

A Local Normal Form Theorem for Infinitary Logic with Unary Quantifiers

H. Jerome Keisler ^{*1} and Wafik Boulos Lotfallah ^{**2}

¹ Department of Mathematics
University of Wisconsin
480 Lincoln Drive
Madison WI 53706

² Dept. of Eng. Math. and Phys.
Cairo University
Cairo, Egypt 11451

Key words normal form, locality, Ehrenfeucht-Fraïssé games, unary quantifiers

MSC (2000) 03C13,03C75,03C80

Copyright line will be provided by the publisher

We prove a local normal form theorem of the Gaifman type for the infinitary logic $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^\omega$ whose formulas involve arbitrary unary quantifiers but finite quantifier rank. We use a local Ehrenfeucht-Fraïssé type game similar to the one in [KL04]. A consequence is that every sentence of $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^\omega$ of quantifier rank n is equivalent to an infinite Boolean combination of sentences of the form $(\exists^{\geq i}y)\psi(y)$ where $\psi(y)$ has counting quantifiers restricted to the $(2^{n-1} - 1)$ -neighborhood of y .

Copyright line will be provided by the publisher

1 Introduction

Two theorems of Hanf and Gaifman show that first order logic has the following locality properties:

(A) (Hanf's locality theorem [Han65]) For every sentence ϕ there is an $r \in \mathbf{N}$ such that two models \mathcal{A}, \mathcal{B} agree on ϕ whenever there is a bijection f from \mathcal{A} to \mathcal{B} such that the r -neighborhood of each element of \mathcal{A} is isomorphic to the r -neighborhood of its image.

(B) (Gaifman's normal form theorem [Ga82]) For each sentence ϕ there is an $r \in \mathbf{N}$ such that ϕ is equivalent to a Boolean combination of sentences of the form

$$(\exists x_1 \dots \exists x_k) \left[\bigwedge_{i < j} \delta(x_i, x_j) > 2r \wedge \bigwedge_i \psi(x_i) \right]$$

where $\delta(x_i, x_j)$ is the natural distance and the quantifiers in $\psi(x)$ are restricted to the r -neighborhood of x .

For finite models, Property (A) has more recently been proved for more powerful logics. [Nur96] showed that (A) holds for the logic $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^\omega$ whose formulas have infinite conjunctions and disjunctions, arbitrary unary quantifiers, and finite quantifier rank. [Imm99] and [Lib00] showed that for both first order logic and $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^\omega$, one can take $r = 2^{n-1} - 1$ where n is the quantifier rank of ϕ .

In this paper we will show that for finite models, a property like (B) also holds for the language $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^\omega$.

We proved in [KL04] that in Property (B) for first order logic one can take $r = 4^{n-1} - 1$, where n is the quantifier rank of ϕ (this improved an earlier result in [LS96]).

* e-mail: keisler@math.wisc.edu

** e-mail: lotfalla@uwalumni.com

Here we will prove a similar normal form theorem for $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^\omega$, but with the best possible bound $r = 2^{n-1} - 1$. Let $\mathcal{L}_{\infty\omega}(\mathbf{C})^\omega$ be the sublanguage of $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^\omega$ with only the counting quantifiers $\exists^{\geq i}x$, $i \in \mathbf{N}$. We prove the following.

(C) Each sentence ϕ of $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^\omega$ of quantifier rank $n > 0$ is equivalent to a finite or infinite Boolean combination of sentences of $\mathcal{L}_{\infty\omega}(\mathbf{C})^\omega$ of the form $(\exists^{\geq i}x)\psi(x)$, where the quantifiers in $\psi(x)$ are restricted to the $(2^{n-1} - 1)$ -neighborhood of x .

To see the similarity with Gaifman's normal form, note that the sentence $(\exists^{\geq i}x)\psi(x)$ in expanded form is

$$(\exists x_1 \dots \exists x_i) \left[\bigwedge_{j < k} x_j \neq x_k \wedge \bigwedge_j \psi(x_j) \right].$$

The normal form (C) is very tractable. For instance, (C) immediately implies the theorem in [Im99] and [Lib00] that (A) holds for $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^\omega$ with the same bound $r = 2^{n-1} - 1$. This corresponds to the result in [HLN99] that strong Gaifman locality implies Hanf locality. The argument there shows that if a logic has Property (B) with a bound on k , then it has Property (A), but with a different value for r .

The sentences in (C) almost preserve quantifier rank. In particular, in the case that all the relation symbols are at most binary, the normal form sentences $(\exists^{\geq i}x)\psi(x)$ can be taken to have quantifier rank at most the quantifier rank n of the original sentence ϕ .

The bijective Ehrenfeucht-Fraïssé game was introduced in [Hel89] and has been a useful tool in studying the logic $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^\omega$ (see also [Hel96]). This game will also play an important role in this paper. As in [Lib00], we will get local conditions which imply that Duplicator has a winning strategy in the bijective Ehrenfeucht-Fraïssé game, and hence that the two models are equivalent with respect to $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^n$. However, in order to get a normal form theorem, we need local conditions that are themselves expressible in $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^\omega$ (and even in $\mathcal{L}_{\infty\omega}(\mathbf{C})^\omega$).

To do this we introduce a local bijective Ehrenfeucht-Fraïssé game in Section 3 which is similar to the shrinking game used in [KL04], and prove that if Duplicator has a winning strategy in the local bijective game then she has one in the ordinary bijective game. Then, in Section 4, we introduce the notion of a localization of a formula of $\mathcal{L}_{\infty\omega}(\mathbf{C})^\omega$ which corresponds in a natural way to the local bijective game. These localized formulas provide the building blocks for the main normal form theorem, which is proved in Section 4. In Section 5 we give an alternative normal form theorem which is even more local in the sense that the quantifiers are restricted to even smaller sets. Finally, in Section 6 we use the normal form theorem to get the result (C) stated above and apply it to the notions of Hanf and Gaifman locality from [HLN99].

2 Basic Definitions

Our main objective is to study finite models, but all of our arguments will also work for a larger class of models. For this reason, we give the reader two options.

Option 1: This is the simpler and more natural option. In this option, all models are understood to be finite, and two formulas are said to be **equivalent** if they agree on all finite models.

Option 2: This option gives the strongest results. A model \mathcal{A} is **countable** if its universe is finite or countably infinite. \mathcal{A} is **locally finite** if each element $a \in \mathcal{A}$ has finite degree, that is, the set of $c \in \mathcal{A}$ such that a and c occur together in some atomic formula true in \mathcal{A} is finite. In this option, all models are understood to be countable and locally finite, and two formulas are said to be **equivalent** if they agree on all countable locally finite models.

