Probabilities of Sentences about Very Sparse Random Graphs

James F. Lynch*
Department of Mathematics and Computer Science, Clarkson University, Potsdam, 13699-5815

ABSTRACT

We consider random graphs with edge probability $\beta n^{-\alpha}$, where n is the number of vertices of the graph, $\beta > 0$ is fixed, and $\alpha = 1$ or $\alpha = (l+1)/l$ for some fixed positive integer l. We prove that for every first-order sentence, the probability that the sentence is true for the random graph has an asymptotic limit.

Key Words: random graphs, first-order logic, asymptotic probability

1. INTRODUCTION

Let S be a set of finite structures such that for every positive integer n, a probability distribution has been defined on the members of S of size n, and let $P \subseteq S$ be a property of some structures in S. A fundamental question is: for large n, what is the probability that a random structure of size n has property P?

Some of the earliest articles that studied this question pertained to graphs and properties such as containment of certain types of subgraphs and connectedness [7]. Since then, this subject has developed into a major area of combinatorics and has had numerous applications in other disciplines, particularly computer science.

While most of the work on random graphs has focused on specific properties, one aspect of it is the determination of classes of properties that have asymptotic probabilities. One of the most productive approaches has been to consider properties describable in certain formal languages.

* Research supported by NSF Grant CCR-8805880. A preliminary version of this article appeared in *Proc. 31st Annual Symp. Foundations of Computer Science*, 1990, pp. 689–696.

Random Structures and Algorithms, Vol. 3, No. 1 (1992)

The first important result in this area was the 0–1 law for first-order properties independently discovered by Y. Glebskiĭ, D. Kogan, M. Liogon'kiĭ, and V. Talanov [9] and R. Fagin [8]. The theorem was originally proven for arbitrary relational structures. When applied to random graphs, it states that if the edge probability p is constant, then any property describable by a first-order sentence has probability asymptotic to 0 or 1.

For a number of years, this result remained relatively isolated, but in the past decade, there has been considerable work extending it to other logics and more general probability distributions. The survey by K. Compton [2] is the best source for most of this material. In particular, S. Shelah and J. Spencer [16] extended the original 0–1 law to the case when $p = n^{-\alpha}$ where $\alpha > 0$ is irrational, or $\alpha > 1$ and $\alpha \neq (l+1)/l$ for any positive integer l. Thus, from the standpoint of first-order logic, the random graph is very well-behaved for all these α. Let us refer to all other α as *exceptional*.

The behavior at the exceptional α is not so easily described. It can be shown that for all such α, there are sentences with an asymptotic probability that is not 0 or 1. Further, $n^{-\alpha}$ is a threshold function for some first-order property if and only if α is exceptional. That is, there is a first-order sentence such that for all $\epsilon > 0$, the asymptotic probability of the sentence is 0 if $p = n^{-\alpha-\epsilon}$, and it is 1 if $p = n^{-\alpha+\epsilon}$.

Shelah and Spencer also showed that for every rational $\alpha < 1$, there is a first-order sentence which does not have an asymptotic probability when $p = n^{-\alpha}$. This article characterizes the behavior of the random graph for all other exceptional α, i.e., $\alpha = 1$ and $\alpha = (l+1)/l$. We will show that for any such α and any $\beta > 0$, every first-order sentence has an asymptotic probability when $p \sim \beta n^{-\alpha}$. In fact, given any first-order sentence, there is an effective procedure for generating a closed formula whose value is the limit. Roughly speaking, the formula is an expression involving $+$, $-$, $\cdot$, $/$, and the "Poisson functions" $x^j e^{-x}/j!$. When $\alpha = 1$, the applications of the Poisson functions may be nested within one another; when $\alpha > 1$, no nesting of the functions occurs.

Another distinction between exceptional $\alpha \geq 1$ and those less than 1 was shown by P. Dolan [5]. He proved that for any rational $\alpha < 1$, there is no effective procedure for separating those sentences whose limiting probability is 0 from those whose limiting probability is 1. Thus n^{-1} may be considered a meta-threshold at which not only do the probabilities of certain individual sentences change, but the properties of the thresholds below and above it differ significantly.

Interestingly, our results grew out of a problem posed in one of the original articles on this subject. In [8], Fagin asked whether every sentence about random unary functions has an asymptotic probability. Here, all of the n^n unary functions on a set of size n are equally likely. Fagin gave an example showing that the limit need not be 0 or 1. The sentence that says that there is no fixed point, i.e., x such that $f(x) = x$, has limiting probability $1/e$. In an earlier article [13], we settled Fagin's question in the affirmative. The methods used in this article are quite similar, especially the case when $\alpha = 1$.

This article consists of the following sections. In Section 2, we give the basic definitions and state the main theorems. In Section 3, we describe the game-theoretic technique that plays a key role in the proofs. In Section 4, we outline the proof for the case $\alpha = 1$. The proof divides naturally into a model-theoretic part and a combinatorial part. The two parts are covered in the next two sections. Then we sketch the proof for the case when $\alpha = (l+1)/l$. This uses many of the same ideas as the first case, but is considerably simpler. In Section 8, we conclude with some generalizations and open problems.

2. DEFINITIONS AND MAIN THEOREMS

Our main theorems pertain to the usual definition of a graph as a structure $G = \langle V, E \rangle$ where E, the set of edges, is an irreflexive, symmetric binary relation on the finite set of vertices V. Unless stated otherwise, we take $V = \{0, 1, \dots, n-1\}$ for some $n \in \omega$. We assume the reader is familiar with the basic definitions of connected, component, path, and cycle (see [11]). If $W \subseteq V$, then $\langle W, E \rangle$ will be an abbreviation for $\langle W, E \cap W \times W \rangle$. More generally, a *rooted graph* is a structure $G = \langle V, E, c_1, \dots, c_i \rangle$ where each $c_j \in V$. The c_j's are referred to as *constants*. If G is a graph, then we can extend it to a rooted graph by choosing values for the constants. Such an extension will be indicated by $G\langle c_1, \dots, c_i \rangle$.

The formal language we will use to describe properties of graphs is the first-order predicate calculus (see [14]). The symbols of this language are the boolean operators $\neg$, $\vee$, $\wedge$, $\Rightarrow$, $\Leftrightarrow$ (not, or, and, implies, if and only if), the quantifiers $\forall$, $\exists$ (for all, there exists), variables x, y, z (sometimes subscripted), the constant symbols $c_1, c_2, \dots$, and the two binary predicates $=$ and $\sim$, where $x \sim y$ means there is an edge between x and y. Lower case Greek letters will represent formulas in this language, and their free variables will be listed in parentheses after them, e.g., $\sigma(x_1, \dots, x_i)$. If $a_1, \dots, a_i \in V$, then $G \models \sigma(a_1, \dots, a_i)$ means σ is true in G with each free occurrence of the free variable x_j replaced by a_j. The *depth* of a formula is the level of nesting of its quantifiers. That is, atomic formulas have depth 0; the depth of $\sigma \vee \tau$ is the maximum of the depths of σ and τ, and similarly for the other boolean operators; and the depth of $(\forall x)\sigma$ is one more than the depth of σ, and similarly for $\exists$.

We use the probability distribution introduced by Erdös and Rényi [7], sometimes referred to as "model A" (see E. Palmer [15]). The function $p : \omega \to [0, 1]$ assigns edges to the random graph independently. That is, for every $n \in \omega$, and every graph G with n vertices and q edges, $\mathrm{pr}(G) = p(n)^q (1 - p(n))^{\binom{n}{2} - q}$. Given a sentence σ, the probability that σ satisfies the random graph on n vertices is $\mathrm{pr}(\sigma, n) = \Sigma_{\substack{G \models \sigma \\ |G| = n}} \mathrm{pr}(G)$.

We will consider $p(n) \sim \beta n^{-\alpha}$, where $\beta > 0$ and either $\alpha = 1$ or $\alpha = (l+1)/l$ for some positive integer l. The limiting probabilities of sentences are the values of sums of terms from the following sets. For $l \geq 1$ let Θ_l be the set of expressions that are products of terms of the form $(\beta^l/a)^j e^{-\beta^l/a}/j!$ or $1 - (\Sigma_{s \leq j} (\beta^l/a)^s/s!) e^{-\beta^l/a}$ where $1 \leq a \in \omega$ and $j \in \omega$. Let Λ_∞ be the smallest set of expressions containing 1 and such that if $\lambda \in \Lambda_\infty$ and $j \in \omega$, then $(\beta\lambda)^j e^{-\beta\lambda}/j! \in \Lambda_\infty$ and $1 - (\Sigma_{s \leq j} (\beta\lambda)^s/s!) e^{-\beta\lambda} \in \Lambda_\infty$, and if $\lambda_1, \lambda_2 \in \Lambda_\infty$ then $\lambda_1 \lambda_2 \in \Lambda_\infty$. Let Θ_∞ be the set of expressions that are products of terms of the form $(\beta^i \lambda/a)^j e^{-\beta^i \lambda/a}/j!$ or $1 - (\Sigma_{s \leq j} (\beta^i \lambda/a)^s/s!) e^{-\beta^i \lambda/a}$, where $\lambda \in \Lambda_\infty$, $1 \leq a \in \omega$, $3 \leq i \in \omega$, and $j \in \omega$.

