Keisler randomization and n-dependent theories

Artem Chernikov

UCLA

"Model Theory and Philosophy of Mathematics", Fudan University, Shanghai, China (via Zoom)

Aug 22, 2021

Joint work with Henry Towsner (University of Pennsylvania).

Continuous logic

- Ben Yaacov, Berenstein, Henson, Usvyatsov "Model theory for metric structures" (earlier variants by Chang-Keisler, Henson, ...).
- Every structure $\mathcal{M} = (M, \ldots)$ is a complete metric space of bounded diameter, with a metric d.
- Signature:
  - function symbols with given moduli of uniform continuity (interpreted as uniformly continuous functions from M^n to M),
  - predicate symbols with given moduli of uniform continuity (interpreted as uniformly continuous functions from M to $[0, 1]$).
- Logical connectives: the set of all continuous functions $[0, 1] \to [0, 1]$, or any subfamily which generates a dense subset (e.g. $\left\{\neg, \frac{x}{2}, \dot{-}\right\}$).
- Quantifiers: "sup" for "$\forall$", "inf" for "$\exists$".
- 0 is "True", 1 is "False".

Keisler randomization, 1

- Assume $\mathcal{M}$ is a first order structure in a language $\mathcal{L}$.
- Given a first-order formula $\varphi(x) \in \mathcal{L}$, what is the probability that a random element from $\mathcal{M}$ satisfies this formula?
- Originally formalized by Keisler in classical logic, later by Ben Yaacov and Keisler in continuous logic.
- Can be thought of as the structure consisting of the random variables on some probability space taking valuee in M; as well as a generalization of the ultraproduct construction, with an ultrafilter replaced by an arbitrary measure.

Keisler randomization, 2

- Let Ω be a set and $(\mathcal{M}_\omega)_{\omega\in\Omega}$ a family of $\mathcal{L}$-structures.
- The product $\prod_{\omega\in\Omega} M_\omega$ consists of all functions $a : \Omega \to \bigcup M_\omega$ with $a(\omega) \in M_\omega$ for all $\omega \in \Omega$. Function symbols and terms of $\mathcal{L}$ are interpreted coordinatewise on $\prod M_\omega$.
- For $\varphi(\bar{x}) \in \mathcal{L}, \bar{a} \in (\prod_\omega M_\omega)^{|\bar{x}|}$ we define a function

$$\langle\varphi(\bar{a})\rangle(\omega) : \omega \in \Omega \mapsto \varphi^{\mathcal{M}_\omega}(\bar{a}(\omega)) \in [0,1].$$

- A *randomization* $\mathcal{M} = \mathcal{M}_{\Omega,\mathcal{F},\mu}$ is a continuous (pre-)structure with two sorts $(\mathrm{M}, \mathcal{A})$ in $\mathcal{L}^R$ s.t.
  - $(\Omega, \mathcal{F}, \mu)$ is a probability algebra and $\mathcal{A} = L_1(\mu) \subseteq [0,1]^\Omega$,
  - $\mathrm{M} \subseteq \prod M_\omega$ is non-empty, closed under function symbols and $\langle P(\bar{a})\rangle \in \mathcal{A}$ for every predicate $P(\bar{x}) \in \mathcal{L}$ and $\bar{a} \in \mathrm{M}^{|\bar{x}|}$.
  - the pseudo-metrics $d(X,Y) = \mathbb{E}(|X-Y|)$ on $\mathcal{A}$ and $d(a,b) = \mathbb{E}\langle d(a,b)\rangle = \int_{\omega\in\Omega} d(a(\omega), b(\omega))d\mu$ on M.
  - $\mathcal{L}^R$ contains the function symbols from $\mathcal{L}$, a function symbol $[\![P(\bar{x})]\!] : \mathrm{M}^{|\bar{x}|} \to \mathcal{A}$ for each predicate $P \in \mathcal{L}$, and the signature $\left\{0, \neg, \frac{x}{2}, \dot{-}\right\}$ on $\mathcal{A}$.

