Infinitary Logics and 0-1 Laws*

Phokion G. Kolaitis †
Computer and Information Sciences
University of California, Santa Cruz
Santa Cruz, CA 95064
e-mail: kolaitis@saturn.ucsc.edu

Moshe Y. Vardi
IBM Research
Almaden Research Center
San Jose, CA 95120
e-mail: vardi@almaden.ibm.com

Abstract. We investigate the infinitary logic $L^{\omega}_{\infty\omega}$, in which sentences may have arbitrary disjunctions and conjunctions, but they involve only a finite number of distinct variables. We show that various fixpoint logics can be viewed as fragments of $L^{\omega}_{\infty\omega}$, and we describe a game-theoretic characterization of the expressive power of the logic. Finally, we study asymptotic probabilities of properties expressible in $L^{\omega}_{\infty\omega}$ on finite structures. We show that the 0-1 law holds for $L^{\omega}_{\infty\omega}$, i.e., the asymptotic probability of every sentence in this logic exists and is equal to either 0 or 1. This result subsumes earlier work on asymptotic probabilities for various fixpoint logics and reveals the boundary of 0-1 laws for infinitary logics.

*A preliminary version of this paper appeared under the title "0-1 laws for infinitary logics" in *Proc. 5th IEEE Symp. on Logic in Computer Science*, June 1990, pp. 156-167.

†During the preparation of this paper this author was partially supported by NSF Grant CCR-8905038

1 Introduction

In recent years the model theory of finite structures has been a meeting point for research in computer science, combinatorics, and mathematical logic. Results and techniques from finite model theory have found interesting applications to several other areas, including database theory [CH82], [Var82] and complexity theory [Ajt83], [Gur84], [Imm86]. One particular direction of research has focused on the asymptotic probabilities of properties expressible in different languages.

In general, if C is a class of finite structures over some vocabulary and if P is a property of some structures in C, then *the asymptotic probability* $\mu(P)$ on C is the limit as $n \to \infty$ of the fraction of the structures in C with n elements which satisfy P, provided that the limit exists. We say that P *is true almost everywhere* on C in case $\mu(P)$ is equal to 1. If $\mu(P) = 0$, then we say that P *is false almost everywhere*. It turns out that many interesting properties on the class G of all finite graphs are either true almost everywhere or false almost everywhere. It is, for example, well known and easy to prove that μ(connectivity)=1, μ(rigidity)=1, while μ(planarity)=0 and μ(l-colorabilty)=0, for $l \geq 2$ [Bol79]. A theorem of Pósa [P$\acute{7}$6] asserts that μ(Hamiltonicity)=1. On the other hand, statements about *cardinalities*, such as "there is an even number of elements" do not have an asymptotic probability.

Fagin [Fag76] and Glebskii et al. [GKLT69] were the first to establish a fascinating connection between logical definability and asymptotic probabilities. More specifically, they showed that if C is the class of all finite structures over some relational vocabulary and if P is any property expressible in first-order logic, then $\mu(P)$ exists and is either 0 or 1. This result, which is known as *the 0-1 law for first-order logic*, became the starting point of a series of investigations aiming in discovering the relationship between expressibility in a logic and asymptotic probabilities. The recent survey by Compton [Com88a] contains an eloquent account of developments in this area.

It is well known that first-order logic has severely limited expressive power on finite structures (cf. [Fag75, AU79, Gai82]). In view of this fact, researchers investigated asymptotic probabilities in logical languages that go beyond first-order logic. Although the 0-1 law fails for second-order logic, it turned out that there are powerful fragments of second-order logic for which the 0-1 law holds. Moreover, the boundary of 0-1 laws for fragments of second-order logic is now understood, through the work of [KS85, Kau87, KV87, PS89, KV90a, PS91, KV92].

The limited expressive power of first-order logic is also due to the absence of any recursion mechanism. Thus, a different direction of investigation pursued the study of 0-1 laws for extensions of first-order logic that allow for fixpoint or iterative constructs. Talanov [Tal81] showed that the 0-1 holds for first-order logic augmented with a *transitive closure* operator. This result was extended by Talanov and Knyazev [TK86], and, independently, by Blass, Gurevich and Kozen [BGK85] who proved that a 0-1 law holds for *positive-fixpoint logic*.[1] Positive-Fixpoint logic is obtained from first-order logic by adding the *least-fixpoint* operator for positive formulas [Mos74, CH82]. It can express properties which are not first-order definable, such as connectivity, acyclicity and 2-colorability. On the other hand the class

[1]To be precise, Talanov and Knyazev's result was in terms of a certain iterative extension of first-order logic, which in particular includes positive-fixpoint logic. The study of 0-1 laws for iterative extensions of first-order logic was further pursued by Knyazev [Kny89].

of positive-fixpoint properties is in general properly contained in PTIME. In particular, "parity" ("there is an even number of elements") is not expressible in positive-fixpoint logic over the class of all finite structures.

In [KV87] we studied the extension of first-order logic that results by adding **while** looping as an iteration construct. This programming query language was introduced by Chandra and Harel [CH82] and, as Abiteboul and Vianu [AV89] showed recently, can be viewed as first-order logic augmented with a *partial-fixpoint* operator for arbitrary first-order formulas. Following [AV89], we use the term *partial-fixpoint properties* for properties expressible in this logic. Partial-fixpoint properties contain all positive-fixpoint properties and are in turn properly contained in the ones computable in PSPACE. Moreover, there are partial-fixpoint properties that are complete for PSPACE.

In [KV87] we announced the 0-1 law for partial-fixpoint logic (we called it there *iterative* logic) and sketched a proof that uses model-theoretic methods similar to the ones employed by Blass, Gurevich and Kozen [BGK85] for positive-fixpoint logic. In particular, the proof uses the compactness theorem of mathematical logic and a model-theoretic characterization of ω-categorical theories due to Engeler [Eng59], Ryll-Nardzewski [RN59], Svenonius [Sve59], and Vaught [Vau61].

Are there logics having higher expressive power than partial-fixpoint logic and possessing the 0-1 law?

Since first-order logic has a finitary syntax, another way to increase its expressive power is to allow for infinitary formation rules. One of the most powerful logics resulting this way is the infinitary logic $L_{\infty\omega}$ which allows for arbitrary disjunctions and conjunctions. The 0-1 law fails, however, for $L_{\infty\omega}$, since "parity" is expressible as a countable disjunction of first-order sentences.

Barwise [Bar77] introduced a family $L^k_{\infty\omega}$, k a positive integer, of infinitary logics that consist of all sentences of $L_{\infty\omega}$ with at most k distinct variables. Although these logics were studied originally on infinite structures, they turn out to have interesting uses in theoretical computer science. They have been investigated on finite structures in their own right in [Kol85, KV90b]. They also underlie much of the work in [Imm82, dR87, LM89, CFI89], although their use there is rather implicit.

We investigate here definability and 0-1 laws for the infinitary languages $L^k_{\infty\omega}$, $k \geq 1$. We show first that every partial-fixpoint property is expressible by a formula of $L^k_{\infty\omega}$, for some $k \geq 1$. This containment is strict, since it is known that the infinitary languages $L^k_{\infty\omega}$, $k \geq 2$, can express non-recursive properties. After this, we establish that the 0-1 law holds for the infinitary logic $L^\omega_{\infty\omega} = \bigcup_{k=1}^{\infty} L^k_{\infty\omega}$. This result on the one hand subsumes the earlier work on 0-1 laws for positive-fixpoint logic and partial-fixpoint logic and on the other reveals the boundary of 0-1 laws for fragments of $L_{\infty\omega}$, since, as mentioned before, "parity" is expressible as a countable disjunction of first-order sentences (a disjunction, however, which involves infinitely many distinct variables).

We supply three different proofs of the 0-1 law for $L^\omega_{\infty\omega}$, each one illuminating the result from a different perspective. The first proof is a generalization of the proofs in [BGK85] for positive-fixpoint logic and in [KV87] for partial-fixpoint logic. This proof is interesting in its use of infinite-model theory to prove a result in finite-model theory (a paradigm established by Fagin [Fag76]). In contrast, our next two proofs are in the spirit of "pure" finite-model theory and they do not appeal to "infinitistic" arguments. One proof is based on a quantifier-elimination method, while the second uses pebble games for infinitary logics.

2 Infinitary Logics

The limited expressive power of first-order logic is due to its finitary syntax and to the absence of any recursion or iteration mechanism. Higher expressive power can be achieved by augmenting the syntax of first-order logic either with infinitary formation rules or with fixpoint operators that act as recursion or iteration constructs. In this section we consider certain infinitary logics, study their properties, and compare them to fixpoint logics.

2.1 Infinitary Logics with a Fixed Number of Variables

Different infinitary logics arise by allowing for infinite disjunctions and conjunctions, or by allowing for infinite strings of quantifiers, or by allowing for both at the same time. We consider the infinitary logic $L_{\infty\omega}$, which is the extension of first-order logic that results by allowing infinite disjunctions and conjunctions in the syntax, while keeping the quantifier strings finite (cf. [BF85]). To illustrate the gain in expressive power, recall the well-known fact that the property "there is an even number of elements" is not expressible by any first-order sentence on finite structures. Let ρ_n be a first-order sentence stating that there are exactly n elements. Then the infinitary sentence $\bigvee_{n=1}^{\infty} \rho_{2n}$ asserts that "there is an even number of elements".

We now define formally the syntax of the infinitary logic $L_{\infty\omega}$.

Definition 2.1: Let σ be a vocabulary consisting of finitely many relational and constant symbols and let $\{v_1, \ldots, v_n\}$ be a countable set of variables. The class $L_{\infty\omega}$ of *infinitary formulas* over σ is the smallest collection of expressions such that

- it contains all first-order fomulas over σ;
- if φ is an infinitary formula, then so is $\neg\varphi$;
- if ψ is an infinitary formula and v_i is a variable, then $(\forall v_i)\varphi$ and $(\exists v_i)\varphi$ are also infinitary formulas;
- if Ψ is a set of infinitary formulas, then $\bigvee \Psi$ and $\bigwedge \Psi$ are also infinitary formulas.[2]

∎

The concept of a *free variable* in a $L_{\infty\omega}$ formula is defined in the same way as for first-order logic. A *sentence* of $L_{\infty\omega}$ is a formula φ of $L_{\infty\omega}$ with no free variables. The semantics of infinitary formulas is a direct extension of the semantics of first-order logic, with $\bigvee \Psi$ interpreted as a disjunction over all formulas in Ψ and $\bigwedge \Psi$ interpreted as a conjunction.

In general, infinitary formulas, even infinitary sentences, may have an infinite number of distinct variables. We now focus attention on fragments of $L_{\infty\omega}$ in which the total number

[2] In mathematical logic, the notation $L_{\kappa\lambda}$, where κ and λ are infinite cardinal numbers, has been used to denote the infinitary logic in which we can form new formulas by taking disjunctions and conjunctions of sets of formulas of cardinality less than κ, and by applying strings of quantifiers of length less than λ. Thus, $L_{\infty\omega} = \cup_\kappa L_{\kappa\omega}$.

of variables is required to be finite. Variables, however, may have an infinite number of occurrences in such formulas.

Definition 2.2: Let k be a positive integer.

- The *infinitary logic with k variables*, denoted by $L^k_{\infty\omega}$, consists of all formulas of $L_{\infty\omega}$ with at most k distinct variables.
- The infinitary logic $L^\omega_{\infty\omega}$ consists of all formulas of $L_{\infty\omega}$ with a finite number of distinct variables. Thus,
$$L^\omega_{\infty\omega} = \bigcup_{k=1}^{\infty} L^k_{\infty\omega}.$$
- We write $L^k_{\omega\omega}$ for the collection of all first-order formulas with at most k variables.

∎

The family $L^\omega_{\infty\omega}$ of the infinitary languages $L^k_{\infty\omega}$, $k \geq 1$, was introduced first by Barwise [Bar77], as a tool for studying positive-fixpoint logic on infinite structures. Since that time, however, these languages have had numerous uses and applications in theoretical computer sciences. Indeed, they underlie much of the work in [Imm82, dR87, LM89, CFI89] and they have also been studied in their own right in [Kol85, KV90b].

We now give some examples that illustrate the expressive power of infinitary logic with a fixed number of variables.

Example 2.3: Cardinalities of Total Orders
Assume that the vocabulary σ consists of a binary relation symbol $<$ and we are considering only the structures in which the interpretation of $<$ is a total order. Let τ_n be a first-order sentence asserting that "there are at least n elements". On arbitrary structures over the vocabulary σ, the sentence τ_n requires n distinct variables. Immerman and Kozen [IK89] pointed out, however, that on total orders τ_n is equivalent to a sentence in $L^2_{\omega\omega}$. For example, τ_4 can be written as

$$(\exists x \exists y)(x < y \wedge (\exists x)(y < x \wedge (\exists y)(x < y))).$$

It follows that on total orders the sentence ρ_n asserting that there are exactly n elements is also in $L^2_{\omega\omega}$, since it is equivalent to $\tau_n \wedge \neg\tau_{n+1}$. As a result, on total orders properties such as "there is an even number of elements", "the universe is finite", etc., are expressible in $L^2_{\infty\omega}$. In general, if P is any set of positive integers, then the property "the cardinality of the total order is a member of P" is expressible in $L^2_{\infty\omega}$, since it is definable by

$$\bigvee_{n \in P} \rho_n.$$

It follows, that $L^\omega_{\infty\omega}$ can express non-recursive properties on total orders. ∎

Example 2.4: Paths and Connectivity
Assume that the vocabulary σ consists of a single binary relation E and let $p_n(x, y)$ be a first-order formula over σ asserting that there is a path of length n from x to y. The obvious way to write $p_n(x, y)$ requires $n + 1$ variables, namely

$$(\exists x_1 \ldots \exists x_{n-1})(E(x, x_1) \wedge E(x_1, x_2) \wedge \ldots \wedge E(x_{n-1}, y)).$$

It is well known, however, that each $p_n(x, y)$ is equivalent to a formula in $L^3_{\omega\omega}$, i.e. a first-order formula with at most three distinct variables x, y, z. To see this, put

$$p_1(x, y) \equiv E(x, y)$$

and assume, by induction on n, that $p_{n-1}(x, y)$ is equivalent to a formula in $L^3_{\omega\omega}$. Then

$$p_n(x, y) \equiv (\exists z)[E(x, z) \wedge (\exists x)(x = z \wedge p_{n-1}(x, y))].$$

It follows that "connectivity" is a property of graphs expressible in $L^3_{\infty\omega}$, since it is given by the formula

$$(\forall x \forall y)(\bigvee_{n=1}^{\infty} p_n(x, y)).$$

Similarly, the property "there is no cycle" is also in $L^3_{\infty\omega}$, since it is definable by:

$$(\forall x)(\bigwedge_{n=1}^{\infty} \neg p_n(x, x)).$$

More generally, if P is *any* set of positive integers, then the property "x and y are connected by a path whose length is a number in P" is expressible in $L^3_{\infty\omega}$ via the formula:

$$\bigvee_{n \in P} p_n(x, y).$$

It follows that $L^\omega_{\infty\omega}$ can express non-recursive properties on finite graphs.
∎

Properties such as "connectivity" and "there is no cycle" are also known to be expressible in positive-fixpoint logic. We consider next extensions of first-order logic with fixpoint formation rules and compare the resulting logics to $L^\omega_{\infty\omega}$.

