On reconstruction of non-$\aleph_0$-categorical continuous theories, groupoids, Skolem functions and the Lelek fan

Conference in honour of Ehud Hrushoski, Fields Institute

Itaï Ben Yaacov

Université Claude Bernard Lyon 1
Institut Camille Jordan

2021-12-13

Continuous first order logic is to Boolean first order logic as $[a, b]$ is to $\{T, F\}$

- Structures are complete bounded metric spaces, $d(x, y)$ replaces $x = y$.
- Formulas are real-valued, uniformly continuous, bounded.
- Connectives are continuous, quantifiers are sup and inf.
- Contains Boolean first order logic: $\{0, 1\}$-valued structures.

Abstract continuous model theory: generalise classical results/tools – and deal with "surprises"

Compactness, Löwenheim-Skolem, stability, $\aleph_0$-stability, Morley, ...	✓
Strongly minimal sets, Baldwin-Lachlan	✗(?)
Omitting types, Ryll-Nardzewski	✓
$\aleph_0$-categorical reconstruction (Coquand, a.k.a. Ahlbrandt-Ziegler)	✓
No 2 models	✗
Skolem functions: "choosing witnesses"	✗
My topic today: non-$\aleph_0$-categorical reconstruction	...

Definition

Let T be an $\aleph_0$-categorical theory (complete in a countable language). Let M be any countable model of T and $G(T) = \mathrm{Aut}(M)$, equipped with the topology of pointwise convergence.

Theorem (Coquand, published by Ahlbrandt & Ziegler ; B. & Kaïchouh)

- *Let T and T' be $\aleph_0$-categorical. Then $G(T) \simeq G(T')$ as topological groups if and only if T and T' are bi-interpretable.*
- *Same, for continuous logic (replace countable with separable).*

Moreover:

- We can characterise groups of the form $G(T)$ (Polish, Roeolcke precompact, ...)
- From $G = G(T)$ we can explicitly reconstruct a theory T' bi-interpretable with T.

Theorem (Coquand, published by Ahlbrandt & Ziegler ; B. & Kaïchouh)

"$\aleph_0$-categorical T (up to bi-interpretation)" equals "Roelcke-precompact Polish group $G = G(T)$".

- Connection between model theory (in T) and dynamics (of $G(T)$) – Ibarlucía, Tsankov, B., and others.
- In one direction: subjects many non-locally-compact Polish groups "of interest" to model-theoretic treatment, e.g., Property (T) for Roelcke-precompact groups (Ibarlucía).
- In the other direction: Ibarlucía's (re-)proof of the preservation of NIP under randomisation: if T is countably/separably categorical and NIP (or stable), then so is T^R.

Theorem (Coquand, published by Ahlbrandt & Ziegler ; B. & Kaïchouh)

"$\aleph_0$-categorical T (up to bi-interpretation)" equals "Roelcke-precompact Polish group $G = G(T)$".

- Connection between model theory (in T) and dynamics (of $G(T)$) – Ibarlucía, Tsankov, B., and others.
- In one direction: subjects many non-locally-compact Polish groups "of interest" to model-theoretic treatment, e.g., Property (T) for Roelcke-precompact groups (Ibarlucía).
- In the other direction: Ibarlucía's (re-)proof of the preservation of NIP under randomisation: if T is countably/separably categorical and NIP (or stable), then so is T^R.

But:

- Specific to $\aleph_0$-categorical theories.

In order to cover non-$\aleph_0$-categorical theories, we need topological groupoids of isomorphisms between countable / separable structures.

Definition

- A groupoid is a set $\mathbb{G}$ equipped with a partial composition law and total inversion map, satisfying appropriate axioms [e.g., all isomorphisms of a small category].
- Its basis is $\mathbb{B}(\mathbb{G}) = \{e \in \mathbb{G} : e^2 = e\}$ [all identity morphisms $\simeq$ all objects].
- The source of $g \in \mathbb{G}$ is $s_g = g^{-1}g \in \mathbb{B}(\mathbb{G})$ [source object].
- The target of $g \in \mathbb{G}$ is $t_g = gg^{-1} \in \mathbb{B}(\mathbb{G})$ [target object].

