subset \( X \subseteq A \) is \( k \)-scattered (w.r.t. \( N \)) if \( a \notin N_k(b) \), for all \( a, b \in X \) with \( a \neq b \).

Let us collect some basic properties of systems of neighbourhoods.

**Lemma 21.** Let \( N \) be a system of neighbourhoods.

(a) If \( b \in N_k(a) \) then \( N_{\zeta(k)}(b) \subseteq N_{\zeta^2(k)}(a) \).
(b) If \( N_k(b) \not\subseteq N_{\zeta(k)}(a) \) then \( N_k(a) \cap N_k(b) = \emptyset \).
(c) \( N_{i(k)}(N_k(a)) \subseteq N_{i^2(k)}(a) \).

d) If \( X \) is \( \xi(k) \)-scattered then
\[
|X \cap N_k(c)| \leq 1, \quad \text{for all } c \in A.
\]

**Proof.** (a) By definition, \( b \in N_k(a) \) implies \( N_k(a) \subseteq N_{\xi(k)}(b) \). In particular, \( a \in N_{\xi(k)}(b) \) which in turn implies \( N_{\xi(k)}(b) \subseteq N_{\xi^2(k)}(a) \).

(b) Suppose that \( c \in N_k(a) \cap N_k(b) \neq \emptyset \). Then \( c \in N_k(a) \) implies \( N_{\xi(k)}(c) \subseteq N_{\xi(k)}(a) \), and \( c \in N_k(b) \) implies \( N_{\xi(k)}(b) \subseteq N_{\xi(k)}(c) \). It follows that \( N_k(b) \subseteq N_{\xi^2(k)}(a) \).

(c) If \( b \in N_k(a) \) and \( c \in N_{\xi(k)}(b) \) then it follows by (a) that
\[
c \in N_{\xi(k)}(b) \subseteq N_{\xi^2(k)}(a).
\]

(d) Let \( a, b \in X \), \( a \neq b \). If \( a \in N_k(c) \) then \( N_k(c) \subseteq N_{\xi(k)}(a) \). Therefore, \( b \notin N_{\xi(k)}(a) \) implies \( b \notin N_k(c) \). \qed

In case of the usual Gaifman metric the distance between two elements is first-order definable. For general metrics this does not need to be the case. Therefore, we add new relations encoding the distances.

**Definition 22.** Let \( \mathcal{A} \) be a structure and \( N \) a system of neighbourhoods.

For \( \bar{a} \subseteq A \) we set
\[
\mathcal{N}_k(\bar{a}) := (\mathcal{A}_{|N_k(\bar{a})}, (D_i)_{i \leq k}, \bar{a})
\]
where \( D_i := \{ (b, c) \mid c \in N_i(b) \} \).

Intuitively the reason why the theorem of Gaifman holds is that elements that are far away cannot be distinguished by atomic formulae. In order to generalise the theorem to other notions of distance we have to require the same property.

**Definition 23.** Let \( \mathcal{A} \) be a structure with system of neighbourhoods \( N \).

(a) We call \( \mathcal{A} \) **globally uniform** w.r.t. \( N \) if, whenever \( \bar{a}, \bar{a}', \bar{b}, \bar{b}' \) are tuples such that
- \( |\bar{a}| = |\bar{a}'| \) and \( |\bar{b}| = |\bar{b}'| \),
- \( \bar{b} \cap N_0(\bar{a}) = \emptyset \) and \( \bar{b}' \cap N_0(\bar{a}') = \emptyset \),
- \( \text{atp}(\bar{a}) = \text{atp}(\bar{a}') \) and \( \text{atp}(\bar{b}) = \text{atp}(\bar{b}') \),
then we have
\[
\text{atp}(\bar{a} \bar{b}) = \text{atp}(\bar{a}' \bar{b}').
\]

(b) If \( \mathcal{A} \) is globally uniform w.r.t. \( N \) and \( k, m < \omega \), we define the **global type** of \( \mathcal{A} \) as
\[
\text{gtp}_{k,m}(\mathcal{A}) := \{ [\text{atp}(\bar{a}), \text{atp}(\bar{b}), \text{atp}(\bar{a} \bar{b})] \mid \bar{a} \in A^k, \bar{b} \in A^m, \bar{b} \cap N_0(\bar{a}) = \emptyset \}.
\]

The next lemma shows that globally uniform structures satisfy our requirement that far away elements are indistinguishable.

