Scientific Report:
Exchange Visit to Helsinki

Tong Wang

Institute for Logic, Language and Computation,
University of Amsterdam

December 6, 2012

1 Purpose of the Visit

I am a student in mathematical logic. I graduated from the Master of Logic
program at the Institute of Logic, Language and Computation (ILLC), Uni-
versity of Amsterdam in September 2012, with Jouko Väänänen and Benedikt
Löwe as supervisors. My Master's Thesis is "An Ehrenfeucht-Fraïssé Game
for $L_{\omega_1\omega}$".

The purpose of my exchange visit to Helsinki was to continue working on
two ongoing projects that I had with Jouko Väänänen. These projects have
been started when Väänänen has given short intensive courses in Amsterdam.
The first project is an investigation of the extent to which second order set
theory can prove (using the Comprehension Axioms) its own categoricity
and consistency. A joint paper on this project has been accepted. The
second project is based on my Master's thesis. In this project we developed
an Ehrenfeucht-Fraïssé game for the infinitary language $L_{\omega_1\omega}$. The goal of
my exchange visit was to solve some of the open problems left in these two
projects. In particular, we planned to prepare a publication on the second
project "An Ehrenfeucht-Fraïssé Game for $L_{\omega_1\omega}$" during this visit.

2 Content of the Visit

I started my exchange visit to Helsinki on October 10th, 2012. I gave a presentation "An New Ehrenfeucht-Fraïssé Game" at the Finite Model Theory Seminar at the University of Tampere on October 18th. I gave a presentation on the same topic at the Logic Seminar at the University of Helsinki on October 24th. We took these chances to discuss our work with other experts in the field, in particular Lauri Hella and Tapani Hyttinen.

I worked with Väänänen to improve the results that we had on the two projects, partly based on the discussions following the aforementioned presentations. We obtained some new results on the Ehrenfeucht-Fraïssé game for $L_{\omega_1\omega}$. I will present these results in the next section. We also wrote a joint paper on this project, which has been submitted to *Mathematical Logic Quarterly*.

3 Main Results

The new results that we have obtained focus on our second project.

3.1 Alternative conceptions of size

In the first of stage of this project we showed that measuring the complexity of formulas by size is *one* way to define an Ehrenfeucht-Fraïssé game for $L_{\omega_1\omega}$. However, our viewpoint naturally suggests the possibility of a plurality of complexity measures, and hence a plurality of Ehrenfeucht-Fraïssé games for $L_{\omega_1\omega}$. In particular, Hyttinen pointed out that we can define the complexity of a conjunction as the supremum of the complexity of the conjuncts plus 1, and similarly for disjunction. Can we find other complexity measures which give rise to an Ehrenfeucht-Fraïssé game for $L_{\omega_1\omega}$? Furthermore, to which extent are these complexity measures *reasonable*?

We shall start by looking at the bare minimum, namely which complexity measures could give rise to an Ehrenfeucht-Fraïssé game. It turns out that if we merely want an adequacy theorem to hold, we can be quite relaxed about the choice of complexity measures.

Definition 1. Let $q : \mathrm{Ord} \to \mathrm{Ord}$, $r : \mathrm{Ord}^2 \to \mathrm{Ord}$ and $t : \mathrm{Ord}^\omega \to \mathrm{Ord}$ be ordinal functions. Consider an abstract complexity measure c defined by q,

r and t:

$$c(\phi) = 1 \text{ if } \phi \text{ is an atomic or negated atomic formula,}$$
$$c(\phi \wedge \psi) = r((c(\phi), c(\psi))),$$
$$c(\phi \vee \psi) = r((c(\phi), c(\psi))),$$
$$c(\exists x \phi) = q(c(\phi)),$$
$$c(\forall x \phi) = q(c(\phi)),$$
$$c(\bigwedge_{i \in \omega} \phi_i) = t(\{c(\phi_i) | i \in \omega\}),$$
$$c(\bigvee_{i \in \omega} \phi_i) = t(\{c(\phi_i) | i \in \omega\}).$$

We say that c is *nice* if q, r and t satisfy the following conditions:

1. If $\gamma > 0$, then $q(\gamma) > \gamma$.

2. If $\gamma_1, \gamma_2 > 0$, then $r(\gamma_1, \gamma_2) > \gamma_1$ and $r(\gamma_1, \gamma_2) > \gamma_2$.

3. If $\{\gamma_i | i \in \omega\}$ is a sequence of non-zero ordinals, then $t(\{\gamma_i | i \in \omega\}) > \gamma_i$ for all $i \in \omega$.

Niceness guarantees that the complexity of a formula is always strictly larger than the complexity of its proper subformulas. Granted this, we can generalize the game EFB_α to any nice complexity measure by performing minor cosmetic surgery. Let EFB_α^c denote the Ehrenfeucht-Fraïssé game for $L_{\omega_1\omega}$ under the complexity measure c. The starting position and the winning condition in the game EFB_α^c remain the same as in EFB_α. The way that the rank of the game is counted is modified. We only give the 'left' moves in the game EFB_α^c here; the 'right' moves are dual.

Finite left splitting move: Player **I** first represents $\mathfrak{A}$ as a union $\mathfrak{A}_1 \cup \mathfrak{A}_2$. He also chooses non-zero ordinals γ_1 and γ_2 such that $r(\gamma_1, \gamma_2) \leq \gamma$. Now the game continues from the position $(\mathfrak{A}_1, \mathfrak{B}, \gamma_1)$ or from the position $(\mathfrak{A}_2, \mathfrak{B}, \gamma_2)$, and player **II** can choose which.

Infinite left splitting move: Player **I** first represents $\mathfrak{A}$ as a countable union $\bigcup_{i \in \omega} \mathfrak{A}_i$. He also chooses a sequence of non-zero ordinals $\{\gamma_i | i \in \omega\}$ such that $t(\{\gamma_i | i \in \omega\}) \leq \gamma$. Now the game continues from the position $(\mathfrak{A}_i, \mathfrak{B}, \gamma_i)$ for some $i \in \omega$, and player **II** can choose which.

Left supplementing move: Player **I** picks an element from each structure $(A, \alpha) \in \mathfrak{A}$. More precisely, **I** chooses a natural number j and a choice function F for $\mathfrak{A}$. He also chooses an ordinal δ such that $q(\delta) \leq \gamma$. Then the game continues from the position $(\mathfrak{A}(F/j), \mathfrak{B}(\star/j), \delta)$.

It is an easy observation that the proof of the adequacy theorem goes through if we substitute size for any nice complexity measure. We have the following generalized adequacy theorem.

Theorem 1. *Let c be a nice complexity measure. Let L be a relational vocabulary, $\mathfrak{A}$ and $\mathfrak{B}$ classes of L-structures, and let α be an ordinal. Then the following are equivalent:*

1. *Player **I** has a winning strategy in the game $\mathrm{EFB}^c_\alpha(\mathfrak{A}, \mathfrak{B})$.*

2. *There is an L-formula ϕ in $L_{\omega_1 \omega}$ with $c(\phi) \leq \alpha$ such that $(\mathfrak{A}, \mathfrak{B}) \models \phi$.*

In particular, size is a nice complexity measure: take q to be the successor function, r the natural sum, and t the infinite natural sum. Another interesting nice complexity measure c_1 can be obtained by taking q to be the successor function, $r(\gamma_1, \gamma_2) = \sup(\gamma_1 + 1, \gamma_2 + 1)$ and $t(\{\gamma_i | i \in \omega\}) = \sup_{i \in \omega}(\gamma_i + 1)$. Note that if we take q to be the successor function and r, t to be the supremum, then the complexity measure is just quantifier rank—but this is not a nice complexity measure.

To determine which of the nice complexity measures are the more natural and reasonable ones is a subtler question. As a general rule of thumb, we would like a complexity measure to carry some meaning, so that we have an idea of what exactly it is measuring. We also want a complexity measure to be invariant under certain logical equivalences, so that it is compatible with our intuition about 'complexity'. Between size and c_1, size has the advantage of having a clear intuition behind it. This is especially so in the propositional case: the size of a propositional formula is just the number of propositional symbols in this formula. The quantifier and infinitary cases are natural extensions of this basic intuition. For c_1 the intuition is not so clear. As for invariance under logical equivalences, both size and c_1 are invariant under permutations of conjuncts and disjuncts. They part ways, however, when we try to count the complexity of repetitions of formulas. Consider atomic formulas $\phi_i, i \in \omega$. Let Θ be the infinite conjunction $\bigwedge_{i \in \omega} \phi_0$. Let Ψ be the infinite conjunction $\bigwedge_{i \in \omega} \phi_i$. Size assigns Θ and Ψ the complexity ω.

The measure c_1 assigns Θ and Ψ the complexity 2. It seems that c_1 gives the more reasonable complexity for Θ, since Θ is no more than a repetition of the same formula ϕ_0. On the other hand size seems to give the more reasonable complexity for Ψ, while c_1 does not reflect the fact that Ψ contains much more information than any finite conjunction. However as long as complexity measures are defined purely syntactically, every complexity measure would assign the same complexity to Θ and Ψ, so there is no complexity measure that assigns the 'right' complexity to Θ and Ψ at the same time. This suggests that there is no point in trying to find *the* reasonable complexity measure. We may indeed adopt a pluralistic view: we shall choose which is the appropriate complexity measure depending on what information we intend to model.

3.2   Extending the game to higher infinitary languages

Another interesting question is whether we can extend our complexity measures and Ehrenfeucht-Fraïssé games from $L_{\omega_1\omega}$ to $L_{\omega_2\omega}$ and even higher infinitary languages.

It turns out that the possibility depends on which complexity measure we choose. For "size", new problem arises when we move from $L_{\omega_1\omega}$ to $L_{\omega_2\omega}$. Suppose we want to define the notion of size for $L_{\omega_2\omega}$-formulas in such a way that it is coherent with our definition of size for $L_{\omega_1\omega}$-formulas. To illustrate the problem, consider the uncountable conjunction $\Phi = \bigwedge_{i\in\omega_1} \phi_i$, where ϕ_i is an atomic formula for all $i \in \omega_1$. What is the right size of Φ? There are the following three desiderata:

1. If θ is an $L_{\omega_1\omega}$-formula, then the size of θ is the same as it is given in Definition **??**.

2. The size of Φ is the supremum of the size of its proper subformulas, namely $s(\Phi) = \sup_{\alpha<\omega_1} s(\bigwedge_{i\in\alpha} \phi_i)$.