$\mathbb{B}(\mathbb{G})$ is a singleton if and only if $(\mathbb{G}, \cdot, {}^{-1})$ is a group.

Definition

A topological groupoid is a groupoid equipped with a topology, such that the composition (on its domain) and inversion are continuous.
It is open if in addition, the source map $s\colon \mathbb{G} \to \mathbb{B}(\mathbb{G})$ is open (equivalently, the target map, equivalent, the composition law).

Theorem (B.)

To every theory T (complete, in a countable language $\mathcal{L}$, in Boolean logic) we can associate a topological groupoid $\mathbb{G}(T)$ such that:

- *$\mathbb{G}(T)$ is an open topological groupoid over the Cantor space: $\mathbb{B}(\mathbb{G}) \simeq 2^{\mathbb{N}}$.*
- *It is a complete bi-interpretation invariant for T:*

  T and T' are bi-interpretable if and only if $\mathbb{G}(T) \simeq \mathbb{G}(T')$.

Moreover, from $\mathbb{G} = \mathbb{G}(T)$, given as a topological groupoid, we can explicitly construct a theory T' that is bi-interpretable with T.

$\aleph_0$-categorical case, reformulated

Say $a = (a_i : i \in \mathbb{N})$ enumerates $M \vDash T$. Then a "codes" M, and as a topological group

$$\begin{array}{ccc} G(T) = Aut(M) & \simeq & \{\mathrm{tp}(a,b) : \mathrm{tp}(a) = \mathrm{tp}(b) \text{ and } \mathrm{dcl}(a) = \mathrm{dcl}(b)\} \subseteq S_{2\times\mathbb{N}}(T) \\ g & \mapsto & \mathrm{tp}(ga, a) \end{array}$$

Group law: $\mathrm{tp}(a,b) \cdot \mathrm{tp}(b,c) = \mathrm{tp}(a,c)$.

$\aleph_0$-categorical case, reformulated

Say $a = (a_i : i \in \mathbb{N})$ enumerates $M \vDash T$. Then a "codes" M, and as a topological group

$$\begin{array}{ccc} G(T) = Aut(M) & \simeq & \{\mathrm{tp}(a,b) : \mathrm{tp}(a) = \mathrm{tp}(b) \text{ and } \mathrm{dcl}(a) = \mathrm{dcl}(b)\} \subseteq \mathrm{S}_{2\times\mathbb{N}}(T) \\ g & \mapsto & \mathrm{tp}(ga, a) \end{array}$$

Group law: $\mathrm{tp}(a,b) \cdot \mathrm{tp}(b,c) = \mathrm{tp}(a,c)$.

General approach: a "good" set D of codes for models $\Longrightarrow$ a topological groupoid

$$\mathbb{G}_D(T) = \{\mathrm{tp}(a,b) : a, b \in D \text{ and } \mathrm{dcl}(a) = \mathrm{dcl}(b)\} \subseteq \mathrm{S}(T).$$

law $\quad \mathrm{tp}(a,b) \cdot \mathrm{tp}(b,c) = \mathrm{tp}(a,c)$

basis $\quad \mathbb{B}_D(T) = \{\mathrm{tp}(a,a) : a \in D\} \simeq \{\mathrm{tp}(a) : a \in D\} = \mathrm{S}_D(T)$

All that's left is to find D...

Definition

Fix a sequence of formulas $\Phi = (\varphi_n)$, such that $\forall x_{<n} \exists y \, \varphi_n(x_{<n}, y)$ is valid. Define $D_\Phi \subseteq M^{\mathbb{N}}$:

$$a = (a_n : n \in \mathbb{N}) \in D_\Phi \iff \varphi_n(a_{<n}, a_n) \text{ for all } n.$$

Assuming Φ is rich enough: D_Φ is a good set of codes for models!

- Every $a \in D_\Phi$ enumerates a model.
- Every countable model is enumerated by a member of D_Φ.
- D_Φ is type-definable, in infinitely many variables.