We fix a finite relational vocabulary ν . Our languages have equality as well as the relation symbols in ν . \mathcal{A}, \mathcal{B} will always denote models of type ν . \mathbf{x} will stand for a tuple (x_1, \dots, x_n) of variables, and \mathbf{a}, \mathbf{b} will stand for tuples in \mathcal{A}, \mathcal{B} , respectively of the same length $|\mathbf{a}| = |\mathbf{b}| = |\mathbf{x}|$. Abusing notation, we also use \mathcal{A} to denote the universe of the model. $\mathcal{A} \cong \mathcal{B}$ means that the models \mathcal{A} and \mathcal{B} are isomorphic. As usual, an occurrence of a variable x in a formula is bound if it is in the scope of a quantifier on x , and free otherwise. When we use the notation $\phi(\mathbf{x})$, it is understood that \mathbf{x} is a list of all the variables which occur free in ϕ . A **sentence** is a formula with no free variables.

If \mathcal{F} is a set of formulas, we write $\mathcal{A} \equiv \mathcal{B} (\mathcal{F})$ if \mathcal{A} and \mathcal{B} agree on all sentences in \mathcal{F} , and we write $(\mathcal{A}, \mathbf{a}) \equiv (\mathcal{B}, \mathbf{b}) (\mathcal{F})$ if $(\mathcal{A}, \mathbf{a})$ and $(\mathcal{B}, \mathbf{b})$ agree on all formulas in \mathcal{F} with at most \mathbf{x} free.

The infinitary logic $\mathcal{L}_{\infty\omega}$ is the set of all formulas that can be built (like the first order formulas) from the atomic formulas using negation, universal and existential quantification, but we also allow infinite conjunctions and disjunctions. The universal and existential quantifiers are merely special cases of more general unary quantifiers $Q_{\mathcal{C}}$, where \mathcal{C} is a class of models in the vocabulary consisting of a finite number k of unary relation symbols and \mathcal{C} is closed under isomorphism. The quantifier $Q_{\mathcal{C}}$ binds the variables y_1, \dots, y_k in the compound formula

$$(Q_{\mathcal{C}}y_1, \dots, y_k)(\psi_1(y_1, \mathbf{x}_1), \dots, \psi_k(y_k, \mathbf{x}_k))$$

and is interpreted as saying that the universe of the model together with the subsets defined by each $\psi_i(y_i, \mathbf{x}_i)$ (where elements of \mathbf{x}_i are treated as parameters) is in the class \mathcal{C} .

The logic $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)$ is the extension of $\mathcal{L}_{\infty\omega}$ where we allow the application of arbitrary unary quantifiers. Both of the logics $\mathcal{L}_{\infty\omega}$ and $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)$ are too powerful to be of any interest in finite model theory. So we consider some weaker fragments of $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)$ obtained by restricting the quantifier rank of a formula.

A formula with no quantifiers is said to have quantifier rank 0. A formula is said to have quantifier rank at most $n + 1$ if it is a Boolean combination of formulas of quantifier rank 0 and formulas of the form $(Q_{\mathcal{C}}y_1, \dots, y_k)(\psi_1(y_1, \mathbf{x}_1), \dots, \psi_k(y_k, \mathbf{x}_k))$ where each ψ_i has quantifier rank at most n . The logic $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^n$ is the fragment of $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)$ where the quantifier rank of formulas is at most n . $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^\omega$ is just the union of all $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^n$ for finite n .

An example of a unary quantifier is the **counting quantifier** $(\exists^{\geq i}x)\phi(x)$ where $i \in \mathbf{N}$, which states the existence of at least i elements with the property ϕ . The logic $\mathcal{L}_{\infty\omega}(\mathbf{C})^n$ is the fragment of $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^n$ where only counting quantifiers are used. $\mathcal{L}_{\infty\omega}(\mathbf{C})^n$ allows the usual existential and universal quantifiers, since $(\exists x)\phi = (\exists^{\geq 1}x)\phi$ and $(\forall x)\phi = \neg(\exists x)\neg\phi$. The union of $\mathcal{L}_{\infty\omega}(\mathbf{C})^n$ for all finite n is the logic $\mathcal{L}_{\infty\omega}(\mathbf{C})^\omega$.

The logic $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^n$ is characterized by the so called **bijective Ehrenfeucht-Fraïssé game**, introduced in [Hel89]. The bijective game is played on two models by two players, Spoiler and Duplicator. Roughly speaking, Spoiler tries to prove that the two models look different, while Duplicator tries to prove that they look alike. By a game **position** we will mean a triple $((\mathcal{A}, \mathbf{a}), (\mathcal{B}, \mathbf{b}), n)$ where $|\mathbf{a}| = |\mathbf{b}|$, and n is a natural number which represents the number of rounds yet to be played. The **bijective game** is defined by recursion on n as follows.

When $n > 0$, the bijective game proceeds from the position $((\mathcal{A}, \mathbf{a}), (\mathcal{B}, \mathbf{b}), n)$ according to the rules:

- (1) Duplicator chooses a bijection $f : \mathcal{A} \rightarrow \mathcal{B}$. (If $|\mathcal{A}| \neq |\mathcal{B}|$, then Spoiler wins.)
- (2) Spoiler chooses an element c in \mathcal{A} .
- (3) The game continues from the new position $((\mathcal{A}, \mathbf{ac}), (\mathcal{B}, \mathbf{bf}(c)), n - 1)$.

When $n = 0$ the game ends. $((\mathcal{A}, \mathbf{a}), (\mathcal{B}, \mathbf{b}), 0)$ is a **winning position** for Duplicator iff $(\mathcal{A}, \mathbf{a})$ and $(\mathcal{B}, \mathbf{b})$ satisfy the same atomic formulas. (In other words, \mathbf{a}, \mathbf{b} are either empty or generate submodels of \mathcal{A}, \mathcal{B} which are isomorphic with an isomorphism mapping \mathbf{a} to \mathbf{b} .)

We write $(\mathcal{A}, \mathbf{a}) \equiv_n (\mathcal{B}, \mathbf{b})$ if Duplicator has a winning strategy in the bijective Ehrenfeucht-Fraïssé game starting from the position $((\mathcal{A}, \mathbf{a}), (\mathcal{B}, \mathbf{b}), n)$.

The importance of the bijective Ehrenfeucht-Fraïssé game stems from the following result, which is proved in [Hel89, Hel96].