3. The size of Φ is at least ω_1.

These three conditions cannot be satisfied simultaneously. Suppose there is a complexity measure s satisfying conditions 1 and 2. Consider any infinite ordinal $\alpha < \omega_1$. By condition 1, we have that $s(\bigwedge_{i\in\alpha} \phi_i) = s(\bigwedge_{i\in\omega} \phi_{p(i)}) = \omega$, where p is a bijection from ω to α. This together with condition 2 renders the size of Φ to be ω, contrary to condition 3.

On the other hand, if we use the complexity measure c_1 discussed in Section 3.1, then it can be extended to $L_{\omega_2\omega}$ and higher infinitary languages straightforwardly.

4   Projected Publication

During the visit to Helsinki I wrote a joint paper "An Ehrenfeucht-Fraïssé Game for $L_{\omega_1\omega}$" with Väänänen. This paper has been submitted to *Mathematical Logic Quarterly* on December 4th. We also submitted the paper in preprint form to arXiv.[1] I attach to this report the submitted paper.

---

[1]The preprint is available at: http://arxiv.org/abs/1212.0108.

An Ehrenfeucht-Fraïssé Game for $L_{\omega_1\omega}$

J. Väänänen[1, 2, *] and **T. Wang**[2, *****]

[1] Department of Mathematics and Statistics, University of Helsinki, Gustaf Hällströmin katu 2b, FIN-00014, Finland.

[2] Institute for Logic, Language and Computation, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands.

In this paper we develop an Ehrenfeucht-Fraïssé game for $L_{\omega_1\omega}$. Unlike the standard Ehrenfeucht-Fraïssé games which are modeled solely after the behavior of quantifiers, this new game also takes into account the behavior of connectives in logic. We prove the adequacy theorem for this game. We also apply the new game to prove complexity results about infinite binary strings.

1 Introduction

One major limitation of the expressive power of first order logic is that many familiar mathematical concepts (e.g. finiteness, connectedness of graphs, and well-foundedness) are not definable. These concepts involve, explicitly or implicitly, the notion of infinity which is difficult to handle in first order logic, partly because first order logic only has *finitary* formulas. One way to transcend this boundary is to consider stronger logics where *infinitary* formulas are also admissible. Infinitary logics made their first appearance in print with the papers of Scott and Tarski [1] and Tarski [2]. Among all infinitary logics, the logic $L_{\omega_1\omega}$ turns out to be an especially interesting example. The logic $L_{\omega_1\omega}$ adds to first order logic countable[1] conjunctions and disjunctions of formulas. It is in a sense the smallest infinitary logic. It strikes a nice balance—$L_{\omega_1\omega}$ is expressive enough to define many familiar mathematical concepts which are not definable in first order logic (including the three aforementioned), while it still enjoys model-theoretic properties such as having the Completeness Theorem [3] and the Interpolation Theorem [4].

Throughout this paper we assume L to be a relational vocabulary. The terms of L are defined as usual.

Definition 1.1 The set of $L_{\omega_1\omega}$-formulas of L is the smallest set closed under the following operations:

1. If t_1 and t_2 are terms, then $t_1 = t_2$ is a formula.

2. If R is an n-place relational symbol and $t_1, \ldots, t_n$ are terms, then $Rt_1 \ldots t_n$ is a formula.

3. If ϕ is a formula, then so is $\neg\phi$.

4. If ϕ and ψ are formulas, then so is $\phi \vee \psi$.

5. If ϕ and ψ are formulas, then so is $\phi \wedge \psi$.

6. If I is a countable set and for every $i \in I$, ϕ_i is a formula, then so is $\bigvee_{i\in I} \phi_i$.

---

[1] In this paper we use 'countable' to mean 'countably infinite'.

7. If I is a countable set and for every $i \in I$, ϕ_i is a formula, then so is $\bigwedge_{i \in I} \phi_i$.

8. If ϕ is a formula and x_n a variable, then $\exists x_n \phi$ is a formula.

9. If ϕ is a formula and x_n a variable, then $\forall x_n \phi$ is a formula.

Rather surprisingly, some questions in the basic model theory of $L_{\omega_1\omega}$ have been left unanswered. In this paper we address one such question, namely the question of separation in $L_{\omega_1\omega}$. Let us fix some notation. We use A, B to denote L-structures and use $\mathfrak{A}$, $\mathfrak{B}$ to denote classes of L-structures. We say that a formula ϕ *separates* a pair of structures A and B if $A \models \phi$ and $B \not\models \phi$. We say that a formula ϕ *separates* two classes of structures $\mathfrak{A}$ and $\mathfrak{B}$, denoted by $(\mathfrak{A}, \mathfrak{B}) \models \phi$, if for all structures $A \in \mathfrak{A}$ we have $A \models \phi$ and for all structures $B \in \mathfrak{B}$ we have $B \not\models \phi$. Let $A \equiv B$ denote that A and B are elementarily equivalent. If α is an ordinal, let $A \equiv^\alpha_{\infty\omega} B$ denote that A and B satisfy the same $L_{\infty\omega}$ [2] sentences up to quantifier rank α. Finally we denote by $A \equiv_{\infty\omega} B$ that A and B are equivalent in $L_{\infty\omega}$. Note that the notions of separation and equivalence are opposite sides of the same coin: two structures can be separated by a formula in a certain language if and only if they are not equivalent in this language. The central question that we try to answer in this paper is the following:

Question 1: Given two structures A and B, is there a formula ϕ in $L_{\omega_1\omega}$ that separates them? More generally, given two classes of structures $\mathfrak{A}$ and $\mathfrak{B}$, is there a formula ϕ in $L_{\omega_1\omega}$ that separates them?

A very useful machinery in logic to study equivalence and separation results is the Ehrenfeucht-Fraïssé game, also known as the back-and-forth game. The Ehrenfeucht-Fraïssé game was first formulated for first order logic by Ehrenfeucht [5, 6], based on the ideas of Fraïssé [7]. Let us denote this game by EF_n. For n a natural number, the game EF_n characterizes equivalence in first order logic up to formulas of quantifier rank n. Karp [3] developed the idea into a characterization of equivalence in infinitary logic using back-and-forth sequences, which was later formulated as a game in [8]. See also [9, Chapter 7]. For α an ordinal, the game EFD_α characterizes equivalence in $L_{\infty\omega}$ up to formulas of quantifier rank α. The logic $L_{\omega_1\omega}$ is intermediate between first order logic and $L_{\infty\omega}$. It is only natural to ask: is there an Ehrenfeucht-Fraïssé game that, standing somehow in between the standard games EF_n and EFD_α, characterizes equivalence in $L_{\omega_1\omega}$? The question of finding such a game has been open since 1970s.

In this paper we introduce an Ehrenfeucht-Fraïssé game for $L_{\omega_1\omega}$. The text is divided as follows. In Section 2 we analyze the reason why the standard game EFD_α fails to characterize equivalence in $L_{\omega_1\omega}$. The idea is roughly that the standard game only models the behavior of quantifiers in logic, and while this is enough for studying separation in $L_{\infty\omega}$, it is not enough for $L_{\omega_1\omega}$ since the latter is defined by explicit reference to the *countability* of conjunctions and disjunctions. An Ehrenfeucht-Fraïssé game for the latter needs to model also the behavior of connectives. Fortunately there already exists an Ehrenfeucht-Fraïssé game for first order logic that is in this spirit, namely the game EFB_n introduced by Hella and Väänänen [10]. We present this game in Section 3. In Section 4 we extend this game to cover $L_{\omega_1\omega}$-formulas and we prove the adequacy theorem for the new game. In Section 5 we consider other alternatives for an Ehrenfeucht-Fraïssé game for $L_{\omega_1\omega}$. In Section 6 we test our new tool on the complexity of information coded by infinite binary strings.

2   The Standard Ehrenfeucht-Fraïssé Game

Let us first recapitulate the standard Ehrenfeucht-Fraïssé game in infinitary logic.

Definition 2.1 Let L be a relational vocabulary and A, B be L-structures. A partial mapping $\pi : A \to B$ is a *partial isomorphism* if it is an isomorphism between $A \restriction \text{dom}(\pi)$ and $B \restriction \text{ran}(\pi)$.

Definition 2.2 Let L be a relational vocabulary and A, B be L-structures. Let α be an ordinal. The game $\text{EFD}_\alpha(A, B)$ has two players. The number α is called the *rank* of the game. The positions in the game are of the form $(\bar{a}, \bar{b}, \gamma)$, where $\bar{a}$ and $\bar{b}$ are tuples in A and B respectively and γ is an ordinal. The game begins from the position $(\emptyset, \emptyset, \alpha)$. Suppose the game is in position $(\bar{a}, \bar{b}, \gamma)$, $\gamma > 0$, there are the following possibilities for the next move in the game.

---

[2] The logic $L_{\infty\omega}$ is the extension of first order logic with arbitrary conjunctions and disjunctions.

1. Player **I** picks an element $c \in A$. He also picks an ordinal $\beta < \gamma$. Player **II** responds by picking an element $d \in B$. The game continues from the position $(\bar{a}c, \bar{b}d, \beta)$.

2. Player **I** picks an element $d \in B$. He also picks an ordinal $\beta < \gamma$. Player **II** responds by picking an element $c \in A$. The game continues from the position $(\bar{a}c, \bar{b}d, \beta)$.

The game terminates in a position $(\bar{a}, \bar{b}, \gamma)$ if $\gamma = 0$. Player **II** wins if $\bar{a} \mapsto \bar{b}$ is a partial isomorphism from A to B. Otherwise player **I** wins.

The following result is essentially due to Karp [3], although it was originally formulated in terms of back-and-forth sequences.

Theorem 2.3 (Karp) *Let L be a relational vocabulary, A and B be L-structures and α be an ordinal. Then the following are equivalent:*

1. $A \equiv^{\alpha}_{\infty\omega} B$.

2. *Player **II** has a winning strategy in* $\mathrm{EFD}_{\alpha}(A, B)$.

The game EFD_{α} provides important information about our Question 1. Make the following simple observation.

Proposition 2.4 *For any formula $\phi \in L_{\omega_1\omega}$, the quantifier rank of ϕ is a countable ordinal.*

Corollary 2.5 *If player **II** has a winning strategy in* $\mathrm{EFD}_{\omega_1}(A, B)$, *then there is no formula in $L_{\omega_1\omega}$ that separates A and B.*

The standard Ehrenfeucht-Fraïssé games EF_n and EFD_α characterize a hierarchy of equivalence relations. They provide in a sense upper bounds and lower bounds for Question 1. The situation can be illustrated by Figure 1.