Theorem 2.1. *Let σ be a sentence in the first-order theory of graphs, and let $p(n) \sim \beta n^{-1}$. Then $\lim_{n\to\infty} \mathrm{pr}(\sigma, n)$ exists and is the value of a finite sum of elements from Θ_∞.*

Theorem 2.2. *Let $p(n) \sim \beta n^{-1}$. Then there is an effective procedure such that, given σ, it generates an expression as in Theorem 2.1 whose value is $\lim_{n\to\infty} \mathrm{pr}(\sigma, n)$. The time complexity of the procedure is bounded by $\exp_\infty(ck)$, for some $c > 0$, where k is the depth of σ, $\exp_\infty(0) = 1$, and $\exp_\infty(n+1) = 2^{\exp_\infty(n)}$.*

Theorem 2.3. *Let σ be a sentence in the first-order theory of graphs, and let $p(n) \sim \beta n^{-(l+1)/l}$ for some positive integer l. Then $\lim_{n\to\infty} \mathrm{pr}(\sigma, n)$ exists and is the value of a finite sum of elements from Θ_l.*

Theorem 2.4. *Let $p(n) \sim \beta^{-(l+1)/l}$ for some positive integer l. Then there is an effective procedure such that, given σ, it generates an expression as in Theorem 2.3 whose value is $\lim_{n\to\infty} \mathrm{pr}(\sigma, n)$. The space complexity of the procedure is polynomial in the depth of σ.*

We will not give proofs of Theorems 2.2 and 2.4 here since they involve automata theoretic concepts that have no bearing on the rest of the article. Methods for proving upper bounds on the complexity of determining whether a sentence has asymptotic probability 1 were developed by E. Grandjean [10]. They also apply to the problem of generating the asymptotic probability. Using a general method of Compton and Henson [3], it can be shown that the upper bound in Theorem 2.2 is the best possible.

3. STRATEGY FOR THE EHRENFEUCHT GAME

To state our main ideas in nontechnical terms, we give a description of the random graph with $p(n) \sim \beta n^{-1}$ from the viewpoint of first-order logic.

- The small cycles are very far apart. That is, the subgraph induced by each one and the vertices close to it looks like a cycle with disjoint trees growing out of it.
- All small trees occur as subgraphs arbitrarily often.

It is well known (see [15, p. 60]) that the structure of the random graph with $p(n) \sim \beta n^{-1}$ depends very strongly on β. When $\beta < 1$, almost surely all of its components are trees or unicyclic, and the largest component is a tree of size on the order of $\log n$. When $\beta = 1$, there are cycles, and the largest component has on the order of $n^{2/3}$ vertices. When $\beta > 1$, there is a unique giant component, and all but $o(n)$ vertices belong to it or to trees of size on the order of $\log n$. These features are not visible to first-order logic, and our proofs are the same for all β.

We then show that, for almost all graphs G, whether or not $G \models \sigma$ is determined by the kinds of small cycles (including their attached trees) that G has. The subgraph induced by all the small cycles and their attached trees will be referred to as $\mathrm{core}(G, r)$, where r is an integer (determined by the depth of σ)

which gives a precise meaning to "small" and "close." We also define the core of a rooted graph to include all its constants and the vertices close to them.

For each $k \in \omega$, a certain equivalence relation, known as k-agreeability, is defined on rooted graphs. Our proof consists essentially of two parts. First, for every k, there is r such that for almost all graphs G^0 and G^1, if core(G^0, r) and core(G^1, r) are k-agreeable, then they are indistinguishable by any sentence σ of depth at most k. That is, if A is any k-agreeability class, then either $G \models \sigma$ for almost all G such that core$(G, r) \in A$ or $G \models \neg\sigma$ for almost all G such that core$(G, r) \in A$. Second, for fixed r, there is a finite set of k-agreeability classes that includes the cores of almost all graphs, and the probability that the core of a random graph belongs to a given k-agreeability class has an asymptotic limit which is in Θ_∞. Theorem 2.1 then follows easily: pr(σ, n) is asymptotic to the sum of pr(core$(G, r) \in A, n)$, taken over all k-agreeability classes A whose members are the cores of graphs that almost surely satisfy σ.

The method we use to show that two graphs G^0 and G^1 are indistinguishable by any sentence of depth k has been quite useful in other articles on limit and 0–1 laws [1, 12, 13, 17]. It is a game-theoretic technique due to A. Ehrenfeucht [6]. The k-round Ehrenfeucht game on G^0 and G^1, $\Gamma_k(G^0, G^1)$, is a game of perfect information with two players I and II. Each round $i = 1, \ldots, k$ of the game results in choosing a value for the constant c_i in each graph. It begins with Player I choosing a value for c_1 in either G^0 or G^1, and Player II choosing a value for c_1 in the other graph. For $g = 0, 1$, let $c_1^g \in V^g$ be the value chosen, where $G^g = \langle V^g, E^g\rangle$. The remaining $k - 1$ rounds in the game are similar; each time Player I chooses $c_i^g \in V^g$ where g is either 0 or 1, and Player II chooses $c_i^{1-g} \in V^{1-g}$. After round k, if the rooted subgraphs induced by the chosen constants are isomorphic, then Player II has won. That is, for all $i, j \in \{1, \ldots, k\}$, $c_i^0 \sim c_j^0$ if and only if $c_i^1 \sim c_j^1$. Player II is said to have a *winning strategy* if he can always choose in such a way that he will win the game. The relevance of the game to this article is shown by the following.

Theorem 3.1 (Ehrenfeucht). *Player II has a winning strategy for $\Gamma_k(G^0, G^1)$ if and only if no first-order sentence of depth at most k can distinguish G^0 from G^1.*

Thus the first part of our proof is to show that there is some r and a strategy for Player II such that for almost all graphs G^0 and G^1 such that core(G^0, r) and core(G^1, r) are k-agreeable, Player II can follow this strategy and win. Taking $r = 3^k$, the strategy is to choose so that at each round i, core$(G^0\langle c_1^0, \ldots, c_i^0\rangle, 3^{k-i})$ and core$(G^1\langle c_1^1, \ldots, c_i^1\rangle, 3^{k-i})$ are $(k-i)$-agreeable. We will show that, for almost all pairs of graphs, if this condition holds after round $i-1$, then no matter how Player I chooses, Player II can choose so that it holds at round i. Since we have assumed it holds at round 0 (the start of the game), it implies that Player II can win.

4. OUTLINE OF THE PROOF FOR THE CASE $\alpha = 1$

Some preliminary definitions are needed first. Let $G = \langle V, E, c_1, \ldots, c_i\rangle$ be a rooted graph. We define a metric: for any $x, y \in V$, $\delta(x, y)$ is the length of the shortest path from x to y. We extend δ to subsets of V: for $X, Y \subseteq V$, $\delta(x, X) = \delta(X, x) = \min\{\delta(x, y) : y \in X\}$ and $\delta(X, Y) = \min\{\delta(x, y) : x \in X \text{ and } y \in Y\}$. For $X \subseteq V$ and $r \geq 0$, $N(X, r) = \{x \in V : \delta(X, x) \leq r\}$ is the *r-neighborhood* of X in G.

The *center* of a component of G is the minimal connected set containing all the cycles and constants in the component. The center of G, center(G), is the union of the centers of all its components. The *radius* of G, radius(G), is $\max\{\delta(\text{center}(G), x) : x \in V\}$. We define an operator *tree* on rooted graphs G and $x \in V$:

$$\text{tree}(G, x) = \langle V_x, E, c_1[x] \rangle ,$$

where

$$V_x = \{y \in V : \delta(\text{center}(G), y) = \delta(\text{center}(G), x) + \delta(x, y)\}$$

and the notation $c_1[x]$ means that the constant c_1 is assigned to x. That is, V_x is the set of all vertices whose (unique) path to center(G) includes x, and tree(G, x) is the tree rooted at x induced by these vertices. Figure 1 illustrates these concepts.

For $j \in \omega$, we define an equivalence relation known as j-morphism, written as $\stackrel{j}{\simeq}$, on rooted graphs. It will be seen that all j-morphic rooted graphs have isomorphic centers and equal radii. We first define this relation on acyclic connected graphs with a single constant c_i, i.e., rooted trees. The definition is by recursion on r, the maximum of their radii.

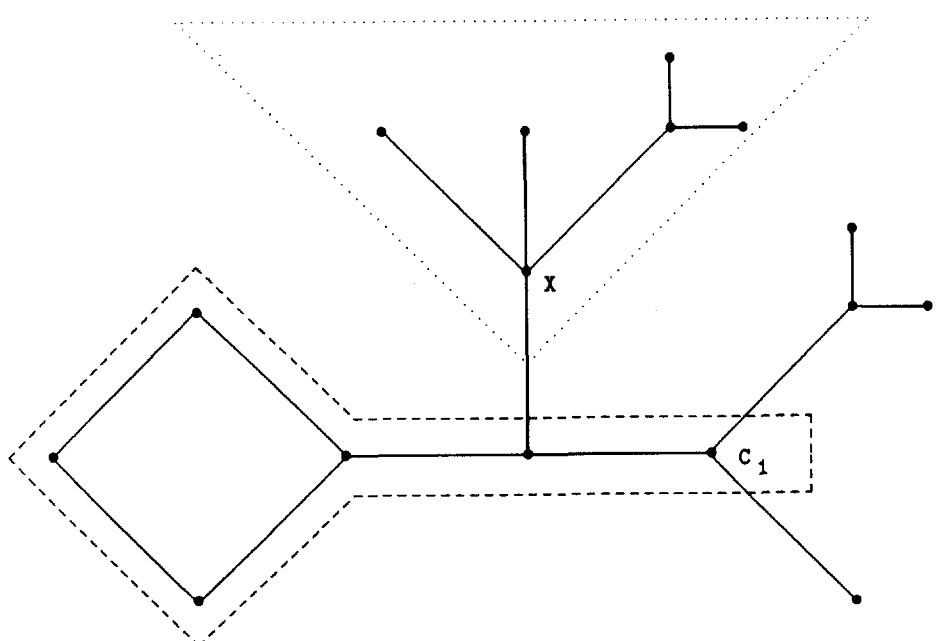

Fig. 1. The vertices in Center(G) are enclosed in: ----. The vertices in tree(G, x) are enclosed in: ····.