2.2 Fixpoint Logics

Let σ be a vocabulary, let S be an n-ary relation symbol not in σ, let $\varphi(x_1, \ldots, x_n, S)$ be a first-order formula over the vocabulary $\sigma \cup \{S\}$, and let $\mathbf{D}$ be a finite structure over σ. The formula φ gives rise to an operator $\Phi(S)$ from n-ary relations on the universe D of $\mathbf{D}$ to n-ary relations on D, where

$$\Phi(T) = \{(a_1, \ldots, a_n) : \mathbf{D} \models \varphi(a_1, \ldots, a_n, T)\},$$

for every n-ary relation T on D.

Every such operator $\Phi(S)$ generates a sequence of *stages* that are obtained by iterating $\Phi(S)$. We will be interested here in the relationship between the stages of the operator and its *fixpoints*.

Definition 2.5: Let $\mathbf{D}$ be a finite structure over the vocabulary σ.

- The stages Φ^m, $m \geq 1$, of Φ on $\mathbf{D}$, are defined by the induction:

$$\Phi^1 = \Phi(\emptyset), \ \Phi^{m+1} = \Phi(\Phi^m).$$

- We say that a relation T on D is a *fixpoint* of the operator $\Phi(S)$ (or, of the formula φ) if $\Phi(T) = T$.

∎

Intuitively, one would like to associate with an operator $\Phi(S)$ the "limit" of its stages. This is possible only when the sequence Φ^m, $m \geq 1$, of the stages "converges", i.e., when there is an integer m_0 such that $\Phi^{m_0} = \Phi^{m_0+1}$ and, hence, $\Phi^{m_0} = \Phi^m$, for all $m \geq m_0$. Notice that in this case Φ^{m_0} is a fixpoint of $\Phi(S)$, since $\Phi^{m_0} = \Phi^{m_0+1} = \Phi(\Phi^{m_0})$. The sequence of stages, however, may not converge. In particular, this will happen if the formula $\varphi(\mathbf{x}, S)$ has no fixpoints. Thus, additional conditions have to be imposed on the formulas considered in order to ensure that the sequence of stages converges.

A formula $\varphi(x_1, \ldots, x_n, S)$ is *positive in S* if every occurrence of S in φ is within an even number of negations. Positivity is a natural syntactic condition that guarantees convergence. Indeed, if $\varphi(\mathbf{x}, S)$ is positive in S, then the associated operator Φ is *monotone* (i.e., if $T_1 \subseteq T_2$, then $\Phi(T_1) \subseteq \Phi(T_2)$) and, as a result, the sequence Φ^m, $m \geq 1$, of stages is increasing. If $\mathbf{D}$ is a finite structure with s elements, then every stage Φ^m has at most s^n elements and, consequently, there is an integer $m_0 \leq s^n$ such that $\Phi^{m_0} = \Phi^m$ for every $m \geq m_0$. Thus, the sequence of stages of $\varphi(\mathbf{x}, S)$ converges to Φ^{m_0}. Moreover, it is easy to verify that Φ^{m_0} is the *least fixpoint of* $\varphi(\mathbf{x}, S)$, i.e., it is a fixpoint of φ with the property that $\Phi^{m_0} \subseteq T$ for every fixpoint T of φ. We write φ^∞ or Φ^∞ to denote the least fixpoint of φ.

Remark 2.6: Although here we are mainly interested in finite structures, we should point out that the stages of a formula can also be defined on infinite structures. This is done by transfinite induction on the ordinals, where at limit stages the operator $\Phi(S)$ is applied to the union of the previously defined stages. A positive formula has a least fixpoint on every infinite structure, which is equal to some transfinite stage of the formula.

The existence of least fixpoints for positive formulas is an instance of a more general result about fixpoints in a lattice-theoretic framework (cf. Tarski [Tar55]). ∎

Positive-fixpoint logic is first-order logic augmented with the *least fixpoint* formation rule for *positive* formulas. The canonical example of a formula of positive-fixpoint logic is provided by the least fixpoint $\varphi^\infty(x, y)$ of the first-order formula

$$E(x, y) \vee (\exists z)(S(x, z) \wedge S(z, y)).$$

In this case $\varphi^\infty(x, y)$ defines the *transitive closure* of the *edge* relation E. It follows that *connectivity* is a property expressible in positive-fixpoint logic, but, as is well known (cf. [Fag75, AU79]), not in first-order logic.

As a fresh example, we consider 2-*colorability.* Using Ehrenfeucht-Fraissé games, it can be proved that this property is not expressible in first-order logic. We now show that 2-colorability on directed graphs without loops is expressible in fixpoint logic. For this, let $\varphi(x, y, S)$ be the first-order formula

$$E(x, y) \vee (\exists z)(\exists w)(E(x, z) \wedge E(z, w) \wedge S(w, y)),$$

where E is a binary relation symbol in the vocabulary σ. It is easy to verify that $\varphi^{\infty}(x, y)$ holds if and only if there is a path of odd length from x to y. It follows that a directed graph $G = (A, E)$ is not 2-colorable if and only if $\exists x \varphi^{\infty}(x, x)$.

The theory of positive-fixpoint logic on infinite structures was developed in Moschovakis [Mos74]. Chandra and Harel [CH82] were the first to focus attention on the collection FP of properties expressible in positive-fixpoint logic on finite structures (*positive-fixpoint properties*). Since that time positive-fixpoint logic has been studied extensively on finite structures and this has resulted to a thorough understanding of its expressive power (cf. [Cha88] for a survey of results in this area). We should remark that often in the literature positive-fixpoint logic is referred to as simply *fixpoint logic.*

Every positive-fixpoint property is computable in polynomial time (in the size of the finite structure), because the sequence of stages converges to the least fixpoint in polynomially many iterations. On the other hand there are PTIME properties, such as "there is an even number of elements", that are not in FP [CH82]. Positive-fixpoint logic can express, however, PTIME-complete properties, for example the *path systems problem* in Cook [Coo74]. Moreover, on *ordered finite structures* (i.e., on finite structures where a binary relation symbol is always interpreted as a total order) we have that FP=PTIME ([Imm86, Var82]).

How can we obtain logics with iteration constructs that are more expressive than positive-fixpoint logic? A more powerful logic results if one iterates arbitrary first-order operators, until a fixpoint is reached (which may never happen). In this case we may have non-terminating computations, unlike positive-fixpoint logic, where the iteration is guaranteed to converge.

Definition 2.7: Let σ be a vocabulary, let S be an n-ary relation symbol not in σ, let $\varphi(x_1, \dots, x_n, S)$ be a first-order formula over the vocabulary $\sigma \cup \{S\}$, let D be a finite structure over σ, and let Φ^m, $m \geq 1$, be the sequence of stages of the associated operator $\Phi(S)$.

If there is an integer m_0 such that $\Phi^{m_0} = \Phi^{m_0+1}$, then we put $\varphi^{\infty} = \Phi^{\infty} = \Phi^{m_0}$; otherwise, we set $\varphi^{\infty} = \Phi^{\infty} = \emptyset$. In the former case we say that φ *converges* on $\mathbf{D}$, and in the latter case we say that φ *diverges* on $\mathbf{D}$. We call φ^{∞} the *partial-fixpoint* of φ on $\mathbf{D}$.

Partial-Fixpoint Logic is first-order logic augmented with the partial-fixpoint formation rule for arbitrary first-order formulas. We write PFP for the collection of all properties definable by formulas of partial-fixpoint logic on finite structures. ∎

Partial-fixpoint logic on finite structures has been investigated by Abiteboul and Vianu [AV89]. In particular, they established that the class PFP of partial-fixpoint properties on finite structures coincides with the class of properties expressible in the language RQL *(ranked query language)*, introduced by Chandra and Harel [CH82] and studied also in [KV87] under the name *iterative logic.* The latter is an extension of first-order logic obtained by adding **while** looping as an iteration construct.

As with positive-fixpoint logic, the syntax of partial-fixpoint logic allows for the interleaving of first-order operations (including negation) with the partial-fixpoint operator. Abiteboul and Vianu [AV89] showed, however, that this does not give rise to a hierarchy of properties and that a single application of the partial-fixpoint operator suffices to generate all PFP properties. An analogous result for fixpoint logic had been obtained by Immerman [Imm86], and Gurevich and Shelah [GS86].

Notice that for every first-order formula $\varphi(x_1, \ldots, x_n, S)$, if $\mathbf{D}$ is a finite structure with s elements, then either the sequence Φ^m, $m \geq 1$, of stages converges or it cycles. Which of the two is the case can be determined by carrying out at most 2^{s^n} iterations of Φ. Thus, the computation of the partial-fixpoint requires space polynomial in the size s of the structure $\mathbf{D}$, since we only have to store one stage at a time and compute the next stage, while making sure that the current level of iteration has not exceeded 2^{s^n}. Notice also that if $\varphi(x_1, \ldots, x_n, S)$ is a positive in S formula, then the partial-fixpoint of φ coincides with the least fixpoint of φ, because the sequence of stages converges. It follows that partial-fixpoint logic is an extension of positive-fixpoint logic.

As a result of the above facts, we have that

$$\text{FP} \subseteq \text{PFP} \subseteq \text{PSPACE}.$$

The class PFP of partial-fixpoint properties is properly contained in PSPACE, since the property of "cardinality is even" is not in PFP [CH82]. On the other hand, it turns out that on *ordered finite structures* PFP = PSPACE, because on such structures partial-fixpoint logic can simulate PSPACE computations. (This was shown by Vardi [Var82] to hold for the class of *while* properties, which is equivalent to PFP [AV89].) Note that this implies that PFP $\not\subseteq$ PTIME, assuming that PTIME $\neq$ PSPACE.

Chandra and Harel [CH82] posed the problem of showing that FP is properly contained in PFP on the class of all finite structures over a vocabulary σ. No progress was made on this problem until recently, when Abiteboul and Vianu [AV91] showed that FP $\neq$ PFP if and only if PTIME $\neq$ PSPACE. Thus, the separation problem for these two fixpoint logics on the class of all finite structures is equivalent to one of the outstanding problems in complexity theory.

Our next result shows that partial-fixpoint logic can be subsumed by the infinitary logic $L^{\omega}_{\infty\omega}$.

Theorem 2.8: *Let σ be a vocabulary, let S be an n-ary relation symbol not in σ, let $\varphi(x_1, \ldots, x_n, S)$ be a first-order formula over the vocabulary $\sigma \cup \{S\}$, and assume that the total number of distinct variables (free and bound) occurring in φ is equal to k. Let $\Phi(S)$ be the operator associated with φ, where*

$$\Phi(T) = \{(a_1, \ldots, a_n) : \mathbf{D} \models \varphi(a_1, \ldots, a_n, T)\},$$

for any n-ary relation T on the universe D of a structure $\mathbf{D}$ over σ. Then

- *For every $m \geq 1$, the stage $\Phi^m(x_1, \ldots, x_n)$ of Φ is definable by a formula of $L^{k+n}_{\omega\omega}$ on all finite structures over σ.*

- *The partial-fixpoint $\varphi^\infty(x_1, \ldots, x_n)$ of $\varphi(x_1, \ldots, x_n, S)$ is definable by a formula of $L^{k+n}_{\infty\omega}$ on all finite structures.*

Proof: Let $y_1, \ldots, y_n$ be n new distinct variables not occurring in φ. We will show, by induction on m, that every stage Φ^m, $m \geq 1$, is expressible by a formula $\varphi^m(x_1, \ldots, x_n)$ of $L^{k+n}_{\omega\omega}$ whose variables are those of φ and $y_1, \ldots, y_n$. The claim is obvious for the first stage $\Phi^1 = \Phi(\emptyset)$. Assume that the induction hypothesis holds for Φ^m. By definition of the stages, we have that

$$\Phi^{m+1}(x_1, \ldots, x_n) \;\equiv\; \varphi(x_1, \ldots, x_n, \Phi^m).$$

At this point, one would like to replace every occurrence of a subformula of the form $S(t_1, \ldots, t_n)$ in $\varphi(x_1, \ldots, x_n, S)$ by the formula $\varphi^m(x_1/t_1, \ldots, x_n/t_n)$, where the latter formula is obtained from $\varphi^m(x_1, \ldots, x_n)$ by substituting t_i for each free occurrence of x_i, $1 \leq i \leq n$. This, however, may increase the total number of variables in the resulting formula beyond any predescribed bounds, since one would have to make the substitutions not to $\varphi^m(x_1, \ldots, x_n)$, but to an equivalent formula (possibly having more variables) in which each t_i can be sustituted for x_i (without changing the meaning of the formula). It turns out, nevertheless, that the above difficulty can be circumvented as follows.