If two structures A and B are not even elementarily equivalent, then obviously there is an $L_{\omega_1\omega}$-formula separating the two because this formulas is already in first order logic. On the other hand, if A and B are $L_{\infty\omega}$-equivalent, then there is no hope of finding an $L_{\omega_1\omega}$-formula separating the two structures. Corollary 2.5 shows that we can improve the upper bound: if two structures are equivalent in $L_{\infty\omega}$ up to formulas of quantifier rank ω_1, then they cannot be separated by any formula in $L_{\omega_1\omega}$.

Therefore, from the point of view of $L_{\omega_1\omega}$ all the excitement concentrates in the interval between $A \equiv B$ and $A \equiv^{\omega_1}_{\infty\omega} B$. However the game EFD_α is not able to tell us much about what happens in this interval. The reason is that although a small number of connectives implies a small quantifier rank (as in Proposition 2.4), there is no implication in the reverse direction. We may very well have a separation formula of very low quantifier rank, yet it contains a large number of conjunctions and disjunctions, and the standard game cannot detect the existence of such a formula. In particular, the standard game does not answer the question whether there is a quantifier-free

$$A \equiv_{\infty\omega} B$$

$$A \equiv^{\omega_1}_{\infty\omega} B$$

$$\exists \phi \in L_{\omega_1\omega} (A \models \phi, B \not\models \phi)?$$

$$A \equiv B$$
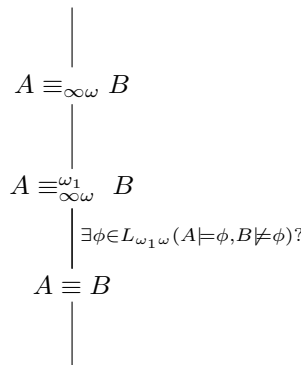
Fig. 1 Hierarchy of equivalence relations.

formula in $L_{\omega_1\omega}$ separating two structures, and still, the elimination of quantifiers in favor of connectives is the oldest application of infinitary logic.[3]

The failure of the standard Ehrenfeucht-Fraïssé game in characterizing equivalence in $L_{\omega_1\omega}$ can be understood in the following way. An Ehrenfeucht-Fraïssé game, meaning a set of rules and winning conditions, corresponds to a way of measuring the complexity of formulas. If a game on the one hand, and a complexity measure on the other hand, form a good match, we would have the adequacy theorem for this game. In the case of the game EFD_α, the complexity of a formula is measured by its quantifier rank. The two types of moves in the game EFD_α model the nature of existential and universal quantifiers respectively. The logic $L_{\omega_1\omega}$, however, is defined by reference to the countability of conjunctions and disjunctions, and quantifier rank simply does not capture this information. Therefore there is a sense in saying that the game EFD_α does not solve the question of separation in $L_{\omega_1\omega}$ because it measures the complexity of $L_{\omega_1\omega}$-formulas in an unsuitable way.

In this light, we need two things in order to develop an Ehrenfeucht-Fraïssé game for $L_{\omega_1\omega}$: (a) a new game and (b) a new measure of the complexity of infinitary formulas. These two components have to be compatible so that we can have the adequacy theorem. Moreover, both of them should take into account not only the behavior of quantifiers, but also of connectives.

3 A Game for First Order Logic

What is a suitable complexity measure for $L_{\omega_1\omega}$-formulas then? We draw inspiration from the game EFB_n, proposed by Hella and Väänänen in [10] as a refinement of the standard Ehrenfeucht-Fraïssé game for first order logic. This game resembles the Ehrenfeucht-Fraïssé game for independence friendly logic, presented in [12]. An innovative feature of the game EFB_n is that it measures the complexity of a formula by making use of our intuition about the length of a formula.

Let L be a relational vocabulary as usual. Henceforth we shall assume all formulas to be in negation normal form.

Definition 3.1 The *size*, denoted by $s(\phi)$, of a formula ϕ in first order logic is defined inductively as follows:

$$s(\phi) = 1 \text{ if } \phi \text{ is an atomic or negated atomic formula},$$
$$s(\phi \wedge \psi) = s(\phi) + s(\psi),$$
$$s(\phi \vee \psi) = s(\phi) + s(\psi),$$
$$s(\exists x\phi) = s(\phi) + 1,$$
$$s(\forall x\phi) = s(\phi) + 1.$$

The size of a first order formula is a positive integer. It is always bigger or equal to the quantifier rank.

To define the game EFB_n we need some notation. We use $x_j, j \in \mathbb{N}$ to denote variables. A variable assignment for a structure A is a finite partial mapping $\alpha : \mathbb{N} \to A$. The finite domain of α is denoted by $\text{dom}(\alpha)$. The game EFB_n is defined on pairs of classes of structures rather than on pairs of structures. We consider classes $\mathfrak{A}$ of structures (A, α), where A is a model and α is a variable assignment on A. We assume that whenever $(A, \alpha), (B, \beta) \in \mathfrak{A}$, then A and B have the same vocabulary, and α and β have the same domain, which we denote by $\text{dom}(\mathfrak{A})$. If α is an assignment on A, $a \in A$ and $j \in \mathbb{N}$, then $\alpha(a/j)$ is the assignment that maps j to a and agrees with α otherwise. If F is a choice function on $\mathfrak{A}$, namely that F is a function defined on $\mathfrak{A}$ such that $F(A, \alpha) \in A$ for all $(A, \alpha) \in \mathfrak{A}$, then $\mathfrak{A}(F/j)$ is defined as $\{(A, \alpha(F(A, \alpha)/j)|(A, \alpha) \in \mathfrak{A}\}$. Finally, $\mathfrak{A}(\star/j) = \{(A, \alpha(a/j))|(A, \alpha) \in \mathfrak{A}, a \in A)\}$.

Definition 3.2 Let L be a relational vocabulary and $\mathfrak{A}_0, \mathfrak{B}_0$ classes of L-structures. Let n be a positive integer. The game $EFB_n(\mathfrak{A}_0, \mathfrak{B}_0)$ has two players. The number n is called the *rank* of the game. The positions in the game are of the form $(\mathfrak{A}, \mathfrak{B}, m)$ where $\mathfrak{A}, \mathfrak{B}$ are classes of L-structures and $m \in \mathbb{N}$. The game begins from position $(\mathfrak{A}_0, \mathfrak{B}_0, n)$. Suppose the game is in position $(\mathfrak{A}, \mathfrak{B}, m)$. There are the following possibilities for the next move in the game.

---

[3] According to Barwise [11], Charles Peirce thought of quantifiers as infinite conjunctions or disjunctions, and this was picked up by Löwenheim, Wittgenstein and others, and used in proof theory by Novikoff already in the 1940s.

Left splitting move:: Player **I** first represents $\mathfrak{A}$ as a union $\mathfrak{A}_1 \cup \mathfrak{A}_2$. He also chooses positive numbers m_1 and m_2 such that $m_1 + m_2 = m$. Now the game continues from the position $(\mathfrak{A}_1, \mathfrak{B}, m_1)$ or from the position $(\mathfrak{A}_2, \mathfrak{B}, m_2)$, and player **II** can choose which.

Right splitting move:: Player **II** first represents $\mathfrak{B}$ as a union $\mathfrak{B}_1 \cup \mathfrak{B}_2$. He also chooses positive numbers m_1 and m_2 such that $m_1 + m_2 = m$. Now the game continues from the position $(\mathfrak{A}, \mathfrak{B}_1, m_1)$ or from the position $(\mathfrak{A}, \mathfrak{B}_2, m_2)$, and player **II** can choose which.

Left supplementing move:: Player **I** picks an element from each structure $(A, \alpha) \in \mathfrak{A}$. More precisely, **I** chooses a natural number j and a choice function F for $\mathfrak{A}$. Then the game continues from the position $(\mathfrak{A}(F/j), \mathfrak{B}(\star/j), m-1)$.

Right supplementing move:: Player **I** picks an element from each structure $(B, \beta) \in \mathfrak{B}$. More precisely, **I** chooses a natural number j and a choice function F for $\mathfrak{B}$. Then the game continues from the position $(\mathfrak{A}(\star/j), \mathfrak{B}(F/j), m-1)$.

The game ends in a position $(\mathfrak{A}, \mathfrak{B}, m)$ and player **I** wins if there is an atomic or a negated atomic formula ϕ such that $(\mathfrak{A}, \mathfrak{B}) \models \phi$. Player **II** wins the game if they reach a position $(\mathfrak{A}, \mathfrak{B}, m)$ such that $m = 1$ and **I** does not win in this position.

The game EFB_n characterizes separation in first order logic up to formulas of size n:

Theorem 3.3 (Hella-Väänänen) *Let L be a relational vocabulary, $\mathfrak{A}$ and $\mathfrak{B}$ be classes of L-structures, and let n be a positive integer. Then the following are equivalent:*

1. *Player **I** has a winning strategy in the game $\text{EFB}_n(\mathfrak{A}, \mathfrak{B})$.*

2. *There is a first order L-formula of size $\leq n$ such that $(\mathfrak{A}, \mathfrak{B}) \models \phi$.*

The proof of this theorem is postponed to the next section, where we prove the adequacy theorem for the generalized game. For now the interested reader is referred to [10].

4   Extending the Game to $L_{\omega_1\omega}$

The game EFB_n sets up a good model in our quest. Our next step is to extend this game to $L_{\omega_1\omega}$. Again, there are two things that we need to do: a) modifying the rules of the game, and b) extending the definition of size to cover $L_{\omega_1\omega}$-formulas. We shall start from the latter.

At a first glance, the most straightforward way of extending Definition 3.1 to infinitary formulas is to leave the clauses for atomic formulas, quantifiers, finite conjunctions and disjunctions unchanged, and define the size of a countable conjunction or disjunction as the supremum of the size of its proper subformulas. However this approach leads to undesirable consequences. Here we encounter the interesting phenomenon that the arithmetic of infinite ordinals behave differently from the finite case. For one thing, addition of infinite ordinals is not commutative. Consider for example a formula ϕ with $s(\phi) = \omega$ and a formula ψ with $s(\psi) = 1$. We would have that $s(\phi \wedge \psi) = \omega + 1$, which is not equal to $s(\psi \wedge \phi) = 1 + \omega = \omega$. This is counterintuitive. Moreover, in this example the formula $\psi \wedge \phi$ has the same size as its proper subformula ϕ, a fact that is fatal to the induction proof of the adequacy theorem. We need a new way of adding up ordinals.

