



**FACULTY
OF MATHEMATICS
AND PHYSICS**
Charles University

MASTER THESIS

Ondřej Ježil

Pseudofinite structures

Department of Algebra

Supervisor of the master thesis: prof. RNDr. Jan Krajíček Ph.D.

Study programme: Mathematics

Study branch: Mathematics for Information
technologies

Prague 2022

I declare that I carried out this master thesis independently, and only with the cited sources, literature and other professional sources. It has not been used to obtain another or the same degree.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

In date
Author's signature

Dedication.

Title: Pseudofinite structures

Author: Ondřej Ježil

Department: Department of Algebra

Supervisor: prof. RNDr. Jan Krajíček Ph.D., Department of Algebra

Abstract: Abstract.

Keywords: key words

Contents

Introduction	2
Preliminaries	3
1 Forcing with random variables and the limit	5
1.1 Setup	5
1.2 The first order limit	5
1.3 The second order limit	7
1.4 The L -closed family F_{rud} and G_{rud}	8
1.5 Different choices of n	8
2 The limit	10
2.1 Basic examples	10
2.1.1 $\mathcal{G}_k = \text{EDGE}_k$	10
2.1.2 Sparse \mathcal{G}_k	10
2.2 Isomorphism closed \mathcal{G}_k	13
3 $F = F_{tree}$	15
3.1 Basic observations	15
3.2 $\mathcal{G}_k = \text{ALL}_k$	16
4 $F = F_{nbtree}$	18
4.1 $\mathcal{G}_k = \text{*PATH}_k$	18
Conclusion	23
Bibliography	24

Introduction

There exist several logical constructions of limits of classes of finite structures such as the ultraproduct construction or using the compactness theorem. The latter used in [Fag76] to prove the 0-1 law for structures over relational vocabularies.

In combinatorics there are also several notions of limits of finite graphs. For example the dense graph limit defined for a sequence of graph $\{G_k\}_{k>0}$ satisfying the condition that

$$t(F, G_n) = \frac{\text{hom}(F, G)}{|G_n|^{|F|}}$$

converges for every fixed connected graph F which provided a framework (see [LS06]) to restate and find new proofs for results in extremal graph theory such as Goodman's theorem relating the number of edges to the number of triangles in a graph. There are other notions of limits of sequences of graphs we recommend to read [NDM13] to the interested reader. Another recent use of limit objects for the results of extremal combinatorics was by Razborov in [Raz07].

These different notions of limits directly or tangentially relate to the concept of pseudofinite structures. For a first order language L , we call an L -structure S pseudofinite, if it satisfies the theory T_f consisting of all sentences true in all finite L -structures.

In this thesis we use the concept of pseudofinite structures to define a limit of a family of finite graphs relative to some computationally restricted class of functions. Instead of studying the density of substructures we study these so called wide limits both generally and by analyzing concrete examples and tying them with complexity-theoretical statements.

The key method we use is arithmetical forcing with random variables which allows us to construct models of (weak) arithmetical theories and by restricting to a language of graphs gives us boolean valued graphs. An important resource for this technique is [Kra10]. In these boolean valued graphs, witnessing of existential quantifiers corresponds to complexity theoretical properties of the class of graphs we are considering.

(TODO: Comment on structure of chapters.)

Preliminaries

In this chapter we recall a few important notions which we will use in the next chapter to define the central construction we study. We do not review the notions formally but always provide a reference for the reader unfamiliar with these topics. Throughout this thesis we assume a basic knowledge of mathematical logic, model theory and measure theory. Important concept for us are the nonstandard models of true arithmetic.

The ambient model \mathcal{M}

We call a model of $\text{Th}(\mathbb{N})$ nonstandard if it is not isomorphic to \mathbb{N} . Every model of $\text{Th}(\mathbb{N})$ of course contains an initial segment isomorphic to \mathbb{N} so we can view nonstandard models as those which also contain ‘infinite natural numbers’, if we assume $\mathcal{M} \models \text{Th}(\mathbb{N})$ contains \mathbb{N} as an initial segment then the elements $\mathcal{M} \setminus \mathbb{N}$ are called nonstandard. We recommend the introduction of [Kay91] for a review of this topic. In the appendix of [Kra10] there is an explicit ultrapower construction of a model $\mathcal{M} \models \text{Th}(\mathbb{N})$ which is \aleph_1 -saturated.

This \aleph_1 -saturated model \mathcal{M} is used throughout this thesis and we call it the ambient model of arithmetic. For our applications we just need to know that the model is nonstandard and the following property holds because of the \aleph_1 -saturation.

Property. If $\{a_k\}_{k \geq 0}$ is a sequence with elements in \mathbb{N} then there is an element $t \in \mathcal{M} \setminus \mathbb{N}$ and a sequence $\{b_k\}_{k < t} \in \mathcal{M}$ with $a_k = b_k$ for all $k \in \mathbb{N}$.

By overspill in \mathcal{M} if some definable property P holds for any a_k with high enough index then there is also some nonstandard $s < t$ such that b_s satisfies the property P .

Nonstandard analysis

The reader can refer to [Gol14] for more formal treatment of topics discussed in this section including proofs. To use the method of forcing with random variables we need to consider the concept of so-called \mathcal{M} -rationals. To define them we start by simply adjoining all negative elements to the semiring \mathcal{M} to obtain the integral domain $\overline{\mathcal{M}}$. \mathcal{M} -rationals are then simply the ordered field of fractions $\text{Frac}(\overline{\mathcal{M}})$ which we denote $\mathbb{Q}^{\mathcal{M}}$.

There is a canonical injection $\mathbb{Q} \hookrightarrow \mathbb{Q}^{\mathcal{M}}$ whose image consist exactly of the ‘standard fractions’. We call a $q \in \mathbb{Q}^{\mathcal{M}}$ finite if there is a standard k such that $|q| < \frac{k}{1}$ otherwise we call it infinite. We call $q \in \mathbb{Q}^{\mathcal{M}}$ infinitesimal if q is infinite. One can check that $\mathbb{Q}^{\mathcal{M}}$ fulfills the axioms of hyperreal numbers which can be used as an alternative foundation to the concepts of mathematical analysis. The following is an important results which we use throughout the thesis.

Theorem. Let $q \in \mathbb{Q}^{\mathcal{M}}$ finite. Then there exist unique $r \in \mathbb{R}$ and an infinitesimal $m \in \mathbb{Q}^{\mathcal{M}}$ such that $q = r + m$.

We use the notation $\text{st}(q) := r$ and call r it the **standard part** of q .

The following result characterizes convergence of a sequence of rational numbers in the language of nonstandard analysis.

Theorem. Let $\{c_k\}_{k \geq 0}$ be a sequence of rational numbers. Then

$$\lim_{k \rightarrow \infty} c_k = r \in \mathbb{R}$$

if and only if for any nonstandard $n \in \mathcal{M}$ we have that $\text{st}(c_n) = r$.

We close this section with two inequalities heavily used in the proofs throughout the thesis.