Replace every occurrence of a subformula of the form $S(t_1, \ldots, t_n)$ in $\varphi(x_1, \ldots, x_n, S)$ by the expression

$$(\exists y_1 \ldots \exists y_n)[(y_1 = t_1 \wedge \ldots \wedge y_n = t_n) \wedge$$
$$(\exists x_1 \ldots \exists x_n)(x_1 = y_1 \wedge \ldots \wedge x_n = y_n \wedge \tilde{\varphi}^m(x_1, \ldots, x_n))].$$

The resulting expression yields a formula $\varphi^{m+1}(x_1, \ldots, x_n)$ of $L^{k+n}_{\omega\omega}$ (whose variables are those of φ and $y_1, \ldots, y_n$) that defines Φ^{m+1} uniformly on all finite structures.

It is now easy to show that on finite structures the partial-fixpoint φ^∞ of the formula $\varphi(x_1, \ldots, x_n, S)$ is expressible by a formula of $L^{k+n}_{\infty\omega}$. Recall that φ^∞ is equal to some stage Φ^{m_0} such that $\Phi^{m_0} = \Phi^{m_0+1}$, if such a stage exists , or equal to $\emptyset$ otherwise. Thus,

$$\varphi^\infty(\mathbf{x}) \;\equiv\; \bigvee_{m=1}^{\infty} [(\forall \mathbf{x})(\varphi^{m+1}(\mathbf{x}) \leftrightarrow \varphi^m(\mathbf{x}))] \wedge \varphi^m(\mathbf{x}).$$

∎

The preceding Theorem 2.8 constitutes an extension of an earlier result to the effect that on every fixed structure the infinitary logic $L^\omega_{\infty\omega}$ can express every positive-fixpoint formula. That result appeared in print in [Bar77, Imm82], but actually goes back to the unpublished Ph.D. thesis of A. Rubin [Rub75].

Corollary 2.9: *$L^\omega_{\infty\omega}$ can express every partial-fixpoint property on finite structures. Thus, on finite structures*

$$\mathrm{FP} \subseteq \mathrm{PFP} \subseteq L^\omega_{\infty\omega}.$$

As mentioned earlier, $L^\omega_{\infty\omega}$ on finite graphs can capture non-recursive properties. Since every PFP property is recursive (actually in PSPACE), it follows that the inclusion $\mathrm{PFP} \subseteq L^\omega_{\infty\omega}$ is proper.

2.3 $L^k_{\infty\omega}$-Equivalence and Pebble Games

Two structures are *equivalent* in some logic L if no sentence of L distinguishes them. This is a central concept in every logic and plays an important role in model theory. We discuss here equivalence in the infinitary logics with a fixed number of variables.

Definition 2.10: Let $\mathbf{A}$ and $\mathbf{B}$ be two structures over the vocabulary σ and let k be a positive integer.

- Assume that $a_1, \ldots, a_m$ and $b_1, \ldots, b_m$ are finite sequences of distinct elements from the universes of $\mathbf{A}$ and $\mathbf{B}$ respectively, where $1 \leq m \leq k$. We write

$$(\mathbf{A}, a_1, \ldots, a_m) \equiv^k_{\infty\omega} (\mathbf{B}, b_1, \ldots, b_m)$$

  to denote that for every formula $\varphi(u_1, \ldots, u_m)$ of $L^k_{\infty\omega}$ with free variables among $u_1, \ldots, u_m$ we have that

$$\mathbf{A}, a_1, \ldots, a_m \models \varphi(u_1, \ldots, u_m) \quad \text{if and only if} \quad \mathbf{B}, b_1, \ldots, b_m \models \varphi(u_1, \ldots, u_m).$$

- We say that $\mathbf{A}$ is $L^k_{\infty\omega}$-*equivalent* to $\mathbf{B}$, and we write $\mathbf{A} \equiv^k_{\infty\omega} \mathbf{B}$, if $\mathbf{A}$ and $\mathbf{B}$ satisfy the same sentences of $L^k_{\infty\omega}$.

- We say that $\mathbf{A}$ is $L^k_{\omega\omega}$-*equivalent* to $\mathbf{B}$, and we write $\mathbf{A} \equiv^k_{\omega\omega} \mathbf{B}$, if $\mathbf{A}$ and $\mathbf{B}$ satisfy the same sentences of first-order logic with k variables.

∎

The connection between definability in $L^k_{\infty\omega}$ and the equivalence relation $\equiv^k_{\infty\omega}$ is described by the following two propositions.

Proposition 2.11: *Let σ be a vocabulary, let $\mathcal{F}$ be the class of all finite structures over σ, and let k be a positive integer. Then every equivalence class of the equivalence relation $\equiv^k_{\infty\omega}$ on $\mathcal{F}$ is definable by a sentence of $L^k_{\infty\omega}$.*

Proof: Observe first that the equivalence relation $\equiv^k_{\infty\omega}$ on $\mathcal{F}$ has only countably many equivalence classes. This is because there are only countably many non-isomorphic finite structures and every equivalence class is a union of isomorphism classes. Let $C_0, C_1, \ldots$ be an enumeration of the equivalence classes of $\equiv^k_{\infty\omega}$ on $\mathcal{F}$.

Let $i \geq 0$. For all $j \neq i$, there is a sentence ψ_j of $L^k_{\infty\omega}$ such that ψ_j holds for the structures in C_i, but fails for the structures in C_j. Thus, the countable conjunction $\bigwedge_{j \neq i} \psi_j$ is a sentence of $L^k_{\infty\omega}$ that is satisfied exactly by the structures in C_i. ∎

Proposition 2.12: *Let $\mathcal{C}$ be a class of finite structures over the vocabulary σ and let k be a positive integer. Then the following statements are equivalent:*

1. *The class $\mathcal{C}$ is $L^k_{\infty\omega}$-definable, i.e. there is a sentence φ of $L^k_{\infty\omega}$ such that for any finite structure $\mathbf{A}$ over σ we have that*

$$\mathbf{A} \in \mathcal{C} \Longleftrightarrow \mathbf{A} \models \varphi.$$

2. If $\mathbf{A}$ *and* $\mathbf{B}$ *are finite structures over* σ *such that* $\mathbf{A} \in \mathcal{C}$ *and* $\mathbf{A} \equiv^k_{\infty\omega} \mathbf{B}$*, then* $\mathbf{B} \in \mathcal{C}$*.*

Proof: The direction (1) $\Rightarrow$ (2) follows from the definition of $\equiv^k_{\infty\omega}$.

For the other direction, assume that statement (2) holds for the class $\mathcal{C}$. The preceding Proposition 2.11 implies that for each $i \geq 0$ there is a sentence Ψ_i of $L^k_{\infty\omega}$ that defines the i-th equivalence class C_i of $\equiv^k_{\infty\omega}$ on finite structures. We now claim that the countable disjunction $\bigvee_{C_i \subseteq \mathcal{C}} \Psi_i$, which is a sentence of $L^k_{\infty\omega}$, defines the class $\mathcal{C}$. Assume that $\mathbf{A}$ is a structure in $\mathcal{C}$ and let C_l be the equivalence class of $L^k_{\infty\omega}$ to which $\mathbf{A}$ belongs. Then, by hypothesis, we have that $\mathcal{C}_l \subseteq \mathcal{C}$. Thus, $\mathbf{A} \models \bigvee_{C_i \subseteq \mathcal{C}} \Psi_i$, since $\mathbf{A} \models \Psi_l$. Conversely, assume that $\mathbf{A}$ is a finite structure satisfying Φ_l, where $C_l \subseteq \mathcal{C}$. Then $\mathbf{A} \in C_l$, so $\mathbf{A} \in \mathcal{C}$. ∎

Remark 2.13: Note that the proof of Proposition 2.12 used only the property that $L^k_{\infty\omega}$ is closed under countable conjunctions and disjunctions. Thus, the proposition can be generalized to any logic that has this closure property. In particular, the proposition can be applied to any logic that is closed under finite conjunctions and disjunctions and has a finite number of equivalence relations. For example, if we consider the fragment L_r of first-order logic consisting of first-order sentences of quantifier depth r, then the analogous version of the above Proposition 2.12 holds for L_r. The reason for this is that, by Fraissé's theorem [Fra54], the relation of elementary equivalence on L_r has finitely many equivalence classes. ∎

It is known that $\equiv^k_{\infty\omega}$-equivalence can be characterized in terms of an *infinitary k-pebble game*. This game was implicit in Barwise [Bar77] and was described in detail in Immerman [Imm82].

Definition 2.14: Assume that $\mathbf{A}$ and $\mathbf{B}$ are two structures over the vocabulary σ and let $c_1, \ldots, c_l$ and $d_1, \ldots, d_l$ be the interpretations of the constant symbols of σ on $\mathbf{A}$ and $\mathbf{B}$, respectively.

The *k-pebble game* between Players I and II on the structures $\mathbf{A}$ and $\mathbf{B}$ has the following rules:

Player I chooses one of the two structures $\mathbf{A}$ and $\mathbf{B}$ and places a pebble on one of its elements. Player II responds by placing a pebble on an element of the other structure. Player I chooses again one of the two structures and the game continues this way until k pebbles have been placed on each structure.

Let a_i and b_i, $1 \leq i \leq k$, be the elements of $\mathbf{A}$ and $\mathbf{B}$ respectively picked by the two players in the i-th move. Player I wins the game at this point if one of the following two conditions holds:

- Two of the pebbles are on the same element of one of the structures, while the corresponding two pebbles are on different elements of the other structure, i.e., $a_i = a_j$ and $b_i \neq bj$ (or $a_i \neq a_j$ and $b_i = b_j$), for some i and j such that $1 \leq i < j \leq k$.
- The previous condition fails and the mapping h with

$$h(a_i) = b_i, \ 1 \leq i \leq k,$$

and

$$h(c_j) = d_j, \ 1 \leq j \leq l,$$

is *not* an isomorphism between the substructures of $\mathbf{A}$ and $\mathbf{B}$ with universes

$$\{a_1, \ldots, a_k\} \cup \{c_1, \ldots, c_l\}$$

and

$$\{b_1, \ldots, b_k\} \cup \{d_1, \ldots, d_l\}$$

respectively.

If both conditions fail, then Player I removes some corresponding pairs of pebbles and the game resumes until again k pebbles have been placed on each structure. Player II wins the game if he can continue playing "forever", i.e. if Player I can never win a round of the game. ∎

In the preceding definition we have described in a rather informal the concept "Player II wins the k-pebble game on $\mathbf{A}$ and $\mathbf{B}$". The concept of a *winning strategy* for Player II in the k-pebble game is formalized in what follows (cf. also [Bar77, Imm82]).

Definition 2.15: Let $\mathbf{A}$ and $\mathbf{B}$ be two structures over the vocabulary σ and let $c_1, \ldots, c_l$ and $d_1, \ldots, d_l$ be the interpretations of the constant symbols of σ on $\mathbf{A}$ and $\mathbf{B}$ respectively.

A *partial isomorphism* between $\mathbf{A}$ and $\mathbf{B}$ is a function h such that its domain is a finite subset of the universe of $\mathbf{A}$ containing the elements $c_1, \ldots, c_l$ of $\mathbf{A}$, its range is a finite subset of the universe of $\mathbf{B}$ containing the elements $d_1, \ldots, d_l$ of $\mathbf{B}$, $h(c_j) = d_j,\ 1 \leq j \leq l$, and such that h is an isomorphism between the substructures of $\mathbf{A}$ and $\mathbf{B}$ with universes the domain and range of h respectively. ∎

Definition 2.16: Let k be a positive integer, let $\mathbf{A}$ and $\mathbf{B}$ be two structures over the vocabulary σ, and let $c_1, \ldots, c_l$ and $d_1, \ldots, d_l$ be the interpretations of the constant symbols of σ on $\mathbf{A}$ and $\mathbf{B}$ respectively.

We say that *Player II has a winning strategy in the k-pebble game on* $\mathbf{A}$ *and* $\mathbf{B}$ if there is a non-empty family $\mathcal{H}$ of partial isomorphisms between $\mathbf{A}$ and $\mathbf{B}$ such that

- $\mathcal{H}$ is *closed under subfunctions*: if $f \in \mathcal{H}$ and $\{(c_1, d_1), \ldots, (c_l, d_l)\} \subseteq g \subseteq f$ (as sets of ordered pairs), then $g \in \mathcal{H}$.
- $\mathcal{H}$ *has the back and forth property up to* k: if $f \in \mathcal{H}$ and $|f| < k + l$, then for any element $a \in \mathbf{A}$ (respectively $b \in \mathbf{B}$) there is an element b in $\mathbf{B}$ (respectively $a \in \mathbf{A}$) such that the function $f \cup \{(a, b)\}$ is in $\mathcal{H}$.

∎

The crucial connection between k-pebble games and $L^k_{\infty\omega}$-equivalence is provided by the following result, which is due to Barwise [Bar77] (cf. also [Imm82]). We include here a detailed proof of this result, since only a hint for the proof is given in [Bar77].

Theorem 2.17: *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be two structures over the vocabulary* σ*, and let* k *be a positive integer. The following are equivalent:*

1. $\mathbf{A} \equiv^k_{\infty\omega} \mathbf{B}$.

2. *Player II has a winning strategy for the k-pebble game on* $\mathbf{A}$ *and* $\mathbf{B}$.