The operation that we turn to is the *natural sum*. The natural sum is defined using the Cantor normal form of ordinals.

Definition 4.1 Let γ_1 and γ_2 be ordinals. One can represent γ_1 and γ_2 uniquely in the form

$$\gamma_1 = \omega^{\alpha_1} \cdot k_1 + \ldots + \omega^{\alpha_n} \cdot k_n$$

$$\gamma_2 = \omega^{\alpha_1} \cdot j_1 + \ldots + \omega^{\alpha_n} \cdot j_n$$

where $\alpha_1 > \ldots > \alpha_n$ is a sequence of ordinals, $k_1, \ldots, k_n$ and $j_1, \ldots, j_n$ are natural numbers satisfying $k_i + j_i > 0$ for all i. Define the *natural sum* of γ_1 and γ_2, denoted by $\gamma_1 \# \gamma_2$, of γ_1 and γ_2 as

$$\gamma_1 \# \gamma_2 = \omega^{\alpha_1} \cdot (k_1 + j_1) + \ldots + \omega^{\alpha_n} \cdot (k_n + j_n).$$

The natural sum is commutative and associative. We also introduce the natural sum for countable sequences of ordinals.

Definition 4.2 Let $\{\gamma_i | i \in \omega\}$ be a sequence of ordinals. Let S_n denote the natural sum of the first n items in the sequence:

$$S_n = \gamma_0 \# \ldots \# \gamma_{n-1}.$$

Define the *infinite natural sum*, denoted by $\sum_{i\in\omega}^{\#} \gamma_i$, of the sequence $\{\gamma_i | i \in \omega\}$ as

$$\sum_{i\in\omega}^{\#} \gamma_i = \sup_{n\in\omega} S_n.$$

We are now ready to define the notion of size for $L_{\omega_1\omega}$-formulas. Again assume all formulas to be in negation normal form.

Definition 4.3 The *size*, denoted by $\mathrm{s}(\phi)$, of a formula ϕ in $L_{\omega_1\omega}$ is defined inductively as follows:

$$\mathrm{s}(\phi) = 1 \text{ if } \phi \text{ is an atomic or negated atomic formula,}$$
$$\mathrm{s}(\phi \wedge \psi) = \mathrm{s}(\phi)\#\mathrm{s}(\psi),$$
$$\mathrm{s}(\phi \vee \psi) = \mathrm{s}(\phi)\#\mathrm{s}(\psi),$$
$$\mathrm{s}(\exists x\phi) = \mathrm{s}(\phi) + 1,$$
$$\mathrm{s}(\forall x\phi) = \mathrm{s}(\phi) + 1,$$
$$\mathrm{s}(\bigwedge_{i\in\omega} \phi_i) = \sum_{i\in\omega}^{\#} \mathrm{s}(\phi_i),$$
$$\mathrm{s}(\bigvee_{i\in\omega} \phi_i) = \sum_{i\in\omega}^{\#} \mathrm{s}(\phi_i).$$

If I is any countable set and for each $i \in I$, ϕ_i is a formula, we define $\mathrm{s}(\bigwedge_{i\in I} \phi_i)$ to be $\mathrm{s}(\bigwedge_{i\in\omega} \phi_{p(i)})$, where p is a bijection from ω to I, and similarly for disjunction. In the next proposition we show that this definition does not depend on the choice of p. The size of an $L_{\omega_1\omega}$-formula is always a countable ordinal. The following are straightforward properties of size.

Proposition 4.4 1. *If ϕ is a first order formula, then Definition 3.1 and Definition 4.3 assign it the same size.*

2. *If ϕ is a proper subformula of ψ, then $\mathrm{s}(\phi) < \mathrm{s}(\psi)$.*

3. *The size of a conjunction is invariant under permutations of the conjuncts. More precisely, if ϕ and ψ are formulas then $\mathrm{s}(\phi \wedge \psi) = \mathrm{s}(\psi \wedge \phi)$. If $\phi_i, i \in \omega$ is a sequence of formulas and p is a permutation of ω, then $\mathrm{s}(\bigwedge_{i\in\omega} \phi_i) = \mathrm{s}(\bigwedge_{i\in\omega} \phi_{p(i)})$. The same applies to disjunctions.*

P r o o f. 1. When restricted to finite ordinals, the natural sum coincides with the usual ordinal sum.

2. Every formula has a non-zero size. If γ_1 and γ_2 are non-zero ordinals, we always have $\gamma_1 < \gamma_1 \# \gamma_2$ and $\gamma_2 < \gamma_1 \# \gamma_2$. Hence the size of a conjunction is always greater than both of the conjuncts. The same applies to disjunctions. The quantifier case is obvious.

3. The natural sum of ordinals is commutative. For the infinitary case, it suffices to observe that every finite sum S_n of the sequence $\{\mathrm{s}(\phi_i) | i \in \omega\}$ is subsumed by a finite sum of the sequence $\{\mathrm{s}(\phi_{p(i)}) | i \in \omega\}$.   □

Now we move on to the other half of our plan, namely adjusting the rules of the game EFB_n for $L_{\omega_1\omega}$. We add, in addition to the four types of moves in EFB_n, two new types of moves: the infinite left splitting move and the infinite right splitting move. These two types of moves are supposed to model infinite conjunctions and disjunctions.

Definition 4.5 Let L be a relational vocabulary and $\mathfrak{A}_0$, $\mathfrak{B}_0$ classes of L-structures. Let α be a countable ordinal. The game $\mathrm{EFB}_\alpha(\mathfrak{A}_0, \mathfrak{B}_0)$ has two players. The number α is called the *rank* of the game. The positions in the game are of the form $(\mathfrak{A}, \mathfrak{B}, \gamma)$ where $\mathfrak{A}, \mathfrak{B}$ are classes of L-structures and γ is an ordinal. The game begins from position $(\mathfrak{A}_0, \mathfrak{B}_0, \alpha)$. Suppose the game is in position $(\mathfrak{A}, \mathfrak{B}, \gamma)$. There are the following possibilities for the next move in the game.

Finite left splitting move:: Player **I** first represents $\mathfrak{A}$ as a union $\mathfrak{A}_1 \cup \mathfrak{A}_2$. He also chooses non-zero ordinals γ_1 and γ_2 such that $\gamma_1 \# \gamma_2 \leq \gamma$. Now the game continues from the position $(\mathfrak{A}_1, \mathfrak{B}, \gamma_1)$ or from the position $(\mathfrak{A}_2, \mathfrak{B}, \gamma_2)$, and player **II** can choose which.

Finite right splitting move:: Player **I** first represents $\mathfrak{B}$ as a union $\mathfrak{B}_1 \cup \mathfrak{B}_2$. He also chooses non-zero ordinals γ_1 and γ_2 such that $\gamma_1 \# \gamma_2 \leq \gamma$. Now the game continues from the position $(\mathfrak{A}, \mathfrak{B}_1, \gamma_1)$ or from the position $(\mathfrak{A}, \mathfrak{B}_2, \gamma_2)$, and player **II** can choose which.

Infinite left splitting move:: Player **I** first represents $\mathfrak{A}$ as a countable union $\bigcup_{i \in \omega} \mathfrak{A}_i$. He also chooses a sequence of non-zero ordinals $\{\gamma_i | i \in \omega\}$ such that $\sum_{i \in \omega}^{\#} \gamma_i \leq \gamma$. Now the game continues from the position $(\mathfrak{A}_i, \mathfrak{B}, \gamma_i)$ for some $i \in \omega$, and player **II** can choose which.

Infinite right splitting move:: Player **I** first represents $\mathfrak{B}$ as a countable union $\bigcup_{i \in \omega} \mathfrak{B}_i$. He also chooses a sequence of non-zero ordinals $\{\gamma_i | i \in \omega\}$ such that $\sum_{i \in \omega}^{\#} \gamma_i \leq \gamma$. Now the game continues from the position $(\mathfrak{A}, \mathfrak{B}_i, \gamma_i)$ for some $i \in \omega$, and player **II** can choose which.

Left supplementing move:: Player **I** picks an element from each structure $(A, \alpha) \in \mathfrak{A}$. More precisely, **I** chooses a natural number j and a choice function F for $\mathfrak{A}$. He also chooses an ordinal $\delta < \gamma$. Then the game continues from the position $(\mathfrak{A}(F/j), \mathfrak{B}(\star/j), \delta)$.

Right supplementing move:: Player **I** picks an element from each structure $(B, \beta) \in \mathfrak{B}$. More precisely, **I** chooses a natural number j and a choice function F for $\mathfrak{B}_m$. He also chooses ordinal $\delta < \gamma$. Then the game continues from the position $(\mathfrak{A}(\star/j), \mathfrak{B}(F/j), \delta)$.

The game ends in a position $(\mathfrak{A}, \mathfrak{B}, \gamma)$ and player **I** wins if there is an atomic or a negated atomic formula ϕ such that $(\mathfrak{A}, \mathfrak{B}) \models \phi$. Player **II** wins the game if they reach a position $(\mathfrak{A}, \mathfrak{B}, \gamma)$ such that $\gamma = 1$ and **I** does not win in this position.

The game EFB_α characterizes separation in $L_{\omega_1\omega}$ up to size α. The following theorem is the central result in this paper.

Theorem 4.6 (Adequacy Theorem for EFB_α) *Let L be a relational vocabulary,[4] $\mathfrak{A}$ and $\mathfrak{B}$ classes of L-structures, and let α be a countable ordinal. Then the following are equivalent:*

1. *Player **I** has a winning strategy in the game $\mathrm{EFB}_\alpha(\mathfrak{A}, \mathfrak{B})$.*

2. *There is an L-formula ϕ in $L_{\omega_1\omega}$ of size $\leq \alpha$ such that $(\mathfrak{A}, \mathfrak{B}) \models \phi$.*

P r o o f. We prove this theorem by induction on the rank α. When $\alpha = 1$ the proposition is obvious. Suppose the proposition is true for all ordinals $\alpha < \gamma$. Now consider the case $\alpha = \gamma$.