Theorem (Bernoulli's inequality). Let $y \in \mathcal{M}$ and $x \in \mathbb{Q}^{\mathcal{M}}, x \geq -1$, then

$$(1 + x)^y \geq 1 + yx.$$

Theorem (Exponential inequality). Let $y \in \mathcal{M}$ and $x \in \mathbb{Q}^{\mathcal{M}}, x \geq 0$, then

$$\left(1 - \frac{x}{y}\right)^y \leq e^{-x}.$$

Total NP search problems and polynomial oracle time

1. Forcing with random variables and the limit

1.1 Setup

Our goal in this chapter is to provide a definition of a limit of a set an infinite set of finite graphs in which arbitrarily large graphs occur.

The following definition makes our requirements of such a class of graphs precise.

Definition 1.1.1. Let $\{\mathcal{G}_k\}_{k>0}$ be a sequence of finite sets of finite graphs. We call it a **wide sequence** if the following hold.

- There is an increasing sequence of positive whole numbers $\{g_k\}_{k>0}$ such that the underlying set of each $G \in \mathcal{G}_k$ is $\langle g_k \rangle$.
- $\lim_{k \rightarrow \infty} |\mathcal{G}_k| = \infty$

The second condition guarantees that \mathcal{G}_n is an infinite set for $n > \mathbb{N}$. Many interesting classes of graphs form a wide sequence if we restrict the vertex-sets to $\langle g_k \rangle$, where $\{g_k\}_{k>0}$ can be taken as the increasing sequence of all cardinalities in such a class.

Example 1.1.2. TODO: Add some examples!

1.2 The first order limit

Let \mathcal{M} be the \aleph_1 -saturated model of true arithmetic discussed in the previous chapter and let \mathcal{G}_k be a wide sequence of graphs and $\Omega := \mathcal{G}_n$ for $n \in \mathcal{M} \setminus \mathbb{N}$.

The model \mathcal{M} treats all its elements (including those which represent sets) as “finite objects” which lets us define uniform probability even on sets which are infinite from the set-theoretical perspective.

Definition 1.2.1. Let $\mathcal{A} := \{A \in \mathcal{M}; A \subseteq \Omega\}$ the set of all subsets of Ω represented by an element in \mathcal{M} .

We define the **counting measure** as the uniform probability of A when we sample Ω uniformly, so we have

$$A \in \mathcal{A} \rightarrow |A| / |\Omega|,$$

the counting measure takes values in \mathcal{M} -rationals.

One can check that \mathcal{A} is a boolean algebra, but not a σ -algebra as it is not closed under all countable unions. Indeed all singleton sets are part of \mathcal{A} but the set of all elements with standardly many predecessors in Ω is not in \mathcal{A} .

Definition 1.2.2. We call an \mathcal{M} -rational **infinitesimal** if it is smaller than all standard fractions $\frac{1}{k}, k \in \mathbb{N}$.

Define an ideal in \mathcal{A} as $\mathcal{I} := \{A \in \mathcal{A}; |A| / |\Omega| \text{ is infinitesimal}\}$. Define the boolean algebra $\mathcal{B} := \mathcal{A}/\mathcal{I}$. The induced measure on \mathcal{B} is a real-valued measure and can be written as

$$\mu(A/\mathcal{I}) = \text{st}(|A| / |\Omega|).$$

We can also check, that now μ is a measure in the ordinary sense and that \mathcal{B} is an σ -algebra. In fact the following key lemma holds.

Lemma 1.2.3. \mathcal{B} is a complete boolean algebra.

Now we define a \mathcal{B} -valued arithmetical model through which we define the \mathcal{B} -valued first order limit of \mathcal{G}_k relative to a family of arithmetical functions.

Definition 1.2.4. Let $L \subseteq L_{all}$ and let F be a non-empty set of functions in \mathcal{M} . We call it an **L -closed family** if it satisfies the following:

- The domain of any function in F is Ω and the range is \mathcal{M} .
- F is closed under all L -functions and contains all L constants, where the L -functions are interpreted by composition

$$f(\alpha_1, \dots, \alpha_k)(\omega) := f(\alpha_1(\omega), \dots, \alpha_k(\omega)),$$

for $k \in \mathbb{N}$, $f \in L$ k -ary and $\alpha_1, \dots, \alpha_k \in F$.

Note that while every $\alpha \in F$ is represented by some element in \mathcal{M} this need not be the case for the whole family F .

Definition 1.2.5. Let F be an L -closed family for some $L \subseteq L_{all}$. Then $K(F)$ will denote a \mathcal{B} -valued L -structure defined as follows.

The universe of $K(F)$ is F . The boolean evaluations of L -sentences are defined by the following inductive conditions:

- $\llbracket \alpha = \beta \rrbracket := \{\omega \in \Omega; \alpha(\omega) = \beta(\omega)\} / \mathcal{I}$.
- $\llbracket R(\alpha_1, \dots, \alpha_k) \rrbracket := \{\omega \in \Omega; R(\alpha_1, \dots, \alpha_k)\} / \mathcal{I}$ for any k -ary L -relation R .
- $\llbracket - \rrbracket$ commutes with \wedge, \vee, \neg .
- $\llbracket (\exists x)A(x) \rrbracket := \bigvee_{\alpha \in F} \llbracket A(\alpha) \rrbracket$.
- $\llbracket (\forall x)A(x) \rrbracket := \bigwedge_{\alpha \in F} \llbracket A(\alpha) \rrbracket$.

Finally, using $K(F)$ we can define the first order limit of \mathcal{G}_k using the following notions.

Definition 1.2.6. We call a function $\alpha \in F$ an F -vertex if $\alpha : \Omega \rightarrow \langle g_n \rangle$.

We define a \mathcal{B} -valued graph $\lim_{k \rightarrow n}^F \mathcal{G}_k$ as an $\{\Gamma\}$ -structure, where Γ is a binary relation symbol, with universe $\{\alpha \in F; \alpha \text{ is an } F\text{-vertex}\}$ and Γ -sentences being evaluated by the following inductive conditions:

- $\llbracket \alpha = \beta \rrbracket := \{\omega \in \Omega; \alpha(\omega) = \beta(\omega)\} / \mathcal{I}$.

- $\llbracket \Gamma(\alpha, \beta) \rrbracket := \{\omega \in \Omega; E_G(\alpha, \beta)\} / \mathcal{I}$.
- $\llbracket - \rrbracket$ commutes with \wedge, \vee, \neg .
- $\llbracket (\exists x)A(x) \rrbracket := \bigvee_{\alpha \in F} \llbracket A(\alpha) \rrbracket$.
- $\llbracket (\forall x)A(x) \rrbracket := \bigwedge_{\alpha \in F} \llbracket A(\alpha) \rrbracket$.

1.3 The second order limit

While we can find a truth value of a sentence in the language of graphs in the limit $\lim_F \mathcal{G}_n$, we will encounter situations where this is not sufficient to analyze the wide sequence $\{\mathcal{G}_k\}_{k>0}$.

In Chapter 3 we will investigate how does existence of large cliques correspond to the size of cliques in the limit graph. But we cannot just measure the set-theoretical cardinality of any such clique, for specific n we could very well have $\text{card}(\lfloor \log n \rfloor) = \text{card}(\lfloor \frac{n}{2} \rfloor)$ but from the point of view of complexity theory cliques of size $\lfloor \log n \rfloor$ and $\lfloor \frac{n}{2} \rfloor$ are dramatically different. In other words, our goal is also to have means to count the number elements of subsets or relations with values in (random variables in) \mathcal{M} .

