Model Theory and Extremal Combinatorics: Structure, Enumeration, and 0-1 Laws

Caroline Terry

UIC

Thesis Defense

What is in my thesis?

- Chapters 1,2,3: Notation and Introduction.
- Chapter 4: Metric Spaces (Joint with D. Mubayi).
- Chapter 5: Multigraphs (Joint with D. Mubayi).
- Chapter 6: Hereditary $\mathcal{L}$-properties.
- Chapter 7: Examples.
- Chapter 8: An Application of Model Theoretic Ramsey Theory (Joint with M. Malliaris).

Logical 0-1 laws

Fix a finite first-order language $\mathcal{L}$. For each n suppose $K(n)$ is a set of $\mathcal{L}$-structures with universe $[n] = \{1, \ldots, n\}$. Set $K = \bigcup_{n \in \mathbb{N}} K(n)$.

Definition

We say K has a 0-1 *law* if for every $\mathcal{L}$-sentence ϕ,

$$\mu(\phi) = \lim_{n \to \infty} \frac{|\{G \in K(n) : G \models \phi\}|}{|K(n)|}$$

is either 0 or 1.

The *almost sure theory* of K is $T_{as}(K) = \{\phi \in \mathcal{L} : \mu(\phi) = 1\}$.

If K has a 0-1 law, then $T_{as}(K)$ is complete.

Chapter 4: Discrete Metric Spaces

This is joint work with D. Mubayi.

Definition

Let $r \geq 2$ be an integer.

1. $M_r(n)$ is the set of metric spaces with underlying set $[n]$ and distances all in $[r]$.
2. $\mathcal{L}_r = \{R_1, \ldots, R_r\}$ where each R_i is a binary relation symbol.

Every $G \in M_r(n)$ is naturally an $\mathcal{L}_r$-structure:
for all $a, b \in G$, interpret

$$G \models R_i(a,b) \Leftrightarrow d(a,b) = i \text{ in } G.$$

Question

Does $M_r := \bigcup_{n \in \mathbb{N}} M_r(n)$ have a 0-1 law?

Answer for r even

Assume $r \geq 2$ is even. We now define a special subfamily of $M_r(n)$.

Definition

Let $C_r(n) = \{G \in M_r(n) : \text{for all } a \neq b \in [n], d(a,b) \in \{\frac{r}{2}, \ldots r\}\}$.

Theorem (Mubayi, T.)

$$\lim_{n \to \infty} \frac{|C_r(n)|}{|M_r(n)|} = 1.$$

Let $C_r = \bigcup_{n \in \mathbb{N}} C_r(n)$. A standard argument shows C_r has a 0-1 law.

Corollary (Mubayi, T.)

M_r has a 0-1 law and $T_{as}(M_r) = T_{as}(C_r)$.

Idea: Precise structure theorem + 0-1 law for $C_r \Rightarrow$ new 0-1 law for M_r.

How do we prove the precise structure theorem?

Key tool: approximate structure and enumeration.

Definition

Given $\delta > 0$ and two elements $G, G' \in M_r(n)$, we say G and G' are *δ-close* if $\left|\left\{(a,b) \in [n]^2 : d^G(a,b) \neq d^{G'}(a,b)\right\}\right| \leq \delta n^2$.

Let $C_r^\delta(n) = \{G \in M_r(n) : G \text{ is } \delta\text{-close to an element of } C_r(n)\}$.

Theorem (Mubayi, T.)

Structure: for all $\delta > 0$, there is $\beta > 0$ such that for large n,

$$\frac{|M_r(n) \setminus C_r^\delta(n)|}{|M_r(n)|} \leq 2^{-\beta\binom{n}{2}}.$$

Enumeration: $|M_r(n)| = |C_r(n)|(1 + 2^{o(n^2)}) = (\frac{r}{2} + 1)^{\binom{n}{2}+o(n^2)}$.

