Intersecting sets in probability spaces and Shelah's classification

Artem Chernikov (joint with Henry Towsner)

UMD / UCLA

14th Panhellenic Logic Symposium

Thessaloniki, Greece

July 1, 2024

- We will discuss some results about intersection patterns for families of events in probability spaces.
- On the face of it, the theorems can be stated with no mention of mathematical logic, but curiously this topic is closely intertwined with the developments in logic both historically and in active current work. We will overview some of the history, and present some new work.

Intersections in a sequence of sets of positive measure

- **Basic fact.** For every real $\varepsilon > 0$ and $n \in \mathbb{N}$ there exists some $N \in \mathbb{N}$ satisfying the following. If $(X, \mathcal{B}, \mu)$ is a probability space and $A_i \in \mathcal{B}, \mu(A_i) \geq \varepsilon$ for $1 \leq i \leq N$, then $\mu\left(\bigcap_{i \in I} A_i\right) > 0$ for some $I \subseteq [N] = \{1, \ldots, N\}$ with $|I| = n$.
- There exist more precise infinitary/density versions (Bergelson's lemma in dynamics).
- We present a slightly more conceptual proof than necessary, as a warm up for what comes later.
- Suppose that we knew that the random variables $\mathbf{1}_{A_i} : X \to \{0, 1\}$ were independent. Then of course $\mu\left(\bigcap_{i \in [n]} A_i\right) = \prod_{i \in [n]} \mu(A_i) \geq \varepsilon^n > 0$.
- We will reduce to this case. Assume that for some fixed $\varepsilon > 0$ and n, no $N \in \mathbb{N}$ satisfies the claim.

Homogenizing the sequence

- Using Ramsey's theorem, we can homogenize our sequence arbitrarily well: e.g., we could assume that for any fixed $\delta > 0$ and k, $\mu(A_{i_1} \cap \ldots \cap A_{i_k}) \approx^{\delta} \mu(A_{j_1} \cap \ldots \cap A_{j_k})$ for any $i_1 < \ldots < i_k$, $j_1 < \ldots < j_k$, and similarly for the measure of arbitrary Boolean combinations of the A_i's.
- Using a compactness argument (e.g. taking Loeb measure on an ultraproduct of counterexamples), we can thus find some large probability space $(X, \mathcal{B}, \mu)$ and sets $A_i \in \mathcal{B}, \mu(A_i) \geq \varepsilon$ for $i \in \mathbb{N}$, still intersection of any n of them has measure 0, so that the sequence of random variables $(\mathbf{1}_{A_i} : i \in \mathbb{N})$ is *spreadable*.

de Finetti's theorem

- **Definition.** A sequence of $[0,1]$-valued random variables $(\xi_i : i \in \mathbb{N})$ is *spreadable* if for every $n \in \mathbb{N}$ and $i_1 < \ldots < i_n, j_1 < \ldots < j_n$ we have $(\xi_{i_1}, \ldots, \xi_{i_n}) =^{\mathrm{dist}} (\xi_{j_1}, \ldots, \xi_{j_n})$.
- For example, every i.i.d. (independent, identically distributed) sequence of random variables is spreadable. The converse holds "up to mixing":
- **de Finetti's theorem.** If an infinite sequence of random variables $(\xi_i : i \in \mathbb{N})$ on $(X, \mathcal{B}, \mu)$ is spreadable then there exists a probability space $(X', \mathcal{B}', \mu')$, a Borel function $f : [0,1]^2 \to [0,1]$ and a collection of $\mathrm{Uniform}[0,1]$ *i.i.d.* random variables $\{\zeta_i : i \in \mathbb{N}\} \cup \{\zeta_\emptyset\}$ on X' so that

$$(\xi_i : i \in \mathbb{N}) =^{\mathrm{dist}} (f(\zeta_i, \zeta_\emptyset) : i \in \mathbb{N}).$$

Exchangeable vs spreadable sequences

- More precisely, de Finetti obtained this conclusion under a stronger assumption that the sequence $(\xi_i : i \in \mathbb{N})$ is *exchangeable*, that is for any $n \in \mathbb{N}$, any permutation $\sigma \in \mathrm{Sym}(n)$ and $i_1 < \ldots < i_n$ we have $(\xi_{i_1}, \ldots, \xi_{i_n}) =^{\mathrm{dist}} (\xi_{i_{\sigma(1)}}, \ldots, \xi_{j_{\sigma(n)}})$.
- And then Ryll-Nardzewski proved that exchangeable is equivalent to spreadable.
- Curiously, Ryll-Nardzewski has a well-known theorem in model theory, but here he worked as a probabilist. It turns out that this result is connected to a central notion in modern model theory!