Definition 1.3.1. Let $L \subseteq L_{all}$, we call a set of functions $G \subseteq \mathcal{M}$ an F -closed **functional family** if every $\Theta \in G$ assigns to every $\omega \in \Omega$ a function $\Theta_\omega \in \mathcal{M}$ and after we define

$$\Theta(\alpha)(\omega) := \begin{cases} \Theta_\omega(\alpha(\omega)) & \alpha(\omega) \in \text{dom}(\Theta_\omega) \\ 0 & \text{otherwise,} \end{cases}$$

we have that for every $\alpha \in F$ and $\Theta \in G$ we have $\Theta(\alpha) \in F$.

We call $\Theta \in G$ a (graph) G -**relation** if for every $\omega \in \Omega$ we have for some $k > 0$ that $\text{dom } \Theta_\omega \supseteq \langle g_n \rangle^k$ and $\Theta_\omega : \text{dom } \Theta_\omega \rightarrow \{0, 1\}$.

Definition 1.3.2. Let $L \subseteq L_{all}$, F an L -closed family and G an F -compatible functional family. We define the L^2 -structure $K(F, G)$ as a two sorted L -structure with sorts F and G interpreting L -sentences as $K(F)$ and treating the sort G as follows. First for equality we let

$$\llbracket \Theta = \Xi \rrbracket := \{\omega \in \Omega; \Theta_\omega = \Xi_\omega\} / \mathcal{I}$$

and for the second order quantifiers we have the following inductive clauses

- $\llbracket (\exists X)A(X) \rrbracket := \bigvee_{\Theta \in G} \llbracket A(\Theta) \rrbracket$
- $\llbracket (\forall X)A(X) \rrbracket := \bigwedge_{\Theta \in G} \llbracket A(\Theta) \rrbracket$.

If there is a $\Gamma \in G$ such that for every $\alpha, \beta \in F$ we have

$$\Gamma(\alpha, \beta)(\omega) := \chi_{E_\omega}(\alpha(\omega), \beta(\omega)),$$

where χ_{E_ω} is the characteristic function of E_ω , we call $K(F, G)$ the **underlying arithmetic of a second order wide limit**.

We define $\lim_{F,n}^G \{\mathcal{G}_k\}_{k>0}$ as the L^2 -substructure with universe consisting of all F -vertices and all G -relations. By abuse of notation we will mostly use the notation $\lim_F^G \mathcal{G}_n$.

1.4 The L -closed family F_{rud} and G_{rud}

Throughout this thesis we will mostly work with the L -closed family F_{rud} which ties the properties of $\lim_F \mathcal{G}_n$ with decision tree complexity.

After we choose the sequence $\{\mathcal{G}_k\}_{k>0}$ and $n > \mathbb{N}$ we again put $\Omega := \mathcal{G}_n$ and define F_{rud} as follows.

Definition 1.4.1. We define a **decision tree** to be a binary tree $T \in \mathcal{M}$ with a labelling of vertices and edges ℓ . The non-leaf vertices are labeled by pairs of numbers (u, v) , where $u, v \in \langle g_n \rangle$ and each edge is labeled either by 1 or 0. Each leaf vertex is then labeled by some element of \mathcal{M} .

Each $\omega \in \Omega$ uniquely determines a path in (T, ℓ) by interpreting the vertex labels as “is $(u, v) \in E_\omega$?” and the edge labels as true (1) and false (0). The path then uniquely determines an output.

We define F_{rud} to be the set of all functions computed by some (T, ℓ) of depth at most $n^{1/t}$.

One can verify that F_{rud} is an L -closed family for any $L \subseteq L_{all}$.

The definition of G_{rud} is a bit more involved. The functionals in it will be computed by tuples of elements from F_{rud} in the following sense.

Definition 1.4.2. Let $\hat{\beta} = (\beta_0, \dots, \beta_{m-1}) \in \mathcal{M}$ be a m -tuple of elements in F_{rud} , for any $\alpha \in F_{rud}$ and $\omega \in \Omega$ we define

$$\hat{\beta}(\omega) = \begin{cases} \beta_{\alpha(\omega)}(\omega) & \alpha(\omega) < m \\ 0 & \text{otherwise.} \end{cases}$$

Definition 1.4.3. The family G_{rud} consists of all functionals Θ such that there is $m \in \mathbb{N}$ and some $\hat{\beta} = (\beta_0, \dots, \beta_{m-1})$ that computes it.

Lemma 1.4.4. G_{rud} is (F_{rud}) -compatible.

Proof. By induction in \mathcal{M} we have that all the depth of all the trees is bounded by $n^{1/t}$ for some $t > \mathbb{N}$.

If we take some $\Theta \in G_{rud}$ and $\alpha \in F_{rud}$ we can compute $\Theta(\alpha)$ also by a tree in F_{rud} by concatenating the trees computing α and β_i s. \square

1.5 Different choices of n

Even though we generally pose no requirements on $n > \mathbb{N}$ there are examples of wide sequences for which the limit is sensitive on the choice of the non-standard number n .

Example 1.5.1. Let

$$G_k := \begin{cases} \{(\langle k \rangle, E); |E| = 2, E(0, 1)\} & k \text{ even} \\ \{(\langle k \rangle, E); |E| = 1, \neg E(0, 1)\} & k \text{ odd.} \end{cases}$$

Let $n > \mathbb{N}$ then

$$\lim_{F_{rud}} \mathcal{G}_{2n+1} \llbracket \Gamma(0, 1) \rrbracket = \mathbf{0}, \quad (1.1)$$

but

$$\lim_{F_{rud}} \mathcal{G}_{2n}[\Gamma(0, 1)] = \mathbf{1}. \quad (1.2)$$

Even though the concrete wide limits we will investigate in the following chapters do not depend on the specific n , it is important to note that we cannot generally remove the parameter n from the definition of the limit.

2. The limit

2.1 Basic examples

2.1.1 $\mathcal{G}_k = \text{EDGE}_k$

We consider the classes of graphs

$$\text{EDGE}_k := \{(k, E); |E| = 1\},$$

and we let $\mathcal{G}_k = \text{EDGE}_k$ and $F = F_{\text{tree}}$.

Intuitively one should not be able to find the edge on a significant i.e. non-infinitesimal fraction of samples with a tree that is allowed to explore only an infinitesimal fraction of edges.

Theorem 2.1.1. We will prove that

$$\llbracket (\exists x)(\exists y)\Gamma(x, y) \rrbracket = \mathbf{0}.$$

Proof. Let $T \in F_{\text{tree}}$ be a tree of depth $n^{1/t}$, for some $t > \mathbb{N}$ that outputs a pair of numbers less than n .

Start from the root of T and always choose the path that corresponds to an edge not existing. At the end we obtain some answer, that gives us a set of at most $2 \cdot n^{1/t} + 2$ vertices. Now we can find at least:

$$\binom{n - 2n^{1/t} - 2}{2} = \frac{(n - 2n^{1/t} - 2)(n - 2n^{1/t} - 3)}{2} \quad (2.1)$$

different $\omega \in \Omega$ such that $T(\omega)$ is not an edge in ω .