Outline of 0-1 law in Chapter 4

Approximate Structure and Enumeration $+$ Ad-hoc arguments

$\Downarrow$

Exact Structure and Enumeration ($C_r(n)$ takes over)

$+$

0-1 law for less complicated family ($C_r(n)$)

$\Downarrow$

0-1 law for complicated family ($M_r(n)$)

Examples where one can apply this strategy:

1. K_ℓ-free graphs (Kolaitis-Prömel-Rothschild)
2. T_k-free digraphs (Kühn-Osthus-Townsend-Zhao) + (Koponen)
3. Triangle-free 3-uniform hypergraphs (Balogh-Mubayi) + (Koponen)
4. Discrete metric spaces (Mubayi-T. – chapter 4)

Focus on Approximate Structure and Enumeration

Approximate structure and enumeration results ________.

- are important tools in proofs of certain 0-1 laws.
- are of independent interest in extremal combinatorics.
- have been proven for lots of combinatorial objects (e.g. graphs, digraphs, hypergraphs, colored hypergraphs, metric spaces).
- often have similar proofs using combination of:

1. extremal results
2. supersaturation results
3. stability theorems
4. graph removal lemmas
5. regularity lemmas
6. hypergraph containers theorem (Balogh-Morris-Samotij, Saxton-Thomason).

Question

Question

Is there a way to view these results (and their proofs) as examples of a general theorem (and its proof)?

Chapters 6: yes.

Main Ingredients:

- Hypergraph containers theorem (Balogh-Morris-Samotij, Saxton-Thomason).
- Triangle Removal for $\mathcal{L}$-structures (Aroskar-Cummings).
- Many combinatorics papers which have made the pattern of proof clear. Particularly recent work using the hypergraph containers theorem.

Contents of Chapter 6

Chapter 6 contains versions of the following for $\mathcal{L}$-structures:

1. Definitions of extremal structures and asymptotic density.
2. Approximate enumeration theorem in terms of asymptotic density.
3. Supersaturation.
4. Hypergraph containers theorem.
5. Definition of stability theorem.
6. Formal version of

$$\text{stability} + \text{extremal structure} \Rightarrow \text{approximate structure.}$$

Goal for today: 1, 2, 5, 6.

Chapter 6: Hereditary $\mathcal{L}$-properties

Let $\mathcal{L}$ be a finite relational language and $\mathcal{H}$ a class of finite $\mathcal{L}$-structures. $\mathcal{H}$ has the *hereditary property* if $A \in \mathcal{H}$ and $B \subseteq_{\mathcal{L}} A$ implies $B \in \mathcal{H}$.

Definition

$\mathcal{H}$ is a *hereditary $\mathcal{L}$-property* if it has the hereditary property and is closed under isomorphism.

In the appropriate language, most of the results we want to generalize are for hereditary $\mathcal{L}$-properties.

Enumeration and structure of hereditary properties in the setting of graphs and other combinatorial structures have been studied in combinatorics.

Setup

For the rest of the talk,

- $\mathcal{L}$ is a fixed, finite relational language.
- $r \geq 2$ is the maximum arity of the relation symbols in $\mathcal{L}$.
- Assume $\mathcal{H}$ is a hereditary $\mathcal{L}$-property.
- For each n, $\mathcal{H}_n$ is the set of elements in $\mathcal{H}$ with domain $[n]$.
- For all n, $\mathcal{H}_n \neq \emptyset$.

Questions

1. $|\mathcal{H}_n| = ??$
2. What is the approximate structure of $\bigcup_{n\in\mathbb{N}} \mathcal{H}_n$?

$\mathcal{L}_\mathcal{H}$-structures

Definition

$S_r(\mathcal{H})$ is the set of complete, quantifier-free $\mathcal{L}$-types $p(x_1, \ldots, x_r)$ s.t. for each $i \neq j$, $x_i \neq x_j \in p(\bar{x})$ and $p(\bar{x})$ is realized in some element of $\mathcal{H}$.