**Lemma 24.** Let \( \mathcal{A} \) and \( \mathcal{B} \) be globally uniform structures such that
\[
\text{gtp}_{k,m}(\mathcal{A}) = \text{gtp}_{k,m}(\mathcal{B}), \quad \text{for all } k, m < \omega.
\]

Let \( \bar{a}_0, \bar{a}_1 \subseteq A \) and \( \bar{b}_0, \bar{b}_1 \subseteq B \) be tuples such that
\[
N_i(\bar{b}_1) \cap N_{i^2}(\bar{a}_1) = \emptyset, \quad \text{for both } i.
\]

If we have
\[
\mathcal{N}_i(\bar{a}_0) \equiv_n \mathcal{N}_i(\bar{b}_0) \quad \text{and} \quad \mathcal{N}_i(\bar{a}_1) \equiv_n \mathcal{N}_i(\bar{b}_1)
\]
then \( \mathcal{N}_i(\bar{a}_0 \bar{a}_1) \equiv_n \mathcal{N}_i(\bar{b}_0 \bar{b}_1) \).

**Proof.** For a tuple \( \bar{c} \subseteq A \), we set \( \bar{c} := \bar{c} \cap N_i(\bar{a}_i) \), and analogously for tuples \( \bar{d} \subseteq B \). We claim that
\[
(J_k)_{k \leq n} : \mathcal{N}_i(\bar{a}_0 \bar{a}_1) \equiv_n \mathcal{N}_i(\bar{b}_0 \bar{b}_1),
\]
where
\[ J_k := \{ \bar{c} \mapsto \bar{d} \mid |\bar{c}| = |\bar{d}| = n - k, (\mathcal{N}_r(\bar{a}_i), \bar{c}_i) \equiv_k (\mathcal{N}_r(\bar{b}_i), \bar{d}_i) \text{ for both } i \}. \]

By definition, we have \( \langle \rangle \mapsto \langle \rangle \in J_0 \neq \emptyset \) and the back-and-forth property is easily verified. Hence, we only need to show that every \( \bar{c} \mapsto \bar{d} \in J_0 \) is a partial isomorphism.

By definition of \( J_0 \), we have atp(\(\bar{c}_0\)) = atp(\(\bar{d}_0\)) and atp(\(\bar{c}_1\)) = atp(\(\bar{d}_1\)). Furthermore, by Lemma 21(c), \( N_r(\bar{b}_1) \cap N_{g^2(\phi)}(\bar{a}_1) = \emptyset \) implies that
\[ N_r(\bar{b}_1) \cap N_0(N_r(\bar{a}_1)) = \emptyset. \]

Hence, we have \( \bar{c}_1 \cap N_0(\bar{c}_0) = \emptyset \) and \( \bar{d}_1 \cap N_0(\bar{d}_0) = \emptyset \). Since \( \mathfrak{A} \) and \( \mathfrak{B} \) are globally uniform and their global types coincide it follows that atp(\(\bar{c}_0\bar{c}_1\)) = atp(\(\bar{d}_0\bar{d}_1\)). Consequently, \( \bar{c} \mapsto \bar{d} \) is a partial isomorphism. \( \Box \)

After these preparations we can prove an analogue of the theorem of Gaifman for globally uniform systems of neighbourhoods.

**Definition 25.**

(a) A sentence \( \psi \) is basic local if it is of the form
\[ \exists \bar{x} \left( \bar{x} \text{ is } r\text{-scattered}^* \land \bigwedge_i \psi(N_r(x_i))(x_i) \right), \]
where \( \psi(N_r(x))(x) \) denotes the relativisation of \( \psi \) to \( N_r(x) \), i.e., the formula obtained from \( \psi \) by replacing every quantifier \( Q \) with \( Q \in [\exists, \forall] \) by a relativised quantifier \( (Q y \in N_r(x)) \). Formally, we replace all subformulae of the form \( \exists y \vartheta \) or \( \forall y \vartheta \) by, respectively, \( \exists y(y \in N_r(x) \land \vartheta) \) and \( \forall y(y \in N_r(x) \rightarrow \vartheta) \).

(b) A sentence \( \psi \) is basic global if it is of the form
\[ \exists \exists y \left( y \cap N_0(\bar{x}) = \emptyset \land \psi(\bar{x}, \bar{y}) \right) \]
where \( \psi \) is quantifier-free.

**Lemma 26.** If \( \mathfrak{A} \) and \( \mathfrak{B} \) are globally uniform structures that satisfy the same basic global sentences then we have
\[ gtp_{k,m}(\mathfrak{A}) = gtp_{k,m}(\mathfrak{B}), \quad \text{for all } k, m < \omega. \]

**Theorem 27.** Let \( \mathfrak{A} \) and \( \mathfrak{B} \) be globally uniform structures such that
\[ gtp_{k,m}(\mathfrak{A}) = gtp_{k,m}(\mathfrak{B}), \quad \text{for all } k, m < \omega. \]
If \( \mathfrak{A} \) and \( \mathfrak{B} \) satisfy the same basic local sentences then \( \mathfrak{A} \equiv_{\omega} \mathfrak{B} \).