A few words about model theory

- Morley's theorem: for a countable theory T, if it has only one model of some uncountable cardinality (up to isomorphism), then it has only one model of every uncountable cardinality.
- Morley's conjecture: for a countable theory T, the number of its models of size κ is non-decreasing on uncountable κ.
- In his solution of Morley's conjecture, Shelah isolated the importance of the class of *stable theories* and had developed a lot of machinery to analyze models of stable theories. (Stability was rediscovered many times in various contexts, e.g. by Grothendieck in his work on Banach spaces, also in dynamics as WAP systems, in theoretical machine learning as Littlestone dimension)
- In particular, probability algebras are stable! (In continuous logic.)

Ryll-Nardzewski = model theoretic stability of probability algebras

- Implicit in Ryll-Nardzewski ("every indiscernible sequence is totally indiscernible"), Ben Yaacov, a more general version by Hrushovski (proved using array de Finetti, will discuss in a minute), and Tao gave a short elementary proof:

Fact

For any $0 \leq p < q \leq 1$ there exists N satisfying: if $(X, \mathcal{B}, \mu)$ is a probability space, and $A_1, \ldots, A_n, B_1, \ldots, B_n \in \mathcal{B}$ satisfy $\mu(A_i \cap B_j) \geq q$ and $\mu(A_j \cap B_i) \leq p$ for all $1 \leq i < j \leq n$, then $n \leq N$.

- Many applications: Hrushovski's work on approximate subgroups, Tao's algebraic regularity lemma, work in topological dynamics by Tsankov, Ibarlucia, ...
- Motivated by model theoretic applications, we are interested in *higher arity* generalizations of these results.

Intersections in multi-parametric families of events

Theorem (C., Towsner)

For every finite bipartite graph $H = (V_0, W_0, E_0)$ and $\varepsilon \in (0, 1]$ there exists a finite bipartite graph $G = (V, W, E)$ and $\delta > 0$ (depending only on H and ε) satisfying the following. Assume that $(X, \mathcal{B}, \mu)$ is a probability space, and for every $(v, w) \in V \times W$ a measurable set $A_{v,w} \in \mathcal{B}$ so that: for any $(v, w) \in E, (v', w') \notin E$ we have $\mu(A_{v,w}) - \mu(A_{v',w'}) \geq \varepsilon$. Then there exists an induced subgraph $H' = (V', W', E')$ of G (i.e. $V' \subseteq V, W' \subseteq W$ and $E' = E \cap (V' \times W')$) isomorphic to H so that:

$$\mu\left(\left(\bigcap_{(v,w)\in E'} A_{v,w}\right) \cap \left(\bigcap_{(v,w)\in(V'\times W')\setminus E'} X \setminus A_{v,w}\right)\right) \geq \delta.$$

Intersecting multi-parametric families of events

- More generally, this holds for partite hypergraphs of any arity instead of graphs.
- For us, the question is motivated by *Keisler randomizations* of first-order structures and whether they preserve NIP (Ben Yaacov, related to work of Talagrand on VC dimension for functions) and its higher arity generalization *n-dependence* (where Ben Yaacov's analytic proof for $n = 1$ doesn't seem to generalize).
- An overall strategy is similar to our proof for sequences of events, but each of the steps becomes significantly harder:
  - extract a counterexample "exchangeable" for *ordered* bipartite graphs using structural Ramsey theory,
  - improve to exchangeability *without the ordering* using stability of probability algebras,
  - use a (generalization of) the Aldous-Hoover presentation theorem for exchangeable arrays to pass to i.i.d. random variables.

Structural Ramsey theory

- Let $\mathcal{K}$ be a class of finite $\mathcal{L}_0$-structures, where $\mathcal{L}_0$ is a relational language (for example, finite graphs). For $A, B \in K$, let $\binom{B}{A}$ be the set of all $A' \subseteq B$ s.t. $A' \cong A$.
- $\mathcal{K}$ is *Ramsey* if for any $A, B \in K$ and $k \in \omega$ there is some $C \in K$ s.t. for any coloring $f : \binom{C}{A} \to k$, there is some $B' \in \binom{C}{B}$ s.t. $f \restriction \binom{B'}{A}$ is constant.
- The usual Ramsey theorem means: the class of finite linear orders is Ramsey. The subject of structural Ramsey theory started with:
- [Nesétril, Rödl], [Abramson, Harrington] For any $k \in \mathbb{N}_{\geq 1}$, the class of all finite ordered (partite) k-hypergraphs is Ramsey.