Proposition 2.1 *The following are equivalent.*

$$\begin{aligned} (\mathcal{A}, \mathbf{a}) &\equiv_n (\mathcal{B}, \mathbf{b}), \\ (\mathcal{A}, \mathbf{a}) &\equiv (\mathcal{B}, \mathbf{b}) (\mathcal{L}_{\infty\omega}(\mathbf{C})^n), \\ (\mathcal{A}, \mathbf{a}) &\equiv (\mathcal{B}, \mathbf{b}) (\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^n). \end{aligned}$$

We say that a formula ϕ is **Boolean over** a set of formulas \mathcal{F} if ϕ is equivalent to a (possibly infinite) Boolean combination of formulas from \mathcal{F} .

Corollary 2.2 *Let \mathcal{F} be a set of formulas. Then every formula $\phi(\mathbf{x})$ in $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^n$ is Boolean over \mathcal{F} if and only if*

$$(\mathcal{A}, \mathbf{a}) \equiv (\mathcal{B}, \mathbf{b}) (\mathcal{F}) \text{ implies } (\mathcal{A}, \mathbf{a}) \equiv_n (\mathcal{B}, \mathbf{b}).$$

Inspired by Corollary 2.2 we will search for sufficient local conditions for Duplicator to win the bijective Ehrenfeucht-Fraïssé game. If moreover these conditions are expressible in $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^\omega$, we get a local normal form for $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^\omega$. Our main tool for doing this will be the local bijective game introduced in the next section.

3 The Local Bijective Game

Before defining this game we need some notation concerning neighborhoods and distances in relational models.

The **Gaifman graph** over a model \mathcal{A} is the graph over the universe of \mathcal{A} whose edges are the pairs (c, d) of distinct elements of \mathcal{A} such that both c and d occur in an atomic sentence which holds in $(\mathcal{A}, \mathbf{a})$ for some \mathbf{a} . The Gaifman graph over \mathcal{A} is undirected and has no loops.

If $c, d \in \mathcal{A}$ we let $\delta(c, d)$ be the natural distance between c and d in the Gaifman graph over \mathcal{A} , i.e. the length of the shortest path connecting c and d . If there is no such path, then $\delta(c, d) = \infty$. For tuples \mathbf{a}, \mathbf{a}' in \mathcal{A} we also define $\delta(\mathbf{a}, \mathbf{a}')$ to be the minimum distance between elements of \mathbf{a} and elements of \mathbf{a}' .

If \mathbf{a} has length $k > 0$ and $r \in \mathbf{N}$, the r -**neighborhood** $\mathcal{N}_{\mathcal{A}}^r(\mathbf{a})$ **around** \mathbf{a} is the submodel of \mathcal{A} generated by the elements within distance $\leq r$ from one of a_1, \dots, a_k , with all of a_1, \dots, a_k as distinguished constants. When $k = 0$, i.e. \mathbf{a} is the empty sequence, we define $\mathcal{N}_{\mathcal{A}}^r(\mathbf{a})$ to be the whole model \mathcal{A} .

Since \mathcal{A} is assumed to be finite or locally finite, every neighborhood $\mathcal{N}_{\mathcal{A}}^r(\mathbf{a})$ around a nonempty tuple \mathbf{a} is finite.

The **local bijective game** is defined as follows.

From the position $((\mathcal{A}, \mathbf{a}), (\mathcal{B}, \mathbf{b}), n)$, where $n > 0$, the game proceeds with the following rules:

- (1) Duplicator chooses a bijection $f : \mathcal{N}_{\mathcal{A}}^{2^{n-1}}(\mathbf{a}) \longrightarrow \mathcal{N}_{\mathcal{B}}^{2^{n-1}}(\mathbf{b})$.
- (2) Spoiler chooses an element c in $\mathcal{N}_{\mathcal{A}}^{2^{n-1}}(\mathbf{a})$.
- (3) The game continues from the new position $((\mathcal{A}, \mathbf{ac}), (\mathcal{B}, \mathbf{bf}(c)), n - 1)$.

When $n = 0$ the game ends as usual, with Duplicator winning if the models agree on all atomic formulas.

We write $(\mathcal{A}, \mathbf{a}) \equiv_n^L (\mathcal{B}, \mathbf{b})$ if Duplicator has a winning strategy in the local bijective game starting from the position $((\mathcal{A}, \mathbf{a}), (\mathcal{B}, \mathbf{b}), n)$.

By a **winning move** at $((\mathcal{A}, \mathbf{a}), (\mathcal{B}, \mathbf{b}), n)$ we mean a bijection $f : \mathcal{N}_{\mathcal{A}}^{2^{n-1}}(\mathbf{a}) \longrightarrow \mathcal{N}_{\mathcal{B}}^{2^{n-1}}(\mathbf{b})$ such that $(\mathcal{A}, \mathbf{ac}) \equiv_{n-1}^L (\mathcal{B}, \mathbf{bf}(c))$ for all $c \in \mathcal{A}$. Thus when $n > 0$, $(\mathcal{A}, \mathbf{a}) \equiv_n^L (\mathcal{B}, \mathbf{b})$ if and only if there is a winning move at $((\mathcal{A}, \mathbf{a}), (\mathcal{B}, \mathbf{b}), n)$.

If $(\mathcal{A}, \mathbf{ac}) \equiv_n^L (\mathcal{B}, \mathbf{bd})$, we will say that the elements c and d have the same n -**local bijective type relative to** \mathbf{a}, \mathbf{b} . If, moreover, \mathbf{a} and \mathbf{b} are empty, then we say that c and d have the same n -**local simple bijective type**.

Lemma 3.1 *Let \mathbf{a}, \mathbf{a}' be tuples in \mathcal{A} and \mathbf{b}, \mathbf{b}' be tuples in \mathcal{B} with $|\mathbf{a}| = |\mathbf{b}| > 0$ and $|\mathbf{a}'| = |\mathbf{b}'| > 0$. Suppose $(\mathcal{A}, \mathbf{a}) \equiv_n^L (\mathcal{B}, \mathbf{b})$, $(\mathcal{A}, \mathbf{a}') \equiv_n^L (\mathcal{B}, \mathbf{b}')$, and both $\delta(\mathbf{a}, \mathbf{a}'), \delta(\mathbf{b}, \mathbf{b}') > 2^n$. Then $(\mathcal{A}, \mathbf{aa}') \equiv_n^L (\mathcal{B}, \mathbf{bb}')$.*

Proof: The proof is by induction on n . The lemma holds for $n = 0$ because by hypothesis no atomic sentence holding in $(\mathcal{A}, \mathbf{aa}')$ involves elements from both \mathbf{a} and \mathbf{a}' , and similarly for \mathcal{B} .