If $r = 0$, then both graphs are one-vertex rooted trees, and we define all such graphs to be j-morphic if their root is the same constant. Clearly their centers are isomorphic.

Now assume G^0 and G^1 are rooted trees with the same constant, say c_1, and maximum radius $r > 0$. Let $C_0, \ldots, C_{c-1}$ be an enumeration of all j-morphism classes with center $\{c_1\}$ and radius less than r. For $g = 0, 1$ and $a < c$ let

$$S_a^g = \{x \in V^g : \delta(c_1^g, x) = 1 \text{ and } \mathrm{tree}(G^g, x) \in C_\alpha\} \tag{4.1}$$

Then $G^0 \overset{j}{\doteq} G^1$ if and only if for all such a,

$$|S_a^0| = |S_a^1| \text{ or } |S_a^0|, |S_a^1| > j .$$

Again, j-morphic graphs have equal radii.

A special case that needs to be pointed out is when $|S_a^0| = 0$ for all $a < c$. In other words, G^0 is simply a one-vertex tree. However, we will not consider it j-morphic to the one-vertex tree in C_0 because G^0 provides more information about the neighborhood of c_1: there are no other vertices in it. Thus we place G^0 in a new j-morphism class, say C_1, of radius 1.

Figure 2 is an example of two rooted trees of radius 3 that are 2-morphic.

We now extend the definition of j-morphism to rooted graphs G^0 and G^1 with constants $c_1, \ldots, c_i$ whose centers are Z^0 and Z^1, respectively. If there exists an isomorphism f from $\langle Z^0, E^0, c_1^0, \ldots, c_i^0 \rangle$ onto $\langle Z^1, E^1, c_1^1, \ldots, c_i^1 \rangle$ such that for each $x \in Z^0$, $\mathrm{tree}(G^0, x) \overset{j}{\doteq} \mathrm{tree}(G^1, f(x))$, then we say that $G^0 \overset{j}{\doteq} G^1$ via f. Since all members of a given j-morphism class have isomorphic centers and equal radii, we may speak of the isomorphism type of the class and the radius of the class.

In the sequel, $C_0, C_1, \ldots$ will always be an enumeration of acyclic j-morphism classes with center $\{c_1\}$ such that $\mathrm{radius}(C_a) \leq \mathrm{radius}(C_b)$ for $a < b$. Each such class C_b, $b > 0$, can be identified by certain sequences of natural numbers less than or equal to $j + 1$. Suppose $\mathrm{radius}(C_b) \leq \mathrm{radius}(C_c)$, and $G = \langle V, E, c_1 \rangle$ is any member of C_b. For $a < c$ let S_a be as in (4.1). Then $\vec{s} = \langle s_0, \ldots, s_{c-1} \rangle$ is a

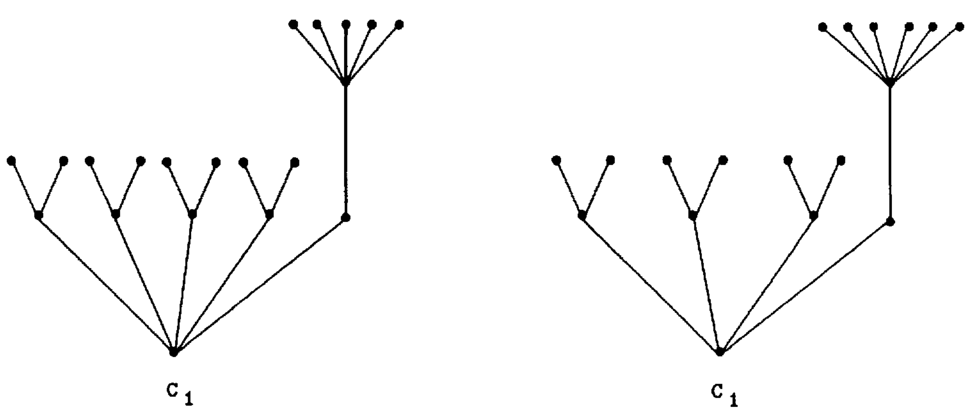

Fig. 2. Two 2-morphic rooted trees of radius 3.

characteristic vector for C_b if for each $a < c$

$$s_a = \begin{cases} |S_a| & \text{if } |S_a| \leq j \\ j+1 & \text{if } |S_a| > j \end{cases}$$

Clearly the values of s_a do not depend on the choice of G, and given any such $\vec{s}$, there is a unique class C_b that corresponds to $\vec{s}$. Note also that C_1 corresponds to any vector of all 0's.

Definition 4.2. For a rooted graph $G = \langle V, E, c_1, \ldots, c_i \rangle$ and $r \in \omega$, $\text{core}(G, r) = \langle N(X, r), E, c_1, \ldots, c_i \rangle$ where X is the union of all cycles in G of size at most $2r + 1$ and all constants.

Definition 4.3. Let G be a graph and $r \in \omega$. G is *r-simple* if each component of $\text{core}(G, r)$ contains at most one cycle.

Definition 4.4. Let G be a graph, $i, j, r \in \omega$, $i \geq 1$. G is (i, j, r)*-rich* if for every rooted tree T with center $\{c_i\}$ and radius at most r, and $c_1, \ldots, c_{i-1} \in V$, there exists $c_i \in V$ such that $N(c_i, r) \cap \text{core}(G\langle c_1, \ldots, c_{i-1}\rangle, r) = \emptyset$ and $\langle N(c_i, r), E, c_i \rangle \stackrel{j}{\simeq} T$.

Definition 4.5. Let G^0 and G^1 be rooted graphs and $j \in \omega$. Let $D_0, \ldots, D_{d-1}$ be an enumeration of all j-morphism classes that are represented by some component of G^g, $g = 0, 1$. For such g and $a < d$, let

$$T_a^g = \{X : X \text{ is a component of } G^g \text{ and } X \in D_a\} \tag{4.6}$$

Then G^0 and G^1 are *j-agreeable* if for all $a < d$,

$$|T_a^0| = |T_a^1| \text{ or } |T_a^0|, |T_a^1| > j\,.$$

As mentioned above, the proof of our main theorem has two major divisions. The first is model theoretic.

Theorem 4.7. *Let* $1 \leq k \in \omega$ *and* G^0, G^1 *be graphs. If* G^0 *and* G^1 *are* $(i, k-i, 3^{k-i})$*-rich for* $1 \leq i \leq k$, *and* $\text{core}(G^0, 3^k)$ *is k-agreeable to* $\text{core}(G^1, 3^k)$, *then* $\Gamma_k(G^0, G^1)$ *is a win for Player II.*

The next three theorems are the combinatorial part.

Theorem 4.8. *For every* $r \in \omega$, $\lim_{n\to\infty} \text{pr}(G \textit{ is r-simple}, n) = 1$.

The significance of Theorem 4.8 is that for fixed k, almost all graphs have $\text{core}(G, 3^k)$ belonging to one of the finitely many k-agreeability classes whose components each have a single cycle of size at most $2 \cdot 3^k + 1$.

Theorem 4.9. *For every* $i, j, r \in \omega$, $i \geq 1$, $\lim_{n\to\infty} \text{pr}(G \textit{ is } (i, j, r)\textit{-rich}, n) = 1$.

Theorem 4.10. *For every $j, r \in \omega$ and every j-agreeability class A whose components have a single cycle, there is a $\theta \in \Theta_\infty$ such that* $\lim_{n\to\infty} \mathrm{pr}(\mathrm{core}(G, r) \in A, n) = \theta$.

Theorem 4.10 easily extends to arbitrary j-agreeability classes. By Theorem 4.8, if A has a component with more than one cycle, then the limit is 0.

5. MODEL-THEORETIC RESULTS

Lemma 5.1. *Let $G^0 \overset{j}{\simeq} G^1$ be rooted trees with center $\{c_1\}$. Then for $h \le j$ and $u \in \omega$,* $\mathrm{core}(G^0, u) \overset{h}{\simeq} \mathrm{core}(G^1, u)$.

Proof. We use induction on u. If $u = 0$ then for $g = 0, 1$ $\mathrm{core}(G^g, 0) = \langle \{c_1^g\}, \emptyset, c_1^g\rangle$, so we are done.