Infinite limits of Ramsey classes

- Given a $\mathcal{K}$ Ramsey class of finite structures, there exists a unique (up to isomorphism) countable structure $\widetilde{K}$ (called the *Fraïssé limit* of $\mathcal{K}$) so that the class of its finite substructures is precisely $\mathcal{K}$ and $\widetilde{K}$ is *homogeneous*, i.e. if K_0 and K_1 are finite substructures of $\widetilde{K}$ and $f : K_0 \to K_1$ is an isomorphism, then f extends to an automorphism of the whole structure $\widetilde{K}$.
- E.g., if $\mathcal{K}$ is the class of all graphs, its limit $\widetilde{K}$ is the countable Rado's random graph; and if $\mathcal{K}$ is the class of finite linear orders, then its limit is $(\mathbb{Q}, <)$.
- Understanding which structures are Ramsey is by now a big and active subject, with connections to model theory, topological dynamics (Ramsey property of $\mathcal{K}$ is equivalent to the extreme amenability of the group $\mathrm{Aut}(\widetilde{K})$ — via the Kechris-Pestov-Todorcevic correspondence), etc.

Finding an "exchangeable" counterexample

- For any $k \in \mathbb{N}_{\geq 1}$, using that the class of all finite ordered (partite) k-hypergraphs is Ramsey, we let $\mathcal{OH}_k$ denote its *Fraïssé limit*.
- Assuming that the conclusion of the theorem fails, by Ramsey and compactness (model theoretic jargon: extracting a generalized indiscernible) we can find some large probability space $(X, \mathcal{B}, \mu)$, $0 < r < s < 1$ and sets $A_{v,w} \in \mathcal{B}$ for all v, w vertices of $\mathcal{OH}_2 = (E; V, W)$ so that:
- $(v, w) \in E \implies \mu(A_{(v,w)}) \geq s$,
- $(v, w) \notin E \implies \mu(A_{(v,w)}) \leq r$,
- for any two isomorphic (as ordered bipartite graphs) substructures H_1, H_2 of $\mathcal{OH}_2$, $(1_{A_{v,w}} : v, w \in H_1) =^{\mathrm{dist}} (1_{A_{v,w}} : v, w \in H_2)$.

Higher exchangeability theory

- This indiscernibility guarantees certain "exchangeability" in the probabilistic sense. Exchangeability theory: exchangeable sequences [de Finetti] and arrays [Aldous-Hoover-Kallenberg] of random variables can be presented "up to mixing" using i.i.d. random variables, and we need a certain generalization to relational structures.
- More recent work on generalizations to exchangeable structures (Ackerman, Freer, Patel, Kruckman, Crane, Towsner, Tsankov):

Exchangeable random structures

- Let $\mathcal{L}' = \{R'_1, \ldots, R'_{k'}\}$, R'_i a relation symbol of arity r'_i. By a *random $\mathcal{L}'$-structure* we mean a (countable) collection of random variables

$$\left(\xi^i_{\bar{n}} : i \in [k'], \bar{n} \in \mathbb{N}^{r'_i}\right)$$

  on some probability space $(\Omega, \mathcal{F}, \mu)$ with $\xi^i_{\bar{n}} : \Omega \to \{0, 1\}$.
- Let now $\mathcal{L} = \{R_1, \ldots, R_k\}$ be another relational language, with R_i a relation symbol of arity r_i, and let $\mathcal{M} = (\mathbb{N}, \ldots)$ be a countable $\mathcal{L}$-structure with domain $\mathbb{N}$. We say that a random $\mathcal{L}'$-structure $\left(\xi^i_{\bar{n}} : i \in [k'], \bar{n} \in \mathbb{N}^{r'_i}\right)$ is *$\mathcal{M}$-exchangeable* if for any two finite subsets $A = \{a_1, \ldots, a_\ell\}, A' = \{a'_1, \ldots, a'_\ell\} \subseteq \mathbb{N}$

$$\mathrm{qftp}_{\mathcal{L}}\left(a_1, \ldots, a_\ell\right) = \mathrm{qftp}_{\mathcal{L}}\left(a'_1, \ldots, a'_\ell\right) \implies$$
$$\left(\xi^i_{\bar{n}} : i \in [k'], \bar{n} \in A^{r'_i}\right) =^{\mathrm{dist}} \left(\xi^i_{\bar{n}} : i \in [k'], \bar{n} \in (A')^{r'_i}\right).$$