Assume the lemma holds for $n - 1$ and let f and f' be winning moves for $((\mathcal{A}, \mathbf{a}), (\mathcal{B}, \mathbf{b}), n)$ and $((\mathcal{A}, \mathbf{a}'), (\mathcal{B}, \mathbf{b}'), n)$ respectively. It suffices to show that $f \cup f'$ is a winning move for $((\mathcal{A}, \mathbf{aa}'), (\mathcal{B}, \mathbf{bb}'), n)$. Note first that

$$\mathcal{N}_{\mathcal{A}}^{2^{n-1}}(\mathbf{a}) \cap \mathcal{N}_{\mathcal{A}}^{2^{n-1}}(\mathbf{a}') = \emptyset \text{ and } \mathcal{N}_{\mathcal{B}}^{2^{n-1}}(\mathbf{b}) \cap \mathcal{N}_{\mathcal{B}}^{2^{n-1}}(\mathbf{b}') = \emptyset.$$

Therefore $f \cup f'$ is a bijection. Let $c \in \mathcal{N}_{\mathcal{A}}^{2^{n-1}}(\mathbf{a})$. Then $(\mathcal{A}, \mathbf{a}c) \equiv_{n-1}^L (\mathcal{B}, \mathbf{b}f(c))$ and $\delta(\mathbf{a}c, \mathbf{a}'), \delta(\mathbf{b}f(c), \mathbf{b}') > 2^{n-1}$. It is clear that $(\mathcal{A}, \mathbf{a}') \equiv_n^L (\mathcal{B}, \mathbf{b}')$ implies $(\mathcal{A}, \mathbf{a}') \equiv_{n-1}^L (\mathcal{B}, \mathbf{b}')$. Now by the induction hypothesis we have $(\mathcal{A}, \mathbf{a}'c) \equiv_{n-1}^L (\mathcal{B}, \mathbf{b}'f(c))$. A similar argument holds when $c \in \mathcal{N}_{\mathcal{A}}^{2^{n-1}}(\mathbf{a}')$. Therefore $f \cup f'$ is a winning move. ■

Theorem 3.2 (1) If $\mathcal{A} \equiv_n^L \mathcal{B}$, then $\mathcal{A} \equiv_n \mathcal{B}$.

(2) If $(\mathcal{A}, \mathbf{a}) \equiv_n^L (\mathcal{B}, \mathbf{b})$ and $\mathcal{A} \equiv_n^L \mathcal{B}$, then $(\mathcal{A}, \mathbf{a}) \equiv_n (\mathcal{B}, \mathbf{b})$.

Proof: (1) is a special case of (2). We prove (2) by induction on n . The case $n = 0$ is trivial.

Assume $n > 0$ and the result holds for $n - 1$. Suppose $(\mathcal{A}, \mathbf{a}) \equiv_n^L (\mathcal{B}, \mathbf{b})$ and $\mathcal{A} \equiv_n^L \mathcal{B}$.

Claim: There is a bijection $f : \mathcal{A} \rightarrow \mathcal{B}$ such that $(\mathcal{A}, \mathbf{a}c) \equiv_{n-1}^L (\mathcal{B}, \mathbf{b}f(c))$ for all $c \in \mathcal{A}$.

Once this claim is established, the proof is completed as follows. It is clear that $\mathcal{A} \equiv_n^L \mathcal{B}$ implies $\mathcal{A} \equiv_{n-1}^L \mathcal{B}$. Then by the inductive hypothesis, $(\mathcal{A}, \mathbf{a}c) \equiv_{n-1}^L (\mathcal{B}, \mathbf{b}f(c))$ for all $c \in \mathcal{A}$. So Duplicator wins the game $(\mathcal{A}, \mathbf{a}) \equiv_n (\mathcal{B}, \mathbf{b})$ by choosing the bijection f .

Proof of Claim: Since $(\mathcal{A}, \mathbf{a}) \equiv_n^L (\mathcal{B}, \mathbf{b})$, Duplicator has a winning move g at $((\mathcal{A}, \mathbf{a}), (\mathcal{B}, \mathbf{b}), n)$ for the local bijective game. In the case that \mathbf{a} and \mathbf{b} are empty, the claim holds with $f = g$. Assume hereafter that \mathbf{a} and \mathbf{b} are nonempty. Duplicator needs to find a bijection

$$h : \mathcal{A} \setminus \mathcal{N}_{\mathcal{A}}^{2^{n-1}}(\mathbf{a}) \rightarrow \mathcal{B} \setminus \mathcal{N}_{\mathcal{B}}^{2^{n-1}}(\mathbf{b})$$

such that $f = g \cup h$ satisfies the Claim. The existence of g implies that the two neighborhoods $\mathcal{N}_{\mathcal{A}}^{2^{n-1}}(\mathbf{a})$ and $\mathcal{N}_{\mathcal{B}}^{2^{n-1}}(\mathbf{b})$ have the same number of elements of each $(n - 1)$ -local bijective type relative to \mathbf{a}, \mathbf{b} . On the other hand, since $\mathcal{A} \equiv_n^L \mathcal{B}$, \mathcal{A} and \mathcal{B} must have the same number of elements of each $(n - 1)$ -local simple bijective type.

Consequently, since the neighborhoods $\mathcal{N}_{\mathcal{A}}^{2^{n-1}}(\mathbf{a})$ and $\mathcal{N}_{\mathcal{B}}^{2^{n-1}}(\mathbf{b})$ are finite and the whole models \mathcal{A} and \mathcal{B} are finite or countable, $\mathcal{A} \setminus \mathcal{N}_{\mathcal{A}}^{2^{n-1}}(\mathbf{a})$ and $\mathcal{B} \setminus \mathcal{N}_{\mathcal{B}}^{2^{n-1}}(\mathbf{b})$ must also have the same number of elements of each $(n - 1)$ -local simple bijective type. Since these elements are far enough from \mathbf{a} and \mathbf{b} , from Lemma 3.1, $\mathcal{A} \setminus \mathcal{N}_{\mathcal{A}}^{2^{n-1}}(\mathbf{a})$ and $\mathcal{B} \setminus \mathcal{N}_{\mathcal{B}}^{2^{n-1}}(\mathbf{b})$ must also have the same number of elements of each $(n - 1)$ -local bijective type relative to \mathbf{a}, \mathbf{b} .

Now Duplicator just builds h from the different bijections between the equally sized subsets. ■

The converse of Theorem 3.2 is also true, with a minor adjustment in quantifier rank.

Theorem 3.3 Let m be the max of 2 and the arity of each relation symbol of ν .

(1) If $\mathcal{A} \equiv_{n+m-2} \mathcal{B}$, then $\mathcal{A} \equiv_n^L \mathcal{B}$.