Definition

$\mathcal{L}_\mathcal{H} = \{R_p(x_1, \ldots, x_r) : p(x_1, \ldots, x_r) \in S_r(\mathcal{H})\}$.

Notation: $V^{\underline{\ell}} = \{(a_1, \ldots, a_\ell) \in V^\ell : a_i \neq a_j \text{ each } i \neq j\}$.

Definition

Let V be a set. An $\mathcal{L}_\mathcal{H}$-structure M with domain V is an $\mathcal{L}_\mathcal{H}$-*template* if

- For all $\bar{a} \in V^{\underline{r}}$, there an $R_p(\bar{x}) \in \mathcal{L}_\mathcal{H}$ such that $M \models R_p(\bar{a})$.
- For all $p, q \in S_r(\mathcal{H})$, if $p(x_1, \ldots, x_r) = q(x_{\mu(1)}, \ldots, x_{\mu(r)})$ for some permutation μ of $[r]$, then for all $(a_1, \ldots, a_r) \in V^{\underline{r}}$,

$$M \models R_p(a_1, \ldots, a_r) \text{ if and only if } M \models R_q(a_{\mu(1)}, \ldots, a_{\mu(r)}).$$

Example

$\mathcal{L} = \{E(x, y)\}$ and $\mathcal{H}$ is the class of all finite triangle-free graphs.

$p(x, y) =$ the complete q.f. type containing $E(x, y) \wedge E(y, x) \wedge x \neq y$.

$q(x, y) =$ the complete q.f. type containing $\neg E(x, y) \wedge \neg E(y, x) \wedge x \neq y$.

Then $S_r(\mathcal{H}) = \{p(x, y), q(x, y)\}$ and $\mathcal{L}_{\mathcal{H}} = \{R_p(x, y), R_q(x, y)\}$.

Subpatterns

Suppose M is an $\mathcal{L}_{\mathcal{H}}$-template with domain V.

Definition

An $\mathcal{L}$-structure N is a *full subpattern* of M if $dom(N) = V$ and for all $\{a_1, \ldots, a_r\} \in \binom{V}{r}$,

$$\text{if } p(x_1, \ldots, x_r) = qftp^N(a_1, \ldots, a_r) \text{ then } M \models R_p(a_1, \ldots, a_r).$$

In this case, write $N \trianglelefteq_p M$.

Observe any $\mathcal{L}$-structure N with domain V is determined by $qftp^N(a_1, \ldots, a_r)$ for all $\{a_1, \ldots, a_r\} \in \binom{V}{r}$.

Example.

Extremal Structures

Definition

An $\mathcal{L}_{\mathcal{H}}$-template M is called $\mathcal{H}$*-random* if $N \trianglelefteq_p M$ implies $N \in \mathcal{H}$.

- $\mathcal{R}([n], \mathcal{H})$ is the set of $\mathcal{H}$-random $\mathcal{L}_{\mathcal{H}}$-templates with domain $[n]$.
- $sub(M) = |\{N : N \trianglelefteq_p M\}|$.
- $\mathrm{ex}(n, \mathcal{H}) = \max\{sub(M) : M \in \mathcal{R}([n], \mathcal{H})\}$.

Definition

$M \in \mathcal{R}([n], \mathcal{H})$ is *extremal* if $sub(M) = \mathrm{ex}(n, \mathcal{H})$.

- $\mathcal{R}_{\mathrm{ex}}([n], \mathcal{H})$ is the set of extremal M in $\mathcal{R}([n], \mathcal{H})$.

Results: Enumeration

Recall $\mathrm{ex}(n, \mathcal{H}) = \max\{sub(M) : M \in \mathcal{R}([n], \mathcal{H})\}$.

The *asymptotic density* of $\mathcal{H}$ is $\pi(\mathcal{H}) := \lim_{n\to\infty} \mathrm{ex}(n, \mathcal{H})^{1/\binom{n}{r}}$.