The probability that any of those graphs is sampled is

$$\frac{\binom{n - 2n^{1/t} - 2}{2}}{|\mathcal{G}_k|} = \frac{(n - 2n^{1/t} - 2)(n - 2n^{1/t} - 3)}{n(n - 1)} \quad (2.2)$$

$$= \left(1 - \frac{2n^{1/t} - 2}{n}\right) \cdot \left(1 - \frac{2n^{1/t} - 3}{n - 1}\right) \quad (2.3)$$

$$\geq \left(1 - \frac{2n^{1/t} - 2}{n}\right)^2 \quad (2.4)$$

$$\geq 1 - \frac{4n^{1/t} - 4}{n}. \quad (2.5)$$

And one can clearly see that $\text{st}(1 - \frac{4n^{1/t} - 4}{n}) = 1$. This proves that the boolean value we are considering is $\mathbf{0}$ since we can the two witnesses for x and y into a tree that could find an edge with depth $n^{1/t}$ for some $t > \mathbb{N}$. \square

2.1.2 Sparse \mathcal{G}_k

One can see that in Theorem 2.1.1 the exact shape of graphs in \mathcal{G}_k does not play a crucial role. If \mathcal{G}_k consisted of all graphs on k vertices containing say

exactly one triangle, or any other fixed subgraph of constant size, and no other edges, we would still find that the non-existence is valid in the limit graph.

A more general case would be to consider a family of graphs in which there is an infinitesimally small chance that two independent uniformly random vertices have an edge between. However, this is not sufficient.

Example 2.1.2. Let \mathcal{G}_k consist of those graphs on the vertex-set k which contain the edge $E(0, 1)$ and then has exactly one other edge.

As k increases, the number of edges get smaller than any standard positive fraction. But

$$\llbracket (\exists x)(\exists y)E(x, y) \rrbracket = \mathbf{1},$$

as witnessed by x being the constant 0 and y the constant 1 both of which are computable by a tree of depth 0.

One can see that having distinguished vertices can ruin the sparseness implying the non-existence of edges in the limit graph. We want to distinguish from this situation by considering the sequences \mathcal{G}_k satisfying the following natural definition.

Definition 2.1.3. We say that $\{\mathcal{G}_k\}_{k=0}^\infty$ is **isomorphism closed**, if there is k_0 such that for every $k > k_0$ if we have that $G_1 \in \mathcal{G}_k$, $V_{G_2} = g_k$ and $G_1 \cong G_2$ then $G_2 \in \mathcal{G}_k$.

Theorem 2.1.4. Let an isomorphism closed \mathcal{G}_k have the following property. There is a sequence $\{b_k\}_k$ and for big enough k , a uniformly sampled 2-element $\{u, v\} \subseteq g_k$ and every $G \in \mathcal{G}_k$ we have

$$\Pr[E_G(u, v)] \leq b_k,$$

and some k_0 such that $\lim_{k \rightarrow \infty} k^{1/k_0} b_k = 0$. Then

$$\llbracket (\exists x)(\exists y)\Gamma(x, y) \rrbracket = \mathbf{0}.$$

Proof. Let us define the number $c_{u,v} := |\{G \in \mathcal{G}_k; E_G(u, v)\}|$, which is the number of graphs G in \mathcal{G}_k satisfying $E_G(u, v)$. Of course $c_{u,u} = 0$ for every u .

Claim: Let $u \neq v, u' \neq v'$ be vertices, then $c_{u,v} = c_{u',v'}$.

proof of claim: Let $\rho := (uu')(vv')$ be a permutation with cycles (uu') and (vv') . We can let ρ act on \mathcal{G}_k by sending G to a graph $\rho(G)$ which renames the edges coordinate-wise.

The fact that \mathcal{G}_k is isomorphism closed implies that ρ restricts to a bijection:

$$\rho' : \{G \in \mathcal{G}_k; E_G(u, v), \neg E_G(u', v')\} \rightarrow \{G \in \mathcal{G}_k; E_G(u', v'), \neg E_G(u, v)\}$$

which proves the claim. □

Now we define a matrix with entries

$$a_{G, \{u, v\}} := \begin{cases} 1 & E_G(u, v) \\ 0 & \text{otherwise} \end{cases}$$

where the rows are indexed by one of $|\mathcal{G}_k|$ many graphs in \mathcal{G}_k and the columns are indexed by the $\frac{k(k-1)}{2}$ many 2-element sets of numbers in k . We take any distinct

vertices u, v and define $p := \Pr_{G \in \mathcal{G}_k}[E_G(u, v)] = \frac{c_{u,v}}{|\mathcal{G}_k|}$, by the claim the choice of u, v does not matter.

The assumption from the statement is equivalent to the equality

$$\sum_{\{u,v\}} a_{G,\{u,v\}} \leq \frac{k(k-1)}{2} b_k$$

for every G . We combine this with the claim and the definition of p to get

$$\frac{k(k-1)}{2} |\mathcal{G}_k| p = \sum_{\{u,v\}} \sum_{G \in \mathcal{G}_k} a_{G,\{u,v\}} \quad (2.6)$$

$$= \sum_{G \in \mathcal{G}_k} \sum_{\{u,v\}} a_{G,\{u,v\}} \quad (2.7)$$

$$\leq |\mathcal{G}_k| \frac{k(k-1)}{2} b_k \quad (2.8)$$

which implies

$$p \leq b_k.$$

Now let $k := n$ and let $T \in F_{tree}$ be a tree of depth $n^{1/t}$ for some $t > \mathbb{N}$, where every leaf of T is labeled by some edge. Walk down the tree T from the root by answering negatively to every edge, this gives us a set E_T of all edges T inspected or outputed and $|E_T| \leq n^{1/t} + 1$.

Now we just need to prove that the probability T find an edge is infinitesimally small. This is enough to prove the theorem, since the trees computing any two witnesses for x and y in the statement can be combined to construct T and if any tree T succeeds with only infinitesimally small probability, no random vertices can witness an edge on a set of non-zero measure.

We use the fact that $p \leq b_n$ to derive

$$\Pr_{G \in \mathcal{G}_n} [T \text{ finds an edge}] \leq \sum_{\{u,v\} \in E_T} \Pr_{G \in \mathcal{G}_n} [E_G(u, v)] \quad (2.9)$$

$$= \sum_{\{u,v\} \in E_T} \frac{c_{u,v}}{|\mathcal{G}_n|} \quad (2.10)$$

$$\leq \sum_{\{u,v\} \in E_T} p \quad (2.11)$$

$$= (n^{1/t} + 1)p \quad (2.12)$$

$$\leq (n^{1/t} + 1)b_k \quad (2.13)$$

$$\leq n^{1/k_0} b_k \quad (2.14)$$

$$\approx 0, \quad (2.15)$$

which proves the theorem. \square

The assumption $\lim_{k \rightarrow \infty} k^{1/k_0} b_k = 0$ for some k_0 may seem unintuitive at first. However, it precisely reflects what is “sparse” for the trees in T_{tree} . The following example shows that without the assumption the theorem fails.

Example 2.1.5. Let \mathcal{G}_k consist of all graphs on the vertex set $\langle k \rangle$ with precisely $\lceil \frac{k(k-1)}{2 \log k} \rceil$ edges.