($1 \Longrightarrow 2$) Suppose **I** has a winning strategy in the game $\mathrm{EFB}_\gamma(\mathfrak{A}, \mathfrak{B})$. We look at the first move in this strategy. Depending on which type of move it is, there are the following possibilities.

Case 1. Player **I** first makes a finite left splitting move. He represents $\mathfrak{A}$ as a union $\mathfrak{A}_1 \cup \mathfrak{A}_2$. He chooses non-zero ordinals γ_1 and γ_2 such that $\gamma_1 \# \gamma_2 \leq \gamma$. Now since this is a winning strategy, both $(\mathfrak{A}_1, \mathfrak{B}, \gamma_1)$ and $(\mathfrak{A}_2, \mathfrak{B}, \gamma_2)$ are winning positions for **I**. Note that both γ_1 and γ_2 are strictly smaller than γ. By the induction hypothesis, there are L-formulas ψ and θ in $L_{\omega_1\omega}$ such that $\mathrm{s}(\psi) \leq \gamma_1$, $\mathrm{s}(\theta) \leq \gamma_2$, $(\mathfrak{A}_1, \mathfrak{B}) \models \psi$ and $(\mathfrak{A}_2, \mathfrak{B}) \models \theta$. Let ϕ be the formula $\psi \vee \theta$. Then we have that $\mathrm{s}(\phi) = \mathrm{s}(\psi) \# \mathrm{s}(\theta) \leq \gamma_1 \# \gamma_2 \leq \gamma$, and that $(\mathfrak{A}, \mathfrak{B}) \models \phi$.

Case 2. Player **I** first makes a finite right splitting move. He represents $\mathfrak{B}$ as a union $\mathfrak{B}_1 \cup \mathfrak{B}_2$. He chooses non-zero ordinals γ_1 and γ_2 such that $\gamma_1 \# \gamma_2 \leq \gamma$. Now since this is a winning strategy, both $(\mathfrak{A}, \mathfrak{B}_1, \gamma_1)$ and

---

[4] Note that L is not necessarily countable.

$(\mathfrak{A}, \mathfrak{B}_2, \gamma_2)$ are winning positions for **I**. Note that both γ_1 and γ_2 are strictly smaller than γ. By the induction hypothesis, there are formulas ψ and θ such that $s(\psi) \leq \gamma_1$, $s(\theta) \leq \gamma_2$, $(\mathfrak{A}, \mathfrak{B}_1) \models \psi$ and $(\mathfrak{A}, \mathfrak{B}_2) \models \theta$. Let ϕ be the formula $\psi \wedge \theta$. Then we have that $s(\phi) = s(\psi) \# s(\theta) \leq \gamma_1 \# \gamma_2 \leq \gamma$, and that $(\mathfrak{A}, \mathfrak{B}) \models \phi$.

Case 3. Player **I** first makes an infinite left splitting move. He represents $\mathfrak{A}$ as a union $\bigcup_{i \in \omega} \mathfrak{A}_i$. He chooses ordinals $\{\gamma_i | i \in \omega\}$ such that $\sum_{i \in \omega}^{\#} \gamma_i \leq \gamma$. Note that every γ_i is strictly smaller than γ. Now since this is a winning strategy, for every $i \in \omega$ $(\mathfrak{A}_i, \mathfrak{B}, \gamma_i)$ is a winning position for him. By the induction hypothesis, there are L-formulas ϕ_i in $L_{\omega_1\omega}$ such that $s(\phi_i) \leq \gamma_i$ and $(\mathfrak{A}_i, \mathfrak{B}) \models \phi_i$ for all $i \in \omega$. Let ϕ be the formula $\bigvee_{i \in \omega} \phi_i$. It is an $L_{\omega_1\omega}$-formula. We have that $s(\phi) = \sum_{i \in \omega}^{\#} s(\phi_i) \leq \sum_{i \in \omega}^{\#} \gamma_i \leq \gamma$, and that $(\mathfrak{A}, \mathfrak{B}) \models \phi$.

Case 4. Player **I** first makes an infinite right splitting move. He represents $\mathfrak{B}$ as a union $\bigcup_{i \in \omega} \mathfrak{B}_i$. He chooses ordinals $\{\gamma_i | i \in \omega\}$ such that $\sum_{i \in \omega}^{\#} \gamma_i \leq \gamma$. Note that every γ_i is strictly smaller than γ. Now since this is a winning strategy, for every $i \in \omega$ $(\mathfrak{A}, \mathfrak{B}_i, \gamma_i)$ is a winning position for him. By the induction hypothesis, there are L-formulas ϕ_i in $L_{\omega_1\omega}$ such that $s(\phi_i) \leq \gamma_i$ and $(\mathfrak{A}, \mathfrak{B}_i) \models \phi_i$ for all $i \in \omega$. Let ϕ be the formula $\bigwedge_{i \in \omega} \phi_i$. It is an $L_{\omega_1\omega}$-formula. We have that $s(\phi) = \sum_{i \in \omega}^{\#} s(\phi_i) \leq \sum_{i \in \omega}^{\#} \gamma_i \leq \gamma$, and that $(\mathfrak{A}, \mathfrak{B}) \models \phi$.

Case 5. Player **I** first makes a left supplementing move. He chooses a natural number j and a choice function F for $\mathfrak{A}$. He also chooses an ordinal $\delta < \gamma$. The game continues from $(\mathfrak{A}(F/j), \mathfrak{B}(\star/j), \delta)$. Since this is a winning strategy for **I**, the position $(\mathfrak{A}(F/j), \mathfrak{B}(\star/j), \delta)$ is a winning position for him too. By the induction hypothesis, there is a formula ψ such that $s(\psi) \leq \delta$, and $(\mathfrak{A}(F/j), \mathfrak{B}(\star/j)) \models \psi$. Let ϕ be the formula $\exists x_j \psi$. Then we have that $s(\phi) = s(\psi) + 1 \leq \delta + 1 \leq \gamma$, and that $(\mathfrak{A}, \mathfrak{B}) \models \phi$.

Case 6. Player **I** first makes a right supplementing move. He chooses a natural number j and a choice function F for $\mathfrak{B}$. He also chooses an ordinal $\delta < \gamma$. The game continues from $(\mathfrak{A}(\star/j), \mathfrak{B}(F/j), \delta)$. Since this is a winning strategy for **I**, the position $(\mathfrak{A}(\star/j), \mathfrak{B}(F/j), \delta)$ is a winning position for him too. By the induction hypothesis, there is a formula ψ such that $s(\psi) \leq \delta$, and $(\mathfrak{A}(\star/j), \mathfrak{B}(F/j)) \models \psi$. Let ϕ be the formula $\forall x_j \psi$. Then we have that $s(\phi) = s(\psi) + 1 \leq \delta + 1 \leq \gamma$, and that $(\mathfrak{A}, \mathfrak{B}) \models \phi$.

Now for the converse direction $(2 \implies 1)$. Suppose there is an $L_{\omega_1\omega}$-formula ϕ of size $\leq \gamma$ such that $(\mathfrak{A}, \mathfrak{B}) \models \phi$. Depending on the shape of ϕ, there are the following possibilities.

Case 1. The formulas ϕ is an atomic formula. By definition player **I** wins the game $\text{EFB}_1(\mathfrak{A}, \mathfrak{B})$.

Case 2. The formula ϕ is $\psi \vee \theta$. Let $\mathfrak{A}_1$ be the class of structures $(A, \alpha) \in \mathfrak{A}$ such that $(A, \alpha) \models \psi$, $\mathfrak{A}_2$ be the class of structures $(A, \alpha) \in \mathfrak{A}$ such that $(A, \alpha) \models \theta$. Since $(A, \alpha) \models \phi$ for all $(A, \alpha) \in \mathfrak{A}$, we have $\mathfrak{A} = \mathfrak{A}_1 \cup \mathfrak{A}_2$. Moreover since $(B, \beta) \not\models \phi$ for all $(B, \beta) \in \mathfrak{B}$, we have $(\mathfrak{A}_1, \mathfrak{B}) \models \psi$ and $(\mathfrak{A}_2, \mathfrak{B}) \models \theta$. Finally, as $s(\phi) \leq \gamma$, there are ordinals γ_1 and γ_2 such that $s(\psi) \leq \gamma_1$, $s(\theta) \leq \gamma_2$, and $\gamma_1 \# \gamma_2 \leq \gamma$. By the induction hypothesis, **I** has a winning strategy in both $(\mathfrak{A}_1, \mathfrak{B}, \gamma_1)$ and $(\mathfrak{A}_2, \mathfrak{B}, \gamma_2)$. Therefore he has a winning strategy in $(\mathfrak{A}, \mathfrak{B}, \gamma)$ by first making a finite left splitting move, and then follow the winning strategy in $(\mathfrak{A}_1, \mathfrak{B}, \gamma_1)$ or $(\mathfrak{A}_2, \mathfrak{B}, \gamma_2)$.

Case 3. The formula ϕ is $\psi \wedge \theta$. This case is dual to case 2 and we omit the details.

Case 4. The formula ϕ is $\bigvee_{i \in \omega} \phi_i$. For each i, let $\mathfrak{A}_i$ be the class of structures $(A, \alpha) \in \mathfrak{A}$ such that $(A, \alpha) \models \phi_i$. Since $(A, \alpha) \models \phi$ for all $(A, \alpha) \in \mathfrak{A}$, we have $\mathfrak{A} = \bigcup_{i \in \omega} \mathfrak{A}_i$. Moreover since $(B, \beta) \not\models \phi$ for all $(B, \beta) \in \mathfrak{B}$, we have $(\mathfrak{A}_i, \mathfrak{B}) \models \phi_i$ for every i. Finally, as $s(\phi) \leq \gamma$, there are ordinals $\{\gamma_i | i \in \omega\}$ such that $s(\phi_i) \leq \gamma_i$ and $\sum_{i \in \omega}^{\#} \gamma_i \leq \gamma$. By the induction hypothesis, **I** has a winning strategy in $(\mathfrak{A}_i, \mathfrak{B}, \gamma_i)$ for every i. Therefore he has a winning strategy in $(\mathfrak{A}, \mathfrak{B}, \gamma)$ by first making an infinite left splitting move, and then follow the winning strategy in some $(\mathfrak{A}_i, \mathfrak{B}, \gamma_i)$.

Case 5. The formula ϕ is $\bigwedge_{i \in \omega} \phi_i$. This case is dual to case 4 and we omit the details.