Theorem

For all hereditary $\mathcal{L}$-properties $\mathcal{H}$, $\pi(\mathcal{H})$ exists.

Theorem

If $\mathcal{H}$ is a hereditary $\mathcal{L}$-property, then the following hold.

1. *If $\pi(\mathcal{H}) > 1$, then $|\mathcal{H}_n| = \pi(\mathcal{H})^{\binom{n}{r}+o(n^r)}$.*
2. *If $\pi(\mathcal{H}) \leq 1$, then $|\mathcal{H}_n| = 2^{o(n^r)}$.*

Distance Between First-Order Structures

Definition

Suppose $\mathcal{L}_0$ is a finite relational language with maximum arity r. If M and N are finite $\mathcal{L}_0$-structures with domain V.

- For each $1 \leq \ell \leq r$, set

$$\text{dist}^\ell(M, N) = \frac{|\{\bar{a} \in V^{\underline{\ell}} : qftp^M(\bar{a}) \neq qftp^N(\bar{a})\}|}{|V|^\ell}.$$

- Set $\text{dist}(M, N) = \sum_{\ell=1}^{r} \text{dist}^\ell(M, N)$.
- M and N are δ-*close* if $\text{dist}(M, N) \leq \delta$.

This is basically the same as a definition of Aroskar-Cummings.

Stability and Structure

Definition

$\mathcal{H}$ has a *stability theorem* if for all $\delta > 0$ there is $\epsilon > 0$ such that for sufficiently large n the following holds. For all $M \in \mathcal{R}([n], \mathcal{H})$

$$sub(M) \geq \text{ex}(n, \mathcal{H})^{1-\epsilon} \Rightarrow M \text{ is } \delta\text{-close to an element of } \mathcal{R}_{ex}([n], \mathcal{H}).$$

Let $E(n, \mathcal{H}) = \{G \in \mathcal{H}_n : G \trianglelefteq_p M, \text{ some } M \in \mathcal{R}_{\text{ex}}([n], \mathcal{H})\}$.
Let $E^\delta(n, \mathcal{H}) = \{G \in \mathcal{H}_n : G \text{ is } \delta\text{-close to some } G' \in E(n, \mathcal{H})\}$.

Theorem

Suppose $\pi(\mathcal{H}) > 1$ and $\mathcal{H}$ has a stability theorem. Then for all $\delta > 0$, there is a $\beta > 0$ such that for sufficiently large n,

$$\frac{|\mathcal{H}_n \setminus E^\delta(n, \mathcal{H})|}{|\mathcal{H}_n|} \leq 2^{-\beta\binom{n}{r}}.$$

- Ashwini Aroskar and James Cummings, *Limits, regularity and removal for finite structures*, arXiv:1412.808v1 [math.LO], 2014.
- József Balogh, Robert Morris, and Wojciech Samotij, *Independent sets in hypergraphs*, J. Amer. Math. Soc. **28** (2015), no. 3, 669–709. MR 3327533
- József Balogh and Dhruv Mubayi, *Almost all triangle-free triple systems are tripartite*, Combinatorica **32** (2012), no. 2, 143–169. MR 2927636
- Phokion G. Kolaitis, Hans J. Prömel, and Bruce L. Rothschild, *K_{l+1}-free graphs: asymptotic structure and a* 0-1 *law*, Trans. Amer. Math. Soc. **303** (1987), no. 2, 637–671. MR 902790 (88j:05016)
- Vera Koponen, *Asymptotic probabilities of extension properties and random l-colourable structures*, Ann. Pure Appl. Logic **163** (2012), no. 4, 391–438. MR 2876836
- Daniela Kühn, Deryk Osthus, Timothy Townsend, and Yi Zhao, *On the structure of oriented graphs and digraphs with forbidden tournaments or cycles*, arxiv:1404.6178 [math.CO], 2014.
- David Saxton and Andrew Thomason, *Hypergraph containers*, Inventiones mathematicae **201** (2015), no. 3, 925–992 (English).