Now assume $u > 0$ and the Lemma holds for all values less than u. Let $r = \mathrm{radius}(G^0)$ and $\mathrm{radius}(C_a) < r$ for all $a < c$. For $g = 0, 1$ and $a < c$, let S_a^g be as defined in Equation (4.1). Let $D_0, \ldots, D_{d-1}$ be an enumeration of all h-morphism classes with center $\{c_1\}$ and radius less than u. For $g = 0, 1$ and $b < d$ let

$$T_b^g = \{x \in \mathrm{core}(G^g, u) : \delta(c_1^g, x) = 1 \text{ and } \mathrm{tree}(\mathrm{core}(G^g, u), x) \in D_b\}\ .$$

Clearly $T_b^g \subseteq \bigcup_{a<c} S_a^g$. Now for every $x \in V^g$ such that $\delta(c_1^g, x) = 1$, $\mathrm{tree}(\mathrm{core}(G^g, u), x) = \mathrm{core}(\mathrm{tree}(G^g, x), u - 1)$, so by induction, for every $a < c$,

$$S_a^g \subseteq T_b^g \text{ for } g = 0, 1 \text{ or } S_a^g \cap T_b^g = \emptyset \text{ for } g = 0, 1\ .$$

Let $I = \{a < c : S_a^g \subseteq T_b^g\}$. Then

$$|T_b^g| = \sum_{a\in I} |S_a^g| \text{ for } g = 0, 1\ .$$

Since $G^0 \overset{j}{\simeq} G^1$, for all $a < c$,

$$|S_a^0| = |S_a^1| \text{ or } |S_a^0|, |S_a^1| > j\ .$$

Therefore, for all $b < d$,

$$|T_b^0| = |T_b^1| \text{ or } |T_b^0|, |T_b^1| > j \ge h$$

so $\mathrm{core}(G^0, u) \overset{h}{\simeq} \mathrm{core}(G^1, u)$. ∎

Lemma 5.2. *For $g = 0, 1$ let $G^g = \langle V^g, E^g, c_1^g\rangle$ and $H^g = \langle W^g, F^g, c_1^g\}$ be rooted trees such that $V^g \cap W^g = \{c_1^g\}$. If $G^0 \overset{j}{\simeq} G^1$ and $H^0 \overset{j}{\simeq} H^1$, then $G^0 \cup H^0 \overset{j}{\simeq} G^1 \cup H^1$.*

Proof. Let $J^g = G^g \cup H^g$. Clearly $\mathrm{center}(J^g) = \{c_1^g\}$. Let r be the maximum of the radii of G^0 and H^0, let S_a^g be as defined in Equation (4.1), and

$$T_a^g = \{x \in W^g : \delta(c_1^g, x) = 1 \text{ and } \mathrm{tree}(H^g, x) \in C_a\}$$
$$U_a^g = \{x \in V^g \cup W^g : \delta(c_1^g, x) = 1 \text{ and } \mathrm{tree}(J^g, x) \in C_a\}$$

By assumption,

$$|S_a^0| = |S_a^1| \text{ or } |S_a^0], |S_a^1| > j\,,$$

and similarly for T_a^g. Since $U_a^g = S_a^g \cup T_a^g$ and $S_a^g \cap T_a^g = \emptyset$,

$$|U_a^0| = |U_a^1| \text{ or } |U_a^0|, |U_a^1| > j\,,$$

and $J^0 \overset{j}{\simeq} J^1$. ∎

Lemma 5.3. *For $g = 0, 1$ let $G^g = \langle V^g, E^g \rangle$ be a tree, $x^g \in V^g$ such that $G^0\langle c_1[x^0]\rangle \overset{j}{\simeq} G^1\langle c_1[x^1]\rangle$ and let $y^g \not\in V^g$. Then $\langle V^0 \cup \{y^0\},\ E^0 \cup \{(x^0, y^0)\},\ c_1[y^0]\rangle \overset{j}{\simeq} \langle V^1 \cup \{y^1\},\ E^1 \cup \{(x^1, y^1)\},\ c_1[y^1]\rangle$.*

Proof. Immediate. ∎

Lemma 5.4. *Let $G^0 \overset{j}{\simeq} G^1$ be rooted graphs. Then for $h \le j$ and $r \in \omega$, $\mathrm{core}(G^0, r) \overset{h}{\simeq} \mathrm{core}(G^1, r)$.*

Proof. For $g = 0, 1$ let $Z^g = \mathrm{center}(G^g)$ and say that $G^0 \overset{j}{\simeq} G^1$ via f. For $x \in \mathrm{center}(\mathrm{core}(G^g, r))$, let

$$W_x^g = \{y \in \mathrm{core}(G^g, r) : \delta(\mathrm{center}(\mathrm{core}(G^g, r)), y) = \delta(x, y)\} \text{ and}$$
$$H_x^g = \langle W_x^g, E^g, c_1[x]\rangle, \text{ i.e., } H_x^g = \mathrm{tree}(\mathrm{core}(G^g, r), x)$$

We need to show $H_x^0 \overset{h}{\simeq} H_{f(x)}^1$.

Let $X_x^g = W_x^g \cap Z^g$ and $J_x^g = \langle X_x^g, E^g, c_1[x]\rangle$. By the definition of center, for every $y \in X_x^g$, the path from x to y is in Z^g, so J_x^g is a tree. Also, f is an isomorphism from J_x^0 onto $J_{f(x)}^1$. For $y \in X_x^g$, let

$$Y_{xy}^g = \{z \in W_x^g : \delta(X_x^g, z) = \delta(y, z)\} \text{ and}$$
$$K_{xy}^g = \langle Y_{xy}^g, E^g \rangle$$

For any such y, $X_x^g \cap Y_{xy}^g = \{y\}$. Now $K_{xy}^g\langle c_1[y]\rangle = \mathrm{core}(\mathrm{tree}(G^g, y), r - \delta(x, y))$ and by assumption, $\mathrm{tree}(G^0, y) \overset{j}{\simeq} \mathrm{tree}(G^1, f(y))$. Therefore by Lemma 5.1, $K_{xy}^0\langle c_1[y]\rangle \overset{h}{\simeq} K_{f(x)f(y)}^1\langle c_1[f(y)]\rangle$. Since J_x^0 and $J_{f(x)}^1$ are isomorphic and $H_x^g = J_x^g \cup_{y \in X_x^g} K_{xy}^g$, by repeated application of Lemmas 5.2 and 5.3, $H_x^0 \overset{h}{\simeq} H_{f(x)}^1$. ∎

Lemma 5.5. *Let $G^0 \overset{j}{\simeq} G^1$ where $G^g = \langle V^g, E^g, c_1^g, \ldots, c_{i-1}^g\rangle$ for $g = 0, 1$. Then for every $c_i^0 \in V^0$, there is $c_i^1 \in V^1$ such that $G^0\langle c_i^0\rangle \overset{j-1}{\simeq} G^1\langle c_i^1\rangle$.*

Proof. Let $Z^g = \text{center}(G^g)$ and say that $G^0 \overset{j}{\simeq} G^1$ via f. We use induction on $\delta(Z^0, c_i^0)$.

If $\delta(Z^0, c_i^0) = 0$, then we choose $c_i^1 = f(c_i^0)$, and $G^0\langle c_i^0\rangle \overset{j}{\simeq} G^1\langle c_i^1\rangle$ by definition, so by Lemma 5.4, we are done.

Now assume $\delta(Z^0, c_i^0) > 0$ and the Lemma holds for all smaller values. Let P^0 be the unique path from Z^0 to c_i^0, x^0 be the unique vertex in $P^0 \cap Z^0$, and y^0 be the unique vertex in P^0 such that $\delta(x^0, y^0) = 1$. Let $x^1 = f(x^0)$ and $r = \text{radius}(G^0)$. Taking c such that $\text{radius}(C_a) < r$ for all $a < c$, we put

$$S_a^g = \{ y \in \text{tree}(G^g, x^g) : \delta(x^g, y) = 1 \text{ and } \text{tree}(G^g, y) \in C_a \}$$

for $g = 0, 1$ and $a < c$. Then

$$|S_a^0| = |S_a^1| \text{ or } |S_a^0], |S_a^1| > j$$

for all $a < c$. Say $y^0 \in S_b^0$. Then there is some $y^1 \in \text{tree}(G^1, x^1)$ such that $\delta(x^1, y^1) = 1$ and $y^1 \in S_b^1$, i.e., $\text{tree}(G^0, y^0) \overset{j}{\simeq} \text{tree}(G^1, y^1)$. Now $\delta(y^0, c_i^0) = \delta(x^0, c_i^0) - 1$, so by induction there is $c_i^1 \in \text{tree}(G^1, y^1)$ such that $\text{tree}(G^0, y^0)\langle c_i^0\rangle \overset{j-1}{\simeq} \text{tree}(G^1, y^1)\langle c_i^1\rangle$.

Let P^1 be the unique path from x^1 to c_i^1. Then $\text{center}(G^g\langle c_i^g\rangle) = Z^g \cup P^g$, and we can extend f to an isomorphism from $Z^0 \cup P^0$ onto $Z^1 \cup P^1$ in the obvious way. It remains to show that for all $x \in Z^0 \cup P^0$, $\text{tree}(G^0\langle c_i^0\rangle, x) \overset{j-1}{\simeq} \text{tree}(G^1\langle c_i^1\rangle, f(x))$.

It is easily seen that for $x \in Z^g - \{x^g\}$, $\text{tree}(G^g\langle c_i^g\rangle, x) = \text{tree}(G^g, x)$ and therefore by Lemma 5.1 (or 5.4), we are done. For $x \in P^g - \{x^g\}$, $\text{tree}(G^g\langle c_i^g\rangle, x) = \text{tree}(\text{tree}(G^g, y^g)\langle c_i^g\rangle, x)$, and again we are done.

Finally, we dispose of the case $x = x^0$. Let

$$T_a^g = \{ y \in \text{tree}(G^g\langle c_i^g\rangle, x^g) : \delta(x^g, y) = 1 \text{ and } \text{tree}(G^g\langle c_i^g\rangle, y) \in C_a \}\,.$$

For all $a \neq b$, $T_a^g = S_a^g$, and $T_b^g = S_b^g - \{y^g\}$, so for all $a < c$,

$$|T_a^0| = |T_a^1| \text{ or } |T_a^0|, |T_a^1| > j - 1\,,$$

and this completes the proof. ■

Proof of Theorem 4.7. By induction on $i = 0, \ldots, k$ we show that Player II can choose so that $\text{core}(G^0\langle c_1^0, \ldots, c_i^0\rangle, 3^{k-i})$ is $(k-i)$-agreeable to $\text{core}(G^1\langle c_1^1, \ldots, c_i^1\rangle, 3^{k-i})$.

By assumption, this is true for $i = 0$. Now assume $i > 0$ and it holds for $i - 1$. Let Player I choose $c_i^0 \in V^0$.