Proof: Let $c_1, \ldots, c_l$ and $d_1, \ldots, d_l$ be the interpretations of the constant symbols of σ on $\mathbf{A}$ and $\mathbf{B}$, respectively. Assume first that $\mathbf{A} \equiv^k_{\infty\omega} \mathbf{B}$. We have to show that there is a family $\mathcal{H}$ of partial isomorphisms on $\mathbf{A}$ and $\mathbf{B}$ that provides a winning strategy for Player II in the k-pebble game.

The desired family $\mathcal{H}$ consists of all partial isomorphisms between $\mathbf{A}$ and $\mathbf{B}$ such that the following hold:

- The domain of h is a set of the form $\{c_1, \ldots, c_l, a_1, \ldots, a_m\}$ and the range of h is a set of the form $\{d_1, \ldots, d_l, b_1, \ldots, b_m\}$, where $m \leq k$.
- $h(c_j) = d_j$ for all $j \leq m$, and $h(a_i) = b_i$, for all $i \leq m$.
- $(\mathbf{A}, a_1, \ldots, a_m) \equiv^k_{\infty\omega} (\mathbf{B}, b_1, \ldots, b_m)$.

We show now that $\mathcal{H}$ has the required properties:

1. $\mathcal{H}$ is non-empty, because $\mathbf{A} \equiv^k_{\infty\omega} \mathbf{B}$ and, thus, the function h with $h(c_j) = d_j$, $1 \leq j \leq l$, is a member of $\mathcal{H}$.

2. It is clear from the definitions that $\mathcal{H}$ is closed under subfunctions.

3. It remains to show that $\mathcal{H}$ has the back and forth property up to k. Assume that $f \in \mathcal{H}$ and $|f| = m + l < k + l$. Then there are sequences of distinct elements $a_1, \ldots, a_m$ in A and $b_1, \ldots, b_m$ in B such that $h(a_i) = b_i$, $1 \leq i \leq m$ and

$$(\mathbf{A}, a_1, \ldots, a_m) \equiv^k_{\infty\omega} (\mathbf{B}, b_1, \ldots, b_m).$$

There are two parts in the back-and-forth property up to k: in the "forth" part we have to show that for every element a in $\mathbf{A}$ there is an element b in $\mathbf{B}$ such that $f \cup \{a, b\}$ is in $\mathcal{H}$, while in the "back" part we have to show that for every b in $\mathbf{B}$ there is an a in $\mathbf{A}$ such that $f \cup \{a, b, \}$ is in $\mathcal{H}$.

We claim that for any element a in A that is different from $a_1, \ldots, a_m$ there is an element b in B that is different from $b_1, \ldots, b_m$ and is such that

$$(\mathbf{A}, a_1, \ldots, a_m, a) \equiv^k_{\infty\omega} (\mathbf{B}, b_1, \ldots, b_m, b).$$

Assume that no such $b \in B$ exists for a certain $a \in A$. Then for every $b \in B$ that is different from $b_1, \ldots, b_m$ there is a formula $\psi_b(v_1, \ldots, v_m, v)$ of $L^k_{\infty\omega}$ over σ such that

$$(\mathbf{A}, a_1, \ldots, a_m, a) \models \psi_b(v_1, \ldots, v_m, v)$$

and

$$(\mathbf{B}, b_1, \ldots, b_m, b) \not\models \psi_b(v_1, \ldots, v_m, v).$$

Hence,

$$(\mathbf{A}, a_1, \ldots, a_m) \models (\exists v) \left((v_1 \neq v) \wedge \ldots \wedge (v_m \neq v) \wedge \bigwedge_{b \in B} \psi_b(v_1, \ldots, v_m, v) \right),$$

and, at the same time,

$$(\mathbf{B}, b_1, \ldots, b_l) \not\models (\exists v)\left((v_1 \neq v) \wedge \ldots \wedge (v_m \neq v) \wedge \bigwedge_{b \in B} \psi_b(v_1, \ldots, v_m, v)\right).$$

But this is a contradiction, since

$$(\exists v)\left((v_1 \neq v) \wedge \ldots \wedge (v_m \neq v) \wedge \bigwedge_{b \in B} \psi_b(v_1, \ldots, v_m, v)\right)$$

is an $L^k_{\infty\omega}$-formula and $(\mathbf{A}, a_1, \ldots, a_m) \equiv^k_{\infty\omega} (\mathbf{B}, b_1, \ldots, b_m)$.

This concludes the argument for the "forth" part of the back-and-forth-property. The "back" part is analogous, using an infinitary conjunction over elements of A.

Assume now that Player II has a winning strategy in the k-pebble game on $\mathbf{A}$ and $\mathbf{B}$. Let $\mathcal{H}$ be a family of partial isomorphisms providing Player II with a winning strategy.

We will show, by induction on the construction of $L^k_{\infty\omega}$ formulas, that if $\psi(v_1, \ldots, v_m)$ is a formula of $L^k_{\infty\omega}$ whose variables are among $v_1, \ldots, v_k$ and whose free variables are among $v_1, \ldots, v_m$, then the following property $(*)$ holds:

$(*)$ For all $h \in \mathcal{H}$ with $|h| \geq l + m$ and for any elements $\{a_1, \ldots, a_m\}$ (not necessarily distinct) from the domain of h, we have

$$(\mathbf{A}, a_1, \ldots, a_m) \models \psi(v_1, \ldots, v_m) \quad \text{if and only if} \quad (\mathbf{B}, h(a_1), \ldots, h(a_m) \models \psi(v_1, \ldots, v_m)).$$

After property $(*)$ is established, we will be able to infer that $\mathbf{A} \equiv^k_{\infty\omega} \mathbf{B}$ by applying this property to sentences of $L^k_{\infty\omega}$ and to an arbitrary member of the non-empty family $\mathcal{H}$.

The base case in the induction (atomic formulas and inequalities) is obvious. The inductive steps for negation $\neg$, infinitary disjunction $\bigvee$, and infinitary conjuction $\bigwedge$ are straightforward using the induction hypothesis.

Assume that the formula $\psi(v_1, \ldots, v_m)$ is of the form $(\exists v)\chi(v_1, \ldots, v_m, v)$ and that the property $(*)$ holds for the formula $\chi(v_1, \ldots, v_m, v)$. Let h be a partial isomorphism in $\mathcal{H}$ such that $|h| \leq l + m$. We have to show that if $a_1, \ldots, a_m$ are arbitrary elements (not necessarily distinct) from the domain of h, then

$$\mathbf{A}, a_1, \ldots, a_m \models (\exists v)\chi(v_1, \ldots, v_m, v) \quad \Longleftrightarrow \quad \mathbf{B}, h(a_1), \ldots, h(a_m) \models (\exists v)\chi(v_1, \ldots, v_m, v).$$

Notice that, by our assumption about the variables of ψ, we must have that v is a variable v_j, for some j such that $1 \leq j \leq k$. We now distinguish two cases, namely the case where $j > m$ and the case where $j \leq m$.

If $j > m$, then it must also be the case that $m < k$. Let $a \in A$ be such that

$$(\mathbf{A}, a_1, \ldots, a_m, a) \models \chi(v_1, \ldots, v_m, v_j).$$

Consider the subfunction h^* of h with domain the set $\{c_1, \ldots, c_l, a_1, \ldots, a_m\}$. Notice that h^* is a member of $\mathcal{H}$, since $\mathcal{H}$ is closed under subfunctions. By the "forth" part of the back and forth property applied to h^* and a, there is an element $b \in B$ such that $h^* \cup \{(a,b)\}$ is in $\mathcal{H}$. By applying the induction hypothesis to $\chi(v_1, \dots, v_m, v)$ and to $h^* \cup \{(a,b)\}$, we infer that

$$(\mathbf{B}, h(a_1), \dots, h(a_m), b) \models \chi(v_1, \dots, v_m, v_j)$$

and, hence,

$$(\mathbf{B}, h(a_1), \dots, h(a_m)) \models (\exists v_j)\chi(v_1, \dots, v_m, v_j).$$

The other direction is proved in a similar way using the "back" part of the back and forth property up to k for the family $\mathcal{H}$.

Finally, assume that $j \leq m$. In this case, the free variables of the formula χ are among the variables $v_1, \dots, v_m$ and

$$(\mathbf{A}, a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_m) \models (\exists v_j)\chi(v_1, \dots, v_m).$$

Let g be the subfunction of h with domain the set

$$\{c_1, \dots, c_l, a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_m\}.$$

Observe that $|g| \leq l + m - 1 < l + k$ and that g is a member of $\mathcal{H}$, since $\mathcal{H}$ is closed under subfunctions. Let $a \in \mathbf{A}$ be such that

$$\mathbf{A}, a_1, \dots, a_{i-1}, a, a_{i+1}, \dots, a_m \models \chi(v_1, \dots, v_m).$$

By the "forth part" of the back and forth property applied to g and a, there is an element $b \in B$ such that $g \cup \{(a,b)\}$ is in $\mathcal{H}$. By applying the induction hypothesis to $\chi(v_1, \dots, v_m)$ and to $g \cup \{(a,b)\}$, we infer that

$$(\mathbf{B}, g(a_1), \dots, g(a_{i-1}), b, g(a_{i+1}), \dots, g(a_m)) \models \chi(v_1, \dots, v_m)$$

and, hence,

$$(\mathbf{B}, h(a_1), \dots, h(a_m)) \models (\exists v_j)\chi(v_1, \dots, v_m),$$

since $g(a_i) = h(a_i)$ for $i \neq j$ and the satisfaction relation depends only on the free variables of a formula. The other direction is proved in a similar way using the "back" part of the back and forth property up to k for the family $\mathcal{H}$. ∎

Remark 2.18:

The preceding Theorem 2.17 and the above remarks should be contrasted with C. Karp's [Kar65] characterization of equivalence in the infinitary logic $L_{\infty\omega}$ (cf. also [BF85]). According to this result, two structures $\mathbf{A}$ and $\mathbf{B}$ satisfy the same sentences of $L_{\infty\omega}$ if and only if there is a family $\mathcal{H}$ of partial isomorphisms between $\mathbf{A}$ and $\mathbf{B}$ such that $\mathcal{H}$ has the back and forth property (with no cardinality restrictions on the size of the partial isomorphisms). It should be pointed out $\mathcal{H}$ does not have to possess the closure under subfunctions property, which was used in a critical way in establishing the direction (2) $\Rightarrow$ (1) of the preceding Theorem 2.17.

For $L_{\infty\omega}$, the proof goes through without the closure under subfunctions property, because in the case of existential quantification one can rename variables and assume that the existentially quantified variable v is not one of the variables $v_1, \dots, v_m$. ∎

The preceding Theorem 2.17 holds for arbitrary finite or infinite structures **A** and **B**. In the case of infinite structures, the infinitary syntax of $L^k_{\infty\omega}$ plays a crucial role in the proof. On the other hand, close scrutiny of the proof reveals that if both **A** and **B** are finite structures, then one can restrict attention to the first-order sentences of $L^k_{\infty\omega}$.

Corollary 2.19: *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be two finite structures over the vocabulary* σ *and let* k *be a positive integer. The following are equivalent:*

1. $\mathbf{A} \equiv^k_{\omega\omega} \mathbf{B}$.
2. $\mathbf{A} \equiv^k_{\infty\omega} \mathbf{B}$.
3. *Player II has a winning strategy for the* k*-pebble game on* $\mathbf{A}$ *and* $\mathbf{B}$.

Proof: The argument used in the proof of Theorem 2.17 goes here through virtually unchanged. One need only observe that in showing the implication (1) $\Rightarrow$ (3) the conjunctions over the universes A and B are finite and, thus, the resulting formulas are in $L^k_{\omega\omega}$. ∎

The preceding Corollary 2.19 yields the following *normal-form theorem* for sentences of $L^k_{\infty\omega}$ on finite structures.

Corollary 2.20: *Let* σ *be a vocabulary and let* k *be a positive integer. Every sentence of* $L^k_{\infty\omega}$ *is equivalent on finite structures over* σ *to a countable disjunction of countable conjunctions of* $L^k_{\omega\omega}$*-sentences.*

Proof: From Proposition 2.11 and Corollary 2.19 it follows that the $\equiv^k_{\infty\omega}$-equivalence class of a finite structure **A** can be defined by the conjunction $\bigwedge \Psi_{\mathbf{A}}$ of the set $\Psi_{\mathbf{A}}$ of all sentences of $L^k_{\omega\omega}$ that are true on **A**. As a result, every sentence ψ of $L^k_{\infty\omega}$ is equivalent on finite structures to $\bigwedge_{\mathbf{A}\models\psi} \Psi_{\mathbf{A}}$. ∎

As a consequence of Proposition 2.12 and Theorem 2.17, we get a game-theoretic characterization of definability in the logics $L^k_{\infty\omega}$, $k \geq 1$, for classes of finite structures.

Proposition 2.21: *Let* $\mathcal{C}$ *be a class of finite structures over the vocabulary* σ *and let* k *be a positive integer. Then the following statements are equivalent:*

1. *The class* $\mathcal{C}$ *is* $L^k_{\infty\omega}$*-definable.*
2. *If* $\mathbf{A}$ *and* $\mathbf{B}$ *are finite structures over* σ *such that* $\mathbf{A} \in \mathcal{C}$ *and Player II has a winning strategy for the* k*-pebble game on* $\mathbf{A}$ *and* $\mathbf{B}$*, then* $\mathbf{B} \in \mathcal{C}$.

The preceding results provide tools for establishing that certain properties are not expressible in infinitary logic with a finite number of variables. More specifically, in order to establish that a property Q is not expressible by any formula of $L^\omega_{\infty\omega}$ on finite structures it is enough to show that for any $k \geq 1$ there are finite structures $\mathbf{A}_k$ and $\mathbf{B}_k$ such that $\mathbf{A}_k \models Q$, $\mathbf{B}_k \not\models Q$, and Player II has a winning strategy for the k-pebble game on $\mathbf{A}_k$ and $\mathbf{B}_k$. Moreover, Proposition 2.21 guarantees that this method is also *complete*, i.e., if Q is not expressible in $L^k_{\infty\omega}$, then such structures $\mathbf{A}_k$ and $\mathbf{B}_k$ must exist.