**Proof.** It is sufficient to show that \( \mathfrak{A} \equiv_{\omega} \mathfrak{B} \), for all \( m < \omega \). Fix \( m < \omega \) and set \( \varrho(k) := \zeta^k(0) \). We define
\[ J_k := \{ \bar{a} \mapsto \bar{b} \mid \bar{a} \in A^{m-k}, \bar{b} \in B^{m-k}, \mathcal{N}_{\varrho(\delta k)}(\bar{a}) \equiv_{mk+1} \mathcal{N}_{\varrho(\delta k)}(\bar{b}) \}. \]
We claim that \( (J_k)_{k \leq m} \) is a back-and-forth system for \( \mathfrak{A} \) and \( \mathfrak{B} \). Since \( \langle \rangle \mapsto \langle \rangle \in J_m \) it then follows that \( \mathfrak{A} \equiv_{\omega} \mathfrak{B} \). Clearly, every \( \bar{a} \mapsto \bar{b} \in J_0 \) is a partial isomorphism. Therefore, we only need to prove the back-and-forth property.

By symmetry it is sufficient to consider the forth property. Let \( \bar{a} \mapsto \bar{b} \in J_k \) and \( c \in A \). We distinguish two cases.

First, suppose that \( N_{\varrho(\delta k-\delta)}(c) \subseteq N_{\varrho(\delta k)}(\bar{a}) \). Since
\[ \mathcal{N}_{\varrho(\delta k)}(\bar{a}) \equiv_{mk+1} \mathcal{N}_{\varrho(\delta k)}(\bar{b}) \]
we can find some \( d \in N_{\varrho(\delta k)}(\bar{b}) \) such that
\[ (\mathcal{N}_{\varrho(\delta k)}(\bar{a}), c) \equiv_{mk} (\mathcal{N}_{\varrho(\delta k)}(\bar{b}), d). \]
It follows that
\[ \mathcal{N}_{\varrho(\delta (k-1))}(\bar{a}c) \equiv_{mk} \mathcal{N}_{\varrho(\delta (k-1))}(\bar{bd}) \]
which implies that \( \bar{a}c \mapsto \bar{bd} \in J_{k-1} \).
It remains to consider the case that $N_{\varphi(8k-6)}(c) \nsubseteq N_{\varphi(8k)}(\bar{a})$. Then we have

$$N_{\varphi(8k-6)}(c) \cap N_{\varphi(8k-6)}(\bar{a}) = \emptyset,$$

by Lemma 21(b). If we find an element $d \in B$ such that

$$N_{\varphi(8k-8)}(d) \cap N_{\varphi(8k-6)}(\bar{b}) = \emptyset$$

and

$$\mathcal{M}_{\varphi(8k-8)}(c) \equiv_m (k-1)+1 \mathcal{M}_{\varphi(8k-8)}(d)$$

then it follows by Lemma 24 that

$$\mathcal{M}_{\varphi(8k-1)}(\bar{a}c) \equiv_m (k-1)+1 \mathcal{M}_{\varphi(8k-1)}(\bar{bd}).$$

Consequently, $\bar{a}c \mapsto \bar{bd} \in J_{k-1}$ and we are done.

In order to find a suitable element $d$ let $\psi(x)$ be a formula such that

$$(\mathfrak{B}, (D_i)_{i<\varphi(8k)}) \models \psi(d) \iff \mathcal{M}_{\varphi(8k-8)}(c) \equiv_m (k-1)+1 \mathcal{M}_{\varphi(8k-8)}(d),$$

and set

$$\partial_n(x_0, \ldots, x_{n-1}) := \text{‘}x \text{’ is } (\varphi(8k-3))-\text{scattered} \land \bigwedge_{i<n} \psi(x_i),$$

and

$$\chi_\lambda(\bar{y}) := \exists x_0 \ldots \exists x_{\lambda-1}\left( \bigwedge_{i<\lambda} N_{\varphi(8k-6)}(x_i) \subseteq N_{\varphi(8k-4)}(\bar{y}) \land \partial_\lambda(\bar{x}) \right).$$

Note that the quantifier rank of $\psi$ is bounded by $m(k-1)+1$ and that of $\chi_\lambda$ by $\lambda + m(k-1) + 1$.

Let $\kappa$ be the maximal finite cardinal such that

$$(\mathfrak{A}, (D_i)_{i<\varphi(8k)}) \models \exists x_0 \ldots \exists x_{\kappa-1} \chi_\lambda(\bar{x})$$

(if no such cardinal exists we set $\kappa := \omega$), and let $\lambda$ be the maximal finite cardinal such that

$$\mathcal{M}_{\varphi(8k)}(\bar{a}) \models \chi_\lambda(\bar{a}).$$

Note that $\lambda \leq |\bar{a}| = m - k$, by Lemma 21(d).

Let $\bar{e}$ be some $\varphi(8k-3)$-scattered sequence of length $\lambda$, with $N_{\varphi(8k-6)}(e_i) \subseteq N_{\varphi(8k-4)}(\bar{a})$ such that every $e_i$ satisfies $\psi$. We claim that $\bar{e}c$ is $\varphi(8k-3)$-scattered.