Then we claim that

$$\llbracket (\exists x)(\exists y)\Gamma(x, y) \rrbracket = \mathbf{1}.$$

Let α and β be vertices computed by the tree of the same shape which inspects a set of any $n^{1/t}$ distinct edges for some $t > \mathbb{N}$. If it finds an edge we define α and β in any way so they are the distinct vertices incidental with this edge. Otherwise we let $\alpha(\omega) = \beta(\omega) = 0$.

Let T be a tree of the same shape, that computes the pair $\{\alpha, \beta\}$ then we can compute the probability where such a tree fails as the fraction of all graphs which have no edges that T inspects. Let $m = \binom{n}{2}$. We get

$$\Pr_{G \in \mathbb{G}_n} [T \text{ fails}] = \frac{\binom{m - n^{1/t}}{\lceil \frac{n(n-1)}{2 \log n} \rceil}}{\binom{m}{\lceil \frac{n(n-1)}{2 \log n} \rceil}} \quad (2.16)$$

$$= \frac{(m - n^{1/t})!}{\frac{\lceil \frac{n(n-1)}{2 \log n} \rceil! (m - \lceil \frac{n(n-1)}{2 \log n} \rceil - n^{1/t})!}{m!}} \quad (2.17)$$

$$= \frac{(m - n^{1/t})! (m - \lceil \frac{n(n-1)}{2 \log n} \rceil)!}{m! (m - \lceil \frac{n(n-1)}{2 \log n} \rceil - n^{1/t})!} \quad (2.18)$$

$$= \prod_{i=0}^{n^{1/t}-1} \frac{m - \lceil \frac{n(n-1)}{2 \log n} \rceil - i}{m - i} \quad (2.19)$$

$$= \left(1 - \frac{\lceil \frac{n(n-1)}{2 \log n} \rceil}{\frac{n(n-1)}{2}} \right)^{n^{1/t}} \quad (2.20)$$

$$\leq \left(1 - \frac{\lceil \frac{n(n-1)}{2 \log n} \rceil}{\frac{n(n-1)}{2}} \right)^{n^{1/t}} \quad (2.21)$$

$$\leq \left(1 - \frac{1}{\log n} \right)^{n^{1/t}} \quad (2.22)$$

And for any standard k we have

$$\left(1 - \frac{1}{\log n} \right)^{n^{1/t}} \leq \left(1 - \frac{1}{\log n} \right)^{k \cdot \log n} \quad (2.23)$$

$$\leq (e^{-\frac{1}{\ln 2}})^k. \quad (2.24)$$

So $\text{st}(\Pr_{G \in \mathbb{G}_n} [T \text{ fails}]) = 0$ and we get

$$\mu(\llbracket (\exists x)(\exists y)\Gamma(x, y) \rrbracket) \geq \mu(\llbracket \Gamma(\alpha, \beta) \rrbracket) \quad (2.25)$$

$$= \text{st}(1 - \Pr_{G \in \mathbb{G}_n} [T \text{ fails}]) \quad (2.26)$$

$$= 1. \quad (2.27)$$

2.2 Isomorphism closed \mathcal{G}_k

So far the measure of every truth value we encountered was either 0 or 1. Is there a sequence \mathcal{G}_k whose F_{tree} -limit and an $\{E\}$ -sentence φ such that

$0 < \llbracket \varphi^\Gamma \rrbracket < 1$? As in the case of edge existence for a limit of sparse graphs, it is not hard to come up with an example if we allow distinguishing elements in \mathcal{G}_k .

Example 2.2.1. Let

$$\mathcal{G}_k = \{(\langle k \rangle, E); E \text{ has exactly two edges, one of them being } \{0, 1\}\},$$

and let $F = F_{tree}$. Then $\mu(\llbracket (\exists x)(\exists y)\Gamma(x, y) \rrbracket) = \frac{1}{2}$.

Proof. Let T_0 be a tree that always outputs 0 and T_1 be a tree that always outputs 1. We can prove that $\llbracket \Gamma(0, 1) \rrbracket \geq \llbracket \Gamma(\alpha, \beta) \rrbracket$ for any α, β . \square

For the case of isomorphism closed \mathcal{G}_k we prove that every $\{E\}$ -sentence has truth value either **0** or **1**. We start with the existential case.

Example 2.2.2. NOT TRUE! Consider all graphs with one edge, or one non-edge. $\mu(\llbracket (\exists x)(\exists y)\Gamma(x, y) \rrbracket) = \frac{1}{2}$.

Theorem 2.2.3. Let $\varphi_0(\bar{x})$ be an open $\{E\}$ -formula, and let \mathcal{G}_k be isomorphism closed and categorical then

$$\llbracket (\exists \bar{x})\varphi_0^\Gamma(\bar{x}) \rrbracket \in \{\mathbf{0}, \mathbf{1}\}.$$

Proof. Let \bar{T} be a tuple of trees computing $\bar{\alpha}$ such that

$$p := \mu(\llbracket \varphi_0(\bar{\alpha}) \rrbracket) > \mathbf{0}.$$

We want to iterate \bar{T} to amplify the probability of success. Let

$$A = \{\omega \in \Omega; \omega \models \varphi_0(\bar{\alpha}(\omega))\}$$

and $\pi \in S_{|\mathcal{G}_n|}$ with maximal $|\pi[A] \setminus A|$. Let $\pi(\bar{T})$ be a tuple of relabeled trees computing $\pi(\bar{\alpha})$. \square

3. $F = F_{tree}$

3.1 Basic observations

Example 3.1.1. Let

$$\mathcal{G}_k = \{([k], E); E \text{ consists of exactly one } n/2\text{-clique}\},$$

and $F = F_{tree}$.

We will prove that for every $t > \mathbb{N}$:

$$\llbracket \Gamma \text{ has an } n^{1/t}\text{-clique} \rrbracket \quad (3.1)$$

$$= \llbracket (\exists \Lambda)(\forall u)(\forall v)((u, v \leq n^{1/t}) \rightarrow \Gamma(\Lambda(u), \Lambda(v))) \wedge (\Lambda : [n^{1/t}] \hookrightarrow \mathcal{M}) \rrbracket \quad (3.2)$$

$$= \bigvee_{\Lambda} \bigwedge_u \bigwedge_v \llbracket (u \neq v < n^{1/t}) \rightarrow (\Gamma(\Lambda(u), \Lambda(v)) \wedge \Lambda(u) \neq \Lambda(v)) \rrbracket \quad (3.3)$$

$$= \mathbf{1}. \quad (3.4)$$

For $j \in [n^{1/t}]$ let Λ_j to be a tree of depth $j \cdot (n^{1/t})^2$ which first tries to find an edge $1 \leftrightarrow k$ for $k \in [n^{1/t}]$ if it fails than it tries to find $2 \leftrightarrow k$ and so on. Once it finds some edge (i, k) , then it starts again but from $i + 1$ until it finds the first j elements of Δ_ω and responds with the j -th element. Since j is always bounded by $n^{1/t}$, Λ really sends F to F .

Example 3.1.2. Let

$$\mathcal{G}_k = \{([k], E); E \text{ consists of exactly one edge}\}$$

and $F = F_{tree}$.