The following examples illustrate the use of k-pebble games in deriving lower bounds for expressibility.

Example 2.22: Cliques, Even Cardinality, and Finiteness

Assume that the vocabulary σ consists of a binary relation symbol E. Let k be a positive integer, let $\mathbf{K}_k$ be the complete graph with k nodes (the k-*clique*), and let $\mathbf{K}$ be a complete graph with more than k nodes.

It is quite obvious that Player II has a winning strategy for the k-pebble game between $\mathbf{K}_k$ and $\mathbf{K}$. The family $\mathcal{H}$ of partial isomorphisms consists of all 1-1 mappings between substructures of $\mathbf{K}_k$ and $\mathbf{K}$ each with l elements, $0 \leq l \leq k$.

The immediate consequences of this fact are:

- For any fixed k, the property "there are exactly k elements" can not be expressed on finite graphs by any formula of $\bigcup_{m<k} L^m_{\infty\omega}$; in other words, this property requires at least k variables (cf. also [Imm82] for a different proof of this fact).
- The property "there is an even number of nodes" can not be expressed on finite graphs by any formula of $L^\omega_{\infty\omega}$. (This should be contrasted with the earlier Example 2.3 concerning the expressive power of $L^\omega_{\infty\omega}$ on total orders.)

It follows that the infinitary logic $L_{\infty\omega}$ has strictly higher expressive power than the infinitary logic $L^\omega_{\infty\omega}$.
∎

Example 2.23: Hamiltonian Graphs (Immerman [Imm82], de Rougemont [dR87])
Let $\mathbf{D}_m$, $m \geq 1$, be the totally disconnected graph with m elements, let $\mathbf{C}_n$, $n \geq 1$, be the cycle with n elements, and let $\mathbf{A}_{m,n}$ be the *product graph* of $\mathbf{D}_m$ and $\mathbf{C}_n$, i.e., the vertex set of $\mathbf{A}_{m,n}$ is the union of the vertex sets of $\mathbf{D}_m$ and $\mathbf{C}_n$, while the set of edges of $\mathbf{A}_{m,n}$ consists of the edges of $\mathbf{C}_n$ and edges between every vertex of $\mathbf{D}_m$ and every vertex of $\mathbf{C}_n$. It easy to see that

- $\mathbf{A}_{m,n}$ is Hamiltonian if and only if $m \leq n$.
- Player II has a winning strategy for the k-pebble game on $\mathbf{A}_{k,k}$ and $\mathbf{A}_{k+1,k}$, for every $k \geq 1$.

It follows that Hamiltonicity is not expressible by any formula of $L^\omega_{\infty\omega}$. This was established first in [Imm82]; the above proof is from [dR87]. ∎

Example 2.24: Eulerian Graphs
Recall that an undirected graph is *Eulerian* if and only if every vertex has even degree. Let $\mathbf{B}_k = (V, E)$ be the undirected graph with vertex set $\{a, b, c_1, \ldots, c_k\}$ and edges

$$(a, c_1), \ldots, (a, c_k), (b, c_1), \ldots, (b, c_k).$$

It is clear that $\mathbf{B}_k$ is Eulerian if and only if k is an even number. Moreover, Player II has an obvious winning strategy for the k-pebble game on $\mathbf{B}_k$ and $\mathbf{B}_{k+1}$. Thus, the property of being Eulerian is a polynomial-time property that is not expressible by any formula of $L^\omega_{\infty\omega}$.
∎

In each of the preceding examples the winning strategy of Player II was quite obvious. More sophisticated applications of pebble games have appeared in several places in the literature, including [Imm82, CFI89, LM89, KV90b], where this method has been used successfully to establish limitations of the expressive power of various logics.

Problem: We conclude this section by presenting an open problem. We showed earlier that 2-colorability is a property expressible in fixpoint logic and, consequently, it is definable by a sentence of $L^{\omega}_{\infty\omega}$. It is not known, however, whether or not 3-colorability is expressible in $L^{\omega}_{\infty\omega}$.

3 0-1 Laws for Infinitary Logics

Let σ be a vocabulary consisting of finitely many relation symbols *only* and let C be the class of all finite structures over σ with universe an initial segment $\{1, 2, ..., n\}$ of the integers for some $n \geq 1$.

If P is a property of (some) structures in C, then the *(labeled) asymptotic probabilty* $\mu(P)$ on C is defined to be equal to the limit as $n \to \infty$ of the fraction of structures in C of cardinality n which satisfy P, provided this limit exists. If L is a logic, we say that the 0-1 *law holds for* L in case $\mu(P)$ exists and is equal to 0 or 1 for every property P expressible in the logic L.

In the past, 0-1 laws for various logics L were proved by establishing first a *transfer theorem for* L of the following kind:

> *There is an infinite structure* $\mathbf{R}$ *over the vocabulary* σ *such that for any property* P *expressible in* L *we have:*

$$\mathbf{R} \textit{ satisfies } P \iff \mu(P) = 1 \textit{ on } C.$$

This method was discovered by Fagin [Fag76] in his proof of the 0-1 law for first-order logic on finite structures. It was also used later in [BGK85] to establish the 0-1 law for positive-fixpoint logic and in [KV87, KV90a] to show that the 0-1 law holds iterative logic (partial-fixpoint logic) and for certain fragments of second-order logic.

It turns out that there is a countable structure $\mathbf{R}$ over the vocabulary σ that satisfies the above equivalence for all these logics. Moreover, this structure $\mathbf{R}$ is unique up to isomorphism. We call $\mathbf{R}$ the *countable random structure over the vocabulary* σ. The random structure $\mathbf{R}$ is characterized by an infinite set of *extension axioms*, which, intuitively, assert that every *type* can be *extended* to any other possible type. The precise definitions are as follows.

Definition 3.1: Let σ be a vocabulary consisting of relation symbols only.

- If $\mathbf{x} = (x_1, ..., x_m)$ is a sequence of distinct variables, then a *type* $t(\mathbf{x})$ *in the variables* $\mathbf{x}$ *over* σ is the conjunction of all the formulas in a *maximally consistent* set S of equalities $x_i = x_j$, inequalities $x_i \neq x_j$, atomic formulas in the variables $\mathbf{x}$, and negated atomic formulas in the variables $\mathbf{x}$.

- Let z be a variable that is different from all the variables in $\mathbf{x}$. We say that *a type* $t(\mathbf{x}, z)$ *extends the type* $s(\mathbf{x})$ if every conjunct of $s(\mathbf{x})$ is also a conjunct of $t(\mathbf{x}, z)$.

- With each pair $s(\mathbf{x})$ and $t(\mathbf{x}, z)$ of types such that t extends s we associate a first-order *extension axiom* $\tau_{s,t}$ stating that

$$(\forall \mathbf{x})(s(\mathbf{x}) \rightarrow (\exists z)t(\mathbf{x}, z)).$$

∎

Let T be the set of all extension axioms. The theory T was studied by Gaifman [Gai64], who showed, using a *back and forth* argument, that any two countable models of T are isomorphic (*T is an ω-categorical theory*). Fagin [Fag76] realized that the extension axioms are relevant to the study of probabilities on finite structures and proved that on the class C of all finite structures over a finite vocabulary σ

$$\mu(\tau_{s,t}) = 1$$

for any extension axiom $\tau_{s,t}$. The equivalence between truth on $\mathbf{R}$ and almost sure truth on C (and consequently the 0-1 law for first-order logic on finite structures) follows from these two results by an application of the compactness theorem.

In proving the 0-1 law for positive-fixpoint logic, Blass, Gurevich and Kozen [BGK85] used the 0-1 law for first-order logic together with a well known model-theoretic characterization of ω-categorical theories, due to Engeler [Eng59], Ryll-Nardzewski [RN59], Svenonius [Sve59], and Vaught [Vau61]. This characterization asserts that a set Σ of first-order sentences has a unique (up to isomorphism) countable model if and only if for every n there are only finitely many inequivalent first-order formulas with n free variables in the models of Σ. In [KV87] we obtained the 0-1 law for iterative logic (partial-fixpoint logic) by employing a model theoretic argument similar to the one in [BGK85] for positive-fixpoint logic. We give a sketch of this argument next.

Let $\varphi(x_1, \ldots, x_n, S)$ be a first-order formula such that $x_1, \ldots, x_n$ are its free variables and S is a n-ary relation symbol not in the vocabulary σ. Let Φ be the operator associated with $\varphi(x_1, \ldots, x_n, S)$, and let Φ^m be the m-th stage of Φ, $m \geq 1$. Recall that, by Theorem 2.8, each stage Φ^m of Φ is definable by a formula $\varphi^m(x_1, \ldots, x_n)$ of first-order logic. Since the random structure $\mathbf{R}$ is a model of the ω-categorical theory T of all extension axioms, it follows that there are only finitely many inequivalent first-order formulas with n free variables over $\mathbf{R}$. Thus, there are integers $N < N'$ such that

$$\mathbf{R} \models (\forall x_1) \ldots (\forall x_n)(\varphi^N(x_1, \ldots, x_n) \leftrightarrow \varphi^{N'}(x_1, \ldots, x_n)).$$

Let N, N' be the smallest such integers. Note that if $N' = N + 1$, then φ converges on $\mathbf{R}$ in N stages. Otherwise, φ diverges on $\mathbf{R}$, and we have that

$$\mathbf{R} \not\models (\forall x_1) \ldots (\forall x_n)(\varphi^N(x_1, \ldots, x_n) \leftrightarrow \varphi^{N+1}(x_1, \ldots, x_n)).$$

In the first case φ^∞ is equivalent to φ^N, and in the second case φ^∞ is equivalent to $false$. Thus, φ^∞ is equivalent to a first-order formula, and this fact is witnessed by first-order sentences. The transfer theorem for first-order logic implies that the witness sentences that hold in $\mathbf{R}$ are true almost everywhere on the class C of all finite structures over σ. Thus, partial-fixpoint logic collapses to first-order logic on almost all finite-structures. The 0-1 law for partial-fixpoint logic is now obtained immediately from the 0-1 law for first-order logic. Moreover, the transfer theorem extends to partial-fixpoint logic as well.

We note that the proof of the 0-1 law for certain fragments of second-order logic in [KV87, KV90a] used among others the 0-1 law for first-order logic, the compactness theorem, and further model-theoretic properties of the logics considered.

In what follows here we show that the 0-1 law holds for the infinitary logic $L^{\omega}_{\infty\omega}$ on finite structures and give three different proofs. The first proof extends the proofs in [BGK85, KV87] in their use of ω-categoricity and their appeal to the 0-1 law for first-order logic. In contrast, neither of the other two proofs employs any "infinitistic" methods. The first of these proofs uses a *quantifier-elimination* method, while the second one uses the pebble games of the previous section and their relation to $L^k_{\infty\omega}$-equivalence. These proofs do not assume the 0-1 law for first-order logic, they do not involve the random structure $\mathbf{R}$ or any other infinite structure, and they do not make use of compactness or of any of its consequences. Moreover, the 0-1 law is derived directly without establishing a transfer theorem first.

The results reported here on the one hand subsume the earlier ones in [Fag76], [BGK85], and [KV87], and on the other hand provide a unifying treatment of 0-1 laws for first-order logic and its extensions with fixpoint operators or infinitary syntax.

3.1 0-1 Laws via a Transfer Theorem

The proof of the 0-1 law for partial-fixpoint logic described earlier used the fact that partial-fixpoint logic collapses to first-order logic on the random structure $\mathbf{R}$, and furthermore, this collapse is witnessed by a first-order sentence. Now it is easy to see that $L^k_{\infty\omega}$ also collapses to $L^k_{\omega\omega}$ over the random structure $\mathbf{R}$, since, by the aforementioned characterization of ω-categorical theories, every infinite disjunction of $L^k_{\omega\omega}$-formula has only a finite number of nonequivalent disjuncts. What is not immediately obvious is that this collapse is witnessed by a first-order sentence. Nevertheless, this turns out to be the case.

Lemma 3.2: *For every $k > 0$, there is a first-order sentence ψ_k such that*

1. *$\mathbf{R} \models \psi_k$, and*
2. *for any $L^k_{\infty\omega}$-formula φ, there is an $L^k_{\omega\omega}$-formula φ' such that $\psi_k \models \varphi \leftrightarrow \varphi'$.*

Proof: For technical convenience assume that only the variables $x_1, \ldots, x_k$ are used in formulas of $L^k_{\infty\omega}$.

Since $\mathbf{R}$ is a model of the ω-categorical theory T of all extension axioms, it follows that there are only finitely many inequivalent formulas of $L^k_{\omega\omega}$ over $\mathbf{R}$. Let $\alpha_1, \ldots, \alpha_m$ be representatives from the equivalence classes of formulas. Note that this collection of formulas must express the atomic formulas and be closed, up to equivalence, under negation, disjunction, and existential quantification. That is, for each atomic formula $p(x_1, \ldots, x_k)$ there is some α_i such that $\mathbf{R} \models (\forall x_1 \ldots \forall x_k)(p(x_1, \ldots, x_k) \leftrightarrow \alpha_i)$. Also, for each α_i there exists $\alpha_{i'}$ such that $\mathbf{R} \models (\forall x_1 \ldots \forall x_k)(\neg \alpha_i \leftrightarrow \alpha_{i'})$. Similarly, for each α_i, α_j there exists α_k such that $\mathbf{R} \models (\forall x_1 \ldots \forall x_k)(\alpha_i \vee \alpha_j \leftrightarrow \alpha_k)$. Finally, for each α_i and variable x_j there exists α_l such that $\mathbf{R} \models (\forall x_1 \ldots \forall x_k)(\exists x_j \alpha_i \leftrightarrow \alpha_l)$.