Case I. $N(c_i^0, 3^{k-1}) \cap \text{core}(G^0\langle c_1^0, \ldots, c_{i-1}^0\rangle, 3^{k-i}) = \emptyset$. Then $N(c_i^0, 3^{k-i})$ is acyclic and is a component of $\text{core}(G^0\langle c_1^0, \ldots, c_i^0\rangle, 3^{k-i}) = \text{core}(G^0\langle c_1^0, \ldots, c_{i-1}^0\rangle, 3^{k-i}) \cup N(c_i^0, 3^{k-i})$. Since G^1 is $(i, k-i, 3^{k-i})$-rich, there is $c_i^1 \in V^1$ such that $N(c_i^1, 3^{k-i}) \cap \text{core}(G^1\langle c_1^1, \ldots, c_{i-1}^1\rangle, 3^{k-1}) = \emptyset$ and $\langle N(c_i^0, 3^{k-i}), E^0, c_i^0\rangle \overset{k-i}{\simeq} \langle N(c_i^1, 3^{k-i}), E^1, c_i^1\rangle$. Now $\text{core}(\text{core}(G^g\langle c_1^g, \ldots, c_{i-1}^g\rangle,$

$3^{k-i+1}), 3^{k-i}) = \mathrm{core}(G^g\langle c_1^g, \ldots, c_{i-i}^g\rangle, 3^{k-i})$, so by the induction assumption and Lemma 5.4, $\mathrm{core}(G^0\langle c_1^0, \ldots, c_i^0\rangle, 3^{k-i}) \overset{k-i}{\simeq} \mathrm{core}(G^1\langle c_1^1, \ldots, c_i^1\rangle, 3^{k-i})$.

Case II. $N(c_i^0, 3^{k-i}) \cap \mathrm{core}(G^0\langle c_1^0, \ldots, c_{i-1}^0\rangle, 3^{k-i}) \neq \emptyset$, say it contains y. Let X be the union of all cycles of size at most $2 \cdot 3^{k-i} + 1$ and $\{c_1^0, \ldots, c_{i-1}^0\}$. Then for any $x \in N(c_i^0, 3^{k-i})$, $\delta(X, x) \leq \delta(X, y) + \delta(y, c_i^0) + \delta(c_i^0, x) \leq 3^{k-i+1}$, so $N(c_i^0, 3^{k-i}) \subseteq \mathrm{core}(G^0\langle c_1^0, \ldots, c_{i-1}^0\rangle, 3^{k-i+1})$. Therefore by the induction assumption and Lemma 5.5, there is $c_i^1 \in \mathrm{core}(G^1\langle c_1^1, \ldots, c_{i-1}^1\rangle, 3^{k-i+1})$ such that $\mathrm{core}(G^0\langle c_1^0, \ldots, c_{i-1}^0\rangle, 3^{k-i+1})\langle c_i^0\rangle \overset{k-i}{\simeq} \mathrm{core}(G^1\langle c_1^1, \ldots, c_{i-1}^1\rangle, 3^{k-i+1})\langle c_i^1\rangle$. It is easy to see that $N(c_i^1, 3^{k-i}) \subseteq \mathrm{core}(G^1\langle c_1^1, \ldots, c_{i-1}^1\rangle, 3^{k-i+1})$, so again by Lemma 5.4, $\mathrm{core}(G^0\langle c_1^0, \ldots, c_i^0\rangle, 3^{k-i}) \overset{k-i}{\simeq} \mathrm{core}(G^1\langle c_1^1, \ldots, c_i^1\rangle, 3^{k-i})$. ∎

6. COMBINATORIAL PROOFS

Theorems 4.9 and 4.10 require some preliminary results first. Theorem 4.8 can be dealt with now.

Proof of Theorem 4.8. $G = \langle V, E\rangle$ is not r-simple if there exist two cycles X_1 and X_2 in G and a path P from X_1 and X_2 such that $|X_1|, |X_2|, |P| \leq 2r+1$. Now $|E \cap (X_1 \cup X_2 \cup P)| \geq |X_1 \cup X_2 \cup P| + 1$ and $|X_1 \cup X_2 \cup P| \leq 6r+1$, so $\mathrm{pr}(G$ is not r-simple$) \leq \Sigma_{s \leq 6r+1} n^s(\beta n^{-1})^{s+1} \to 0$ as $n \to \infty$. ∎

The proofs of Theorems 4.9 and 4.10 rely on the following combinatorial lemmas. Lemmas 6.1 and 6.2 are extensions of the inclusion–exclusion principle and Bonferroni's inequalities. They are slight generalizations of Lemmas 7.2 and 7.3 in [13]. Lemma 6.3 is identical to Lemma 7.4 in [13]. Let A and I be finite sets where a probability measure pr is defined on A. For every $i \in I$, let B_i be a collection of properties of members of A, say the elements of B_i are $C_{i0}, C_{i1}, \ldots$ where each $C_{ia} \subseteq A$.

Consider any family of sets $\vec{S} = \{S_i : i \in I\}$ such that each $S_i \subseteq B_i$, i.e., it is a set of properties. Let

$$E^{\geq}(\vec{S}) = \bigcap_{i \in I}\Big(\bigcap_{C_{ia} \in S_i} C_{ia}\Big)$$

$$E^{=}(\vec{S}) = E^{\geq}(\vec{S}) - \bigcup_{i \in I}\Big(\bigcup_{C_{ia} \in B_i - S_i} C_{ia}\Big).$$

That is, $E^{\geq}(\vec{S})$ is the set of elements in A that have all the properties in each S_i, and $E^{=}$ is the set of elements in A that have exactly those properties in each S_i. Let $\vec{s} = \langle s_i : i \in I\rangle$ be a sequence of nonnegative integers. Let $L(\vec{s}) = \Sigma_{\vec{S}}\, \mathrm{pr}(E^{\geq}(\vec{S}))$ where the sum is taken over all $\vec{S}$ such that $|S_i| = s_i$. For $J \subseteq I$ let $M(J, \vec{s}) = \bigcup_{\vec{S}} E^{=}(\vec{S})$ where the union is over all $\vec{S}$ such that $|S_i| = s_i$ for $i \in J$ and $|S_i| \geq s_i$ for $i \in I - J$. Thus $M(J, \vec{s})$ is the set of elements in A with exactly s_i properties in B_i for $i \in J$ and at least s_i properties in B_i for $i \in I - J$. We put $\Sigma(\vec{s})$ for $\Sigma_{i \in I} s_i$, and for any other sequence $\vec{t} = \langle t_i : i \in I\rangle$, $\vec{t} \geq \vec{s}$ means $t_i \geq s_i$ for all $i \in I$.

Lemma 6.1. *If $s_i > 0$ for all $i \in I - J$, then*

$$\mathrm{pr}(M(J, \vec{s})) = \sum_{\vec{t} \geq \vec{s}} (-1)^{\Sigma(\vec{t}) - \Sigma(\vec{s})} \prod_{i \in J} \binom{t_i}{s_i} \times \prod_{i \in I - J} \binom{t_i - 1}{s_i - 1} \times L(\vec{t})$$

Lemma 6.2. *If $s_i > 0$ for $i \in I - J$ and $v \geq \Sigma(\vec{s})$ then*

$$\sum_{\substack{\vec{t} \geq \vec{s} \\ \Sigma(\vec{t}) \geq v}} (-1)^{\Sigma(\vec{t}) - v} \prod_{i \in J} \binom{t_i}{s_i} \times \prod_{i \in I - J} \binom{t_i - 1}{s_i - 1} \times L(\vec{t}) \geq 0$$

Lemma 6.3. *Let $u : \omega \times \omega \to [0, \infty)$ and $\langle u_m : m \in \omega \rangle$ be a sequence in $[0, \infty)$ satisfying the following*:

(i) $\sum_{m \in \omega} (-1)^m u(n, m)$ *converges for all $n \in \omega$.*

(ii) $\sum_{m \in \omega} (-1)^m u_m$ *converges.*

(iii) $\lim_{n \to \infty} u(n, m) = u_m$ *for all $m \in \omega$.*

(iv) $\sum_{m \geq v} (-1)^{m-v} u(n, m) \geq 0$ *for all $n, v \in \omega$.*

Then $\lim_{n \to \infty} \sum_{m \in \omega} (-1)^m u(n, m) = \sum_{m \in \omega} (-1)^m u_m$.

Before proceeding with the proofs of Theorems 4.9 and 4.10, we give several definitions and establish some conventions that we will use. We have already defined the probability that the random graph with n vertices satisfies a sentence σ, i.e., $\mathrm{pr}(\sigma, n)$. We now extend this to formulas with free variables, say $\sigma(x_1, \ldots, x_i)$. For $n \geq i$, $\mathrm{pr}(\sigma(x_1, \ldots, x_i), n) = \Sigma_{\substack{G \models \sigma(a_1, \ldots, a_i) \\ |G| = n}} \mathrm{pr}(G)$ where $a_1, \ldots, a_i$ are any distinct elements of the vertex set $\{0, 1, \ldots, n-1\}$. Obviously, this definition does not depend on the choice of $a_1, \ldots, a_i$. An open formula is a formula without quantifiers. We will also use conditional probabilities. If σ and τ are two formulas (possibly with free variables), then $\mathrm{pr}(\sigma \mid \tau, n) = \mathrm{pr}(\sigma, n) / \mathrm{pr}(\tau, n)$.

Some abbreviations we will use are $\varphi_r(x_1, \ldots, x_i)$ where $G \models \varphi_r(x_1, \ldots, x_i)$ if and only if for every y there is at most one path from y to $\{x_1, \ldots, x_i\}$ of length $\leq r$, and $T_r(x_1, \ldots, x_i; y)$ for the operator tree $(\mathrm{core}(G_1 \langle c_1[x_1], \ldots, c_i[x_i] \rangle, r), y)$.