If $e_i \in N_{\varphi(8k-3)}(c)$, for some $i$, then $N_{\varphi(8k-3)}(c) \subseteq N_{\varphi(8k-2)}(e_i)$. But there is some index $i$ such that $e_i \in N_{\varphi(8k-3)}(a_i)$ which implies

$$N_{\varphi(8k-2)}(e_i) \subseteq N_{\varphi(8k-1)}(a_i).$$

Hence, we have $N_{\varphi(8k-3)}(c) \subseteq N_{\varphi(8k-1)}(a_i)$ in contradiction to our assumption on $c$.

Similarly, if $c \in N_{\varphi(8k-3)}(e_i)$, for some $i$, then $e_i \in N_{\varphi(8k-4)}(a_i)$, for some $i$, implies $N_{\varphi(8k-3)}(e_i) \subseteq N_{\varphi(8k-2)}(a_i)$, and it follows that $c \in N_{\varphi(8k-2)}(a_i)$. Therefore, we have $N_{\varphi(8k-1)}(c) \subseteq N_{\varphi(8k)}(a_i)$ which again contradicts our assumption on $c$.

We have shown that $\bar{e}c$ is a $\varphi(8k-3)$-scattered sequence every element of which satisfies $\psi$. This implies $\kappa \geq \lambda + 1$.

Since $\mathfrak{A}$ and $\mathfrak{B}$ satisfy the same basic local sentences it follows that

$$(\mathfrak{B}, (D_i)_{i<\varphi(8k)}) \models \exists x_0 \ldots \exists x_{\kappa+1} \partial_{\lambda+1}(\bar{x}).$$

By definition of $J_k$ we further have

$$\mathcal{M}_{\varphi(k)}(\bar{b}) \models \chi_\lambda(\bar{b}) \land \lnot \chi_{\lambda+1}(\bar{b}),$$

since the quantifier rank of this formula is bounded by

$$\lambda + 1 + m(k-1) + 1 \leq m - k + 1 + m(k-1) + 1 \leq mk + 1.$$ 

Let $\bar{e} \in B^{k+1}$ be a sequence satisfying $\partial_{\lambda+1}$. There must be some index $i$ such that $N_{\varphi(8k-6)}(e_i) \nsubseteq N_{\varphi(8k-4)}(\bar{b})$. By Lemma 21(b), it follows that

$$N_{\varphi(8k-6)}(e_i) \cap N_{\varphi(8k-6)}(\bar{b}) = \emptyset.$$
Thus, we have found an element \( d := e \) such that
\[
N_{\emptyset(8k-8)}(d) \cap N_{\emptyset(8k-6)}(b) = \emptyset
\]
and
\[
\forall_{\emptyset(8(k-1))}(c) \iff m(k-1)+1 \forall_{\emptyset(8(k-1))}(d). \quad \Box
\]

**Corollary 28.** On the class of globally uniform structures every first-order sentence is equivalent to a boolean combination of basic local and basic global sentences.

### 7. The theorem of Gaifman for linearly \( \Sigma \)-uniform structures

The requirement of global uniformity is a rather strong one. Essentially it only covers structures that can be obtained from a sparse structure by an interpretation. In particular, linear orders are not globally uniform. In order to extend the theorem of Gaifman to linearly ordered structures we therefore try to weaken our assumptions by considering structures that are globally uniform only with respect to some relations.

**Definition 29.** Let \( \mathfrak{A} \) be a \((\Sigma \cup \Xi)\)-structure where \( \Xi \) contains a binary relation symbol \( \leq \in \Xi \).

- (a) A system of neighbourhoods \( N \) for \( \mathfrak{A} \) is **linear** \( \text{w.r.t. } \leq \) if \( \leq^\mathfrak{A} \) is a linear preorder on \( A \) such that every set \( N_k(a) \) is convex \( \text{w.r.t. } \leq \).
- (b) Let \( N \) be a linear system of neighbourhoods for \( \mathfrak{A} \). For \( a, b \in A \) and \( k < \omega \), we define
  \[
  H_k(a, b) := \bigcup \{ N_k(c) \mid a \leq c \leq b \text{ or } b \leq c \leq a \}.
  \]
- (c) We call \( \mathfrak{A} \) **linearly \( \Sigma \)-uniform** \( \text{w.r.t. } N \) if
  - \( N \) is a linear system of neighbourhoods,
  - the \( \Sigma \)-reduct \( \mathfrak{A}|_{\Sigma} \) is globally uniform \( \text{w.r.t. } N \), and
  - there is a number \( \beta \) such that, for all \( n < \omega \), every pair \( a, a' \in A \), and all tuples \( \bar{b}, \bar{c} \in (H_\beta(a, a'))^n \),
    \[
    \text{atp}(\bar{b}) = \text{atp}(\bar{c}) \quad \text{implies} \quad \text{atp}(\bar{b}/A \setminus H_{\beta}(a, a')) = \text{atp}(\bar{c}/A \setminus H_{\beta}(a, a')).
    \]

(All types are with respect to the full signature \( \Sigma \cup \Xi \).)