Keisler randomization, 3

- Given a randomization $\mathcal{L}^R$-pre-structure $\mathcal{M} = (\mathrm{M}, \mathcal{A})$, its completion (the metric completion of the quotient by elements at distance 0) is an $\mathcal{L}^R$-structure $\widehat{\mathcal{M}} = (\widehat{\mathrm{M}}, \widehat{\mathcal{A}})$.
- When $\mathrm{M} = \prod M_\omega$, $\mathcal{A} = [0,1]^\Omega$, $\mu := \mathcal{U}$ is an ultrafilter on Ω, then $\widehat{\mathcal{A}} = [0,1]$ and $\widehat{\mathcal{M}}$ is naturally identified with the ultraproduct $\prod \mathcal{M}_\omega / \mathcal{U}$.
- We would like to axiomatize (and find a model companion) for the theory of randomizations.
- A randomization $(\mathrm{M}, \mathcal{A})$ is *full* if $\forall a \neq b \in \mathrm{M}, X \in \mathcal{A} \exists c \in \mathrm{M}$ s.t. $c(\omega) = a(\omega)$ for all $\omega \in \Omega$ with $X(\omega) = 1$, $c(\omega) = b(\omega)$ for all ω with $X(\omega) = 0$, and $c(\omega)$ is arbitrary otherwise.
- $(\mathrm{M}, \mathcal{A})$ is *atomless* if $\mathcal{F}$ is an atomless algebra.
- Ex: let $\mathcal{M}$ be a structure, $(\Omega, \mathcal{F}, \mu)$ an atomless probability space, and $\mathrm{M} \subseteq M^\Omega$ consits of all functions $a : \Omega \to M$ taking at most countably many values in M, each on a measurable set. Then the corresponding $(\mathrm{M}, \mathcal{A})$ is a full atomless randomization.

Keisler randomization, 4

Fact (Ben Yaacov)

1. *For a fixed language* $\mathcal{L}$*, there exists a continuous theory* T_0^R *so that: an* $\mathcal{L}^R$*-structure is a model of* T_0^R *if and only if it is isomorphic to* $(\widehat{\mathrm{M}}, \widehat{\mathcal{A}})$ *for some full atomless randomization* $(\mathrm{M}, \mathcal{A})$*; and for every* $\varphi(\bar{x}) \in \mathcal{L}$ *and* $\bar{a} \in \mathrm{M}^{\bar{x}}$ *we have* $\langle\varphi(\bar{a})\rangle = [\![\varphi(\bar{a})]\!]$*.*
2. *For an* $\mathcal{L}$*-theory* T*, let* $T^R := T_0^R \cup \{[\![\varphi]\!] = 0 : \varphi \in T\}$*. Then* T^R *eliminates quantifiers down to the formulas of the form* $\mathbb{E}[\![\varphi(\bar{x})]\!]$ *with* $\varphi(\bar{x}) \in \mathcal{L}$*.*
3. *The types in* $S_n\left(T^R\right)$ *are in bijection with regular Borel probability measures on the space* $S_n(T)$*. In particular if* T *is complete, then so is* T^R*.*

Shelah's classification

- Classification theory: Shelah's dividing lines express limitations on definable *binary* relations, by forbidding certain finitary combinatorial configurations (stability, NIP, simplicity, see Baldwin's talk).
- Often on the tame case, obtain consequences of the form: types (over infinite sets) in more than one variable are controlled by unary types, up to a "small error" (e.g. stationarity of non-forking in stable theories, up to algebraic closure).
- Emerging "n-classification theory": types in any number of variables are controlled by types in at most n-variables, up to a "small error".
- Here we focus on *n-dependence* introduced by Shelah:

N-dependent theories

- Given an $(n+1)$-ary relation $E \subseteq \prod_{1 \leq i \leq n+1} X_i$ and $d \in \mathbb{N}$, we write $\mathrm{VC}_n(E) \leq d$ if there do not exist sets $A_i \subseteq X_i$ with $|A_i| > d$ for $1 \leq i \leq n$ and $b_S \in X_{n+1}$ for $S \subseteq \prod_{1 \leq i \leq n} A_i$ so that
$$(a_1, \ldots, a_n, b_S) \in E \iff (a_1, \ldots, a_n) \in S$$
for all $(a_1, \ldots, a_n) \in \prod_{1 \leq i \leq n} A_i$.
- Write $\mathrm{VC}_n(E) < \infty$ and say E is *n-dependent* if $\mathrm{VC}_n(E) \leq d$ for some $d \in \mathbb{N}$.
- A theory T is *n-dependent* if every formula $\varphi(x_1, \ldots, x_{n+1})$, with x_i a tuple of variables, defines an n-dependent relation in any model of T.

N-dependent theories: basic facts and examples

- The case $n = 1$ corresponds to NIP.
- The property $\mathrm{VC}_n < \infty$ is preserved under permutations of variables and Boolean combinations, and n-dependence of a theory is witnessed by formulas with all but one variable singletons.
- Examples of n-dependent theories:
  - For $n \geq 2$, the theory of the generic n-hypergraph is strictly n-dependent (i.e. n-dependent, but not $(n-1)$-dependent).
  - [C., Hempel] For each $n \geq 2$, there exist strictly n-dependent pure groups.
  - [Cherlin, Hrushovski] Smoothly approximable structures are 2-dependent.
  - [C., Hempel] For $n \geq 2$, non-degenerate n-linear forms on vector spaces over NIP fields are strictly n-dependent.
  - Conjecturally, there are no strictly n-dependent (pure) fields for $n \geq 2$.