For each $b \in \omega$, using recursion on $r = \mathrm{radius}(C_b)$, we define the expression $\lambda_b \in \Lambda_\infty$ associated with the j-morphism class C_b. Its value will be the probability that $T_r(x_1, \ldots, x_i; y) \in C_b$ for any $x_1, \ldots, x_i, y$. We start with $\lambda_0 = 1$. Now assume that λ_a has been defined for all a such that $\mathrm{radius}(C_a) < \mathrm{radius}(C_b)$. It will be convenient to define λ_b in terms of an arbitrary characteristic vector $\vec{s} = \langle s_0, \ldots, s_{c-1} \rangle$ of C_b. For $a < c$, let

$$\xi_{ba} = \begin{cases} (\beta\lambda_a)^{s_a} e^{-\beta\lambda_a} / s_a! & \text{if } s_a \leq j \text{ and } \mathrm{radius}(C_a) < \mathrm{radius}(C_b) \\ 1 - (\Sigma_{s \leq j} (\beta\lambda_a)^s / s!) e^{-\beta\lambda_a} & \text{if } s_a = j + 1 \text{ and } \mathrm{radius}(C_a) < \mathrm{radius}(C_b) \\ 1 & \text{if } \mathrm{radius}(C_a) \geq \mathrm{radius}(C_b)\,. \end{cases}$$

Then $\lambda_b = \Pi_{a<c} \xi_{ba}$, and λ_b does not depend on the choice of $\vec{s}$. For example, when $j=2$, the expression associated with the class of radius 1 containing the rooted tree with two vertices is $\beta e^{-\beta}$, and the expression associated with the class containing the instances in Figure 2 is

$$[1-(1+\beta\lambda_2+\beta_2\lambda_2^2/2)e^{-\beta\lambda_2}]\beta\lambda_4 e^{-\beta\lambda_4}$$

where

$$\begin{aligned}
\lambda_2 &= (\beta\lambda_1)^2 e^{-\beta\lambda_1/2}\,, \\
\lambda_1 &= e^{-\beta}\,, \\
\lambda_4 &= \beta\lambda_3 e^{-\beta\lambda_3}, \text{ and} \\
\lambda_3 &= 1-(1+\beta+\beta^2/2)e^{-\beta}\,.
\end{aligned}$$

Lemma 6.4. *For any $i, r \in \omega$, $\lim_{n\to\infty} \mathrm{pr}(\varphi_r(x_1,\ldots,x_i)) = 1$.*

Proof. We show that $\lim_{n\to\infty} \mathrm{pr}(\neg\varphi_r(x_1,\ldots,x_i)) = 0$. Thus suppose there are two paths P_1 and P_2 from some y to $\{x_1,\ldots,x_i\}$ such that $|P_1|, |P_2| \le r+1$. The union of these two paths contains a path P_3 of length $u \le 2r-1$ from some x_k to some $z \not\in \{x_1,\ldots,x_i\}$ such that $z \sim w$ for some $w \in \{x_1,\ldots,x_i) \cup P_3$. For each u, the number of choices for vertices of P_3 is $i(n-i)\cdots(n-i-u+1)$, the number of choices for w is $i+u$, and the probability that there are edges joining the vertices in P_3 and $z \sim w$ is $(\beta n^{-1})^{u+1}$. Therefore, summing over $u = 1,\ldots,2r-1$, the probability that there is such a path P_3 is $O(n^{-1})$. ∎

The next lemma states that for any sequence of vertices $x_1,\ldots,x_i$, the probabilities that each $T_r(x_1,\ldots,x_i;x_k)$ are in a particular j-morphism class C_{b_k} are almost independent.

Lemma 6.5. *Let $i, r \in \omega$, $K \subseteq \{1,\ldots,i\}$ and suppose that for each $k \in K$, b_k is such that* radius$(C_{b_k}) \le r$. *Let $\rho(x_1,\ldots,x_i)$ be a consistent open formula. Then*

$$\lim_{n\to\infty} \mathrm{pr}\Big(\bigwedge_{k\in K} T_r(x_1,\ldots,x_i;x_k) \in C_{b_k} \,\Big|\, \rho(x_1,\ldots,x_i), n\Big) = \prod_{k\in K} \lambda_{b_k}\,.$$

Proof. We use induction on r. When $r=0$, $T_r(x_1,\ldots,x_i;x_k) = \langle\{x_k\}, \emptyset, c_1[x_k]\rangle$ and $\lambda_{b_k}=1$, and the result is immediate.

Now assume $r>0$ and the result has been proven for all smaller values. For each b such that radius$(C_b) \le r$, let $\vec{s}_b = \langle s_{b0},\ldots,s_{b,c-1}\rangle$ be a characteristic vector of C_b. Then for each $k \in K$,

$$\varphi_r(x_1,\ldots,x_i) \wedge T_r(x_1,\ldots,x_i;x_k) \in C_{b_k}$$

is equivalent to

$$\varphi_r(x_1,\ldots,x_i) \wedge \bigwedge_{a<c} \tau_{b_k a}(x_1,\ldots,x_i;x_k)$$

where $\tau_{ba}(x_1, \dots, x_i; x_k)$ is

$$(\exists y_0)\cdots(\exists y_{s-1})\Big(\bigwedge_{g,h<s} y_g \neq y_h \wedge \bigwedge_{k<s} (x_k \sim y_h \wedge T_{r-1}(x_1, \dots, x_i; y_h) \in C_a)$$
$$\wedge (\forall y_s)(x_k \sim y_s \wedge T_{r-1}(x_1, \dots, x_i; y_s) \in C_a \Rightarrow \bigvee_{h<s} y_h = y_s)\Big)$$

if $s_{ba} = s \leq j$, or

$$(\exists y_0)\cdots(\exists y_j)\Big(\bigwedge_{g,h\leq j} y_g \neq y_h \wedge \bigwedge_{h\leq j} (x_k \sim y_h \wedge T_{r-1}(x_1, \dots, x_i; y_h) \in C_a)\Big)$$

if $s_{ba} = j+1$.

Let $I = \{(k, a) : k \in K \text{ and } a < c\}$. For the remainder of this proof, we will take $\vec{s}$ to be the concatenation of all $\vec{s}_{b_k}$ where $k \in K$, i.e., $\vec{s} = \langle s_{b_k a} : (k, a) \in I \rangle$. By Lemma 6.4 we will be done if we prove

$$\lim_{n\to\infty} \mathrm{pr}\Big(\bigwedge_{(k,a)\in I} \tau_{b_k a}(x_1, \dots, x_i; x_k) \,|\, \rho(x_1, \dots, x_i) \wedge \varphi_r(x_1, \dots, x_i), n\Big) = \prod_{(k,a)\in I} \xi_{b_k a} \,. \tag{6.6}$$

Let $J = \{(k, a) \in I : s_{b_k a} \leq j\}$. Then by Lemma 6.1,

$$\mathrm{pr}\Big(\bigwedge_{(k,a)\in I} \tau_{b_k a}(x_1, \dots, x_i; x_k) \,|\, \rho(x_1, \dots, x_i) \wedge \varphi_r(x_1, \dots, x_i), n\Big)$$
$$= \sum_{m \geq \Sigma(\vec{s})} (-1)^{m-\Sigma(\vec{s})} u(m, n)\,,$$

where

$$u(m, n) = \sum_{\substack{\vec{t} > \vec{s} \\ \Sigma(\vec{t}) = m}} \Big[\prod_{(k,a)\in J} \binom{t_{ka}}{s_{b_k a}} \times \prod_{(k,a)\in I-J} \binom{t_{ka}-1}{j}$$
$$\times (n-i)(n-i-1)\cdots(n-i-m+1) \Big/ \prod_{(k,a)\in I} t_{ka}!$$
$$\times \mathrm{pr}\Big(\bigwedge_{(k,a)\in I} \Big(\bigwedge_{h<t_{ka}} (x_k \sim y_h^{ka}$$
$$\wedge T_{r-1}(x_1, \dots, x_i, \dots, y_{h'}^{k'a'}, \dots; y_h^{ka}) \in C_a)\Big)$$
$$|\rho(x_1, \dots, x_i) \wedge \varphi_r(x_1, \dots, x_i), n\Big)\Big]$$

(We add the superscripts to each variable y to indicate that all y_h^{ka} and $y_h^{k'a'}$ are distinct when $(k, a) \neq (k', a')$. In $T_{r-1}(x_1, \dots, x_i, \dots, y_{h'}^{k'a'}, \dots; y_h^{ka})$, k', a', h' range over all $(k', a') \in I$ and $h' < t_{k'a'}$.) Let

$$u_m = \sum_{\substack{\vec{t}>\vec{s} \\ \Sigma(\vec{t})=m}} \prod_{(k,a)\in J} (s_{b_ka}!(t_{ka}-s_{b_ka})!)^{-1}(\beta\lambda_a)^{t_{ka}}$$

$$\times \prod_{(k,a)\in I-J} (j!(t_{ka}-j-1)!t_{ka})^{-1}(\beta\lambda_a)^{t_{ka}} .$$

Equation (6.6) will follow from applying Lemma 6.3 to $u(m, n)$ and u_m. Condition (i) is obvious since $u(m, n) = 0$ for $m > n$. To show condition (ii), for $(k, a) \in J$, letting $s = s_{b_ka}$,

$$\sum_{t\geq s} (-1)^{t-s}(s!(t-s)!)^{-1}(\beta\lambda_a)^t = (\beta\lambda_a)^s e^{-\beta\lambda_a}/s! = \xi_{b_ka}$$

and for $(k, a) \in I - J$,

$$\sum_{t>j} (-1)^{t-j-1}(j!(t-j-1)!t)^{-1}(\beta\lambda_a)^t$$

$$= \sum_{t>j} (-1)^{t-j-1}\left[\sum_{s\leq j} (-1)^{j-s}\binom{t}{s}\right](\beta\lambda_a)^t/t!$$

$$= 1 - \sum_{t\leq j}\left[\sum_{s<t} (-1)^{t-s}\binom{t}{s}\right](\beta\lambda_a)^t/t! - \sum_{t>j}\left[\sum_{s\leq j} (-1)^{t-s}\binom{t}{s}\right](\beta\lambda_a)^t/t!$$

$$= 1 - \sum_{s\leq j}\left[\sum_{t\geq s} (-1)^{t-s}\binom{t}{s}(\beta\lambda_a)^t/t!\right]$$

$$= 1 - \sum_{s\leq j} (\beta\lambda_a)^s e^{-\beta\lambda_a}/s!$$

$$= \xi_{b_ka}$$

Therefore $\Sigma_{m\geq\Sigma(\vec{s})} (-1)^{m-\Sigma(\vec{s})}u_m = \Pi_{(k,a)\in I}\, \xi_{b_ka}$.