We call the above sentences the *closure axioms*. Let ψ_k be the conjunction of all the closure axioms. Notice that ψ_k is a sentence of $L^k_{\omega\omega}$, because there are only finitely many inequivalent formulas of $L^k_{\omega\omega}$. Moreover, it is clear that $\mathbf{R} \models \psi_k$. It remains to prove the second claim. We show that if ψ_k is taken as an axiom, then each formula of $L^k_{\infty\omega}$ is equivalent to one of the α_i's. The proof is by induction on the structure of formulas of $L^k_{\infty\omega}$, assuming that formulas are built using negation, infinite disjunctions, and existential quantification.

Let φ be a formula of $L^k_{\infty\omega}$. If φ is an atomic formula, then by the closure axioms it is equivalent to an α_i. If φ is $\neg\rho$ or $\exists x_j \rho$, then by the induction hypothesis ρ is equivalent to an α_i, and, by the closure axioms, φ is equivalent to an α_i. Finally, if φ is the infinite disjunction $\bigvee \varphi_j$ of $L^k_{\omega\omega}$-formulas, then by the induction hypothesis each φ_j is equivalent to some α_i, and by the closure axioms φ is equivalent to some α_i. ∎

The transfer theorem for $L^\omega_{\infty\omega}$ follows from Lemma 3.2.

Theorem 3.3: *If φ is a sentence of $L^\omega_{\infty\omega}$, then the following are equivalent:*

1. $\mu(\varphi) = 1$.
2. $\mathbf{R} \models \varphi$.

Proof: Let φ be a sentence of $L^k_{\infty\omega}$. By Lemma 3.2 and the 0-1 law for first-order logic we have that $\mu(\psi_k) = 1$ and, for any $L^k_{\infty\omega}$-sentence φ, there is an $L^k_{\omega\omega}$-sentence φ' such that $\psi_k \models \varphi \leftrightarrow \varphi'$. If $\mathbf{R} \models \varphi$, then $\mathbf{R} \models \varphi'$, and by the 0-1 law for first-order logic we have that $\mu(\varphi') = 1$. It follows that $\mu(\varphi) = 1$. If $\mathbf{R} \not\models \varphi$, then $\mathbf{R} \models \neg\varphi$, and $\mu(\varphi) = 0$. ∎

The 0-1 law for $L^\omega_{\infty\omega}$ is an immediate consequence of the transfer theorem.

Theorem 3.4: *The 0-1 law holds for the infinitary logic $L^\omega_{\infty\omega}$, i.e., if ψ is a sentence of $L^\omega_{\infty\omega}$, then the asymptotic probability $\mu(\psi)$ exists and is equal to either 0 or 1.*

Remark 3.5: The 0-1 law for $L^\omega_{\infty\omega}$ has also certain immediate applications to definability theory. For example, the property "there is an even number of elements" is not expressible in $L^\omega_{\infty\omega}$, because it does not have an asymptotic probability. This fact was obtained earlier in Example 2.22 using k-pebble games. ∎

Remark 3.6: There is an extensive literature on 0-1 laws for first-order logic on restricted classes of finite structures (cf. [Com88a] for a survey of results in this area). The method developed in Lemma 3.2 and Theorem 3.3 applies to arbitrary classes $\mathcal{C}$ of finite structures and yields the 0-1 law for $L^\omega_{\infty\omega}$ on $\mathcal{C}$, provided the set of first-order sentences with probability 1 on $\mathcal{C}$ is an $\aleph_0$-categorical theory. This is, for example, the case with the class of partial orders investigated by Compton [Com88b], and the class of K_{l+1}-free graphs investigated by Kolaitis, Prömel, and Rothschild [KPR87].[3] ∎

3.2 0-1 Laws via Quantifier-Elimination

Glebskii et al. [GKLT69] proved the 0-1 law for first-order logic, independently of Fagin [Fag76], using what amounts to a certain quantifier-elimination method. We use here a different quantifier-elimination method that has its origin in Grandjean's work [Gra83] on the computational complexity of the 0-1 law for first-order logic.

[3] We thank an anonymous referee for making the observations contained in this remark.

Definition 3.7: If k is a positive integer, then we write θ_k for the conjunction of all extension axioms $\tau_{s,t}$ with at most k variables. ∎

Notice that each θ_k is a sentence of $L^k_{\omega\omega}$, i.e., it is a first-order sentence with at most k distinct variables.

Theorem 3.8: *Let k and m be two positive integers such that $m \leq k$. If $s(x_1, \ldots, x_m)$ is a type over the vocabulary σ and $\varphi(x_1, \ldots, x_m)$ is a formula of $L^k_{\infty\omega}$ with free variables among $\mathbf{x} = (x_1, \ldots, x_m)$, then exactly one of the following two statements holds:*

1. $\theta_k \models (\forall \mathbf{x})(s(\mathbf{x}) \to \varphi(\mathbf{x}))$.
2. $\theta_k \models (\forall \mathbf{x})(s(\mathbf{x}) \to \neg\varphi(\mathbf{x}))$.

Proof: This theorem will be proved by induction on the construction of formulas in $L^k_{\infty\omega}$ whose variables are among $x_1, \ldots, x_k$ and whose free variables are among $x_1, \ldots, x_m$, simultaneously for all $m \leq k$ and for all types $s(x_1, \ldots, x_m)$. A crucial use of the extension axioms will be made in the case where the formula $\varphi(\mathbf{x})$ starts with an existential quantifier.

The base case of the induction (equalities and atomic formulas) and the induction step for the negation ($\neg$) are obvious. Assume that $\varphi(x_1, \ldots, x_m)$ is an infinitary conjunction $\bigwedge \Psi$ of formulas $\psi(x_1, \ldots, x_m)$ of $L^k_{\infty\omega}$. By induction hypothesis, for each $\psi \in \Psi$ either

- $\theta_k \models (\forall \mathbf{x})(s(\mathbf{x}) \to \psi(\mathbf{x}))$

or

- $\theta_k \models (\forall \mathbf{x})(s(\mathbf{x}) \to \neg\psi(\mathbf{x}))$.

If there is a formula $\psi \in \Psi$ such that $\theta_k \models (\forall \mathbf{x})(s(\mathbf{x}) \to \neg\psi(\mathbf{x}))$, then

$$\theta_k \models (\forall \mathbf{x})(s(\mathbf{x}) \to \neg \bigwedge \Psi(\mathbf{x}));$$

otherwise,

$$\theta_k \models (\forall \mathbf{x})(s(\mathbf{x}) \to \bigwedge \Psi(\mathbf{x})).$$

Assume next that $\varphi(x_1, \ldots, x_m)$ is the formula $(\exists z)\psi(x_1, \ldots, x_m, z)$ and that the induction hypothesis holds for $\psi(x_1, \ldots, x_m, z)$. If

$$\theta_k \models (\forall \mathbf{x})(s(\mathbf{x}) \to \neg(\exists z)\psi(\mathbf{x}, z)),$$

then (2) holds for $\varphi(x_1, \ldots, x_m)$. Otherwise,

$$\theta_k \not\models (\forall \mathbf{x})(s(\mathbf{x}) \to \neg(\exists z)\psi(\mathbf{x}, z)).$$

We will show that in the latter case

$$\theta_k \models (\forall \mathbf{x})(s(\mathbf{x}) \to (\exists z)\psi(\mathbf{x}, z).$$

Notice that, by our assumption about the variables of $\varphi(x_1, \ldots, x_m)$, we must have that the variable z is the variable x_j, for some j such that $1 \leq j \leq k$. We now distinguish two cases, namely the case where $j > m$ and the case where $j \leq m$.

Case 1: $j > m$, which means that the variable z is different from all the variables $x_1, \ldots, x_m$. Notice that in this case m must be less than k.

Since $\theta_k \not\models (\forall \mathbf{x})(s(\mathbf{x}) \rightarrow \neg(\exists z)\psi(\mathbf{x}, z))$, there is a structure $\mathbf{D}$ over σ such that $\mathbf{D} \models \theta_k$ and

$$\mathbf{D} \models (\exists \mathbf{x})(s(\mathbf{x}) \wedge (\exists z)\psi(\mathbf{x}, z)).$$

Let $a_1, \ldots, a_m, b$ be elements of the universe D of $\mathbf{D}$ such that

$$\mathbf{D} \models s(a_1, \ldots, a_m) \wedge \psi(a_1, \ldots, a_m, b).$$

Let $t(x_1, \ldots, x_m, z)$ be the unique type such that $\mathbf{D} \models t(a_1, \ldots, a_m, b)$, i.e., $t(x_1, \ldots, x_m, z)$ is the conjunction of all equalities, inequalities, atomic formulas, and negated atomic formulas in the variables $x_1, \ldots, x_m, z$ satisfied by $a_1, \ldots, a_m, b$. Notice that the type $t(x_1, \ldots, x_m, z)$ extends the type $s(x_1, \ldots, x_m)$. We also have that

$$\mathbf{D} \models (\exists x_1 \ldots \exists x_m)(\exists z)(t(x_1, \ldots, x_m, z) \wedge \psi(x_1, \ldots, x_m, z)).$$

By applying the induction hypothesis to the formula $\psi(x_1, \ldots, x_m, z)$ of $L^k_{\infty\omega}$ and to the type $t(x_1, \ldots, x_m, z)$, we infer that

$$\theta_k \models (\forall \mathbf{x})(\forall z)(t(\mathbf{x}, z) \rightarrow \psi(\mathbf{x}, z)).$$

Since the type t is an extension of the type s and θ_k is the conjunction of all extension axioms with at most k variables, it follows that

$$\theta_k \models (\forall \mathbf{x})(s(\mathbf{x}) \rightarrow (\exists z)t(\mathbf{x}, z)).$$

We can now conclude that

$$\theta_k \models (\forall \mathbf{x})(s(\mathbf{x}) \rightarrow (\exists z)\psi(\mathbf{x}, z)).$$

Case 2: $j \leq m$, which means that the variable z is the variable x_j for some $j \leq m$. In this case, the free variables of the formula ψ are among the variables $x_1, \ldots, x_m$ and, moreover, we have that there is a structure $\mathbf{D}$ over σ such that $\mathbf{D} \models \theta_k$ and

$$\mathbf{D} \models (\exists x_1 \ldots \exists x_m)(s(x_1, \ldots, x_m) \wedge (\exists x_j)\psi(x_1, \ldots, x_m)).$$

Let $a_1, \ldots, a_m$ be elements of the universe D of $\mathbf{D}$ such that

$$\mathbf{D}, a_1, \ldots, a_m \models s(x_1, \ldots, x_m) \wedge (\exists x_j)\psi(x_1, \ldots, x_m),$$

let $\mathbf{x}^*$ be the sequence of variables $x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_m$, and let $s^*(\mathbf{x}^*)$ be the unique type such that $\mathbf{D} \models s^*(a_1, \ldots, a_{j-1}, a_{j+1}, \ldots, a_m)$. Then there is an element b of the universe D of $\mathbf{D}$ such that

$$\mathbf{D} \models s^*(a_1, \ldots, a_{j-1}, a_{j+1}, \ldots, a_m) \wedge \psi(a_1, \ldots, a_{j-1}, b, a_{j+1}, \ldots, a_m).$$

Let $t^*(x_1, \ldots, x_j, \ldots, x_m)$ be the unique type such that $\mathbf{D} \models t^*(a_1, \ldots, a_{j-1}, b, a_{j+1} \ldots, a_m)$. Notice that the type $t^*(x_1, \ldots, x_m)$ extends the type $s^*(\mathbf{x}^*)$ and that

$$\mathbf{D} \models (\exists x_1 \ldots \exists x_m)(t^*(x_1, \ldots, x_m) \wedge \psi(x_1, \ldots, x_m)).$$

By applying the induction hypothesis to the formula $\psi(x_1,\ldots,x_m)$ of $L^k_{\infty\omega}$ and to the type $t^*(x_1,\ldots,x_m)$, we infer that

$$\theta_k \models (\forall x_1 \ldots \forall x_m)(t^*(x_1,\ldots,x_m) \rightarrow \psi(x_1,\ldots,x_m)).$$

Since the type $t^*(x_1,\ldots,x_m)$ is an extension of the type $s^*(\mathbf{x}^*)$ and θ_k is the conjunction of all extension axioms with at most k variables, it follows that

$$\theta_k \models (\forall \mathbf{x}^*)(s^*(\mathbf{x}^*) \rightarrow (\exists x_j)t^*(x_1,\ldots,x_m)).$$

We can now conclude that

$$\theta_k \models (\forall \mathbf{x}^*)(s^*(\mathbf{x}^*) \rightarrow (\exists x_j)\psi(x_1,\ldots,x_m))$$

and, consequently,

$$\theta_k \models (\forall \mathbf{x})(s(\mathbf{x}) \rightarrow (\exists x_j)\psi(x_1,\ldots,x_m)),$$

since $(\forall \mathbf{x})(s(\mathbf{x}) \rightarrow s^*(\mathbf{x}^*))$ is valid.
∎

Corollary 3.9: *If ψ is a sentence of $L^k_{\infty\omega}$, then either $\theta_k \models \psi$ or $\theta_k \models \neg\psi$. As a result, if $\mathbf{A}$ and $\mathbf{B}$ are two models of θ_k, then $\mathbf{A} \equiv^k_{\infty\omega} \mathbf{B}$.*

The first-order version of the preceding Theorem 3.8 was obtained by Grandjean [Gra83], while Immerman [Imm82] established Corollary 3.9 for sentences of $L^k_{\omega\omega}$.

We now have all the machinery needed to establish the 0-1 law.