(2) If $(\mathcal{A}, \mathbf{a}) \equiv_{n+m-2} (\mathcal{B}, \mathbf{b})$, then both $\mathcal{A} \equiv_n^L \mathcal{B}$ and $(\mathcal{A}, \mathbf{a}) \equiv_n^L (\mathcal{B}, \mathbf{b})$.

Proof: (1) is the special case of (2) when \mathbf{a} and \mathbf{b} are empty. Also, in (2) we note that $(\mathcal{A}, \mathbf{a}) \equiv_n (\mathcal{B}, \mathbf{b})$ trivially implies $\mathcal{A} \equiv_n \mathcal{B}$, so it suffices to prove that $(\mathcal{A}, \mathbf{a}) \equiv_{n+m-2} (\mathcal{B}, \mathbf{b})$ implies $(\mathcal{A}, \mathbf{a}) \equiv_n^L (\mathcal{B}, \mathbf{b})$. To prove this in the non-trivial case $n > 0$, we must show that the bijection chosen by Duplicator in the bijective game already includes the bijection from $\mathcal{N}_{\mathcal{A}}^{2^{n-1}}(\mathbf{a})$ to $\mathcal{N}_{\mathcal{B}}^{2^{n-1}}(\mathbf{b})$ needed in the local game. This, however, follows from the fact that (even in the regular Ehrenfeucht-Fraïssé game) distances up to 2^{n-1} from \mathbf{a} and \mathbf{b} must be preserved by Duplicator to win a game with $n + m - 2$ moves. ■

Corollary 3.4 Suppose each relation symbol in ν is at most binary.

(1) $\mathcal{A} \equiv_n \mathcal{B}$ if and only if $\mathcal{A} \equiv_n^L \mathcal{B}$.

(2) $(\mathcal{A}, \mathbf{a}) \equiv_n (\mathcal{B}, \mathbf{b})$ if and only if both $\mathcal{A} \equiv_n^L \mathcal{B}$ and $(\mathcal{A}, \mathbf{a}) \equiv_n^L (\mathcal{B}, \mathbf{b})$.

4 Local Normal Forms

In this section we will introduce a fragment $\mathcal{L}_{(loc)}(\mathbf{C})(\mathbf{x})^n$ of $\mathcal{L}_{\infty\omega}(\mathbf{C})^\omega$, called the set of **local normal formulas**, which contains formulas defining each local bijective type.

We will use the bounded quantifier notation

$$(\exists^{\geq i} y \in \mathcal{N}^r(\mathbf{x})) \psi(\mathbf{x}, y) \text{ for } (\exists^{\geq i} y) (\delta(\mathbf{x}, y) \leq r \wedge \psi(\mathbf{x}, y)).$$

Definition 4.1 $\mathcal{L}_{(loc)}(\mathbf{C})(\mathbf{x})^0$ is the set of all quantifier-free formulas $\psi(\mathbf{x})$.

$\mathcal{L}_{(loc)}(\mathbf{C})(\mathbf{x})^{n+1}$ is the set of all Boolean combinations of formulas of the form

$$\left(\exists^{\geq i} y \in \mathcal{N}^{2^n}(\mathbf{x})\right) \psi(\mathbf{x}y),$$

where y is not one of the variables in \mathbf{x} , and $\psi(\mathbf{x}y) \in \mathcal{L}_{(loc)}(\mathbf{C})(\mathbf{x}y)^n$.

We put $\mathcal{L}_{(loc)}(\mathbf{C})^n = \mathcal{L}_{(loc)}(\mathbf{C})(\mathbf{x})^n$ where \mathbf{x} is the empty sequence.

When \mathbf{x} is nonempty, each local normal formula $\phi(\mathbf{x})$ in $\mathcal{L}_{(loc)}(\mathbf{C})(\mathbf{x})^n$ has only bounded quantifiers. In the case that \mathbf{x} is empty and $n > 0$, the local normal sentences in $\mathcal{L}_{(loc)}(\mathbf{C})^n$ are Boolean combinations of sentences of the form $(\exists^{\geq i} y)\psi(y)$, with no bound on the outer quantifiers.

Lemma 4.2 (1) $\mathcal{L}_{(loc)}(\mathbf{C})(\mathbf{x})^n$ is contained in $\mathcal{L}_{\infty\omega}(\mathbf{C})^\omega$.

(2) If each relation symbol in ν is at most binary, then $\mathcal{L}_{(loc)}(\mathbf{C})(\mathbf{x})^n$ is contained in $\mathcal{L}_{\infty\omega}(\mathbf{C})^n$.

Proof: Let m be the maximum of 2 and the largest arity of a relation symbol in ν . Then $\delta(\mathbf{x}, y) \leq 2^n$ can be expressed in first order logic by a formula of quantifier rank $n + m - 2$. It follows that $\mathcal{L}_{(loc)}(\mathbf{C})(\mathbf{x})^n$ is a set of formulas in $\mathcal{L}_{\infty\omega}(\mathbf{C})^{n+m-2}$. This proves (1), and we get (2) by taking $m = 2$. ■

The next theorem shows that the fragment $\mathcal{L}_{(loc)}(\mathbf{C})(\mathbf{x})^n$ of local normal formulas captures the equivalence relation \equiv_n^L .

Theorem 4.3 (1) $\mathcal{A} \equiv \mathcal{B}$ ($\mathcal{L}_{(loc)}(\mathbf{C})^n$) if and only if $\mathcal{A} \equiv_n^L \mathcal{B}$.

(2) $(\mathcal{A}, \mathbf{a}) \equiv (\mathcal{B}, \mathbf{b})$ ($\mathcal{L}_{(loc)}(\mathbf{C})(\mathbf{x})^n$) if and only if $(\mathcal{A}, \mathbf{a}) \equiv_n^L (\mathcal{B}, \mathbf{b})$.

Proof: (1) is a special case of (2). We prove (2) by induction on n .

The result is trivial for $n = 0$. Assume that $n > 0$ and (2) holds for $n - 1$. Then for each \mathbf{x} , each $(n - 1)$ -local bijective type relative to \mathbf{x} is definable by a formula in $\mathcal{L}_{(loc)}(\mathbf{C})(\mathbf{x}y)^{n-1}$. Therefore the following statements are equivalent:

- $(\mathcal{A}, \mathbf{a}) \equiv (\mathcal{B}, \mathbf{b})$ ($\mathcal{L}_{(loc)}(\mathbf{C})(\mathbf{x})^n$).
- $\mathcal{N}_{\mathcal{A}}^{2^{n-1}}(\mathbf{a})$ and $\mathcal{N}_{\mathcal{B}}^{2^{n-1}}(\mathbf{b})$ have the same number of elements of each $(n - 1)$ -local bijective type relative to \mathbf{a}, \mathbf{b} .
- Duplicator has a winning move at $((\mathcal{A}, \mathbf{a}), (\mathcal{B}, \mathbf{b}), n)$.
- $(\mathcal{A}, \mathbf{a}) \equiv_n^L (\mathcal{B}, \mathbf{b})$.