Theorem 3.1.3. Let $\varphi = (\forall \bar{x})\varphi_0(\bar{x})$ be a universal $\{E\}$ -sentence, such that

$$\lim_{k \rightarrow \infty} \Pr_{G \in \mathcal{G}_k} (G \models \varphi) = 1.$$

Then φ is valid in the b.v. structure.

Proof. From \aleph_1 -saturation of \mathcal{M} and our assumption, we know that for each $m \in \mathbb{N}$ there exists a $k_0 \in \mathbb{N}$ such that

$$\mathcal{M} \models (\forall k > k_0) \left(\Pr_{G \in \mathcal{G}_k} (G \models \varphi) > 1 - 1/m \right).$$

Therefore, since $n > \mathbb{N}$, we have that $\text{st}(\Pr_{G \in \mathcal{G}_n} (G \models \varphi)) = 1$ and therefore $\llbracket \varphi_0(\bar{\alpha}) \rrbracket = \mathbf{1}$ for each tuple $\bar{\alpha}$ in F .

Therefore

$$\llbracket \varphi \rrbracket = \bigwedge_{\bar{\alpha}} \llbracket \varphi_0(\bar{\alpha}) \rrbracket \quad (3.5)$$

$$= \bigwedge_{\bar{\alpha}} \mathbf{1} \quad (3.6)$$

$$= \mathbf{1}. \quad (3.7)$$

□

Theorem 3.1.4. Let $F = F_{tree}$. Let $\varphi_0(x_0, \dots, x_{l-1})$ be a q.f. $\{E\}$ -formula. Let $0 < p \leq 1$, consider subset $A \subseteq [g_k]^l$ such that for all $\bar{a} \in A$

$$\Pr_{G \in \mathcal{G}_k} (G \models \varphi_0(\bar{a})) \geq p$$

and

$$\{\{G \models \varphi_0(\bar{a})\} \subseteq \mathcal{G}_k; \bar{a} \in A\} \text{ are mutually independent.}$$

moreover let A_k be the set with the largest cardinality that has this property.

If $\lim_{k \rightarrow \infty} |A_k| = \infty$, then $\llbracket (\exists \bar{x}) \varphi_0(\bar{x}) \rrbracket = \mathbf{1}$.

Proof. Let $\bar{x} = (x_0, \dots, x_{l-1})$. Let $T_{\bar{a}}$ be a tree of some standard depth d , that tests whether $G \models \varphi_0(\bar{a})$.

From \aleph_1 -saturation of \mathcal{M} we have $n' > \mathbb{N}$ many tuples $\bar{a}_0, \dots, \bar{a}_{n'-1} \in A_n$, such that $\Pr_{G \in \mathcal{G}_k} (G \models \varphi_0(\bar{a}_i)) \geq p$, we can assume $n' < n^{1/t_0}$ for some $t_0 > \mathbb{N}$.

For $j \in [l]$ construct a tree T_j inductively as follows: Start with $T_{\bar{a}_0}$. Replace the label of every accepting leaf by $(\bar{a}_0)_j$ and remove the label of every rejecting leaf. Call this tree T_j^0 . Assume we have already constructed T_j^m . Construct T_j^{i+1} by appending $T_{\bar{a}_{m+1}}$ to every undefined leaf, relabeling every satisfied leaf to $(\bar{a}_{i+1})_j$ and removing labels from every rejecting leaf. We will define T_j as $T_j^{n'}$ with undefined leaves labeled by 0. (This can be done, because all instances of induction are in $\text{Th}(\mathbb{N})$.) Note that $\text{dp}(T_j) = d \cdot n' < n^{1/t}$ for some $t > \mathbb{N}$.

Call $\bar{\alpha}$ the tuple computed by T_0, \dots, T_{l-1} . We will prove that probability of $\bar{\alpha}$ being a witness to $\varphi_0(\bar{x})$ is 1. For each \bar{a}_i we have, that the probability of $G \models \varphi_0(\bar{a}_i)$ is at least p . The mutual independence of $\{G \models \varphi_0(\bar{a}_i); i \in [n']\}$ and the construction of T_j implies that T_j has a probability of $(1-p)^{n'}$ of failing, which is obviously almost 0. \square

Example 3.1.5. Let

$$\mathcal{G}_k = \{([k], E); E \text{ at least one edge, and may have exactly } k/2 \text{ more from start}\}$$

and let $F = F_{tree}$. Then $\mu(\llbracket (\exists x)(\exists y)\Gamma(x, y) \rrbracket) = \frac{1}{2}$.

3.2 $\mathcal{G}_k = \text{ALL}_k$

Theorem 3.2.1 (Everything exists). Let $\varphi(\bar{x}, \bar{y}) = \bigwedge_{i=0}^{m-1} \psi_i(\bar{x}, \bar{y}) \wedge \bigwedge_{i=0}^{l-1} \vartheta_i(\bar{y})$, where ψ_i, ϑ_i are literals and ψ_i are not of the form $(y_i = y_j)^b$, $E(y_i, y_j)^b$, $x_i \neq x_i$, $E(x_i, x_i)$, $b \in \{0, 1\}$.

Let $\bar{\beta}$ be a tuple of vertices computed by F_{tree} of the same length as \bar{y} . Then $\llbracket (\exists \bar{x}) \varphi^\Gamma(\bar{x}, \bar{\beta}) \rrbracket = \llbracket \bigwedge_{i=0}^{l-1} \vartheta_i(\bar{\beta}) \rrbracket$, specifically if $l = 0$ then $(\exists \bar{x}) \phi_0(\bar{x}, \bar{\beta})$ is valid in the b.v. graph.

Proof. We will construct one tree T computing the whole tuple of witnesses $\bar{\alpha}$, such a construction can be straightforwardly split into a tuple of tree each computing the specific element.

First we concatenate all the trees used to compute $\bar{\beta}$. At each leave we can now proceed knowing the value of $\bar{\beta}$ at the specific $\omega \in \Omega$. Now we just construct a tree as in Theorem 3.1.4 but searching only over edges not checked previously

and only to fulfill each ψ_i . Luckily we have so far searched only an infinitesimal part of the edges and since we assume $\mathcal{G}_k = \text{ALL}_k$ both of the conditions of the theorem are satisfied, so by analogous argument, we have a tree that finds a witness all of the $\psi_i(\bar{x}, \bar{\beta})$ with probability infinitesimally close to 1.

Therefore

$$\llbracket (\exists \bar{x}) \varphi^\Gamma(\bar{x}, \bar{\beta}) \rrbracket = \llbracket (\exists \bar{x}) \bigwedge_{i=0}^{m-1} \psi_i(\bar{x}, \bar{\beta}) \rrbracket \wedge \llbracket \bigwedge_{i=0}^{l-1} \vartheta_i(\bar{\beta}) \rrbracket \quad (3.8)$$

$$= \llbracket \bigwedge_{i=0}^{l-1} \vartheta_i(\bar{\beta}) \rrbracket. \quad (3.9)$$

□

Corollary 3.2.2. For each $\varphi(\bar{x})$ that is not a tautology in the theory of graphs we have that $\llbracket (\forall \bar{x}) \varphi^\Gamma(\bar{x}) \rrbracket = \mathbf{0}$.