Definition

Fix a sequence of formulas $\Phi = (\varphi_n)$, such that $\forall x_{<n} \exists y\, \varphi_n(x_{<n}, y)$ is valid. Define $D_\Phi \subseteq M^{\mathbb{N}}$:

$$a = (a_n : n \in \mathbb{N}) \in D_\Phi \iff \varphi_n(a_{<n}, a_n) \text{ for all } n.$$

Assuming Φ is rich enough: D_Φ is a good set of codes for models!

- Every $a \in D_\Phi$ enumerates a model.
- Every countable model is enumerated by a member of D_Φ.
- D_Φ is type-definable, in infinitely many variables.
- In fact it is definable in the sense of continuous logic (AKA strictly pro-definable by Hrushovski-Loeser): every projection of D_Φ to finitely many coordinates is definable. (The projection to the first n coordinates is defined $\bigwedge_{k<n} \varphi_k(x_{<k}, x_k)$.)

Definition

Fix a sequence of formulas $\Phi = (\varphi_n)$, such that $\forall x_{<n} \exists y\, \varphi_n(x_{<n}, y)$ is valid. Define $D_\Phi \subseteq M^{\mathbb{N}}$:

$$a = (a_n : n \in \mathbb{N}) \in D_\Phi \iff \varphi_n(a_{<n}, a_n) \text{ for all } n.$$

Assuming Φ is rich enough: D_Φ is a good set of codes for models!

- Every $a \in D_\Phi$ enumerates a model.
- Every countable model is enumerated by a member of D_Φ.
- D_Φ is type-definable, in infinitely many variables.
- In fact it is definable in the sense of continuous logic (AKA strictly pro-definable by Hrushovski-Loeser): every projection of D_Φ to finitely many coordinates is definable. (The projection to the first n coordinates is defined $\bigwedge_{k<n} \varphi_k(x_{<k}, x_k)$.)

Do D_Φ and $\mathbb{G}_{D_\Phi}(T)$ depend on Φ?

Proposition (Uniqueness of D_Φ and $\mathbb{G}_{D_\Phi}(T)$)

If Φ and Ψ are both rich, then there exists a definable bijection $D_\Phi \simeq D_\Psi$. Consequently, $\mathbb{G}_{D_\Phi}(T) \simeq \mathbb{G}_{D_\Psi}(T)$.

$\Longrightarrow$ canonical topological groupoid

$$\mathbb{G}(T) = \mathbb{G}_{D_\Phi}(T) = \{\mathrm{tp}(a,b) : a, b \in D_\Phi \text{ and } \mathrm{dcl}(a) = \mathrm{dcl}(b)\}$$

basis: $$\mathbb{B}(T) = \{\mathrm{tp}(a,a) : a \in D_\Phi\} \simeq \mathrm{S}_{D_\Phi}(T) \simeq \text{Cantor}.$$

It is Polish ($\mathrm{dcl}(a) = \mathrm{dcl}(b)$ is G_δ) and open (since D_Φ is definable).

Theorem (Restated)

The topological groupoid $\mathbb{G}(T)$ is a complete bi-interpretation invariant for T.
Moreover, a theory bi-interpretable with T can be explicitly reconstructed from $\mathbb{G}(T)$.

Now, let's just generalise this to continuous logic.

Recall the construction of D_Φ

Each $\exists y \varphi_n(x_{<n}, y)$ is valid, and

$$a = (a_n : n \in \mathbb{N}) \in D_\Phi \iff \varphi_n(a_{<n}, a_n) \text{ for all } n.$$

This, or something similar, must also work in continuous logic. . . right?

Now, let's just generalise this to continuous logic.

Recall the construction of D_Φ

Each $\exists y \varphi_n(x_{<n}, y)$ is valid, and

$$a = (a_n : n \in \mathbb{N}) \in D_\Phi \iff \varphi_n(a_{<n}, a_n) \text{ for all } n.$$

This, or something similar, must also work in continuous logic. . . right?

WRONG

- Boolean logic: if $\exists y \varphi(x, y)$ is valid, then $\{(a, b) : \varphi(a, b)\}$ is a definable set, that projects onto the first coordinate.
- Continuous logic: if $\inf_y \varphi(x, y) = 0$, then $\{(a, b) : \varphi(a, b) = 0\}$ need not be a definable set, and (in a non-saturated model) the projection need not be onto.