Suppose that \( \mathfrak{A} \) is a \((\Sigma \cup \Xi)\)-structure such that \( \mathfrak{A}|_{\Sigma} \) is globally uniform. We try to simplify the Ehrenfeucht–Fraïssé game by removing all relations in \( \Sigma \) and playing on the resulting reduct. Of course, we have to somehow take into account the relations we deleted. We do so by labelling the elements of \( A \) by the type of their neighbourhoods. Thus, the simplification consists in replacing the relations of \( \Sigma \) by unary predicates.

**Definition 30.** Let \( \mathfrak{A} \) be a \((\Sigma \cup \Xi)\)-structures with linear systems of neighbourhoods. For \( \bar{a} \subseteq A \) and \( b, c \in A \), we define
\[
\forall_{\mathfrak{A}}(\bar{a}) := (\mathfrak{A}|_{N_{k}(\bar{a})}, \quad (D_{l})_{l \leq k}, \quad \bar{a}),
\]
\[
\exists_{\mathfrak{A}}(b, c) := (\mathfrak{A}|_{H_{l}(b, c)}, \quad (D_{l})_{l < k}),
\]
and
\[
\emptyset_{m,k}(\bar{a}) := (\mathfrak{A}|_{\Xi}, \quad (D_{l})_{l \leq k}, \quad (P_{\tau})_{\tau}, \quad \bar{a}),
\]
where
\[
D_{l} := \{ (b, c) \mid c \in N_{l}(b) \},
\]
\[
P_{\tau} := \{ b \in A \mid \text{Th}_m(\forall_{\mathfrak{A}}(b)) = \tau \},
\]
and \( \text{Th}_m(\mathfrak{A}) \) denotes the infinitary first-order theory of \( \mathfrak{A} \) of quantifier rank \( m \). We also set \( \emptyset_{m,k}(\emptyset) := \emptyset_{m,k}(\langle \rangle) \).

We will reduce the game on two structures \( \mathfrak{A} \) and \( \mathfrak{B} \) to a game on the corresponding structures \( \emptyset_{m,k}(\mathfrak{A}) \) and \( \emptyset_{m,k}(\mathfrak{B}) \). Note that the classical theorem of Gaifman can be seen as a reduction of \( \text{EF}_m(\mathfrak{A}, \mathfrak{B}) \) to a game between two structures \( (A, \bar{P}) \) and \( (B, \bar{P}) \) with only unary predicates \( \bar{P} \).

**Definition 31.** Let \( \mathfrak{A} \) and \( \mathfrak{B} \) be \((\Sigma \cup \Xi)\)-structures with linear systems of neighbourhoods and let \( \leq \in \Xi \) be the corresponding preorder.
(a) Two tuples $\vec{a} \in A^n$ and $\vec{b} \in B^n$ are $m$-congruent if
\[ a_i \in N_k(a_i) \iff b_i \in N_k(b_i), \quad \text{for all } i, l < n \text{ and } k \leq m - \max\{i, l\}. \]

(b) Set $Q(n) := \mathbb{Z}^{2^n}(\vec{b})$ and let $m < \omega$ and $\vec{a} \in A^{m-k}$. We define a partition of $\vec{a}$ into several intervals $H_k(a_i, a_l)$ as follows.

The partition is induced by the following equivalence relation $\sim \subseteq [m-k]$. For $l < n < m-k$, we define by induction on $n$
\[ l \sim n \iff \text{there are } i, j < n \text{ with } i \sim j \sim l \text{ such that} \]
\[ N_{Q(m-n-1)}(a_n) \subseteq H_{Q(m-n)}(a_i, a_j). \]

(The induction starts with $n = 1$ and $l = 0$. In this case the above definition reads: $0 \sim 1 \iff N_{Q(m-2)}(a_1) \subseteq H_{Q(m-1)}(a_0, a_0).$)

Further, we set
\[ \sigma(n) := \min\{i \mid i \sim n\} \quad \text{and} \quad S := \{n \mid \sigma(n) = n\}, \]

and we define functions $\mu$ and $v$ by
\[ \mu(n) \sim n \sim v(n) \quad \text{and} \quad a_{\mu(n)} \leq a_n \leq a_{v(n)}, \]

for every $n < m-k$.