Corollary 3.2.3. For each $\varphi(\bar{x}, \bar{y})$ that is not falsifiable by \bar{y} in the theory of graphs we have that $\llbracket (\forall \bar{y}) (\exists \bar{x}) \varphi^\Gamma(\bar{x}) \rrbracket = \mathbf{1}$.

4. $F = F_{nbtree}$

4.1 $\mathcal{G}_k = *PATH_k$

Definition 4.1.1. We define $*PATH_k$ (the pointed paths on k vertices) as the set of all (undirected) graphs G on the vertex set $[k]$, where G is isomorphic to the path on n vertices and $\deg_G(0) = 1$.

Definition 4.1.2. After we fix n , we define F_{nbtree} as the set of all functions computed by some some labeled tree with the following shape:

- Each non-leaf node is labeled by some $v \in [n]$.
- For each $\{u, v\} \subseteq [n]$ and a node N there is an outgoing edge from N labeled A .
- Each leaf is labeled by some $m \in \mathcal{M}_n$.
- The depth of the tree is at most $n^{1/t}$ for some $t > \mathbb{N}$.

Computation of such a tree on a undirected graph G goes as follows. We interpret the non-leaf nodes as questions "what is the neighbour set of v ?" and the edges as answers from our graph G , and thus we follow a path determined by G until we find a vertex for which the answer is not an edge (in which case the computation returns 0) or until we find a leaf, in which case the computation returns the label of the leaf.

We now shift out focus to analysing the ability of trees from F_{nbtree} to find the non-zero degree 1 vertex in $G \in *PATH_n$. We say a tree $T \in F_{nbtree}$ fails at a graph G if $T(G)$ is not a non-zero vertex of degree one in G .

Definition 4.1.3. Let $m \leq n$ and $v \in [w]$ and $U \subseteq [w]$ with $|U| \leq 2$, then we define

$$\mathcal{G}_m^{v?=U} := \{G \in \mathcal{G}_m; N_G(v) = U\},$$

where N_G is the neighbour-set function of G .

Lemma 4.1.4. There are bijections for all nonstandard $m \leq n$ and distinct $u, v, w \in [m] \setminus \{0\}$:

$$\mathcal{G}_m^{v?=\{u,w\}} \cong \mathcal{G}_{m-2} \times [2] \tag{4.1}$$

$$\mathcal{G}_m^{v?=\{u,0\}} \cong \mathcal{G}_{m-2} \tag{4.2}$$

$$\mathcal{G}_m^{0?=\{u\}} \cong \mathcal{G}_{m-1}. \tag{4.3}$$

Proof. (sketch) For (4.1) we can just contract all of u, v, w into one vertex and relabel the rest of the graph, leaving the orientation as a one remaining bit of information. This is obviously reversible and a bijection.

For (4.2) we can do the same, but the orientation is given by 0. \square

Lemma 4.1.5. Let $T \in F_{nbtree}$, with root labeled $v \in [m] \setminus 0$, we have for each $T_{v?=\{u,w\}}$ a tree $\tilde{T}_{v?=\{u,w\}}$ of the same depth, such that

$$P_m(T_{v?=\{u,w\}} \text{ fails} | v? = \{u, w\}) = P_{m-2}(\tilde{T}_{v?=\{u,w\}}). \quad (4.4)$$

For a tree T with the root labeled 0, we have a tree $\tilde{T}_{v?=\{u,w\}}$ of the same depth, such that

$$P_m(T_{v?=\{u\}} \text{ fails} | v? = \{u\}) = P_{m-1}(\tilde{T}_{v?=\{u\}}). \quad (4.5)$$

Proof. (sketch) To construct the tree, we just replace all vertices in labels of $T_{v?=\{u,w\}}$ by there renumbering from the bijection in (4.1).

(TODO: Elaborate) One can then check that the trees $T_{v?=\{u,w\}}$ and $\tilde{T}_{v?=\{u,w\}}$ are isomorphic in a sense that their computation of a graph G and \tilde{G} respectively, \tilde{G} being the corresponding $(m-2)$ -vertex graph, agree with the structure of the path and that correctness of leaves is preserved under the renumbering. Essentially they emulate the same computation but on a smaller graph. \square

Lemma 4.1.6. For all nonstandard $t > \mathbb{N}$, $m \geq n - 2n^{1/t}$ and $k \in [n^{1/t} + 1]$ for all trees $T \in F_{nbtree}$ of depth k we have

$$P_m(T \text{ fails}) \geq \prod_{i=0}^k \left(1 - \frac{2}{m - 2i - 2}\right).$$

Proof. We proceed by induction on k .

$k = 0$: We have that the probability of success of a straight guess is at most $\frac{1}{m-1}$. Therefore

$$P(T \text{ fails}) \geq \left(1 - \frac{1}{m-1}\right) \geq \left(1 - \frac{2}{m-2}\right). \quad (4.6)$$

$(k-1) \Rightarrow k$: First we assume that the root is labeled 0. Then we have

$$P(T \text{ fails}) = \sum_{u \in V \setminus \{0\}} P_{m-1}(0Eu) P_{m-1}(T_{0?=\{u\}} \text{ fails} | 0Eu) \quad (4.7)$$

$$\geq P_{m-1}(T_{0?=\{u\}} \text{ fails} | 0Eu) \quad (4.8)$$

$$= P_{m-1}(\tilde{T}_{0?=\{u\}} \text{ fails}) \quad (4.9)$$

$$\geq \prod_{i=0}^{k-1} \left(1 - \frac{2}{m - 2i - 2}\right) \quad (4.10)$$

$$\geq \prod_{i=0}^k \left(1 - \frac{2}{m - 2i - 2}\right). \quad (4.11)$$

Now we assume that the root is labeled $v \neq 0$. First we notice that

$$P_m(vE0) = \frac{1}{m-1} \quad (4.12)$$

$$P_m(N(V) = 1) = \frac{1}{m-1} \quad (4.13)$$

$$P_m(|N(V) \setminus \{0\}| = 2) = 1 - \frac{2}{m-1}, \quad (4.14)$$

the first two probabilities are obviously $\frac{1}{m-1}$ because they correspond to v being positioned on one of the ends of the non-zero segment which has length $m-1$. The event in (4.14) is the complement of the union of the first two events, which have empty intersestion, giving us that stated probability.