Case 6. The formula ϕ is $\exists x_j \psi$. Since $(A, \alpha) \models \phi$ for all $(A, \alpha) \in \mathfrak{A}$, there is a choice function F for $\mathfrak{A}$ such that $(A, \alpha(F(A, \alpha)/j)) \models \psi$. Thus $(A, \alpha^*) \models \psi$ for all $(A, \alpha^*) \in \mathfrak{A}(F/j)$. On the other hand, we have that $(\mathfrak{B}, \beta^*) \not\models \psi$ for all $(\mathfrak{B}, \beta^*) \in \mathfrak{B}(\star/j)$. Therefore $(\mathfrak{A}(F/j), \mathfrak{B}(\star/j)) \models \psi$. By the induction hypothesis, player **I** has a winning strategy in $(\mathfrak{A}(F/j), \mathfrak{B}(\star/j), s(\psi))$. Note that $s(\phi) = s(\psi) + 1$. Hence **I** has a winning strategy in $(\mathfrak{A}, \mathfrak{B}, \gamma)$ by first making a left supplementing move, chooses the ordinal $s(\psi) < s(\phi) \leq \gamma$ and then follow the winning strategy in $(\mathfrak{A}(F/j), \mathfrak{B}(\star/j), s(\psi))$.

Case 7. The formula ϕ is $\forall x_j \psi$. This case is dual to case 6 and we omit the details. $\square$

Corollary 4.7 *Given two classes of L-structures $\mathfrak{A}$ and $\mathfrak{A}'$, let $\mathfrak{A} \subseteq \mathfrak{A}'$ denote that every structure in $\mathfrak{A}$ is contained in $\mathfrak{A}'$. Let $\mathfrak{A}, \mathfrak{A}', \mathfrak{B}, \mathfrak{B}'$ be classes of L-structures and α be a countable ordinal. Suppose $\mathfrak{A}' \subseteq \mathfrak{A}$ and $\mathfrak{B}' \subseteq \mathfrak{B}$. If player **I** has a winning strategy in $\text{EFB}_\alpha(\mathfrak{A}, \mathfrak{B})$, then he also has a winning strategy in $\text{EFB}_\alpha(\mathfrak{A}', \mathfrak{B}')$.*

P r o o f. Suppose **I** has a winning strategy in $\text{EFB}_\alpha(\mathfrak{A}, \mathfrak{B})$. By Theorem 4.6, there is a formula ϕ with $\text{s}(\phi) \leq \alpha$ such that $(\mathfrak{A}, \mathfrak{B}) \models \phi$. Since $\mathfrak{A}' \subseteq \mathfrak{A}$ and $\mathfrak{B}' \subseteq \mathfrak{B}$, we also have that $(\mathfrak{A}', \mathfrak{B}') \models \phi$. By Theorem 4.6 again, **I** has a winning strategy in $\text{EFB}_\alpha(\mathfrak{A}', \mathfrak{B}')$. $\qquad\square$

The six types of moves in the game EFB_α model the six formula-forming operations in $L_{\omega_1\omega}$ respectively. The finite left splitting move corresponds to finite disjunction, the finite right splitting move to finite conjunction, the infinite left splitting move to infinite disjunction, the infinite right splitting move to infinite conjunction, the left supplementing move to existential quantification, and the right supplementing move to universal quantification. This feature enables us to consider various fragments of the game EFB_α which characterize separation in different fragments of $L_{\omega_1\omega}$. If we forbid infinite splitting moves, what we get is essentially the old game EFB_n. If we only allow splitting moves in the game, the result is a game for the propositional fragment of $L_{\omega_1\omega}$.

Corollary 4.8 *Let* $\text{EFB}_\alpha^P(\mathfrak{A}, \mathfrak{B})$ *be the restriction of the game* $\text{EFB}_\alpha(\mathfrak{A}, \mathfrak{B})$ *where player **I** can only make finite and infinite splitting moves, but no supplementing moves. Then the following are equivalent.*

1. *Player **I** has a winning strategy in the game* $\text{EFB}_\alpha^P(\mathfrak{A}, \mathfrak{B})$.

2. *There is a quantifier-free* $L_{\omega_1\omega}$-*formula* ϕ *of size* $\leq \alpha$ *such that* $(\mathfrak{A}, \mathfrak{B}) \models \phi$.

5   Alternative Conceptions of Size

Let us take some time to reflect on our result. We have shown that measuring the complexity of formulas by size is *a* way to define an Ehrenfeucht-Fraïssé game for $L_{\omega_1\omega}$. However, our viewpoint naturally suggests the possibility of a plurality of complexity measures, and hence a plurality of Ehrenfeucht-Fraïssé games for $L_{\omega_1\omega}$. Can we find other complexity measures which give rise to an Ehrenfeucht-Fraïssé game for $L_{\omega_1\omega}$? Furthermore, to which extent are these complexity measures *reasonable*?

We shall start by looking at the bare minimum, namely which complexity measures could give rise to an Ehrenfeucht-Fraïssé game. It turns out that if we merely want an adequacy theorem to hold, we can be quite relaxed about the choice of complexity measures.

Definition 5.1 Let $q : \text{Ord} \to \text{Ord}$, $r : \text{Ord}^2 \to \text{Ord}$ and $t : \text{Ord}^\omega \to \text{Ord}$ be ordinal functions. Consider an abstract complexity measure c defined by q, r and t:

$$c(\phi) = 1 \text{ if } \phi \text{ is an atomic or negated atomic formula,}$$
$$c(\phi \wedge \psi) = r((c(\phi), c(\psi)),$$
$$c(\phi \vee \psi) = r((c(\phi), c(\psi)),$$
$$c(\exists x\phi) = q(c(\phi)),$$
$$c(\forall x\phi) = q(c(\phi)),$$
$$c(\bigwedge_{i\in\omega} \phi_i) = t(\{c(\phi_i)|i \in \omega\}),$$
$$c(\bigvee_{i\in\omega} \phi_i) = t(\{c(\phi_i)|i \in \omega\}).$$

We say that c is *nice* if q, r and t satisfy the following conditions:

1. If $\gamma > 0$, then $q(\gamma) > \gamma$.

2. If $\gamma_1, \gamma_2 > 0$, then $r(\gamma_1, \gamma_2) > \gamma_1$ and $r(\gamma_1, \gamma_2) > \gamma_2$.

3. If $\{\gamma_i|i \in \omega\}$ is a sequence of non-zero ordinals, then $t(\{\gamma_i|i \in \omega\}) > \gamma_i$ for all $i \in \omega$.

Niceness guarantees that the complexity of a formula is always strictly larger than the complexity of its proper subformulas. Granted this, we can generalize the game EFB_α to any nice complexity measure by performing minor cosmetic surgery. Let EFB_α^c denote the Ehrenfeucht-Fraïssé game for $L_{\omega_1\omega}$ under the complexity measure c. The starting position and the winning condition in the game EFB_α^c remain the same as in EFB_α. The way that the rank of the game is counted is modified. We only give the 'left' moves in the game EFB_α^c here; the 'right' moves are dual.

Finite left splitting move:: Player **I** first represents $\mathfrak{A}$ as a union $\mathfrak{A}_1 \cup \mathfrak{A}_2$. He also chooses non-zero ordinals γ_1 and γ_2 such that $r(\gamma_1, \gamma_2) \leq \gamma$. Now the game continues from the position $(\mathfrak{A}_1, \mathfrak{B}, \gamma_1)$ or from the position $(\mathfrak{A}_2, \mathfrak{B}, \gamma_2)$, and player **II** can choose which.

Infinite left splitting move:: Player **I** first represents $\mathfrak{A}$ as a countable union $\bigcup_{i \in \omega} \mathfrak{A}_i$. He also chooses a sequence of non-zero ordinals $\{\gamma_i | i \in \omega\}$ such that $t(\{\gamma_i | i \in \omega\}) \leq \gamma$. Now the game continues from the position $(\mathfrak{A}_i, \mathfrak{B}, \gamma_i)$ for some $i \in \omega$, and player **II** can choose which.

Left supplementing move:: Player **I** picks an element from each structure $(A, \alpha) \in \mathfrak{A}$. More precisely, **I** chooses a natural number j and a choice function F for $\mathfrak{A}$. He also chooses an ordinal δ such that $q(\delta) \leq \gamma$. Then the game continues from the position $(\mathfrak{A}(F/j), \mathfrak{B}(\star/j), \delta)$.

It is an easy observation that the proof of the adequacy theorem goes through if we substitute size for any nice complexity measure. We have the following generalized adequacy theorem.

Theorem 5.2 *Let c be a nice complexity measure. Let L be a relational vocabulary, $\mathfrak{A}$ and $\mathfrak{B}$ classes of L-structures, and let α be an ordinal. Then the following are equivalent:*

1. *Player **I** has a winning strategy in the game $\mathrm{EFB}_\alpha^c(\mathfrak{A}, \mathfrak{B})$.*

2. *There is an L-formula ϕ in $L_{\omega_1\omega}$ with $c(\phi) \leq \alpha$ such that $(\mathfrak{A}, \mathfrak{B}) \models \phi$.*

In particular, size is a nice complexity measure: take q to be the successor function, r the natural sum, and t the infinite natural sum. Another interesting nice complexity measure c_1 can be obtained by taking q to be the successor function, $r(\gamma_1, \gamma_2) = \sup(\gamma_1 + 1, \gamma_2 + 1)$ and $t(\{\gamma_i | i \in \omega\}) = \sup_{i \in \omega}(\gamma_i + 1)$. Note that if we take q to be the successor function and r, t to be the supremum, then the complexity measure is just quantifier rank—but this is not a nice complexity measure.

