Describing structures and classes of structures

Julia F. Knight

University of Notre Dame

Topics to be discussed

I. Describing specific structures—Scott complexity

II. Characterizing classes—Borel complexity

III. Classification problems—Borel cardinality,
the slippery notion of “useful” invariants

Conventions

1. Structures are countable, with universe ω
2. Languages are countable, usually computable.
3. Classes consist of structures for a fixed language, and are closed under isomorphism.

$L_{\omega_1\omega}$-formulas

In $L_{\omega_1\omega}$, we allow countably infinite disjunctions and conjunctions, but only finite strings of quantifiers.

Sample formulas

1. The following sentence says of a real closed ordered field that it is Archimedean:

$$(\forall x) \bigvee_n x < \underbrace{1 + \cdots + 1}_n$$

2. The following formula says of an element in an Abelian group that it has infinite order.

$$\bigwedge_n \underbrace{x + \cdots + x}_n \neq 0$$

Normal form

Note. For $L_{\omega_1\omega}$-formulas, we do not have prenex normal form. We cannot, in general, bring the quantifiers to the front.

Normal form. We can bring the negations inside. This gives a different normal form, in which the complexity goes up with the alternations of $\bigvee\exists$ and $\bigwedge\forall$.

Complexity of $L_{\omega_1\omega}$-formulas

1. $\varphi(\bar{x})$ is Σ_0 and Π_0 if it is finitary quantifier-free.

2. For a countable ordinal $\alpha > 0$,

   (a) $\varphi(\bar{x})$ is Σ_α if it has form $\bigvee_i (\exists \bar{u}_i)\psi_i(\bar{x}, \bar{u}_i)$, where each ψ_i is Π_{β_i} for some $\beta_i < \alpha$,

   (b) $\varphi(\bar{x})$ is Π_α if it has form $\bigwedge_i (\forall \bar{u}_i)\psi_i(\bar{x}, \bar{u}_i)$, where each ψ_i is Σ_{β_i} for some $\beta_i < \alpha$.

Further terminology. A formula is d-Σ_α if it has form $(\varphi \,\&\, \psi)$, where φ is Σ_α and ψ is Π_α.

Computable infinitary formulas

Computable infinitary formulas are $L_{\omega_1\omega}$-formulas in which the infinite disjunctions and conjunctions are over computably enumerable (c.e.) sets.

We classify these formulas as *computable* Σ_α, *computable* Π_α, for *computable* ordinals α.

Remark: Computable infinitary formulas seem comprehensible.

Complexity of sample formulas

1. The sentence $(\forall x) \bigvee_n x < \underbrace{1 + \cdots + 1}_{n}$ is computable Π_2.

2. The formula $\bigwedge_n \underbrace{x + \cdots + x}_{n} \neq 0$ is computable Π_1.

3. For each $n \geq 1$, there is a computable d-Σ_2 sentence saying of a $\mathbb{Q}$-vector space that it has dimension n.

   We say that there exists an independent n-tuple, and there does not exist an independent $(n+1)$-tuple.

Describing a specific structure

Theorem (Scott, 1965). For each structure $\mathcal{A}$, there is an $L_{\omega_1\omega}$-sentence φ whose countable models are precisely the isomorphic copies of $\mathcal{A}$. (Such a sentence is called a *Scott sentence*.)

Proof sketch. There is a family of formulas $\varphi_{\bar{a}}$ defining the orbits of tuples $\bar{a}$. Then $\varphi = \bigwedge_{\bar{a}} \rho_{\bar{a}}$, where

- $\rho_{\emptyset} = (\forall y) \bigvee_b \varphi_b(y) \ \& \ \bigwedge_b (\exists y)\varphi_b(y)$,
- $\rho_{\bar{a}} = (\forall \bar{u})[\varphi_{\bar{a}}(\bar{u}) \rightarrow ((\forall y) \bigvee_b \varphi_{\bar{a},b}(\bar{u}, y) \ \& \ \bigwedge_b (\exists y)\varphi_{\bar{a},b}(\bar{u}, y))]$, for other $\bar{a}$.

Note: If the formulas $\varphi_{\bar{a}}$ are all Σ_α, then φ is $\Pi_{\alpha+1}$.