**Example.** For the tuple $\vec{a} \in A^5$ in the above diagram, we have $S = \{0, 2, 4\}$.

\[ \sigma^{-1}(0) = \{0, 1, 3\}, \quad \sigma^{-1}(2) = \{2\}, \quad \sigma^{-1}(4) = \{4, 5\}, \]
\[ \mu(0) = \mu(1) = \mu(3) = 0 \quad \text{and} \quad v(0) = v(1) = v(3) = 1, \]
\[ \mu(2) = 2 \quad \text{and} \quad v(2) = 2, \]
\[ \mu(4) = \mu(5) = 5 \quad \text{and} \quad v(4) = v(5) = 4. \]

**Remark.** (a) For all $n$, we have
\[ N_k(a_n) \subseteq H_k(a_{\mu(n)}, a_{v(n)}) \subseteq N_{m-\sigma(n)}(a_{\sigma(n)}). \]

(b) If $\vec{a}$ and $\vec{b}$ are $m$-congruent then both tuples lead to the same $\sigma, S, \mu, \text{and } v$.

(c) Note that each $\sim$-class is of the form $\sigma^{-1}(n)$, for some $n \in S$.

**Theorem 32.** Let $\mathfrak{A}$ and $\mathfrak{B}$ be $(\Sigma \cup \Sigma)$-structures with linearly $\Sigma$-uniform systems of neighbourhoods.

\[ \mathfrak{L}_{m-1, Q(m-1)}(\mathfrak{A}) \equiv_m \mathfrak{L}_{m-1, Q(m-1)}(\mathfrak{B}) \]

implies $\mathfrak{A} \equiv_m \mathfrak{B}$.

**Proof.** To simplify notation we set
\[ H_k[\vec{a}; n] := H_k(a_{\mu(n)}, a_{v(n)}) \quad \text{and} \quad S_k[\vec{a}; n] := S_k(a_{\mu(n)}, a_{v(n)}). \]

Let $\vec{a} \in A^{m-k}$ and $\vec{b} \in B^{m-k}$. We call the map $\vec{a} \mapsto \vec{b}$ good if

- $\vec{a}$ and $\vec{b}$ are $m$-congruent,
- $(S_{Q(k)}[\vec{a}; n], a_{\sigma^{-1}(n)}) \equiv_k (S_{Q(k)}[\vec{b}; n], b_{\sigma^{-1}(n)})$, for all $n \in S$,
- $\mathfrak{L}_{k-1, Q(k-1)}(\mathfrak{A}) \equiv_{m+k} \mathfrak{L}_{k-1, Q(k-1)}(\mathfrak{B})$.

Let
\[ J_k := \{ \vec{a} \mapsto \vec{b} \mid \vec{a} \in A^{m-k}, \vec{b} \in B^{m-k}, \text{\vec{a} \mapsto \vec{b} is good} \}. \]

We claim that $(J_k)_{k \leq m} : \mathfrak{A} \equiv_m \mathfrak{B}$. By assumption we have $(\emptyset \mapsto \emptyset) \in J_m$. For the forth property, suppose that $\vec{a} \mapsto \vec{b} \in J_k$ and $c \in A$. We consider two cases.
First, suppose that \( N_{\varphi(k)}(c) \subseteq H_{\varphi(k)}[\bar{a}; n], \) for some \( n \in S. \)

\[
\{ \bar{a} \mid \sigma^{-1}(n) \} \equiv_k \{ \bar{b} \mid \sigma^{-1}(n) \}
\]
implies that there is some \( d \in H_{\varphi(k)}[\bar{b}; n] \) such that

\[
\{ \bar{a} \mid \sigma^{-1}(n) \} \equiv_k \{ \bar{b} \mid \sigma^{-1}(n) \}. \]

It follows that

\[
\{ \bar{a} \mid \sigma^{-1}(n) \} \equiv_k \{ \bar{b} \mid \sigma^{-1}(n) \}, \] for all \( n \in S. \)

Two reflexive sets \( A \) and \( B \) are distinct if there is \( a \in A \) and \( b \in B \) such that

\[
A = \{ a \mid \sigma^{-1}(n) \} \] and \( B = \{ b \mid \sigma^{-1}(n) \} \) for some \( n \in S. \)

It remains to consider the case that \( N_{\varphi(k)}(c) \not\subseteq H_{\varphi(k)}[\bar{a}; n], \) for all \( n \in S. \) Then we have \( N_{\varphi(k)}(c) \cap H_{\varphi(k)}[\bar{a}; n] = \emptyset, \) for all \( n \in S, \) and

\[
\{ \bar{a} \mid \sigma^{-1}(n) \} \equiv_{m+k} \{ \bar{b} \mid \sigma^{-1}(n) \}, \]

implies that there is some \( d \in B \) such that

\[
\{ \bar{a} \mid \sigma^{-1}(n) \} \equiv_k \{ \bar{b} \mid \sigma^{-1}(n) \}. \]

Therefore, we have

\[
\{ \bar{a} \mid \sigma^{-1}(n) \} \equiv_k \{ \bar{b} \mid \sigma^{-1}(n) \}, \]

and \( \bar{a} \mapsto \bar{b} \in J_{k-1}. \)