Then we have

$$P_m(T \text{ fails}) = P_m(vE0)P_m(T \text{ fails}|vE0) \quad (4.15)$$

$$+ P_m(|N(v) \setminus \{0\}| = 2)P_m(T \text{ fails}| |N(v) \setminus \{0\}| = 2) \quad (4.16)$$

$$+ P_m(|N(v)| = 1)P_m(vE0)P_m(T \text{ fails}| |N(v)| = 1) \quad (4.17)$$

$$\geq P_m(|N(v) \setminus \{0\}| = 2)P_m(T \text{ fails}| |N(v) \setminus \{0\}| = 2) \quad (4.18)$$

$$\geq (1 - \frac{2}{m-1}) \quad (4.19)$$

$$\cdot \sum_{\substack{u, w \in V \setminus \{0\} \\ u \neq w}} P_m(v? = \{u, w\})P_m(T_{v?=\{u, w\}} \text{ fails}|v? = \{u, w\}) \quad (4.20)$$

$$\geq (1 - \frac{2}{m-1})P_m(T_{v?=\{u_0, w_0\}} \text{ fails}|v? = \{u_0, w_0\}) \quad (4.21)$$

$$\geq (1 - \frac{2}{m-1})P_{m-2}(\tilde{T}_{v?=\{u_0, w_0\}} \text{ fails}) \quad (4.22)$$

$$\geq (1 - \frac{2}{m-1}) \prod_{i=0}^{k-1} (1 - \frac{2}{m-2i-4}) \quad (4.23)$$

$$\geq (1 - \frac{2}{m-2}) \prod_{i=1}^k (1 - \frac{2}{m-2i-2}) \quad (4.24)$$

$$\geq \prod_{i=0}^k (1 - \frac{2}{m-2i-2}). \quad (4.25)$$

where in (4.21) we choose u_0, w_0 with the lowest value of

$$P_m(T_{v?=\{u_0, w_0\}}|v? = \{u_0, w_0\}),$$

the bound follows the fact that all $P_m(v? = \{u, w\})$ are the same for distinct non-zero u, w . In (4.22) we use the lemma 4.1.5 and in (4.23) we use the induction hypothesis. \square

Corollary 4.1.7. For a tree $T \in F_{nbtree}$ we have that

$$P_n(T \text{ fails}) \approx 1.$$

Proof. Since T has depth at most $n^{1/t}$ for some $t > \mathbb{N}$ we by the previous lemma that

$$P_n(T \text{ fails}) \geq \prod_{i=0}^{n^{1/t}} \left(1 - \frac{2}{n-2i-2}\right) \quad (4.26)$$

$$\geq \left(1 - \frac{2n^{1/t}}{n-2n^{1/t}-2}\right) \quad (4.27)$$

$$\approx 1. \quad (4.28)$$

\square

Finally we can prove the following theorem.

Theorem 4.1.8.

$$\llbracket (\exists v)(\exists u)(\forall w)((v \neq 0) \wedge (\Gamma(v, u)) \wedge (\Gamma(v, w) \rightarrow u = w)) \rrbracket = \mathbf{0}$$

Proof. Expanding the value of the formula in the statement we get

$$\bigvee_{\alpha} \bigvee_{\beta} \bigwedge_{\gamma} \llbracket (\alpha \neq 0) \wedge (\Gamma(\alpha, \beta)) \wedge (\Gamma(\alpha, \gamma) \rightarrow \beta = \gamma) \rrbracket,$$

to prove it evaluates to $\mathbf{0}$ we need to find for every α, β some γ such that

$$\llbracket (\alpha \neq 0) \wedge (\Gamma(\alpha, \beta)) \wedge (\Gamma(\alpha, \gamma) \rightarrow \beta = \gamma) \rrbracket = \mathbf{0}.$$

For any α, β we define

$$\gamma(\omega) := \begin{cases} v & N(\alpha(\omega)) = \{v\} \\ w & w \in N(\alpha(\omega)) \setminus \{\beta(\omega)\}, \end{cases}$$

such a function can be computed by a tree in F_{nbtree} which we can construct by concatenation of trees computing α and β .

Let T be the tree computing α . Now we proceed by contradiction, let

$$\epsilon := \mu(\llbracket (\alpha \neq 0) \wedge (\Gamma(\alpha, \beta)) \wedge (\Gamma(\alpha, \gamma) \rightarrow \beta = \gamma) \rrbracket) > 0,$$

by definition this means that

$$\epsilon = \text{st}(P_n[(\alpha \neq 0) \wedge (\Gamma(\alpha, \beta)) \wedge (\Gamma(\alpha, \gamma) \rightarrow \beta = \gamma)]) > 0.$$

Expanding the value of the formula in the statement we get

$$\bigvee_{\alpha} \bigvee_{\beta} \bigwedge_{\gamma} \llbracket (\alpha \neq 0) \wedge (\Gamma(\alpha, \beta)) \wedge (\Gamma(\alpha, \gamma) \rightarrow \beta = \gamma) \rrbracket,$$

to prove it evaluates to $\mathbf{0}$ we need to find for every α, β some γ such that

$$\llbracket (\alpha \neq 0) \wedge (\Gamma(\alpha, \beta)) \wedge (\Gamma(\alpha, \gamma) \rightarrow \beta = \gamma) \rrbracket = \mathbf{0}.$$

For any α, β we define

$$\gamma(\omega) := \begin{cases} v & N(\alpha(\omega)) = \{v\} \\ w & w \in N(\alpha(\omega)) \setminus \{\beta(\omega)\}, \end{cases}$$

such a function can be computed by a tree in F_{nbtree} which we can construct by concatenation of trees computing α and β .

Let T be the tree computing α . Now we proceed by contradiction, let

$$\epsilon := \mu(\llbracket (\alpha \neq 0) \wedge (\Gamma(\alpha, \beta)) \wedge (\Gamma(\alpha, \gamma) \rightarrow \beta = \gamma) \rrbracket) > 0,$$

by definition this means that

$$\epsilon = \text{st}(P_n[(\alpha \neq 0) \wedge (\Gamma(\alpha, \beta)) \wedge (\Gamma(\alpha, \gamma) \rightarrow \beta = \gamma)]) > 0.$$

But by definition of γ and Corollary 4.1.7 we have

$$\begin{aligned}
0 &< \epsilon \\
&= \text{st}(P_n[(\alpha \neq 0) \wedge (\Gamma(\alpha, \beta)) \wedge (\Gamma(\alpha, \gamma) \rightarrow \beta = \gamma)]) \\
&\leq \text{st}(P_n[(\alpha \neq 0) \wedge (\Gamma(\alpha, \beta)) \wedge |N(\alpha)| = 1]) \\
&\leq \text{st}(P_n[(\alpha \neq 0) \wedge |N(\alpha)| = 1]) \\
&= \text{st}(P_n[T \text{ does not fail}]) \\
&= 0.
\end{aligned}$$

A contradiction. □

Conclusion

Bibliography

- [Fag76] Ronald Fagin. Probabilities on finite models. *The Journal of Symbolic Logic*, 41(1):50–58, 1976.
- [Gol14] I Goldbring. Lecture notes on nonstandard analysis. *Ucla summer school in logic*, 2014.
- [Kay91] Richard Kaye. *Models of Peano arithmetic*, volume 15 of *Oxford Logic Guides*. The Clarendon Press, Oxford University Press, New York, 1991. Oxford Science Publications.
- [Kra10] Jan Krajíček. *Forcing with random variables and proof complexity*, volume 382. Cambridge University Press, 2010.
- [LS06] László Lovász and Balázs Szegedy. Limits of dense graph sequences. *Journal of Combinatorial Theory, Series B*, 96(6):933–957, 2006.
- [NDM13] Jaroslav Nešetřil and Patrice Ossona De Mendez. A model theory approach to structural limits. *arXiv preprint arXiv:1303.2865*, 2013.
- [Raz07] Alexander A Razborov. Flag algebras. *The Journal of Symbolic Logic*, 72(4):1239–1282, 2007.