Theorem 3.10: *The 0-1 law holds for the infinitary logic $L^\omega_{\infty\omega}$, i.e., if ψ is a sentence of $L^\omega_{\infty\omega}$, then the asymptotic probability $\mu(\psi)$ exists and is equal to either 0 or 1.*

Proof: If ψ is a sentence of $L^k_{\infty\omega}$, for some $k \geq 1$, then, by Corollary 3.9,

$$\theta_k \models \psi \text{ or } \theta_k \models \neg\psi.$$

In the first case we have that $\mu(\psi) = 1$ and in the second $\mu(\neg\psi) = 1$, because $\mu(\theta_k) = 1$. The latter holds, because θ_k is a finite conjunction of extension axioms and, as Fagin [Fag76] showed, $\mu(\tau_{s,t}) = 1$ for each extension axiom $\tau_{s,t}$. ∎

We can also easily derive a *transfer theorem* for each infinitary logic $L^k_{\infty\omega}$, $k \geq 1$.

Theorem 3.11: *Let k be a positive integer, and let $\mathbf{B}$ be a model of θ_k. If ψ is a sentence of $L^k_{\infty\omega}$, then the following are equivalent:*

1. $\mu(\psi) = 1$.
2. $\theta_k \models \psi$.
3. $\mathbf{B} \models \psi$.

Proof: Let ψ be a sentence of $L^k_{\infty\omega}$ such that $\mu(\psi) = 1$. Then $\theta_k \models \psi$, since, otherwise, it would follow from Corollary 3.9 that $\theta_k \models \neg\psi$, which in turn yields that $\mu(\neg\psi) = 1$ and $\mu(\psi) = 0$. It is obvious that if $\theta_k \models \psi$, then $\mathbf{B} \models \psi$. Finally, assume that $\mathbf{B} \models \psi$. Then $\mu(\psi) = 1$, since, otherwise, by the 0-1 law for $L^k_{\infty\omega}$, we would conclude that $\mu(\neg\psi) = 1$, which implies that $\mathbf{B} \models \neg\psi$. ∎

Corollary 3.12: *If* $\mathbf{R}$ *is the countable random structure over the vocabulary* σ *and* ψ *is a sentence of* $L^\omega_{\infty\omega}$, *then*

$$\mu(\psi) = 1 \iff \mathbf{R} \models \psi.$$

Proof: The random structure $\mathbf{R}$ is a model of each θ_k, $k \geq 1$. ∎

Notice that each θ_k has both finite and infinite models; actually, an arbitrary finite structure over σ is a model of θ_k with probability 1. In particular, there are infinitely many countable models satisfying the transfer theorem for the infinitary logic $L^k_{\infty\omega}$. This should be contrasted with the situation in $L^\omega_{\infty\omega}$, where the random structure $\mathbf{R}$ is the unique countable structure satisfying the transfer theorem, since $L^\omega_{\infty\omega}$ includes all the extension axioms, which have a unique countable model.

Although we used earlier the term "quantifier-elimination method", we did not actually justify this terminology. We conclude this section by establishing a *quantifier-elimination* theorem for $L^k_{\infty\omega}$ on models of θ_k, which strengthens Lemma 3.3.

Theorem 3.13: *Let* k *be a positive integer and let* $\varphi(x_1, \ldots, x_m)$ *be a formula of* $L^k_{\infty\omega}$ *with free variables among* $\mathbf{x} = (x_1, \ldots, x_m)$. *Then there is a quantifier-free formula* $\chi(x_1, \ldots, x_m)$ *of* $L^k_{\omega\omega}$ *such that*

$$\theta_k \models (\forall \mathbf{x})(\varphi(\mathbf{x}) \leftrightarrow \chi(\mathbf{x})).$$

Proof: Let X_φ be the set of all types $s(x_1, \ldots, x_m)$ for which there is a structure $\mathbf{D}$ such that

$$\mathbf{D} \models \theta_k \wedge (\exists \mathbf{x})(s(\mathbf{x}) \wedge \varphi(\mathbf{x})).$$

We claim that the required formula $\chi(x_1, \ldots, x_m)$ is

$$\bigvee_{s \in X_\varphi} s(x_1, \ldots, x_m).$$

Notice first that $\chi(x_1, \ldots, x_m)$ is a quantifier-free formula of $L^k_{\omega\omega}$, because the vocabulary σ is finite and, as a result, there are finitely many distinct types in the variables $x_1, \ldots, x_m$.

Moreover, it follows from the definitions that

$$\theta_k \models (\forall x)(\varphi(\mathbf{x}) \rightarrow \chi(\mathbf{x})).$$

For the other direction, let $\mathbf{D}$ be a model of θ_k, let $a_1, \ldots, a_m$ be elements from the universe D of $\mathbf{D}$, and let $s(x_1, \ldots, x_m)$ be a type in the set X_φ such that

$$\mathbf{D} \models s(a_1, \ldots, a_m).$$

By Theorem 3.8, exactly one of the following two statements holds:

1. $\theta_k \models (\forall \mathbf{x})(s(\mathbf{x}) \rightarrow \varphi(\mathbf{x}))$.
2. $\theta_k \models (\forall \mathbf{x})(s(\mathbf{x}) \rightarrow \neg\varphi(\mathbf{x}))$.

Since $s(x_1, \ldots, x_m)$ is a type in the set X_φ, the second statement (2) is ruled out and, hence,

$$\mathbf{D} \models \varphi(a_1, \ldots, a_m).$$

∎

Remark 3.14: Notice that the proof of Theorem 3.10 implies in particular that if a sentence ψ of $L^k_{\infty\omega}$ is true almost everywhere, then there is a first-order sentence that is true almost everywhere and logically implies ψ. Blass and Harary [BH79] showed that there is no first-order sentence that is true almost everywhere and logically implies Hamiltonicity. We can, therefore, conclude that there is no property of $L^\omega_{\infty\omega}$ that is true almost everywhere and logically implies Hamiltonicity. In other words, there is *no* sufficient condition for Hamiltonicity which is expressible in $L^\omega_{\infty\omega}$ and has asymptotic probability equal to 1. This can be viewed as a strengthening of the earlier fact in Example 2.23 that Hamiltonicity is not expressible in $L^\omega_{\infty\omega}$. ∎

3.3 0-1 Laws via Pebble Games

Let k be a fixed positive integer. If $\mathbf{A}$ is a finite structure, then we write $[\mathbf{A}]$ for *the equivalence class* of $\mathbf{A}$ with respect to the equivalence relation $\equiv^k_{\infty\omega}$. In what follows we will show that there is a tight connection between 0-1 laws and the asymptotic probabilities of equivalence classes $[\mathbf{A}]$. Actually, this turns out to be a general fact that holds for arbitrary probability measures.

So far all the results presented here are about the *uniform* probability measures on $\mathcal{C}$, i.e., all structures with n elements carry the same probability. There is, however, a well developed study of random structures under *variable* probability measures. This started with the work of Erdös and Rényi [ER60] and is presented in detail in Bollobás [Bol85]. In general, for each $n \geq 1$ one has a probability measure pr_n on all structures in $\mathcal{C}$ with n elements, where pr_n may be a non-uniform distribution. The *asymptotic probability* $pr(P)$ of a property P (*relative to the probability measures* pr_n, $n \geq 1$) is defined by $pr(P) = \lim_{n\to\infty} pr_n(P)$, provided this limit exists. If L is a logic, then we say that a *0-1 law holds for* L *relative to the measure* pr if for every sentence ψ of L the asymptotic probability $pr(\psi)$ exists and is either 0 or 1. Notice that, strictly speaking, pr is not a probability measure, because it is not countably additive (it is, however, finitely additive).

Spencer and Shelah [SS88] investigated 0-1 laws for first-order logic under variable probability measures on the class of undirected graphs. They obtained a classification of the probability measures for which the first-order 0-1 law holds. We establish next a necessary and sufficient condition for the existence of 0-1 laws for $L^k_{\infty\omega}$ under arbitrary probability measures.

Theorem 3.15: *Let K be a class of finite structures over the vocabulary σ, let k be a positive integer, and let pr_n, $n \geq 1$, be a sequence of probability measures on the structures in K with n elements. Then the following are equivalent:*

1. *The 0-1 law holds for the infinitary logic $L^k_{\infty\omega}$ relative to the measure pr.*

2. *There is an equivalence class C of the equivalence relation $\equiv^k_{\infty\omega}$ such that $pr(C) = 1$.*

Proof: Assume that $pr(C) = 1$ for some equivalence class C of $\equiv^k_{\infty\omega}$ and let ψ be a sentence of $L^k_{\infty\omega}$. If ψ holds for the structures in C, then $pr(\psi) = 1$, because the set of models of ψ contains C. If, on the other hand, ψ fails for the structures in C, then $pr(\neg\psi) = 1$ and, hence, $pr(\psi) = 0$.

In the other direction, we show that if the 0-1 law held for $L^k_{\infty\omega}$ relative to a measure, but every equivalence class C of $\equiv^k_{\infty\omega}$ had probability 0, then we could find a sentence of $L^k_{\infty\omega}$ whose probability is neither 0 nor 1. To see this, let $C_0, C_1, \ldots$ be an enumeration of the equivalence classes of $\equiv^k_{\infty\omega}$ on K, and let C^n_j be the set of n-element structures in C_j. Note that for all $n > 0$ there exists some integer m such that $C^n_{m'} = \emptyset$ for all $m' \geq m$, since there are finitely many n-element structures. Let $m_1 < m_2 < \ldots$ be an increasing sequence such that $C^n_{m'} = \emptyset$ for all $m' \geq m_n$.

We denote by N the set of nonnegative integers, and we denote by $[0, j)$ the set $\{0, \ldots, j-1\}$. For any set $X \subseteq N$, let $C_X = \bigcup_{i\in X} C_i$. We define $pr_n(X)$ (resp., $pr(X)$) to be $pr_n(C_X)$ (resp. $pr(C_X)$). We are going to use the following three properties:

1. If X is finite, then $pr(X) = 0$, since by assumption $pr(C_i) = 0$ for all $i \geq 0$.

2. $pr_n([0, m_n)) = 1$, since $pr_n(N) = 1$ and $C^n_m = \emptyset$ for all $m \geq m_n$.

3. If $X \subseteq [0, m_n)$ and $Z \subseteq N - [0, m_n)$, then $pr_l(X) = pr_l(X \cup Z)$ for all $l \leq n$, since $C^l_m = \emptyset$ for all $m \in Z$ and $l \leq n$.

We construct a set X of integers such that for infinitely many i's we have that $pr_i(X) > 3/4$ and infinitely many i's we have that $pr_i(X) < 1/4$. It follows that $pr(X)$ is undefined. X is constructed in stages. In the i-th stage we define a nonnegative integer n_i and a pair of finite disjoint sets $X_i, Y_i \subseteq [0, m_{n_i})$ such that the following hold:

1. $X_i \subseteq X_{i+1}$, $Y_i \subseteq Y_{i+1}$, and $X_i \cup Y_i = [0, m_{n_i})$;

2. if i is odd, then $pr_{n_i}(X_i) > 3/4$, and if i is even then $pr_{n_i}(X_i) < 1/4$.

The desired set X is simply $\bigcup_i X_i$. We now define the sets X_i and Y_i by induction. For $i = 1$, let $n_1 = 1$, $X_1 = [0, m_1)$, and $Y_1 = \emptyset$. Then $pr_1(X_1) = 1$. Assume inductively, that n_i, X_i and Y_i have been defined. There are two cases now.

Case 1: If i is odd, then $pr_{n_i}(X_i) > 3/4$. Since $pr(X_i) = 0$, there is an integer $q > n_i$ such that $pr_q(X_i) < 1/4$. We let $n_{i+1} = q$, $X_{i+1} = X_i$, and $Y_{i+1} = [0, m_q) - X_i$.

Case 2: If i is even, then $pr_{n_i}(X_i) < 1/4$. Since $pr(Y_i) = 0$, there is an integer $q > n_i$, such that $pr_q([0, m_q) - Y_i) > 3/4$. We let $n_{i+1} = q$, $X_i = [0, m_q) - Y_{i+1}$, and $Y_{i+1} = Y_i$.

Now let $X = \bigcup_i X_i$. It is easy to see that X is disjoint from Y_i, for all $i \geq 1$. Since $X_i \cup Y_i = [0, m_{n_i})$, it follows that $pr_{n_i}(X) = pr_{n_i}(X_i)$. It follows that there are infinitely many i's such $pr_{n_i}(X) > 3/4$, and there infinitely many i's such that $pr_{n_i}(X) < 1/4$. Thus, $pr(X)$ is undefined. Using Proposition 2.11, we can construct a sentence φ_X of $L^k_{\infty\omega}$ that defines the class C_X. It follows that $pr(\varphi_X)$ is undefined. ∎

Remark 3.16:

1. Notice that if the 0-1 law holds for $L^k_{\infty\omega}$ relative to a measure pr, then there is exactly one equivalence class C of $\equiv^k_{\infty\omega}$ such that $pr(C) = 1$. All other equivalence classes of $\equiv^k_{\infty\omega}$ have probability 0.

2. We should also point out that the preceding theorem does not hold in general for arbitrary logics. For example, if L is first-order logic and pr is the uniform measure μ, then the 0-1 law holds for L, but each equivalence class of L has probability 0, because every finite structure is described up to isomorphism by a first-order sentence.

   The crucial property of $L^k_{\infty\omega}$ used in the proof is its closure under infinite conjunctions and disjunctions. Let L_r be the fragment of first-order logic consisting of first-order sentences of quantifier depth r. By Fraissé's theorem [Fra54], the relation of elementary equivalence on L_r has finitely many equivalence classes and, consequently, L_r is closed under arbitrary disjunctions and conjunctions. Thus, the analogous version of the above Theorem 3.15 holds for L_r.

3. Spencer [Spe91] obtained 0-1 laws for first-order logic with respect to the class of undirected graphs relative to certain variable probability measures by examining first-order sentences of fixed quantifier depth and using Ehrenfeucht-Fraissé games. The idea of using games to obtain 0-1 laws seems to originate with the work of Lynch [Lyn80] (cf. also Compton [Com88a]).