The back property follows by symmetry. It therefore remains to show that every \( \bar{a} \mapsto \bar{b} \in J_0 \) is a partial isomorphism. Suppose that \( \mathfrak{A} \models \varphi(c) \) for some \( c \) and \( \varphi \) is a literal. If there is some \( n \in S \) such that \( c \leq H_{\bar{a}}[\bar{a}; n] \) then

\[
\{ \bar{a} \mid \sigma^{-1}(n) \} \equiv_k \{ \bar{b} \mid \sigma^{-1}(n) \}, \]

implies that \( \mathfrak{B} \models \varphi(d) \) where \( d \subseteq \bar{b} \) is the corresponding subtuple of \( \bar{b}. \) Hence, we may assume that \( c = \bar{c}_0 \ldots \bar{c}_l \) for some \( l \in \mathbb{N}, \) and there are distinct numbers \( n_0, \ldots, n_l \in S \) such that

\[
\bar{c}_i \subseteq H_{\bar{a}}[\bar{a}; n_i], \] for \( i \leq l. \)

Since \( \mathfrak{A} \) is a \( k \)-ary \( k \)-uniform we have

\[
\mathfrak{A} \models \varphi(\vec{e}_0, \ldots, \vec{e}_l) \] for all tuples \( \vec{e}_i \subseteq H_{\bar{a}}[\bar{a}; n_i] \) with \( \text{atp}(\vec{e}_i) = \text{atp}(\vec{c}_i). \)

Let \( \alpha_i(\bar{x}, \bar{y}) := \bigwedge \neg N_{\varphi(m-n_i)}(\bar{a}_i) \land \bigwedge_{k \in S \setminus \{ n_i \}} N_{\varphi(m-k)}(\bar{a}_i) = \emptyset. \)

Then we have

\[
\mathfrak{A} \models \forall \bar{x}_0 \ldots \forall \bar{x}_l \left( \bigwedge_{i \leq l} \beta_i(\bar{x}_i) \rightarrow \varphi(\bar{x}_0, \ldots, \bar{x}_l) \right). \]

Since \( \mathfrak{L}_{\varphi}(\bar{a}) \equiv_{m} \mathfrak{L}_{\varphi}(\bar{b}) \) it follows that

\[
\mathfrak{B} \models \forall \bar{x}_0 \ldots \forall \bar{x}_l \left( \bigwedge_{i \leq l} \beta_i(\bar{x}_i) \rightarrow \varphi(\bar{x}_0, \ldots, \bar{x}_l) \right). \]

Consequently, we have \( \mathfrak{B} \models \varphi(\vec{d}_0, \ldots, \vec{d}_l) \) where \( \vec{d}_i \subseteq \bar{b} \) is the subtuple of \( \bar{b} \) corresponding to \( \vec{c}_i. \)

Example. Set \( \varSigma := (\leq, E) \) and \( \Sigma := R. \) We consider two linearly ordered structures \( \mathfrak{A} = (A, \leq, E, \bar{R}) \) and \( \mathfrak{B} = (B, \leq, E, \bar{R}) \) where \( E \) is the successor relation of \( \leq. \) We define the distance \( d(a, b) \) between elements \( a, b \in A \) as their Gaifman distance in the reduct \( \mathfrak{A}|_{E, \bar{R}}, \) i.e., we ignore \( \leq. \) Let \( N_r(a) \) be the \( r \)-neighbourhood of \( a \) with respect to this distance. The structures \( \mathfrak{L}_{m,n}(\bar{a}) \) are labelled linear orders where the colour of an element denotes the type of its \( r \)-neighbourhood. Hence, we have reduced the game on \( \mathfrak{A} \) and \( \mathfrak{B} \) to a simpler game on labelled linear orders.
In order to apply Theorem 32, we partition the signature into

\[ \lambda r = (M_n +, (\lambda r)_{r \in R}) \]

\[ \cdots \to M_2 \to M_1 \to M_0 \to 0 \]

that is, a sequence of homomorphisms \( d_n : M_n \to M_{n-1} \) between modules with \( d_n \circ d_{n+1} = 0 \). We encode such a complex as a structure

\[ \mathcal{C}(M_\bullet, d_\bullet) = \langle C, +, (\lambda r)_{r \in R}, d, \leq \rangle \]

where

- \( C = \bigcup_n M_n \) is the disjoint union of the universes \( M_n \),
- \( + \) is the union of the (graphs of the) addition operations on each \( M_n \),
- \( \lambda r \) is the union of the scalar multiplication operations on each \( M_n \),
- \( d \) is the union of the graphs of the \( d_n \), and
- the preorder \( \leq \) is defined by

\[ a \leq b \iff a \in M_l \text{ and } b \in M_k \text{ for } i \leq k. \]

Recall that the \( n \)-th homology group of \((M_\bullet, d_\bullet)\) is

\[ H_n(M_\bullet, d_\bullet) := \ker d_n / \text{rng} d_{n+1}. \]

We will prove that there does not exist a first-order formula \( \varphi \) that holds in a structure of the form \( \mathcal{C}(M_\bullet, d_\bullet) \) if and only if there exists a maximal index \( n < \omega \) with \( H_n(M_\bullet, d_\bullet) \neq 0 \) and this index \( n \) is even.