■

Putting Theorem 4.3 together with Theorem 3.2 and Corollary 2.2, we get:

Theorem 4.4 (Normal Form Theorem)

(1) Every sentence in $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^n$ is equivalent to a sentence in the set $\mathcal{L}_{(loc)}(\mathbf{C})^n$.

(2) Every formula $\phi(\mathbf{x})$ in $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^n$ is Boolean over the set of sentences $\mathcal{L}_{(loc)}(\mathbf{C})^n$ and the set of formulas $\mathcal{L}_{(loc)}(\mathbf{C})(\mathbf{x})^n$. ■

On one hand, Theorem 4.4 gives uniform local normal forms for sentences and formulas of $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^\omega$, and on the other hand it shows that all unary quantifiers can be defined from the counting localized quantifiers $(\exists^{\geq i} y \in \mathcal{N}^{2^{n-1}}(\mathbf{x}))$ without affecting the expressiveness of $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^\omega$. This gives a simpler proof of a result in [KV95].

Corollary 4.5 (1) The logics $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^\omega$ and $\mathcal{L}_{\infty\omega}(\mathbf{C})^\omega$ are equivalent.

(2) If each relation symbol in ν is at most binary, then the logics $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^n$ and $\mathcal{L}_{\infty\omega}(\mathbf{C})^n$ are equivalent.

5 Cluster Normal Forms

In this section we give an alternative normal form which is even more local.

We start by defining the r -**cluster** $\mathcal{K}_{\mathcal{A}}^r(\mathbf{a})$ to be the subset of \mathcal{A} consisting of all elements within distance $\leq r$ from *every* $a \in \mathbf{a}$. Thus, as sets, $\mathcal{K}_{\mathcal{A}}^r(\mathbf{a}) = \bigcap_{a \in \mathbf{a}} \mathcal{N}_{\mathcal{A}}^r(a)$. In particular, if $|\mathbf{a}| = 1$, $\mathcal{K}_{\mathcal{A}}^r(a) = \mathcal{N}_{\mathcal{A}}^r(a)$, and if \mathbf{a} is the empty sequence, $\mathcal{K}_{\mathcal{A}}^r(\mathbf{a})$ is just the whole model \mathcal{A} .

The **cluster bijective game** is defined as follows.

When $n > 0$, the cluster bijective game proceeds from the position $((\mathcal{A}, \mathbf{a}), (\mathcal{B}, \mathbf{b}), n)$ according to the rules:

- (1) Spoiler picks a subtuple \mathbf{c} of \mathbf{a} (with a corresponding subtuple \mathbf{d} of \mathbf{b}). If \mathbf{a} is nonempty, then \mathbf{c} and $\mathcal{K}_{\mathcal{A}}^{2^{n-1}}(\mathbf{c})$ must be nonempty.
- (2) Duplicator chooses a bijection $f : \mathcal{K}_{\mathcal{A}}^{2^{n-1}}(\mathbf{c}) \longrightarrow \mathcal{K}_{\mathcal{B}}^{2^{n-1}}(\mathbf{d})$.
- (3) Spoiler chooses an element e in $\mathcal{K}_{\mathcal{A}}^{2^{n-1}}(\mathbf{c})$.
- (4) The game continues from the new position $((\mathcal{A}, \mathbf{c}e), (\mathcal{B}, \mathbf{d}f(e)), n - 1)$.

When $n = 0$ the game ends as usual.

We write $(\mathcal{A}, \mathbf{a}) \equiv_n^K (\mathcal{B}, \mathbf{b})$ if Duplicator has a winning strategy in the cluster bijective game starting from the position $((\mathcal{A}, \mathbf{a}), (\mathcal{B}, \mathbf{b}), n)$.

If $(\mathcal{A}, \mathbf{a}c) \equiv_n^K (\mathcal{B}, \mathbf{b}d)$, we will say that the elements c and d have the same n -**cluster type relative to \mathbf{a}, \mathbf{b}** .

Lemma 5.1 *Let $\mathbf{a}, \mathbf{a}', \mathbf{a}''$ be tuples in \mathcal{A} and $\mathbf{b}, \mathbf{b}', \mathbf{b}''$ be tuples in \mathcal{B} with $|\mathbf{a}| = |\mathbf{b}|$, $|\mathbf{a}'| = |\mathbf{b}'|$ and $|\mathbf{a}''| = |\mathbf{b}''|$. Suppose $(\mathcal{A}, \mathbf{a}\mathbf{a}') \equiv_n^K (\mathcal{B}, \mathbf{b}\mathbf{b}')$, $(\mathcal{A}, \mathbf{a}\mathbf{a}'') \equiv_n^K (\mathcal{B}, \mathbf{b}\mathbf{b}'')$, and both $\delta(\mathbf{a}', \mathbf{a}'') > 2^n$. Then $(\mathcal{A}, \mathbf{a}\mathbf{a}'\mathbf{a}'') \equiv_n^K (\mathcal{B}, \mathbf{b}\mathbf{b}'\mathbf{b}'')$.*

Proof: We note that for any corresponding subtuples \mathbf{c} of $\mathbf{a}\mathbf{a}'\mathbf{a}''$ and \mathbf{d} of $\mathbf{b}\mathbf{b}'\mathbf{b}''$ such that $\mathcal{K}_{\mathcal{A}}^{2^{n-1}}(\mathbf{c})$ or $\mathcal{K}_{\mathcal{B}}^{2^{n-1}}(\mathbf{d})$ are nonempty, we must have that \mathbf{c} and \mathbf{d} are subtuples of either $\mathbf{a}\mathbf{a}'$ and $\mathbf{b}\mathbf{b}'$ or $\mathbf{a}\mathbf{a}''$ and $\mathbf{b}\mathbf{b}''$ respectively. ■

Theorem 5.2 (1) *If $\mathcal{A} \equiv_n^K \mathcal{B}$ then $\mathcal{A} \equiv_n \mathcal{B}$.*

(2) *If $(\mathcal{A}, \mathbf{a}) \equiv_n^K (\mathcal{B}, \mathbf{b})$ and $\mathcal{A} \equiv_n^K \mathcal{B}$, then $(\mathcal{A}, \mathbf{a}) \equiv_n (\mathcal{B}, \mathbf{b})$.*

Proof: The proof is similar to that of Theorem 3.2. As before, (1) is a special case of (2). Using induction for (2), assume that $n > 0$, $(\mathcal{A}, \mathbf{a}) \equiv_n^K (\mathcal{B}, \mathbf{b})$, $\mathcal{A} \equiv_n^K \mathcal{B}$, and (2) holds for $n - 1$. We must prove the following

Claim: There is a bijection $f : \mathcal{A} \rightarrow \mathcal{B}$ such that $(\mathcal{A}, \mathbf{a}c) \equiv_{n-1}^K (\mathcal{B}, \mathbf{b}f(c))$ for all $c \in \mathcal{A}$.