Complexity of Scott sentences

Theorem (A. Miller, 1983). If $\mathcal{A}$ has a $\Sigma_{\alpha+1}$ Scott sentence and one that is $\Pi_{\alpha+1}$, then there is one that is d-Σ_α. (If α is a limit ordinal, this is Π_α.)

Theorem (Montalbán, 2015). $\mathcal{A}$ has a $\Pi_{\alpha+1}$ Scott sentence iff the orbits of all tuples are defined by Σ_α formulas.

Alvir-Knight-McCoy. If $\mathcal{A}$ has a computable $\Pi_{\alpha+1}$ Scott sentence, then the orbits are defined by computable Σ_α formulas.

Alvir-Greenberg-Harrison-Trainor-Turetsky. Precise results on possible complexities of Scott sentences and orbits.

Scott sentence for $\mathbb{Z}$

Fact: The additive group of integers has a d-Σ_2 Scott sentence. Moreover, this is optimal.

For the Scott sentence, we take the conjunction of a Π_2 sentence characterizing the torsion-free Abelian groups, a Π_2 sentence saying that for any pair x, y, there is some z that generates both, and a Σ_2 sentence saying that there is some x not divisible by $n > 1$.

This is optimal. The proof involves index set calculations. For a d-Σ^0_2 set S, there is a computable sequence of groups G_n s.t. $G_n \cong \mathbb{Z}$ iff $n \in S$. Similarly, for a set S that is d-Σ^0_2 relative to X, there is an X-computable sequence of groups.

The free group F_n

Theorem (Carson-Harizanov-K-Lange-McCoy-Quinn-Morozov-Safranski-Wallbaum, 2012). For all finite $n \geq 1$, the free group F_n of rank n has a (computable) d-Σ_2 Scott sentence. This is optimal.

Proof sketch. For $n = 1$, $F_1 = \mathbb{Z}$. For $n \geq 2$, we take the conjunction of:

- a Π_2 sentence saying that for every tuple $\bar{y}$, there is an n-tuple $\bar{x}$ that generates $\bar{y}$,
- a Σ_2 sentence saying that there is an n-tuple $\bar{x}$ satisfying no non-trivial relations, s.t. for all n-tuples $\bar{y}$, no "imprimitive" n-tuple of words takes $\bar{y}$ to $\bar{x}$.

Note: Nielsen (1917, 1918) described the primitive tuples of words, and showed that the set of these is computable.

The free group F_∞

Theorem (McCoy-Wallbaum, 2012). The free group F_∞ has a (computable) Π_4 Scott sentence. This is optimal.

K-Saraph, Ho, Raz. Other familiar kinds of finitely generated groups have d-Σ_2 Scott sentences.

Based on the known examples, Ho and I had conjectured that all finitely generated groups have d-Σ_2 Scott sentences, and all computable finitely generated groups have computable d-Σ_2 Scott sentences.

Complexity of orbits

The complexity of an optimal Scott sentence for a structure is connected to the complexity of the orbits.

Theorem (Montalbán, 2015). $\mathcal{A}$ has a $\Pi_{\alpha+1}$ Scott sentence iff the orbits of all tuples in $\mathcal{A}$ are defined by Σ_α formulas.

Theorem (Alvir-K-McCoy, 2020). If $\mathcal{A}$ has a computable $\Pi_{\alpha+1}$ Scott sentence, then the orbits are defined by computable Σ_α formulas.

Scott sentences for finitely generated structures

Theorem (Harrison-Trainor-Ho, 2018). A finitely generated structure $\mathcal{A}$ has a d-Σ_2 Scott sentence iff it is not "self-reflexive; i.e., there is no substructure $\mathcal{B}$ generated by a tuple $\bar{b}$ s.t. the existential formulas true of $\bar{b}$ in $\mathcal{A}$ are all true in $\mathcal{B}$.

Theorem (Alvir-Knight-McCoy,). A finitely generated group has a d-Σ_2 Scott sentence iff for some generating tuple, the orbit is defined by a Π_1 formula iff for all generating tuples, the orbit is defined by a Π_1 formula.

Theorem (Harrison-Trainor-Ho). There is a computable finitely generated group with no d-Σ_2 Scott sentence.