To show (iii), fix $\vec{t} \geq \vec{s}$ such that $\Sigma(\vec{t}) = m$. Then

$$\mathrm{pr}\Big(\bigwedge_{(k,a)\in I}\Big(\bigwedge_{h<t_{ka}} (x_k \sim y_h^{ka} \wedge T_{r-1}(x_1, \ldots, x_i, \ldots, y_{h'}^{k'a'}, \ldots; y_h^{ka}) \in C_a)\Big)$$

$$|\rho(x_1, \ldots, x_i) \wedge \varphi_r(x_1, \ldots, x_i), n) =$$

$$\mathrm{pr}\Big(\bigwedge_{(k,a)\in I}\Big(\bigwedge_{h<t_{ka}} T_{r-1}(x_1, \ldots, x_i, \ldots, y_{h'}^{k'a'}, \ldots; y_h^{ka}) \in C_a\Big)$$

$$|\rho(x_1, \ldots, x_i) \wedge \bigwedge_{(k,a)\in I}\Big(\bigwedge_{h<t_{ka}} x_k \sim y_h^{ka}\Big) \wedge \varphi_r(x_1, \ldots, x_i), n\Big) \tag{6.7}$$

$$\times\, \mathrm{pr}\Big(\bigwedge_{(k,a)\in I}\Big(\bigwedge_{h>t_{ka}} x_k \sim y_h^{ka}\Big)$$

$$|\rho(x_1, \ldots, x_i) \wedge \varphi_r(x_1, \ldots, x_i), n) \tag{6.8}$$

By Lemma 6.4 and the induction assumption, the limit of the probability in (6.7) is $\Pi_{(k,a)\in I}(\beta\lambda_a)^{t_{ka}}$. The probability in (6.8) is equal to n^{-m}, and therefore (iii) holds.

Lastly, (iv) holds by Lemma 6.2. ∎

Proof of Theorem 4.9. We will show that

$$\lim_{n\to\infty} \operatorname{pr}\Big((\exists x_0)\cdots(\exists x_{i-1})\varphi_{3r}(x_0,\ldots,x_{i-1}) \wedge \bigwedge_{k<i} \langle N(\{x_k\}, r), E, c_i[x_k]\rangle \stackrel{j}{\simeq} T\Big) = 1 \tag{6.9}$$

Condition $\varphi_{3r}(x_0,\ldots,x_{i-1})$ implies that there are no cycles of size less than or equal to $2r+1$ within a distance of $2r$ from any x_k, and $\delta(x_h, x_k) > 6r$ for all $h \neq k$. Therefore if $G \models \varphi_{3r}(x_0,\ldots,x_{i-1})$, then for any $c_1,\ldots,c_{i-1} \in V$, there is some k such that $\delta(\{c_1,\ldots,c_{i-1}\}, x_k) > 2r$, and $N(x_k, r) \cap \operatorname{core}(G\langle c_1,\ldots,c_{i-1}\rangle, r) = \emptyset$.

For $m \geq i$ and $t \leq m$ let $\sigma_{tm}(x_0,\ldots,x_{t-1}, y_0,\ldots,y_{m-1})$ be an abbreviation for

$$\begin{aligned}&\{x_0,\ldots,x_{t-1}\} \subseteq \{y_0,\ldots,y_{m-1}\} \wedge \varphi_{3r}(x_0,\ldots,x_{t-1})\\ &\wedge \bigwedge_{k<t} \langle N(\{x_k\}), E, c_i[x_k]\rangle \stackrel{j}{\simeq} T)\,.\end{aligned}$$

Equation (6.9) will follow if we show that $\lim_{n\to\infty} \operatorname{pr}((\exists x_0)\cdots(\exists x_{i-1})\sigma_{im}, n) = u_m$ for some u_m, and $\lim_{m\to\infty} u_m = 1$.

Let $\lambda \in \Lambda_\infty$ be the expression associated with the j-morphism class of T. Then by inclusion–exclusion,

$$\operatorname{pr}((\exists x_0)\cdots(\exists x_{i-1})\sigma_{im}, n) = 1 - \sum_{s<i}\Big[\sum_{s\leq t\leq m} (-1)^{t-s}\binom{t}{s}\binom{m}{t}\operatorname{pr}(\sigma_{tm}, n)\Big]$$

By Lemmas 6.4 and 6.5, $\lim_{n\to\infty} \operatorname{pr}(\sigma_{tm}) = \lambda^t$. Therefore

$$\begin{aligned}\lim_{n\to\infty} \operatorname{pr}((\exists x_0)\cdots(\exists x_{i-1})\sigma_{im}, n) &= 1 - \sum_{s<i}\Big[\sum_{s\leq t\leq m} (-1)^{t-s}\binom{t}{s}\binom{m}{t}\lambda^t\Big]\\ &= 1 - \sum_{s<i}\binom{m}{s}\lambda^s(1-\lambda)^{m-s}\\ &\to 1 \text{ as } m\to\infty\end{aligned}$$

because $\lambda \neq 0$. ∎

Proof of Theorem 4.10. Just as we associated an expression in Λ_∞ with each j-morphism class with center $\{c_1\}$, we now associate an expression with each j-morphism class D whose center is a single cycle. Suppose $|\operatorname{center}(D)| = i$ and $\operatorname{radius}(D) = r$. Then D is characterized by a cycle whose vertices $\{x_1,\ldots,x_i\}$ are colored with elements from $\{0,\ldots,c-1\}$ (the set of indices of the j-morphism classes with center $\{c_1\}$ and radius $\leq r$). The color of vertex x_k is b_k, meaning the tree attached to x_k is in class C_{b_k}. Letting $\operatorname{aut}(D)$ be the automorphism group of this colored cycle, the expression associated with D is

$$\xi = (\beta^i/|\operatorname{aut}(D)|)\prod_{1\leq k\leq i}\lambda_{b_k}\,.$$

ξ is related to the asymptotic probability of the following event. Let $X = \{x_1, \ldots, x_i\}$ and $\rho(X)$ be an open formula that describes all the edges in the center of D. That is, we take some fixed set of edges F on X such that $\langle X, F \rangle$ is isomorphic to center(D), and

$$\rho(X) = \bigwedge_{(x_h, x_k) \in F} x_h \sim x_k \wedge \bigwedge_{(x_h, x_k) \notin F} \neg(x_h \sim x_k)\,.$$

Then

$$\mathrm{pr}(\varphi_r(X) \wedge N(X, r) \in D) =$$
$$(i!/|\mathrm{aut}(D)|)\mathrm{pr}\Big(\varphi_r(X) \wedge \bigwedge_{k<i} T_r(x_1, \ldots, x_i; x_k) \in C_{b_k} \,\big|\, \rho(X)\Big) \times \mathrm{pr}(\rho(X))$$

so by Lemmas 6.4 and 6.5,

$$\mathrm{pr}(N(X, r) \in D) \sim i!\xi n^{-1}$$

More generally, let $\{D_a : a \in K\}$ be a collection of connected cyclic j-morphism classes with $|\mathrm{center}(D_a)| = i_a$ and radius$(D_a) \le r$, ξ_a be the expression associated with D_a, and $\rho_a(X_a)$ be the open formula that describes center(D_a). (The D_a's need not be distinct.) Then for any collection $\{X_a : a \in K \text{ and } |X_a| = i_a\}$ where the X_a's are pairwise disjoint,

$$\mathrm{pr}\Big(\bigwedge_{a\in K} N(X_a, r) \in D_a\Big) \sim \prod_{a \in K} i_a!\xi_a n^{-i_a}\,. \tag{6.10}$$

Let the components of the j-agreeability class A be included in $\{D_a : a \in I\}$. (Here, we assume all the D_a's in I are distinct.) A can be characterized by the vector $\vec{s} = \langle s_a : a \in I \rangle$. Take any $G \in A$ and let T_a be as defined in equation (4.6). Then for all $a \in I$

$$s_a = \begin{cases} |T_a| & \text{if } |T_a| \le j \\ j+1 & \text{if } |T_a| > j\,. \end{cases}$$

Then by Lemma 6.1, letting $J = \{a \in I : s_a \le j\}$,

$$\mathrm{pr}(\mathrm{core}(G, r) \in A, n) = \sum_{m \ge \Sigma(\vec{s})} (-1)^{m - \Sigma(\vec{s})} u(m, n)$$

where

$$u(m, n) = \sum_{\substack{\vec{t} \ge \vec{s} \\ \Sigma(\vec{t}) = m}} \Bigg[\prod_{a\in J} \binom{t_a}{s_a} \times \prod_{a \in I-J} \binom{t_a - 1}{j}$$
$$\times\, n(n-1)\cdots\Big(n+1-\sum_{a\in I} t_a i_a\Big) \Big/ \Big(\prod_{a\in I} t_a!(i_a!)^{t_a}\Big)$$
$$\times\, \mathrm{pr}\Big(\bigwedge_{a\in I}\Big(\bigwedge_{h<t_a} N(X_{ah}, r) \in D_a, n\Big)\Big)\Bigg]$$

Here, we implicitly assume $|X_{ah}| = i_a$ and all X_{ah}'s are pairwise disjoint.