Let $\mathbf{x} = (x_1, \dots, x_k)$. For each subset $S \subseteq \{1, \dots, k\}$, define

$$A_S = \{a \in \mathcal{A} : a \in \mathcal{N}_{\mathcal{A}}^{2^{n-1}}(a_i) \text{ iff } i \in S\}, \quad B_S = \{b \in \mathcal{B} : b \in \mathcal{N}_{\mathcal{B}}^{2^{n-1}}(b_i) \text{ iff } i \in S\}.$$

Note that A_S and B_S partition the models \mathcal{A} and \mathcal{B} respectively.

Using backward induction on $|S|$ (i.e. starting from $|S| = k$ and going downward until $|S| = 0$), one can show that each subset $A_S \subseteq \mathcal{A}$ and its corresponding subset $B_S \subseteq \mathcal{B}$ contain the same number of elements of each $(n - 1)$ -cluster type relative to \mathbf{a}, \mathbf{b} . Lemma 5.1 is used here.

Duplicator now builds a winning move f from the various bijections she can get between the equally sized subsets containing elements of the same cluster type. ■

As before, the converse holds when every relation symbol is at most binary. We now define another fragment of $\mathcal{L}_{\infty\omega}(\mathbf{C})^\omega$, the set $\mathcal{L}_{(clu)}(\mathbf{C})(\mathbf{x})^n$ of **cluster normal formulas**, which contains formulas defining each cluster type. We use bounded quantifiers as before.

Definition 5.3 $\mathcal{L}_{(clu)}(\mathbf{C})(\mathbf{x})^0$ is the set of all quantifier-free formulas $\psi(\mathbf{x})$.

When \mathbf{x} is nonempty, $\mathcal{L}_{(clu)}(\mathbf{C})(\mathbf{x})^{n+1}$ is the set of all Boolean combinations of formulas of the form

$$\left(\exists^{\geq i} y \in \mathcal{K}^{2^n}(\mathbf{z})\right) \psi(\mathbf{z}y),$$

where \mathbf{z} is a nonempty subtuple of \mathbf{x} , y is not one of the variables in \mathbf{z} , and $\psi(\mathbf{z}y) \in \mathcal{L}_{(clu)}(\mathbf{C})(\mathbf{z}y)^n$.

$\mathcal{L}_{(clu)}(\mathbf{C})^{n+1}$ is the set of all Boolean combinations of sentences of the form $(\exists^{\geq i} y)\psi(y)$ where $\psi(y) \in \mathcal{L}_{(clu)}(\mathbf{C})(y)^n$.

When \mathbf{x} is nonempty, each cluster normal formula in $\mathcal{L}_{(clu)}(\mathbf{C})(\mathbf{x})^n$ has only bounded quantifiers. Note that the quantifiers in the cluster normal formulas $\mathcal{L}_{(clu)}(\mathbf{C})(\mathbf{x})^n$ are bounded more strongly than the quantifiers in the local normal formulas $\mathcal{L}_{(loc)}(\mathbf{C})(\mathbf{x})^n$, because when $|\mathbf{x}| > 1$ the cluster $\mathcal{K}^r(\mathbf{x})$ is generally smaller than the neighborhood $\mathcal{N}^r(\mathbf{x})$. The set of cluster normal formulas $\mathcal{L}_{(clu)}(\mathbf{C})(\mathbf{x})^n$ is contained in $\mathcal{L}_{\infty\omega}(\mathbf{C})^\omega$, and if all relation symbols are at most binary, then $\mathcal{L}_{(clu)}(\mathbf{C})(\mathbf{x})^n$ is contained in $\mathcal{L}_{\infty\omega}(\mathbf{C})^n$.

The following theorem shows that the set $\mathcal{L}_{(clu)}(\mathbf{C})(\mathbf{x})^n$ of cluster normal formulas captures the equivalence relation \equiv_n^K .

Theorem 5.4 (1) $\mathcal{A} \equiv \mathcal{B}$ ($\mathcal{L}_{(clu)}(\mathbf{C})^n$) if and only if $\mathcal{A} \equiv_n^K \mathcal{B}$.

(2) $(\mathcal{A}, \mathbf{a}) \equiv (\mathcal{B}, \mathbf{b})$ ($\mathcal{L}_{(clu)}(\mathbf{C})(\mathbf{x})^n$) if and only if $(\mathcal{A}, \mathbf{a}) \equiv_n^K (\mathcal{B}, \mathbf{b})$. ■

As a consequence of Theorem 5.4, we get another normal form theorem.

Theorem 5.5 (Cluster Normal Form Theorem)

(1) Every sentence in $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^n$ is equivalent to a sentence in the set $\mathcal{L}_{(clu)}(\mathbf{C})^n$.

(2) Every formula $\phi(\mathbf{x})$ in $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^n$ is Boolean over the set of sentences $\mathcal{L}_{(clu)}(\mathbf{C})^n$ and the set of formulas $\mathcal{L}_{(clu)}(\mathbf{C})(\mathbf{x})^n$. ■

6 Gaifman and Hanf Locality

In this section we will apply the normal form theorem to the notions of Gaifman and Hanf locality introduced in [HLN99].

Let \mathbf{x} be a nonempty tuple of variables. A formula $\phi(\mathbf{x})$ of $\mathcal{L}_{\infty\omega}(\mathbf{C})^\omega$ is *r-local* if all of its quantifiers are restricted to $\mathcal{N}^r(\mathbf{x})$. Every *r*-local formula has strong Gaifman locality rank at most *r*, that is, whenever $\mathcal{N}_{\mathcal{A}}^r(\mathbf{a}) \cong \mathcal{N}_{\mathcal{B}}^r(\mathbf{b})$ where $|\mathbf{a}| = |\mathbf{b}| = |\mathbf{x}|$, we have $\mathcal{A} \models \phi(\mathbf{a})$ if and only if $\mathcal{B} \models \phi(\mathbf{b})$.