∎

We now return to the uniform measure μ on $L^k_{\infty\omega}$ and give a different proof of the 0-1 law for $L^k_{\infty\omega}$ using the preceding Theorem 3.15 and the characterization of $\equiv^k_{\infty\omega}$ in terms of pebble games.

Theorem 3.17: *Let $\mathcal{C}$ be the class of all finite structures, let k be a positive integer, and let θ_k be the conjunction of all extension axioms with at most k variables. If $\mathbf{A}$ is a finite structure that is a model of θ_k, then $\mu([\mathbf{A}]) = 1$. As a result, the 0-1 law holds for $L^k_{\infty\omega}$ relative to the uniform measure on $\mathcal{C}$.*

Proof: If $\mathbf{A}$ and $\mathbf{B}$ are both models of θ_k, then it is easy to verify that Player II has a winning strategy in the k-pebble game on $\mathbf{A}$ and $\mathbf{B}$. Intuitively, the winning strategy for Player II is provided by the elements of $\mathbf{A}$ and $\mathbf{B}$ witnessing the extension axioms with at most k variables. We now describe this more formally.

Let $c_1, \ldots, c_l$ and $d_1, \ldots, d_l$ be the interpretations of the constant symbols of the vocabulary σ on $\mathbf{A}$ and $\mathbf{B}$, respectively. We have to show that there is a family $\mathcal{H}$ of partial isomorphisms on $\mathbf{A}$ and $\mathbf{B}$ that provides a winning strategy for Player II in the k-pebble game.

The desired family $\mathcal{H}$ is built by starting with the partial isomorphism that maps c_i to d_i, for $1 \le i \le k$, and taking the closure under subfunctions and back-and-forth extensions, where a back-and-forth-extension is defined as follows.

Let h be a member of $\mathcal{H}$ whose domain is the set $\{c_1, \ldots, c_l, a_1, \ldots, a_m\}$, where $m < k$, and let a be an element that is not in the domain of h. Let $s(x_1, \ldots, x_m)$ and $t(x_1, \ldots, x_m, z)$ be types such that $\mathbf{A} \models s(a_1, \ldots, a_m)$ and $\mathbf{A} \models t(a_1, \ldots, a_m, a)$. Since h is a partial isomorphism, we also have $\mathbf{B} \models s(a_1, \ldots, a_m)$. Consider the extension axiom $\tau_{s,t}$, i.e.,

$$(\forall x_1, \ldots, x_m)(s(x_1, \ldots, x_m) \rightarrow (\exists z)t(x_1, \ldots, x_m, z)).$$

Since $\tau_{s,t}$ uses at most k variables and $\mathbf{B} \models \theta_k$, there exists an element b such that $\mathbf{B} \models t(h(a_1), \ldots, h(a_m), b)$. Thus, $h \cup (a, b)$ is a partial isomorphism, which is added to $\mathcal{H}$. This is the "forth" extension; the "back" extension is defined analogously using the fact that $\mathbf{A} \models \theta_k$.

It follows from Theorem 2.17 that if $\mathbf{A}$ and $\mathbf{B}$ are both models of θ_k, then $\mathbf{A} \equiv^k_{\infty\omega} \mathbf{B}$. Since $\mu(\theta_k) = 1$ [Fag76], it follows that $\mu([\mathbf{A}]) = 1$, for any finite structure $\mathbf{A}$ that is a model of θ_k. ∎

4 Concluding Remarks

We established here the 0-1 law for the infinitary logic $L^\omega_{\infty\omega}$ under the uniform probability measure. It is an interesting open problem to investigate 0-1 laws for fixpoint logics or for infinitary logics under variable probability measures. No results in this direction are known at present, but our Theorem 3.15 provides a handle for attacking this problem.

Previous investigations of 0-1 laws for first-order logic and fixpoint logics examined also the computational complexity of the decision problem for the values of the probabilities, namely the complexity of deciding whether the probability of a sentence is 0 or 1 [Gra83], [BGK85], [KV87]. This problem, however, is computationally meaningful only when the logic under consideration has an effective syntax. Thus, this investigation cannot be carried out for the infinitary logics $L^k_{\infty\omega}$, $k \geq 1$.

Acknowledgements. We are grateful to Ron Fagin for many enlightening discussions on 0-1 laws and helpful comments on previous drafts of this paper. In particular, he pointed out to us the connection between games and 0-1 laws and suggested the current formulation of Theorem 3.15. We also wish to thank Jouko Väänänen for many fruitful conversations and e-mail exchanges in which he made, among other things, the comments contained in Remark 2.13. Finally, we wish to thank an anonymous referee for several useful remarks and suggestions, including Remark 3.6.

References

[Ajt83] M. Ajtai. Σ^1_1 formulae on finite structures. *Ann. of Pure and Applied Logic*, 24:1–48, 1983.

[AU79] A. V. Aho and J. D. Ullman. Universality of data retrieval languages. In *Proc. 6th ACM Symp. on Principles of Programming Languages*, pages 110–117, 1979.

[AV89] S. Abiteboul and V. Vianu. Fixpoint extensions of first-order logic and Datalog-like languages. In *Proc. 4th IEEE Symp. on Logic in Computer Science*, pages 71–79, 1989.

[AV91] S. Abiteboul and V. Vianu. Generic computation and its complexity. In *Proc. 23rd ACM Symp. on Theory of Computing*, pages 209–219, 1991.

[Bar77] J. Barwise. On Moschovakis closure ordinals. *Journal of Symbolic Logic*, 42:292–296, 1977.

[BF85] J. Barwise and S. Feferman, editors. *Model-Theoretic Logics*. Springer-Verlag, 1985.

[BGK85] A. Blass, Y. Gurevich, and D. Kozen. A zero–one law for logic with a fixed point operator. *Information and Control*, 67:70–90, 1985.

[BH79] A. Blass and F. Harary. Properties of almost all graphs and complexes. *Journal of Graph Theory*, 3:225–240, 1979.

[Bol79] B. Bollobas. *Graph Theory*. Springer-Verlag, 1979.

[Bol85] B. Bollobas. *Random Graphs*. Academic Press, 1985.

[CFI89] J. Cai, M. Fürer, and N. Immerman. An optimal lower bound on the number of variables for graph identification. In *Proc. 30th IEEE Symp. on Foundations of Computer Science*, pages 612–617, 1989.

[CH82] A. Chandra and D. Harel. Structure and complexity of relational queries. *Journal of Computer and System Sciences*, 25:99–128, 1982.

[Cha88] A. Chandra. Theory of database queries. In *Proc. 7th ACM Symp. on Principles of Database Systems*, pages 1–9, 1988.

[Com88a] K. J. Compton. 0-1 laws in logic and combinatorics. In I. Rival, editor, *NATO Adv. Study Inst. on Algorithms and Order*, pages 353–383. D. Reidel, 1988.

[Com88b] K. J. Compton. The computational complexity of asymptotic problems I: partial orders. *Information and Computation*, 78:108–123, 1988.

[Coo74] S. A. Cook. An observation of time-storage trade-off. *Journal of Computer and System Sciences*, 9:308–316, 1974.

[dR87] M. de Rougemont. Second-order and inductive definability on finite structures. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 33:47–63, 1987.

[Eng59] E. Engeler. A characterization of theories with isomorphic denumerable models. *Notices of the Amer. Math. Soc.*, 6:161, 1959.

[ER60] P. Erdös and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci.*, 7:17–61, 1960.

[Fag75] R. Fagin. Monadic generalized spectra. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 21:89–96, 1975.

[Fag76] R. Fagin. Probabilities on finite models. *Journal of Symbolic Logic*, 41:50–58, 1976.

[Fra54] R. Fraïssé. Sur quelques classifications des systèmes de relations. *Publ. Sci. Univ. Alger. Sér. A*, 1:35–182, 1954.

[Gai64] H. Gaifman. Concerning measures in first-order calculi. *Israel Journal of Mathematics*, 2:1–18, 1964.

[Gai82] H. Gaifman. On local and nonlocal properties. In J. Stern, editor, *Logic Colloquium '81*, pages 105–135. North Holland, 1982.

[GKLT69] Y. V. Glebskii, D. I. Kogan, M. I. Liogonki, and V. A. Talanov. Range and degree of realizability of formulas in the restricted predicate calculus. *Cybernetics*, 5:142–154, 1969.

[Gra83] E. Grandjean. Complexity of the first–order theory of almost all structures. *Information and Control*, 52:180–204, 1983.

[GS86] Y. Gurevich and S. Shelah. Fixed-point extensions of first-order logic. *Annals of Pure and Applied Logic*, 32:265–280, 1986.

[Gur84] Y. Gurevich. Toward logic tailored for computational complexity. In M. M. Richter et al., editor, *Computation and Proof Theory, Lecture Notes in Mathematics 1104*, pages 175–216. Springer-Verlag, 1984.

[IK89] N. Immerman and D. Kozen. Definability with bounded number of bound variables. *Information and Computation*, 83:121–139, 1989.

[Imm82] N. Immerman. Upper and lower bounds for first-order expressibility. *Journal of Computer and System Sciences*, 25:76–98, 1982.

[Imm86] N. Immerman. Relational queries computable in polynomial time. *Information and Control*, 68:86–104, 1986.

[Kar65] C. Karp. Finite quantifier equivalence. In J. Addison, L. Henkin, and A. Tarski, editors, *The Theory of Models*, pages 407–412. North Holland, 1965.

[Kau87] M. Kaufmann. Counterexample to the 0-1 law for existential monadic second-order logic. CLI Internal Note 32, Computational Logic Inc., December 1987.

[Kny89] V. V. Knyazev. Iterative extensions of first-order logic. *Mathematical Problems in Cybernetics*, pages 123–130, 1989. MR 91a:03062.

[Kol85] Ph. G. Kolaitis. On asymptotic probabilities of inductive queries and their decision problem. In R. Parikh, editor, *Logics of Programs '85, Lecture Notes in Computer Science 193*, pages 153–166. Springer-Verlag, 1985.

[KPR87] Ph. G. Kolaitis, H. J. Prömel, and B. L. Rothschild. K_{l+1}-free graphs: asymptotic structure and a 0-1 law. *Trans. Amer. Math. Soc.*, 303:637–671, 1987.

[KS85] M. Kaufmann and S. Shelah. On random models of finite power and monadic logic. *Discrete Mathematics*, 54:285–293, 1985.

[KV87] Ph. G. Kolaitis and M. Y. Vardi. The decision problem for the probabilities of higher-order properties. In *Proc. 19th ACM Symp. on Theory of Computing*, pages 425–435, 1987.

[KV90a] Ph. G. Kolaitis and M. Y. Vardi. 0-1 laws and decision problems for fragments of second-order logic. *Information and Computation*, 87:302–338, 1990.

[KV90b] Ph. G. Kolaitis and M. Y. Vardi. On the expressive power of Datalog: tools and a case study. In *Proc. 9th ACM Symp. on Principles of Database Systems*, pages 61–71, 1990. Full version to appear in *Journal of Computer and System Sciences.* Full version appeared also as IBM Research Report RJ8010, March 1991.

[KV92] Ph. G. Kolaitis and M. Y. Vardi. 0-1 laws for fragments of second-order logic: an overview. In Y. N. Moschovakis, editor, *Logic from computer science (Proc. of Workshop, 1989)*, pages 265–286, 1992.

[LM89] V. S. Lakshmanan and A. O. Mendelzon. Inductive pebble games and the expressive power of DATALOG. In *Proc. 8th ACM Symposium on Principles of Database Systems*, pages 301–310, 1989.

[Lyn80] J. F. Lynch. Almost sure theories. *Annals of Mathematical Logic*, 18:91–135, 1980.

[Mos74] Y. N. Moschovakis. *Elementary Induction on Abstract Structures.* North Holland, 1974.

[Ṕ76] L. Pósa. Hamiltonian circuits in random graphs. *Discrete Mathematics*, 14:359–364, 1976.

[PS89] L. Pacholski and W. Szwast. The 0-1 law fails for the class of existential second-order Gödel sentences with equality. In *Proc. 30th IEEE Symp. on Foundations of Computer Science*, pages 280–285, 1989.

[PS91] L. Pacholski and W. Szwast. On the 0-1 law for the existential second-order minimal Gödel sentences with equality. In *Proc. 6th IEEE Symp. on Logic in Computer Science*, pages 280–285, 1991.

[RN59] C. Ryll-Nardzewski. On the categoricity in power $\aleph_0$. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astron. Phys.*, 7:239–263, 1959.

[Rub75] A. Rubin. PhD thesis, California Institute of Technology, 1975.

[Spe91] J. Spencer. Threshold spectra via the Ehrenfeucht game. *Discrete Appl. Math.*, 30:235–252, 1991.

[SS88] S. Shelah and J. Spencer. Zero-one laws for sparse random graphs. *J. Amer. Math. Soc.*, 1:97–115, 1988.

[Sve59] L. Svenonius. $\aleph_0$-categoricity in first-order predicate calculus. *Theoria*, 25:82–94, 1959.

[Tal81] V. A. Talanov. Asymptotic solvability of logical formulas. *Combinatorial-Algebraic Methods in Applied Mathematics*, pages 118–126, 1981. MR 85i:03081, ZBL 538.03007.

[Tar55] A. Tarski. A lattice theoretical fixpoint theorem and its applications. *Pacific J. of Mathematics*, 5:285–309, 1955.

[TK86] V. A. Talanov and V. V. Knyazev. The asymptotic truth of infinite formulas. In *Proceedings of All-Union Seminar on Discrete and Applied Mathematics and it Applications*, pages 56–61, 1986. MR 89g:03054.

[Var82] M. Y. Vardi. The complexity of relational query languages. In *Proc. 14th ACM Symp. on Theory of Computing*, pages 137–146, 1982.

[Vau61] R. Vaught. Denumerable models of complete theories. In *Infinitistic Methods*, pages 303–321. Pergamon Press, 1961.