A higher amalgamation condition on the indexing structure

- Let $\mathcal{K}$ be a collection of finite structures in a relational language $\mathcal{L}$.
- For $n \in \mathbb{N}_{\geq 1}$, we say that $\mathcal{K}$ satisfies the *n-disjoint amalgamation property* (*n-DAP*) if for every collection of $\mathcal{L}$-structures $(\mathcal{M}_i = (M_i, \ldots) : i \in [n])$ so that
  - each $\mathcal{M}_i$ is isomorphic to some structure in $\mathcal{K}$,
  - $M_i = [n] \setminus \{i\}$, and
  - $\mathcal{M}_i|_{[n]\setminus\{i,j\}} = \mathcal{M}_j|_{[n]\setminus\{i,j\}}$ for all $i \neq j \in [n]$,

  there exists an $\mathcal{L}$-structure $\mathcal{M} = (M, \ldots)$ isomorphic to some structure in $\mathcal{K}$ such that $M = [n]$ and $\mathcal{M}|_{[n]\setminus\{i\}} = \mathcal{M}_i$ for every $1 \leq i \leq n$.
- We say that an $\mathcal{L}$-structure $\mathcal{M}$ satisfies n-DAP if the collection of its finite induced substructures does.
- Ex.: the generic k-hypergraph $\mathcal{H}_k$ satisfies n-DAP for all n, but $(\mathbb{Q}, <)$ fails 3-DAP.

Presentation for random relational structures

Fact (Crane, Towsner; generalizing Aldous-Hoover-Kallenberg)

Let $\mathcal{L}' = \{R'_i : i \in [k']\}, \mathcal{L} = \{R_i : i \in [k]\}$ *be finite relational languages with all* R'_i *of arity at most* r'*, and* $\mathcal{M} = (\mathbb{N}, \ldots)$ *a countable homogeneous* $\mathcal{L}$*-structure that has n-DAP for all* $n \geq 1$*.*
Suppose that $\left(\xi^i_{\bar{n}} : i \in [k'], \bar{n} \in \mathbb{N}^{r'_i}\right)$ *is a random* $\mathcal{L}'$*-structure that is* $\mathcal{M}$*-exchangeable, such that the relations* R'_i *are symmetric with probability* 1*.*
Then there exists a probability space $(\Omega', \mathcal{F}', \mu')$*,* $\{0,1\}$*-valued Borel functions* $f_1, \ldots, f_{r'}$ *and a collection of* $Uniform[0,1]$ *i.i.d. random variables* $(\zeta_s : s \subseteq \mathbb{N}, |s| \leq r')$ *on* Ω' *so that*

$$\left(\xi^i_{\bar{n}} : i \in [k'], \bar{n} \in \mathbb{N}^{r'_i}\right) =^{dist}$$
$$\left(f_i\left(\mathcal{M}|_{\operatorname{rng}\bar{n}}, (\zeta_s)_{s \subseteq \operatorname{rng}\bar{n}}\right) : i \in [k'], \bar{n} \in \mathbb{N}^{r'_i}\right),$$

where $\operatorname{rng}\bar{n}$ *is the set of its distinct elements, and* $\subseteq$ *denotes "subsequence".*

Step 2: getting rid of the ordering

- Our counterexample is only guaranteed to be $\mathcal{OH}_n$-exchangeable (and the ordering is unavoidable in the Ramsey theorem for hypergraphs) — but the presentation theorem requires n-DAP.
- Using Ryll-Nardzewski inductively, we can show $\mathcal{OH}_n$-exchangeability implies $\mathcal{H}_n$-exchangeability, using that the theory of probability algebras is *stable*!
- Applying the exchangeable presentation to the counterexample, we reduce (modulo some mixing, not hard to take care of) to working with *independent* random variables.

Some questions

- Do there exist infinitary/density versions of these results?
- Apart from *k-partite* k-hypergraphs, which other structures satisfy analogous theorems?
- [Tim Austin] Doesn't work for graphs:
  Let H be the triangle K_3 and let $\epsilon = 1/2$. Consider any $G = (V, E)$. On some probability space (Ω, Σ, μ), let $(\pi_v : v \in V)$ be a process of independent uniform $0, 1$-valued RVs, and consider the events A_{vw} defined by $A_{vw} = \pi_v \neq \pi_w$ if $vw \in E$, and $A_{vw} = \emptyset$ if $vw \notin E$. Then $\mu(A_{vw})$ is equal to $1/2$ if $vw \in E$, but equal to 0 if $vw \notin E$. However, for any induced triangle in G, say with vertices u, v, w, we have $\mu(A_{uv} \cap A_{vw} \cap A_{wu}) = \mu(\pi_u \neq \pi_v \neq \pi_w \neq \pi_u) = \mu(\emptyset) = 0$.
- Apart from n-dependence, what other higher arity tameness notions are preserved under Keisler randomization?

Thank you!

- *Randomizing a model*, H. Jerome Keisler, Advances in Mathematics 143 (1999), no. 1, 124–158.
- *On theories of random variables*, Itaï Ben Yaacov, Israel Journal of Mathematics 194.2 (2013): 957–1012.
- *Hypergraph regularity and higher arity VC-dimension*, Artem Chernikov, Henry Towsner (arXiv:2010.00726)