To determine which of the nice complexity measures are the more natural and reasonable ones is a subtler question. As a general rule of thumb, we would like a complexity measure to carry some meaning, so that we have an idea of what exactly it is measuring. We also want a complexity measure to be invariant under certain logical equivalences, so that it is compatible with our intuition about 'complexity'. Between size and c_1, size has the advantage of having a clear intuition behind it. This is especially so in the propositional case: the size of a propositional formula is just the number of propositional symbols in this formula. The quantifier and infinitary cases are natural extensions of this basic intuition. For c_1 the intuition is not so clear. As for invariance under logical equivalences, both size and c_1 are invariant under permutations of conjuncts and disjuncts. They part ways, however, when we try to count the complexity of repetitions of formulas. Consider atomic formulas $\phi_i, i \in \omega$. Let Θ be the infinite conjunction $\bigwedge_{i \in \omega} \phi_0$. Let Ψ be the infinite conjunction $\bigwedge_{i \in \omega} \phi_i$. Size assigns Θ and Ψ the complexity ω. The measure c_1 assigns Θ and Ψ the complexity 2. It seems that c_1 gives the more reasonable complexity for Θ, since Θ is no more than a repetition of the same formula ϕ_0. On the other hand size seems to give the more reasonable complexity for Ψ, while c_1 does not reflect the fact that Ψ contains much more information than any finite conjunction. However as long as complexity measures are defined purely syntactically, every complexity measure would assign the same complexity to Θ and Ψ, so there is no complexity measure that assigns the 'right' complexity to Θ and Ψ at the same time. This suggests that there is no point in trying to find *the* reasonable complexity measure. We may indeed adopt a pluralistic view: we shall choose which is the appropriate complexity measure depending on what information we intend to model.

6 An Application: the Complexity of Infinite Binary String Properties

In this section we apply the newly defined game EFB_α to solve one instance of Question 1. We focus on information code by *infinite binary strings*. An infinitary binary string t is a function from ω to 2. We can think of a string t simply as a sequence of 0's and 1's. A property $\mathcal{P}$ of infinite binary strings is a subset $\mathcal{P} \subset 2^\omega$.

Infinite binary strings can be studied using the propositional fragment of $L_{\omega_1\omega}$. Let L be the propositional language with countably many propositional symbols $p_i, i \in \omega$. For an infinitary binary string t and a propositional symbol p_i, we say that

$$t \models p_i \text{ if } t(i) = 1,$$

and

$$t \models \neg p_i \text{ if } t(i) = 0.$$

Intuitively, p_i has the intended meaning that 'the i'th position of t is 1', and $\neg p_i$ has the intended meaning that 'the i'th position of t is 0'. We pose the question:

Question 2: Given a string property $\mathcal{P}$, what is the minimal size of a propositional $L_{\omega_1\omega}$-formula defining $\mathcal{P}$?

Note that an infinite binary string t codes a real number. Hence this is essentially a question about the definability of sets of real numbers.

There is a straightforward way to define a string property $\mathcal{P}$. An infinite binary string t can be characterized by the following sentence in $L_{\omega_1\omega}$:

$$\theta_t = \bigwedge_{i \in \omega} p_i^t,$$

where

$$p_i^t = \begin{cases} p_i & \text{if } t(i) = 1, \\ \neg p_i & \text{if } t(i) = 0. \end{cases}$$

A property $\mathcal{P}$ of infinite binary strings can be defined by the following sentence:

$$\Phi_{\mathcal{P}} = \bigvee_{t \in \mathcal{P}} \theta_t.$$

In general $\Phi_{\mathcal{P}}$ is an $L_{\infty\omega}$-sentence. When the cardinality of $\mathcal{P}$ is countable, this sentence is in $L_{\omega_1\omega}$. In this case the sentence $\Phi_{\mathcal{P}}$ has size

$$s(\Phi_{\mathcal{P}}) = \sum_{t \in \mathcal{P}}^{\#} s(\theta_t) = \sum_{i \in \omega}^{\#} \omega = \omega^2.$$

It is easy to see that if the complement of $\mathcal{P}$ in 2^ω is countable, $\mathcal{P}$ can also be defined by an $L_{\omega_1\omega}$-sentence of size ω^2. The question is whether the sentence $\Phi_{\mathcal{P}}$ is minimal.

Regarding each infinite binary string as a 'model', we can play the game EFB_α^P on classes of infinite binary strings. This game solves Question 2 for an interesting class of string properties.

Definition 6.1 A property $\mathcal{P}$ of infinite binary strings is said to be *dense* if it is dense in the space[5] 2^ω with the product topology.

Theorem 6.2 *Let $\mathcal{P}_1$ and $\mathcal{P}_2$ be dense properties of infinite binary strings. Player **II** has a winning strategy in $\text{EFB}_\alpha^P(\mathcal{P}_1, \mathcal{P}_2)$ for all $\alpha < \omega^2$.*

Together with Corollary 4.8, this leads to the following theorem.

Theorem 6.3 *If $\mathcal{P}_1$ and $\mathcal{P}_2$ are dense properties of infinitary binary strings and ϕ is a propositional $L_{\omega_1\omega}$-formula separating $\mathcal{P}_1$ and $\mathcal{P}_2$, then the size of ϕ is at least ω^2.*

Corollary 6.4 *If $\mathcal{P}$ is a string property such that both $\mathcal{P}$ and its complement are dense, then $\mathcal{P}$ cannot be defined by a propositional $L_{\omega_1\omega}$-formula of size less than ω^2.*

If a string property $\mathcal{P}$ is countable, then by Cantor diagonal argument its complement is dense. For a countable dense property $\mathcal{P}$ the standard formula $\Phi_{\mathcal{P}}$ is a minimal formula defining $\mathcal{P}$.

In the remaining of this section we give a proof of Theorem 6.2. Let us first fix some notation. Let $2^{<\omega}$ denote the set of partial functions $\omega \to 2$ with a finite domain. We call a function $g \in 2^{<\omega}$ a *finite segment*. For an infinite binary string $h \in 2^\omega$ and a finite segment $g \in 2^{<\omega}$, we say that h *agrees with* g, or that h *extends* g, if $g \subset h$. A string property $\mathcal{P}$ is dense if and only if for any finite segment $g \in 2^{<\omega}$, there exists $h \in \mathcal{P}$ that extends g.

---

[5] Called the Cantor space.

In the game EFB_α^P the players are not allowed to play supplementing moves, but can only play splitting moves. When player **I** plays an infinite splitting move, he splits a class of structures $\mathfrak{A}$ (or $\mathfrak{B}$) into countably many pieces $\mathfrak{A}_i, i \in \omega$ (or $\mathfrak{B}_i, i \in \omega$). We say that an infinite splitting move is a *proper infinite splitting move* if there are infinitely many pieces $\mathfrak{A}_i$ (or $\mathfrak{B}_i$) that are non-empty. If player **I** splits $\mathfrak{A}$ (or $\mathfrak{B}$) into infinitely many pieces among which only finitely many are non-empty, we call it a *degenerate infinite splitting move*. Strange as it may seem, there is nothing in the rules that forbids **I** playing in this way. A degenerate infinite splitting move is closer in nature to a finite series of finite splitting moves. Henceforth we refer to degenerate infinite splitting move and finite series of finite splitting moves as *generalized finite splitting moves*.

Let α be an ordinal smaller than ω^2. We describe the following strategy for player **II** in the game $\mathrm{EFB}_\alpha^P(\mathcal{P}_1, \mathcal{P}_2)$.

Player II's strategy: Let $(\mathfrak{A}, \mathfrak{B}, \gamma)$ denote a game position. In each round, player **II** makes sure that

$$\exists f \in 2^{<\omega} \Bigg(\forall h \in \mathfrak{A}(h \supset f) \wedge$$
$$\forall h \in \mathfrak{B}(h \supset f) \wedge \tag{1}$$
$$\forall g \in 2^{<\omega} \Big(g \supset f \rightarrow \big(\exists h \in \mathfrak{A}(h \supset g) \wedge \exists h' \in \mathfrak{B}(h' \supset g) \big) \Big) \Bigg).$$

She maintains this condition until she sees an opportunity to finish off the game directly.

Condition (1) says that at any stage of the game, we can always find a finite segment f such that all strings in $\mathfrak{A}$ and $\mathfrak{B}$ agree with f. Moreover, for any finite extension g of f, there are string h and h' in $\mathfrak{A}$ and $\mathfrak{B}$ respectively that agree with g. Therefore if we imagine we 'cut off' the finite segment f, the remaining configuration in the game is almost identical to the starting position. We will specify later what counts as an opportunity for **II** to win the game directly.

Theorem 6.2 is established through a series of lemmas.

Lemma 6.5 *Condition (1) is true in the starting position* $(\mathcal{P}_1, \mathcal{P}_2, \alpha)$.

P r o o f. Let f be the empty function. Then condition (1) is just the density condition. $\square$

Lemma 6.6 *If condition (1) is true in position* $(\mathfrak{A}, \mathfrak{B}, \gamma)$, *then no atomic formula or its negation separates* $\mathfrak{A}$ *and* $\mathfrak{B}$.

P r o o f. There is an atomic formula or its negation separating $\mathfrak{A}$ and $\mathfrak{B}$ if and only if there exists $g \in 2^{<\omega}$ such that $|\mathrm{dom}(g)| = 1$, $g \subset h$ for all $h \in \mathfrak{A}$, and $g \not\subset h'$ for all $h' \in \mathfrak{B}$. We claim that condition (1) guarantees that this situation does not occur. There are the following three possibilities for g, let us consider them one by one. Firstly, if $g \subset f$, then by condition (1) all strings in $\mathfrak{A}$ and $\mathfrak{B}$ agree with g. Secondly, if g disagrees with f, namely if $\mathrm{dom}(g) \subset \mathrm{dom}(f)$ yet $g \not\subset f$, then by condition (1) again all strings in $\mathfrak{A}$ and $\mathfrak{B}$ disagree with g. Finally, suppose $\mathrm{dom}(g) \cap \mathrm{dom}(f) = \emptyset$. The finite segment $g \cup f$ is an extension of f. Therefore by condition (1) there are $h \in \mathfrak{A}$ and $h' \in \mathfrak{B}$ that agree with $g \cup f$ respectively. In particular they agree with g. $\square$

Lemma 6.7 *If player **I** makes a generalized finite splitting move, player **II** can maintain condition (1) to the next round.*

P r o o f. Suppose the game is in position $(\mathfrak{A}, \mathfrak{B}, \gamma)$ and condition (1) holds. Suppose player **I** makes a generalized finite left splitting move. He represents $\mathfrak{A}$ as a union $\mathfrak{A}_1 \cup \ldots \cup \mathfrak{A}_n$ of n pieces. He also picks ordinals $\gamma_1, \ldots, \gamma_n$ such that $\gamma_1 \# \ldots \# \gamma_n \leq \gamma$. We claim that there must exist a finite segment $f' \supset f$ and a piece $\mathfrak{A}_k$ such that

$$\forall g \in 2^{<\omega}\big(g \supset f' \rightarrow \exists h \in \mathfrak{A}_k(h \supset g) \big). \tag{2}$$