The remainder of the proof is similar to the proof of Lemma 6.5. Let

$$u_m = \sum_{\substack{\vec{t} > \vec{s} \\ \Sigma(\vec{t}) = m}} \prod_{a \in J} (s_a!(t_a - s_a)!)^{-1} \xi_a^{t_a} \times \prod_{a \in I-J} (j!(t_a - j - 1)!t_a)^{-1} \xi_a^{t_a} .$$

Checking the conditions of Lemma 6.3, (i) is immediate, and (ii) follows from the same reasoning used in Lemma 6.5. Here, we get

$$\sum_{m \geq \Sigma(\vec{s})} (-1)^{m - \Sigma(\vec{s})} u_m = \prod_{a \in J} \xi_a^{s_a} e^{-\xi_a} / s_a! \times \prod_{a \in I-J} \left(1 - \sum_{s \leq j} \xi_a^s e^{-\xi_a} / s!\right) \in \Theta_\infty .$$

Condition (iii) follows from Equation 6.10, and condition (iv) follows from Lemma 6.2. ■

7. THE CASE $\alpha = (l + 1)/l$

Although this case is not a corollary of the case when $\alpha = 1$, the proof has the same structure, and the important definitions are similar. We shall try to show these analogies by using the same terms wherever possible, with the understanding that now they are defined for $\alpha = (l + 1)/l$, l fixed.

Definition 7.1. For a rooted graph $G = \langle V, E, c_1, \ldots, c_i \rangle$ core(G) is the union of all components in G with at least $l + 1$ vertices and all components containing some constant.

Definition 7.2. Let G be a graph. G is *simple* if all the components of G are trees with at most $l + 1$ vertices.

Definition 7.3. Let G be a graph, $1 \leq i \in \omega$. G is *i-rich* if for every rooted tree T with root $\{c_i\}$ and at most l vertices, and $c_1, \ldots, c_{i-1} \in V$, there exists $c_i \in V$ such that if X is the component of G containing c_i, then $X \cap \text{core}(G\langle c_1, \ldots, c_{i-1} \rangle) = \emptyset$ and $\langle X, E, c_i \rangle$ is isomorphic to T.

Definition 7.4. Let G^0 and G^1 be rooted graphs and $j \in \omega$. Let $D_0, \ldots, D_{d-1}$ be an enumeration of all isomorphism classes of rooted graphs that are represented by some component of G^g, $g = 0, 1$. For such g and $a < d$, let

$$T_a^g = \{X : X \text{ is a component of } G^g \text{ and } X \in D_a\}$$

Then G^0 and G^1 are *j-agreeable* if for all $a < d$,

$$|T_a^0| = |T_a^1| \text{ or } |T_a^0|, |T_a^1| > j .$$

Theorem 7.5. *Let $1 \leq k \in \omega$ and G^0, G^1 be graphs. If G^0 and G^1 are i-rich for $1 \leq i \leq k$, and* core(G^0) *is k-agreeable to* core(G^1), *then $\Gamma_k(G^0, G^1)$ is a win for Player II.*

Theorem 7.6 $\lim_{n\to\infty} \mathrm{pr}(G \textit{ is simple}, n) = 1$.

Theorem 7.7. *For every* $i \in \omega$, $i \geq 1$, $\lim_{n\to\infty} \mathrm{pr}(G \textit{ is i-rich}, n) = 1$.

Theorem 7.8. *For every* $j \in \omega$ *and every* j*-agreeability class* A *whose components are trees with exactly* $l+1$ *vertices, there is a* $\theta \in \Theta_l$ *such that* $\lim_{n\to\infty} \mathrm{pr}(\mathrm{core}(G) \in A, n) = \theta$.

Theorems 7.6 and 7.7 are well-known results of random graph theory [15]. The proofs of Theorems 7.5 and 7.8 are similar to the proofs of Theorems 4.7 and 4.10, respectively.

8. CONCLUSIONS

Our techniques extend to a number of similar classes of structures. For example, our two main theorems also apply to directed graphs and graphs with colored edges. Formally, graphs with colored edges are structures $\langle V, E_1, \ldots, E_h \rangle$ where each E_i is a set of edges on V (the edges colored i). Our language now has binary predicates $\overset{h}{\sim}, \ldots, \overset{h}{\sim}$ for each type of edge. We can permit multiple edges of different colors between the same pair of vertices, or we can require that there is at most one edge of any color between them. In the first case, for each $i = 1, \ldots, h$ the edges in E_i are chosen independently as before. We can even permit different edge probabilities p_i for each i. In the second case, there is an edge between any two vertices with probability p, and its color is chosen with probability $1/h$.

Our methods also work for the other well-known model of random graphs where the number of edges q is fixed for each n, and all graphs with n vertices and q edges are equally likely. This is referred to as "model B" in [15]. Taking $q(n) = \Theta(p(n)n^2)$, we get the same conclusions as in our main Theorems 2.1 through 2.4. We expect that there is a general theorem that implies all of these results as corollaries. However, it may be so awkward to state that its value would be questionable.

Some more interesting problems would be to prove limit laws for structures with relations of degree greater than 2, i.e., relations that are d-tuples for $d > 2$. Our techniques rely heavily on graph-theoretic concepts such as path and cycle, and it is not obvious how to extend them to relations of higher degree.

There are classes of structures that have limit laws for second-order sentences [1]. Here, we may quantify over sets and relations on the universes of the structures in addition to quantifying over elements of the universes. Of course, this gives us much more expressibility, and there are many instances where second-order sentences do not have asymptotic probabilities (see [2]), but for very sparse random graphs, limit laws may still hold.

Another area pertains to structures with built-in relations. That is, for every n, some of the relations on $\{0, 1, \ldots, n-1\}$ have fixed interpretations. The remaining relations are chosen randomly. Limit laws for such structures were studied in [12], but only for constant probabilities. We conjecture that for edge probabilities $\leq n^{-1}$, all sentences pertaining to such structures have an asymptotic probability.

Negative results were given in other articles. In particular, for structures $\langle\{0, 1, \ldots, n-1\}, \leq, R\rangle$ where $\leq$ is a built-in relation with the usual meaning and R is a random binary relation with constant edge probability, it was shown in [4] that there is a sentence that does not have an asymptotic probability. This result has recently been extended to variable edge probabilities by the author in collaboration with P. Dolan and J. Spencer (in preparation).

Finally, we believe that it is still worthwhile to seek limit laws for graphs when the edge probability is $n^{-\alpha}$, α rational, in spite of the negative result in [16]. The counterexample in that article is an extremely complicated sentence. There may be large classes of sentences satisfying some simple syntactic or model-theoretic condition, that have limit laws.

REFERENCES

[1] K. J. Compton, A logical approach to asymptotic combinatorics II: monadic second-order properties, *J. Combinat. Th. Ser. A*, **50**, 110–131 (1989).

[2] K. J. Compton, 0–1 laws in logic and combinatorics, *Proc. 1987 NATO Adv. Study Inst. on Algorithms and Order*, I. Rival, Ed., Reidel, Dordrecht, 1988.

[3] K. J. Compton and C. W. Henson, A uniform method for proving lower bounds on the complexity of logical theories, *Ann. Pure and Applied Logic*, **48**, 1–79 (1990).

[4] K. J. Compton, C. W. Henson, and S. Shelah, Nonconvergence, undecidability, and intractability in asymptotic problems, *Ann. Pure and Applied Logic*, **36**, 207–224 (1987).

[5] P. Dolan, Undecidable statements and random graphs, *Discrete App. Math.*, to appear.

[6] A. Ehrenfeucht, An application of games to the completeness problem for formalized theories, *Fundam. Math.*, **49**, 129–141 (1961).

[7] P. Erdös and A. Rényi, On the evolution of random graphs. *Mag. Tud. Akad. Mat. Kut. Int. Közl.*, **5**, 17–61 (1960).

[8] R. Fagin, Probabilities of finite models, *J. Symbolic Logic*, **41**, 50–58 (1976).

[9] Y. V. Glebskiĭ, D. I. Kogan, M. I. Liogon'kiĭ, and V. A. Talanov, Range and degree of realizability of formulas in the restricted predicate calculus, *Kibernetika (Kiev)*, **2**, 17–28 (1969); English translation, *Cybernetics*, **5**, 142–154 (1972).

[10] E. Grandjean, Complexity of the first-order theory of almost all finite structures, *Inf. Control*, **57**, 180–204 (1983).

[11] F. Harary, *Graph Theory*, Addison-Wesley, Reading, MA 1969.

[12] J. F. Lynch, Almost sure theories, *Ann. Math. Logic*, **18**, 91–135 (1980).

[13] J. F. Lynch, Probabilities of first-order sentences about unary functions, *Trans. AMS*, **287**, 543–568 (1985).

[14] E. Mendelson, *Introduction to Mathematical Logic*, Wadsworth and Brooks/Cole, Monterey, CA, 1987.

[15] E. M. Palmer, *Graphical Evolution*, Wiley, New York, 1985.

[16] S. Shelah and J. Spencer, Zero-one laws for sparse random graphs, *J. AMS*, **1**, 97–115 (1988).

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

Received June 29, 1990