N-dependence in continuous logic

- Stability, NIP, etc. all have natural generalizations in continuous logic.
- Given a function $f : \prod_{1 \leq i \leq n+1} X_i \to [0,1]$ and a countable sequence $\bar{d} = (\bar{d}_{r,s} \in \mathbb{N} : r < s \in \mathbb{Q} \cap [0,1])$, we write $\mathrm{VC}_n(f) \leq \bar{d}$ if for each $r < s \in \mathbb{Q} \cap [0,1]$ there do not exist sets $A_i \subseteq X_i$ with $|A_i| > d_{r,s}$ for $1 \leq i \leq n$ and $b_S \in X_{n+1}$ for $S \subseteq \prod_{1 \leq i \leq n} A_i$ so that

$$(a_1, \ldots, a_n) \in S \implies f(a_1, \ldots, a_n, b_S) \geq s,$$
$$(a_1, \ldots, a_n) \notin S \implies f(a_1, \ldots, a_n, b_S) \leq r.$$

- A function f is *n-dependent*, written $\mathrm{VC}_n(f) < \infty$, if $\mathrm{VC}_n(f) \leq \bar{d}$ for some sequence $\bar{d}$.
- A continuous theory T is *n-dependent* if for every (continuous) formula in $n+1$ tuples of variables, the function from any model of T to $[0,1]$ defined by it is n-dependent.

Randomization and classification

Fact

- *[Ben Yaacov, Keisler] If T is ($\aleph_0$-, super-) stable, then T^R is also ($\aleph_0$-, super-) stable.*
- *[Ben Yaacov] If T is NIP, then T^R is also NIP.*
- *[Ben Yaacov] If T is not NIP, then T^R has TP_2. In particular simplicity is not preserved. But at least:*
- *[Ben Yaacov, C., Ramsey] If T is $NSOP_1$, then T^R is also $NSOP_1$.*

Theorem (C., Towsner)

For every $n \geq 1$, if T is n-dependent, then T^R is also n-dependent.

Preservation of NIP: key point

- Ben Yaacov's proof, using relative quantifier elimination in T^R and that composing NIP functions with continuous functions $[0,1]^k \to [0,1]$ preserves NIP, reduces to showing that $\mathbb{E}[\![\varphi(\bar{x})]\!]$ is NIP assuming φ is NIP, i.e. the average of a "uniformly NIP" family of functions is NIP (the case $n = 1$ of the theorem below).
- Ben Yaacov establishes this by developing elements of the VC-theory for real valued functions (connected to some earlier work of Talagrand and others).

A generalization to n-dependence

Theorem (C., Towsner)

For every $k \in \mathbb{N}_{\geq 1}$ and $\bar{d}$ there exists some $\bar{D}$ satisfying the following.
Assume $f : \prod_{i \in [n+2]} V_i \to [0,1]$ is a function and $(V_{n+2}, \mathcal{F}, \mu)$ a probability space, so that

- *for any fixed $\bar{x} \in \prod_{i \in [n+1]} V_i$, the function $\omega \mapsto f(\bar{x}, \omega)$ is measurable;*
- *for any fixed $\omega \in \Omega$, the function $f_\omega : \bar{x} \mapsto f(\bar{x}, \omega)$ satisfies $\mathrm{VC}_n(f_\omega) \leq \bar{d}$.*

Then the "average" function $f' : \prod_{i \in [n+1]} \to [0,1]$ defined by

$$f'(x_1, \ldots, x_{k+1}) := \int_{\omega \in \Omega} f(x_1, \ldots, x_{k+1}, \omega) d\mu$$

satisfies $\mathrm{VC}_n(f') \leq \bar{D}$.

Generalized indiscernibles, 1

- T is a theory in a language $\mathcal{L}$, $\mathbb{M} \models T$.
- Let I be an $\mathcal{L}'$-structure. Then $\bar{a} = (a_i : i \in I)$, with a_i a tuple in $\mathbb{M}$, is *I-indiscernible* if for all $i_1, \ldots, i_n$ and $j_1, \ldots, j_n$ from I:

$$\mathrm{qftp}_{\mathcal{L}'}(i_1, \ldots, i_n) = \mathrm{qftp}_{\mathcal{L}'}(j_1, \ldots, j_n) \implies$$
$$\mathrm{tp}_{\mathcal{L}}(a_{i_1}, \ldots, a_{i_n}/C) = \mathrm{tp}_{\mathcal{L}}(a_{j_1}, \ldots, a_{j_n}/C).$$

- Say that $(b_j : j \in I)$ is based on $(a_i : i \in I)$ if for any finite set Δ of $\mathcal{L}$-formulas and $(j_0, \ldots, j_n)$ from I there is some $(i_1, \ldots, i_n)$ from I s.t.

$$\mathrm{qftp}_{\mathcal{L}_0}(j_1, \ldots, j_n) = \mathrm{qftp}_{\mathcal{L}_0}(i_1, \ldots, i_n), \text{ and}$$
$$\mathrm{tp}_{\Delta}(b_{j_1}, \ldots, b_{j_n}) = \mathrm{tp}_{\Delta}(a_{i_1}, \ldots, a_{i_n}).$$

- The usual indiscernible sequences correspond to the case when I is a linear order.