Suppose this is not the case for $\mathfrak{A}_1, \ldots, \mathfrak{A}_{n-1}$. Then we have

$$\forall f' \supset f \exists g \in 2^{<\omega} \Big(g \supset f' \wedge \forall h \in 2^\omega (h \supset g \rightarrow h \notin \mathfrak{A}_i) \Big) \tag{3}$$

for $i = 1, \ldots, n-1$. Now let $f' = f$, apply condition (3) to $\mathfrak{A}_1$:

$$\exists g_1 \in 2^{<\omega}\Big(g \supset f \wedge \forall h \in 2^\omega\big(h \supset g_1 \rightarrow h \notin \mathfrak{A}_1\big)\Big).$$

This means that

$$\exists g_1 \in 2^{<\omega}\Big(g \supset f \wedge \forall h \in 2^\omega\big((h \supset g_1 \wedge h \in \mathfrak{A}) \rightarrow h \in \mathfrak{A}_2 \cup \ldots \cup \mathfrak{A}_n\big)\Big).$$

Now let $f' = g_1$, apply condition (3) to $\mathfrak{A}_2$:

$$\exists g_2 \in 2^{<\omega}\Big(g_2 \supset g_1 \wedge \forall h \in 2^\omega\big((h \supset g_2 \wedge h \in \mathfrak{A}) \rightarrow h \in \mathfrak{A}_3 \cup \ldots \cup \mathfrak{A}_n\big)\Big).$$

Iterate this process, we get:

$$\exists g_{n-1} \in 2^{<\omega}\Big(g_{n-1} \supset g_{n-2} \forall h \in 2^\omega\big((h \supset g_{n-1} \wedge h \in \mathfrak{A}) \rightarrow h \in \mathfrak{A}_n\big)\Big).$$

This means that condition (2) is true for $\mathfrak{A}_n$ when we take g_{n-1} for f'. Therefore the previous claim is true.

Now let player **II** pick this piece $\mathfrak{A}_k$ and the corresponding ordinal γ_k. What remains to be done is a cosmetic surgery on $\mathfrak{A}_k$ and $\mathfrak{B}$. We throw away the strings that do not agree with f'. Let

$$\mathfrak{A}' = \{h \in \mathfrak{A}_k | h \supset f'\}$$

and let

$$\mathfrak{B}' = \{h' \in \mathfrak{B} | h' \supset f'\}.$$

By Corollary 4.7, if player **II** has a winning strategy in this new position, she also has one in the old position. It is straightforward to check that condition (1) holds for $\mathfrak{A}'$ and $\mathfrak{B}'$. Let the game continue from the position $(\mathfrak{A}', \mathfrak{B}', \gamma_k)$.

The case where **I** plays a generalized finite right splitting move is similar. □

Lemma 6.8 *Suppose condition (1) is true in position $(\mathfrak{A}, \mathfrak{B}, \gamma)$ with $\gamma < \omega^2$. If player **I** makes a proper infinite splitting move, then player **II** has a winning strategy in the rest of the game.*

P r o o f. Suppose the game is in position $(\mathfrak{A}, \mathfrak{B}, \gamma)$ and condition (1) is true. Suppose player **I** makes a proper infinite splitting move. He represents $\mathfrak{A}$ as a union $\bigcup_{i \in \omega} \mathfrak{A}_i$, among which there are infinitely many non-empty pieces. He also chooses ordinals $\gamma_i, i \in \omega$ such that $\sum_{i \in \omega}^{\#} \gamma_i \leq \gamma < \omega^2$. Note that there can only be finitely many $i \in \omega$ such that γ_i is an infinite ordinal, for otherwise the natural sum $\sum_{i \in \omega}^{\#} \gamma_i$ would be greater or equal to ω^2, which is a contradiction. On the other hand there are infinitely many $i \in \omega$ such that $\mathfrak{A}_i$ is non-empty. Therefore there must exist $k \in \omega$ such that $\mathfrak{A}_k$ is non-empty and γ_k is finite.

If this situation occurs, we ask player **II** to jump out of the strategy of maintaining condition (1) and go on to win the game directly. Let player **II** pick this piece $\mathfrak{A}_k$ and γ_k. The game continues from the position $(\mathfrak{A}_k, \mathfrak{B}, \gamma_k)$. Apply Corollary 4.7 and suppose $\mathfrak{A}_k$ consists of a single string h_A. Note that since γ_k is a finite number, in the rest of the game **I** can only make finite right splitting moves.

Suppose **I** makes a finite right splitting move. He represents $\mathfrak{B}$ as a union $\mathfrak{B}_1 \cup \ldots \cup \mathfrak{B}_n$ of n pieces. Consider the witnessing finite segment f for condition (1). Let $k = \sup(\text{dom}(f))$. For any natural number m, let h_m be the string in $\mathfrak{B}$ such that h_m agrees with h_A on the first $k + m$ elements of ω. By condition (1) this string always exists. In other words, we have

$$\forall m \in \omega \, \exists h_m \in \mathfrak{B}\big(h_m \restriction (k+m) = h_A \restriction (k+m)\big). \tag{4}$$

Among the n pieces $\mathfrak{B}_1, \ldots, \mathfrak{B}_n$, there must be one piece $\mathfrak{B}_l$ that contains h_m for infinitely many $m \in \omega$. Let player **II** pick this piece $\mathfrak{B}_l$. She uses the same strategy if **I** makes another finite right splitting move. She makes sure that at any stage of the remaining game, the piece $\mathfrak{B}_r$ at hand always contain h_m for infinitely many $m \in \omega$. This guarantees that no atomic formula separates $\mathfrak{A}_k$ and $\mathfrak{B}$. Player **II** wins when the game terminates after finitely many rounds. □

Now we can see that the strategy prescribed above is indeed a winning strategy for player **II**. Player **II** hangs on in the game by maintaining condition (1). If her opponent plays a proper infinite splitting move at some point, she goes on to win the rest of the game by appealing to Lemma 6.8. Otherwise she continues with condition (1). This is possible because of Lemma 6.7. This strategy keeps her away from losing by Lemma 6.6. Player **II** will eventually prevail after finitely many rounds. Thus we have finished the proof of Theorem 6.2.

The following are some examples of dense properties of infinite binary strings. By Corollary 6.4 none of these properties can be defined by a propositional $L_{\omega_1\omega}$-formula of size less than ω^2.

Example 6.9 1. $\mathcal{P}_1 = \{f \in 2^\omega | f^{-1}(1) \text{ is finite}\}$ 'f has finitely many ones'.

2. $\mathcal{P}_2 = \{f \in 2^\omega | |f^{-1}(1)| \text{ is an odd number}\}$ 'f has an odd number of ones'.

3. $\mathcal{P}_3 = $ 'f codes a rational number'.

4. $\mathcal{P}_4 = $ 'f codes a decidable set'.

This result is a first step in the application of EFB$_\alpha$. The game suggests many natural questions about the expressive power of $L_{\omega_1\omega}$, for example:

Open Problems:

1. Can we use the game to prove higher complexity results for propositional $L_{\omega_1\omega}$?

2. Can we find applications of the full game instead of the propositional fragment?

3. Can we prove complexity results for other nice complexity measures?

These remain questions for further study.

Finally, we point out that new problems arise when we move from $L_{\omega_1\omega}$ to $L_{\omega_2\omega}$ and even higher infinitary languages. Suppose we want to define the notion of size for $L_{\omega_2\omega}$-formulas in such a way that it is coherent with our definition of size for $L_{\omega_1\omega}$-formulas. To illustrate the problem, consider the uncountable conjunction $\Phi = \bigwedge_{i\in\omega_1} \phi_i$, where ϕ_i is an atomic formula for all $i \in \omega_1$. What is the right size of Φ? There are the following three desiderata:

1. If θ is an $L_{\omega_1\omega}$-formula, then the size of θ is the same as it is given in Definition 4.3.

2. The size of Φ is the supremum of the size of its proper subformulas, namely $s(\Phi) = \sup_{\alpha<\omega_1} s(\bigwedge_{i\in\alpha} \phi_i)$.

3. The size of Φ is at least ω_1.

It turns out that these three conditions cannot be satisfied simultaneously. Suppose there is a complexity measure s satisfying conditions 1 and 2. Consider any infinite ordinal $\alpha < \omega_1$. By condition 1, we have that $s(\bigwedge_{i\in\alpha} \phi_i) = s(\bigwedge_{i\in\omega} \phi_{p(i)}) = \omega$, where p is a bijection from ω to α. This together with condition 2 renders the size of Φ to be ω, contrary to condition 3.

References

[1] D. Scott and A. Tarski, The sentential calculus with infinitely long expressions, Colloquium Mathematicum **6**, 165–170 (1958).
[2] A. Tarski, Remarks on predicate logic with infinitely long expressions, Colloquium Mathematicum **6**, 171–176 (1958).
[3] C. Karp, Languages with Expressions of Infinite Length (North-Holland, Amsterdam, 1964).
[4] E. G. K. Lopez-Escobar, An interpolation theorem for denumerably long formulas, Fundamenta Mathematicae **57**, 253–272 (1965).
[5] A. Ehrenfeucht, Application of games to some problems of mathematical logic, Bulletin de l'Acadámie Polonaise des Science, Série des Sciences Mathématiques, Astronomiques et Physiques Cl. III.,5 **IV**, 35–37 (1957).
[6] A. Ehrenfeucht, An application of games to the completeness problem for formalized theories, Fundamenta Mathematicae **49**, 129–141 (1960/1961).
[7] R. Fraïssé, Sur quelques classifications des relations, basées sur des isomorphismes restreints. ii. application aux relations d'ordre, et construction d'exemples montrant que ces classifications sont distinctes., Publ. Sci. Univ. Alger. Sér. A. **2**, 273–295 (1957).

[8] M. Benda, Reduced products and nonstandard logics, Journal of Symbolic Logic **3**(Sep.), 424–436 (1969).

[9] J. Väänänen, Models and Games, No. 132 in Cambridge Studies in Advanced Mathematics (Cambridge University Press, Cambridge, 2011).

[10] L. Hella and J. Väänänen, The size of a formula as a measure of complexity, `http://arxiv.org/abs/1208.4803`, 2012.

[11] J. Barwise, The role of the omitting types theorem in infinitary logic, Arch. Math. Logik Grundlag **21(1-2)**, 55–68 (1981).

[12] J. Väänänen, On the semantics of informational independence, Log. J. IGPL **10(3):339–352** (2002).