Definition

Let D an interpretable sort in T. It is a universal Skolem sort if it is *"like the set D_Φ"*, i.e., if *"it is easy to construct definable Skolem functions from D"*.

Proposition

If T admits two universal Skolem sorts D and D', then there exists a definable bijection $\sigma\colon D \simeq D'$.

Theorem (B.)

Assume that T admits a universal Skolem sort D. Then it is a set of codes for models, (i.e., $\mathrm{dcl}(a) = \mathrm{dcl}(M)$), and the topological groupoid $\mathbb{G}(T)$ is a complete bi-interpretation invariant for T:

$$\mathbb{G}(T) = \mathbb{G}_D(T) = \{\mathrm{tp}(a, b) : a, b \in D \text{ and } \mathrm{dcl}(a) = \mathrm{dcl}(b)\}$$

basis: $$\mathbb{B}(T) \simeq S_D(T) \simeq \textit{Cantor}.$$

Theorem (B.)

Assume that T admits a universal Skolem sort D. Then $\mathbb{G}(T) = \mathbb{G}_D(T)$ is a complete bi-interpretation invariant for T.

- If T is Boolean, then D_Φ is universal Skolem (case already covered).
- If T is $\aleph_0$-categorical, and $\mathrm{dcl}(a) = \mathrm{dcl}(M)$, then $D_0 = \{b : \mathrm{tp}(a) = \mathrm{tp}(b)\}$ is definable, and $D_0 \times 2^{\mathbb{N}}$ is universal Skolem. Consequently,

$$\mathbb{G}(T) = 2^{\mathbb{N}} \times G(T) \times 2^{\mathbb{N}}.$$

- [J. Muñoz] If T admits a universal Skolem sort, then so does its Keisler randomisation T^R.

But:

- There exist theories which do not admit one (e.g., the theory of $[0, 1]$ equipped with the unary identity predicate and the 0/1 distance).

The problem(s)

In continuous logic, if $\inf_y \varphi(x, y) = 0$, then

- A the set $\{(a, b) : \varphi(a, b) = 0\}$ need not be a definable set, and
- B the projection on the first coordinate need not be onto (in a non-saturated model).

Second (new) solution: solve the problem

The problem(s)

In continuous logic, if $\inf_y \varphi(x, y) = 0$, then

- A the set $\{(a, b) : \varphi(a, b) = 0\}$ need not be a definable set, and
- B the projection on the first coordinate need not be onto (in a non-saturated model).

The solution (almost)

- B allow an error, considering $\{(a, b) : \varphi(a, b) \leq 1\}$.
- A allow a variable error, considering $D = \{(r, a, b) : \varphi(a, b) \leq r\}$. This set D is definable:

  if (r, a, b) is logically close to D (i.e., $\varphi(a, b) \leq r + \varepsilon$),
  then it is metrically close to D (e.g., to $(r + \varepsilon, a, b) \in D$).

The problem(s)

In continuous logic, if $\inf_y \varphi(x, y) = 0$, then

A the set $\{(a, b) : \varphi(a, b) = 0\}$ need not be a definable set, and

B the projection on the first coordinate need not be onto (in a non-saturated model).

The solution (almost)

B allow an error, considering $\{(a, b) : \varphi(a, b) \leq 1\}$.

A allow a variable error, considering $D = \{(r, a, b) : \varphi(a, b) \leq r\}$. This set D is definable:

if (r, a, b) is logically close to D (i.e., $\varphi(a, b) \leq r + \varepsilon$),
then it is metrically close to D (e.g., to $(r + \varepsilon, a, b) \in D$).

New problem

For this to work, r must not bounded, and by compactness, we must allow $r = \infty$.
But. . . with infinite error, the condition $\varphi(a, b) \leq r$ is meaningless.

Definition (Reminiscing of Summer 2004...)

Let X be a set. We define

$$*X = ([0,1] \times X)/\sim = \left\{[\alpha, x] : \alpha \in [0,1],\ x \in X\right\}$$

where we identify $[0, x] = 0$ regardless of x.