Theorem (Harrison-Trainor-Ho). Every finitely generated field has a d-Σ_2 Scott sentence.

$Mod(L)$

For a countable language L, $Mod(L)$ is the set of L-structures with universe ω.

Identifying $Mod(L)$ with Cantor space. For simplicity, we suppose L is a relational language. Let C be a set of new constants representing the natural numbers. Let $(\alpha_n)_{n\in\omega}$ be a list of the atomic sentences $R\bar{a}$, where R is a relation symbol of L and $\bar{a}$ is a tuple from C.

We identify $\mathcal{A} \in Mod(L)$ with the function $f \in 2^\omega$ s.t.

$$f(n) = \begin{cases} 1 & \text{if } \mathcal{A} \models \alpha_n \\ 0 & \text{otherwise} \end{cases}$$

Borel classes

There is a natural topology on Cantor space, generated by the clopen sets $N_p = \{f \in 2^\omega : f \supseteq p\}$, for $p \in 2^{<\omega}$. The *Borel sets* are those in the σ-algebra generated by these N_p.

1. B is $\mathbf{\Sigma_0}$ and $\mathbf{\Pi_0}$ if it is a finite union of basic clopen sets.
2. For a countable ordinal $\alpha > 0$,
   (a) B is $\mathbf{\Sigma}_\alpha$ if $B = \cup_i B_i$, where each B_i is $\mathbf{\Pi}_{\beta_i}$ for some $\beta_i < \alpha$,
   (b) B is $\mathbf{\Pi}_\alpha$ if $B = \cap_i B_i$, where each B_i is $\mathbf{\Sigma}_{\beta_i}$ for some $\beta_i < \alpha$.

Axiomatizing Borel classes

Theorem (Lopez-Escobar, 1965): For $K \subset Mod(L)$, closed under automorphism, K is Borel iff it is axiomatized by a sentence of $L_{\omega_1\omega}$.

Theorem (Vaught, 1974): For $K \subseteq Mod(L)$, closed under automorphism, K is $\mathbf{\Sigma}_\alpha$ (for $\alpha \geq 1$) iff it is axiomatized by a Σ_α sentence.

Vaught's proof involved "Vaught transforms."

Vanden Boom (in his senior thesis at *ND*), gave an effective version.

Vanden Boom

Effective Borel hierarchy. The *effective Borel sets* are obtained from the basic clopen neighborhoods using c.e. unions and intersections.

Theorem (Vanden Boom, 2007). A set $B \subseteq Mod(L)$, closed under isomorphism, is effective Σ_α (for $\alpha \geq 1$) iff it is axiomatized by a computable Σ_α formula.

This, suitably relativized, gives Vaught's Theorem.

Idea of Proof: We imagine building a generic copy $\mathcal{A}^*$ of a structure $\mathcal{A}$. The forcing language is propositional, with propositional variables for the atomic sentences that might be true in $\mathcal{A}^*$. We have predicate formulas that define forcing—these accomplish what the Vaught transforms did.

Borel embeddings and Borel cardinality

Definition (H. Friedman & Stanley, 1989). Let $K \subseteq Mod(L)$, $K' \subseteq Mod(L')$, both closed under isomorphism. A *Borel embedding* of K in K' is a Borel function $\Phi : K \to K'$ s.t. for $\mathcal{A}, \mathcal{B} \in K$, $\mathcal{A} \cong \mathcal{B}$ iff $\Phi(\mathcal{A}) \cong \Phi(\mathcal{B})$.

Notation: We write $K \leq_B K'$ if there is such an embedding. We write $K <_B K'$ if $K \leq_B K'$ and $K' \not\leq_B K$, and we write $K \equiv_B K'$ if $K \leq_B K'$ and $K' \leq_B K$.

Definition. The *Borel cardinality* of K is its $\equiv_B$-class.

On top

Theorem (Lavrov, Maltsev, Mekler, Friedman-Stanley, Marker). The following classes lie on top under $\leq_B$:

1. undirected graphs
2. fields
3. 2-step nilpotent groups
4. linear orderings
5. real closed ordered fields

New result (Paolini and Shelah). Torsion-free Abelian groups also lie on top.