Generalized indiscernibles, 2

- Let $\mathcal{K}$ be a class of finite $\mathcal{L}_0$-structures. 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: the class of finite linear orders is Ramsey.
- [Scow] Let $\mathcal{K}$ be a Fraïssé class of finite structures, and let I be its limit. If $\mathcal{K}$ is *Ramsey*, then for any $\bar{a}$ indexed by I there exists (in $\mathbb{M}$) an I-indiscernible based on it.
- [Nesétril, Rödl], [Abramson, Harrington] For any $k \in \mathbb{N}_{\geq 1}$, the class of all finite ordered (partite) k-hypergraphs is Ramsey (let $\mathcal{OH}_k$ denote its Fraïssé limit).

Step 1: a sufficiently indiscernible witness

- Assuming that the theorem fails, using some analytic arguments and extracting an indiscernible, we can thus find some $r < s, q > t \in [0,1]$ and an $\mathcal{OH}_{n+1}$-indiscernible $\bar{a}$ in some expansion of the language making the measure μ definable so that

$$\mathcal{OH}_{n+1} \models R(g_1, \ldots, g_{n+1}) \implies \\ \mu(\{\omega : f(a_{g_1}, \ldots, a_{g_{n+1}}, \omega) < r\}) \geq q \text{ and} \\ \mathcal{OH}_{n+1} \models \neg R(g_1, \ldots, g_{n+1}) \implies \\ \mu(\{\omega : f(a_{g_1}, \ldots, a_{g_{n+1}}, \omega) < s\}) \leq t.$$

- 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 (parallel to the hypergraph regularity lemma), and we need a certain generalization to relational structures.

Exchangeable random structures

- Let $\mathcal{L}' = \{R'_1, \dots, R'_{k'}\}$, 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 $\mathcal{L} = \{R_1, \dots, R_k\}$ be another relational language, with R_i a relation symbol of arity r_i, and let $\mathcal{M} = (\mathbb{N}, \dots)$ 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 $A = \{a_1, \dots, a_\ell\}, A' = \{a'_1, \dots, a'_\ell\} \subseteq \mathbb{N}$

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

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)

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 ultrahomogeneous* $\mathcal{L}$*-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* V' *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 $\mathcal{OH}_{n+1}$-exchangeable (and the ordering is unavoidable in the Ramsey theorem for hypergraphs) — but the presentation theorem requires n-DAP.
- We show that $\mathcal{OH}_n$-exchangeability implies $\mathcal{H}_n$-exchangeability, using that the theory of probability algebras is *stable*!
- Implicit in [Ryll-Nardzewski], explicit in [Ben Yaacov], a more general result by [Hrushovski] (proved using array de Finetti), and [Tao] gives an elementary proof:

Fact

For any $0 \leq p < q \leq 1$ there exists N satisfying: if $(V, \mathcal{F}, \mu)$ is a probability space, and $A_1, \ldots, A_n, B_1, \ldots, B_n \in \mathcal{F}$ 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 $n \leq N$.

Step 3: finding a common point

- Applying the exchangeable presentation to the counterexample and working with *independent* random variables, we show that for any finite set $S \subseteq \mathcal{OH}_{n+1}$, the following set has positive measure:

$$\bigcap_{\bar{g} \in R\restriction_S} \{\omega \in \Omega : f(a_{g_1}, \ldots, a_{g_{n+1}}, \omega) < r\} \cap \bigcap_{\bar{g} \in \neg R\restriction_S} \Omega \setminus \{\omega \in \Omega : f(a_{g_1}, \ldots, a_{g_{n+1}}, \omega) < s\}.$$

- By saturation we then find $\omega \in \Omega$ so that for all $(n+1)$-tuples $\bar{g}$ in $\mathcal{OH}_{n+1}$ we have:
  - $\bar{g} \in R \implies f(a_{g_1}, \ldots, a_{g_{n+1}}, \omega) < r$,
  - $\bar{g} \notin R \implies f(a_{g_1}, \ldots, a_{g_{n+1}}, \omega) \geq s$.
- This contradicts the assumption $\mathrm{VC}_n(f_\omega) < \infty$.

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.
- *Continuous and random Vapnik-Chervonenkis classes*, Itaï Ben Yaacov, Israel Journal of Mathematics 173 (2009): 309–333.
- *Hypergraph regularity and higher arity VC-dimension*, Artem Chernikov, Henry Towsner (arXiv:2010.00726)