Lemma 6.1 Suppose the tuple \mathbf{x} is nonempty. Every formula in $\mathcal{L}_{(loc)}(\mathbf{C})(\mathbf{x})^n$ is equivalent to a $(2^n - 1)$ -local formula. Similarly, every formula in $\mathcal{L}_{(clu)}(\mathbf{C})(\mathbf{x})^n$ is equivalent to a $(2^n - 1)$ -local formula.

Proof: Adding the distances seen from \mathbf{x} in a formula in $\mathcal{L}_{(loc)}(\mathbf{C})(\mathbf{x})^n$ or $\mathcal{L}_{(clu)}(\mathbf{C})(\mathbf{x})^n$, we get $2^{n-1} + 2^{n-2} + \dots + 1 = 2^n - 1$. ■

Theorem 6.2 (Locality Theorem) Let $n > 0$.

(1) Every sentence in $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^n$ is Boolean over sentences of the form $(\exists^{\geq i} y)\psi(y)$ where $\psi(y)$ is a $(2^{n-1} - 1)$ -local formula in $\mathcal{L}_{\infty\omega}(\mathbf{C})^\omega$.

(2) Every formula $\phi(\mathbf{x})$ in $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^n$ is Boolean over sentences as in (1) above, and $(2^n - 1)$ -local formulas in $\mathcal{L}_{\infty\omega}(\mathbf{C})^\omega$.

Proof: By Theorem 4.4 (or Theorem 5.5) and Lemma 6.1. ■

Part (1) is the result (C) stated in the Introduction. This theorem is weaker than the Normal Form Theorems 4.4 and 5.5, but the statement does not depend on the inductive definition of the fragment $\mathcal{L}_{(loc)}(\mathbf{C})^n$ or $\mathcal{L}_{(clu)}(\mathbf{C})^n$. It complements the results in [HLN99] and [Lib00] on Hanf and Gaifman locality ranks.

Part (1) implies that every sentence ϕ in $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^n$ has Hanf locality rank less than 2^{n-1} . That is, if there is a bijection $f : \mathcal{A} \rightarrow \mathcal{B}$ such that $\mathcal{N}_{\mathcal{A}}^r(c) \cong \mathcal{N}_{\mathcal{B}}^r(f(c))$ for all $c \in \mathcal{A}$ where $r = 2^{n-1} - 1$, then $\mathcal{A} \models \phi$ if and only if $\mathcal{B} \models \phi$.

Part (2) implies that for nonempty tuples \mathbf{x} , every formula $\phi(\mathbf{x})$ in $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^n$ has Gaifman locality rank less than 2^n . That is, if $\mathcal{N}_{\mathcal{A}}^r(\mathbf{a}) \cong \mathcal{N}_{\mathcal{B}}^r(\mathbf{b})$ where $r = 2^n - 1$, then $\mathcal{A} \models \phi(\mathbf{a})$ if and only if $\mathcal{B} \models \phi(\mathbf{b})$.

Both of these bounds were obtained (for finite models) by a different method in [Lib00].

7 Conclusion

We defined the local bijective Ehrenfeucht-Fraïssé game and showed that Duplicator has a winning strategy for that game if and only if she has a winning strategy for the bijective Ehrenfeucht-Fraïssé game. Thus the local bijective game characterizes the equivalence of two models with respect to $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^n$. This leads to a Normal Form Theorem of the Gaifman type for the logic $\mathcal{L}_{\infty\omega}(\mathbf{Q}_u)^\omega$. This normal form uses only counting quantifiers, showing that arbitrary unary quantifiers can be captured by counting quantifiers. Our normal form differs from the normal form in Gaifman's theorem in that it allows infinite Boolean combinations, and the inner formula states that the neighborhoods are distinct rather than disjoint. It has the advantage of preserving quantifier rank, except that unbounded quantifiers are replaced by bounded quantifiers. In the case that every relation in ν is at most binary, the normal forms for a sentence of quantifier rank n will again have quantifier rank at most n .

A consequence of the Normal Form Theorem is the Locality Theorem, which shows that each sentence of quantifier rank $n > 0$ is a Boolean combination of sentences which say that there are at least i elements y which satisfy a $(2^{n-1} - 1)$ -local formula $\psi(y)$. There is a similar result for formulas. The Locality Theorem implies many of the results in the paper [Lib00]. In particular, one can read off the optimum Gaifman locality rank and the optimum Hanf locality rank for formulas of given quantifier rank.

Questions: Is there a normal form for first order logic which is like the one used here?

Can such a normal form give a bound for Gaifman's Theorem for first order logic that is better than the $4^{n-1} - 1$ bound found in [KL04]?

Are there similar normal form theorems for infinitary logics with quantifiers of higher arity?

References:

[FSV95] R. Fagin, L. Stockmeyer, and M. Vardi, On monadic NP vs monadic co-NP, **Information and Computation**, vol. 120 (1995), pp. 78-92.

[Gai82] H. Gaifman, On local and nonlocal properties, J. Stern, editor, **Logic Colloquium '81**, North Holland, Amsterdam (1982), pp. 105-135.

[Han65] W. Hanf, Model-theoretic methods in the study of elementary logic, **The Theory of Models** (J. Addison, L. Henkin, and A. Tarski, Eds.), North Holland, Amsterdam (1965), pp. 132-145.

[Hel89] L. Hella, Definability hierarchies of generalized quantifiers, **Annals of Pure and Applied Logic**, vol. 43 (1989), pp. 235-271.

[Hel96] L. Hella, Logical hierarchies in PTIME, **Information and Computation**, vol. 129 (1996), pp. 1-19.

[HLN99] L. Hella, L. Libkin, and J. Nurmonen, Notions of locality and their logical characterizations over finite models, **J. Symbolic Logic**, vol. 64 (1999), pp. 1751-1773.

[Hod97] W. Hodges, **A Shorter Model Theory**, Cambridge University Press, 1997.

[Imm99] N. Immerman, **Descriptive Complexity**, Springer Verlag, New York, (1999).

[KL04] H.J. Keisler and W.B. Lotfallah, Shrinking games and local formulas, to appear in **Annals of Pure and Applied Logic**.

[KV95] Ph. Kolaitis and J. Väänänen, Generalized quantifiers and pebble games on finite structures, **Annals of Pure and Applied Logic**, vol. 74 (1995), pp. 23-75.

[Lib00] L. Libkin, On counting logics and local properties, In **ACM TOCL**, vol. 1 (1) (2000).

[LS96] S. Lifsches and S. Shelah, Distorted sums of models, Unpublished manuscript (1996).

[Nur96] J. Nurmonen, On winning strategies with unary quantifiers, **J. Logic and Computation**, vol. 6 (1996), pp. 779-798.