Definition

Say $\inf_y \varphi_n(x_{<n}, y) = 0$ for each n, and $\Phi = (\varphi_n)$ is sufficiently rich. Define $D^*_\Phi \subseteq *M^{\mathbb{N}}$ by:

$$[\alpha, a] \in D^*_\Phi \iff \varphi_n(a_{<n}, a_n) \leq 1/n\alpha.$$

- D^*_Φ is definable (same argument as in the previous slide)
- If $[\alpha, a] \in D^*_\Phi$ and $\alpha > 0$ (finite error), then $[\alpha, a]$ codes a model: $\mathrm{dcl}([\alpha, a]) = \mathrm{dcl}(M)$.
- There exists a unique root $0 = [0, a] \in D^*_\Phi$. It codes nothing: $\mathrm{dcl}(0) = \mathrm{dcl}(\varnothing)$.

Definition

Say $\inf_y \varphi_n(x_{<n}, y) = 0$ for each n, and $\Phi = (\varphi_n)$ is sufficiently rich. Define $D^*_\Phi \subseteq {*M}^{\mathbb{N}}$ by:

$$[\alpha, a] \in D^*_\Phi \iff \varphi_n(a_{<n}, a_n) \leq 1/\alpha.$$

Theorem (B.)

The definable set $D^* = D^*_\Phi$ is unique, up to definable bijection. The groupoid $\mathbb{G}^*(T) = \mathbb{G}_{D^*}(T)$ is a complete bi-interpretation invariant for T:

$$\mathbb{G}^*(T) = \mathbb{G}_{D^*}(T) = \{\mathrm{tp}(a, b) : a, b \in D^* \textit{ and } \mathrm{dcl}(a) = \mathrm{dcl}(b)\}$$

basis: $\mathbb{B}^*(T) \simeq \mathrm{S}_{D^*}(T) \simeq$ *the Lelek fan L.*

A quick reminder about the Lelek fan?

- A fan is a connected subset of $*2^{\mathbb{N}}$.
- If the endpoints are dense, then it is a Lelek fan, and is unique up to homeomorphism.

Summary of the 3 constructions

Hypothesis	$\aleph_0$-categorical T D_0 = type of a model	a universal Skolem sort D exists	General case
Groupoid invariant	$G(T) = \mathrm{Aut}(M) = \mathbb{G}_{D_0}(T)$ (group)	$\mathbb{G}(T) = \mathbb{G}_D(T)$	$\mathbb{G}^*(T) = \mathbb{G}_{D^*}(T)$
Basis	$\mathrm{S}_{D_0}(T)$ = Point	$\mathrm{S}_D(T)$ = Cantor	$\mathrm{S}_{D^*}(T)$ = Lelek fan
Reconstruction	recover $\mathrm{Th}(D_0)$	recover $\mathrm{Th}(D)$	recover $\mathrm{Th}(D^*)$

Hypothesis	$\aleph_0$-categorical T D_0 = type of a model	a universal Skolem sort D exists	General case
Groupoid invariant	$G(T) = \mathrm{Aut}(M) = \mathbb{G}_{D_0}(T)$ (group)	$\mathbb{G}(T) = \mathbb{G}_D(T)$	$\mathbb{G}^*(T) = \mathbb{G}_{D^*}(T)$
Basis	$\mathbb{S}_{D_0}(T)$ = Point	$\mathbb{S}_D(T)$ = Cantor	$\mathbb{S}_{D^*}(T)$ = Lelek fan
Reconstruction	recover $\mathrm{Th}(D_0)$	recover $\mathrm{Th}(D)$	recover $\mathrm{Th}(D^*)$

Each case generalises the previous ones

- $\aleph_0$-categorical $\rightsquigarrow$ a Universal Skolem sort D:

$$D = 2^{\mathbb{N}} \times D_0$$

$$\mathbb{G}(T) = 2^{\mathbb{N}} \times G(T) \times 2^{\mathbb{N}}.$$

- Universal Skolem sort $D \rightsquigarrow$ general case

$$D^* = (L \times D)/\sim \quad \Big(= (L \times D_0)/\sim\Big).$$

$$\mathbb{G}^*(T) = \Big(L \times \mathbb{G}(T) \times L\Big)/\sim \quad \text{(almost)}.$$

Thank you