For a contradiction, suppose that \( \varphi \) is a sentence with the desired properties. Let \( R = Q \) and \( M_n := Q^\omega \), for all \( n \). We define a function \( d : Q^\omega \to Q^\omega \) by

\[ d(a)_i = (b)_i \quad \text{where } b_i := \begin{cases} a_{i+1} & \text{if } i \text{ is even}, \\ 0 & \text{otherwise}. \end{cases} \]

Note that \( \ker(d) = \text{rng}(d) \) and the sequence

\[ \cdots \to M_2 \to M_1 \to M_0 \]

is exact. Furthermore, let \( d' : Q^\omega \to Q^\omega \) be the constant map with value 0.

Let \( m \) be the quantifier rank of \( \varphi \), set \( r := \varphi(m - 1) \), and fix a number \( l > 2r + 2^{2m} \). We define two complexes \((M_\bullet, d_\bullet^i)\) and \((M_\bullet, d_\bullet^j)\) by setting

\[ d_n^i := \begin{cases} d' & \text{if } n = 0 \text{ or } n = l + i, \\ d & \text{otherwise}, \end{cases} \quad \text{for } i < 2. \]

\[ \cdots \to M_{l+i} \overset{d^i_n}{\to} M_l \overset{d^n}{\to} \cdots \to M_2 \overset{d}{\to} M_1 \overset{d}{\to} M_0 \overset{d}{\to} 0. \]

It follows that

\[ H_n(M_\bullet, d_\bullet^i) = \begin{cases} Q^\omega & \text{if } n = 0 \text{ or } n = l + i, \\ 0 & \text{otherwise}. \end{cases} \]

By assumption on \( \varphi \) we therefore have \( \mathcal{C}(M_\bullet, d_\bullet^i) \models \varphi \iff \mathcal{C}(M_\bullet, d_\bullet^j) \not\models \varphi \).

In order to apply Theorem 32, we partition the signature into \( \Sigma := \{\leq\} \) and \( \Sigma := \{+, (\lambda r)_{r \in R}, d\} \). We define a system of neighbourhoods \( N \) by setting

\[ N_r(a) := \bigcup \{M_i \mid n - r \leq i \leq n + r\} \]

where \( n \) is the index such that \( a \in M_n \). With these definitions a structure of the form \( \mathcal{C}(M_\bullet, d_\bullet) \) becomes linearly \( \Sigma \)-uniform.

Let \( a \) be an element of the first complex and \( b \) an element of the second one. Suppose that \( a \in M_l \) and \( b \in M_j \). If

- \( l = j < 1 - r \) or
- \( l = j - 1 > r \) or
- \( r < i < l - r \) and \( r < j < l + 1 - r \)

\[ M_j \not\models \varphi \iff M_l \not\models \varphi. \]
then we have
\[ H_r(a) \cong H_r(b). \]

It follows that \( L_{m-1,r}(C(M_d, d_0^i)) \) consists of a coloured linear preorder where in the middle part there are only two colours: the zero elements of each \( M_i \) have one colour and all other elements have the second colour. Furthermore, we obtain the structure \( L_{m-1,r}(C(M_d, d_1^i)) \) from the first one by inserting a copy of \( M_i \) that is also coloured this way. Since the middle part consists of more than \( 2r + 2^{2m} - 2r \geq 2^{2m} \) copies of \( M_i \) it follows that Duplicator has a winning strategy for the \( 2m \)-round Ehrenfeucht–Fraïssé game between these structures. By Theorem 32, it follows that
\[ C(M_d, d_0^i) \equiv_m C(M_d, d_1^i). \]
A contradiction.

8. Conclusion

We have investigated tools to simplify Ehrenfeucht–Fraïssé games on non-sparse structures. In the first part of the paper we have presented several simple ways to decompose a game on two structures into games on certain substructures. Technically the main idea behind these constructions was the colouring of tuples by their external type.

In the second part of the paper we have tried to generalise the theorem of Gaifman to non-sparse structures. In particular, we aimed at covering well-known examples from the literature which successfully employed locality-based Ehrenfeucht–Fraïssé arguments. By introducing the notions of global uniformity and linear uniformity we were able to do so for the case of linearly ordered structures. We conclude this article by mentioning two important cases which still remain open.

Open problem. Extend the theorem of Gaifman such that it covers

(a) trees (with ordering),
(b) Presburger Arithmetic and algebraically closed fields.

References