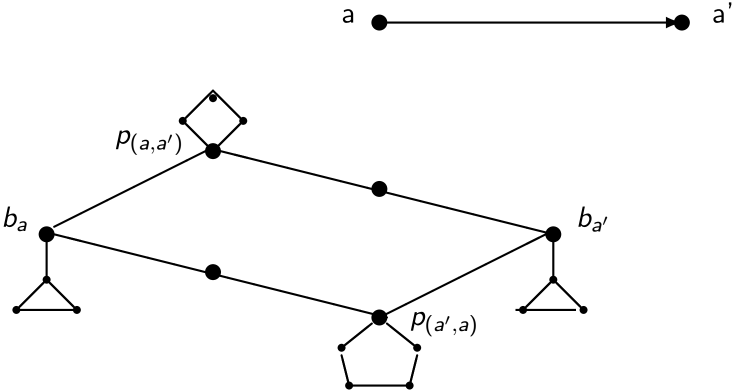
Embedding $Mod(L)$ in undirected graphs

Theorem (Lavrov, 1963). $Mod(L) \leq_B$ undirected graphs.

There are slightly different embeddings due to Marker, Nies. We follow Marker.

Start with the case where L has just one binary relation symbol–$Mod(L)$ is the class of directed graphs. The embedding Φ takes a directed graph $(A, \rightarrow)$ to an undirected graph $(B, -)$ with a special point b_a representing each $a \in A$ and a special point $p_{(a,a')}$ representing each ordered pair (a, a'). The following picture shows how the embedding works.

Picture



More

To see that Φ is $1-1$ on isomorphism types, we note that there is a copy of $\mathcal{A}$ defined in $\Phi(\mathcal{A})$—the definition uses finitary existential formulas.

To embed $Mod(L)$ in undirected graphs, we use more special points and more n-gons.

Interpretations

Definition. An *interpretation* of $\mathcal{A}$ in $\mathcal{B}$ is a sequence of formulas that define a set D of tuples in $\mathcal{B}$, relations R_i^* on D corresponding to the basic relations R_i of $\mathcal{A}$, and a congruence relation $\sim$ s.t. $(D, R_i^*)/_\sim \cong \mathcal{A}$.

For familiar examples such as the interpretation of $\mathbb{Z}$ in $\mathcal{N}$ or $\mathbb{Q}$ in $\mathbb{Z}$, $D \subseteq \mathcal{B}^n$ for some n. Montalbán defined a more general kind of interpretation, in which $D \subseteq \mathcal{B}^{<\omega}$.

Facts:

1. If D, $\sim$, $\not\sim$, R_i^* and $\neg R_i^*$ are defined by computable Σ_1 formulas, then there is a Turing operator Φ that takes copies of $\mathcal{B}$ to copies of $\mathcal{A}$.
2. If D, $\sim$, and R_i^* are defined by formulas of $L_{\omega_1\omega}$, then there is a Borel operator Φ that takes copies of $\mathcal{B}$ to copies of $\mathcal{A}$.

fields $\leq_B$ 2-step nilpotent groups

Maltsev, 1960. Let Φ take each field F to its Heisenberg group $H(F)$, which consists of matrices

$$\begin{bmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{bmatrix},$$

for $a, b, c \in F$.

Maltsev gave finitary existential formulas that define a copy of F in $H(F)$ with an arbitrary non-commuting pair as parameters.

Theorem (Alvir-Calvert-Goodman-Harizanov-K-Miller-Morozov-Soskova-Weisshaar). There are finitary existential formulas, with no parameters, that, for all fields F, effectively interpret F in $H(F)$.

graphs $\leq_B$ linear orderings

Friedman and Stanley defined an embedding Φ of graphs in linear orderings. The proof that Φ is $1-1$ does not involve a definition or interpretation, at least, not of the usual kind. For a graph G, let $L(G)$ be the corresponding linear ordering.

Theorem (Harrison-Trainor-Montalbán, 2020; K-Soskova-Vatev, 2020). There do not exist $L_{\omega_1\omega}$-formulas that, for all graphs G, define an interpretation of G in $L(G)$.

Below the top

The following classes lie strictly below the top under $\leq_B$, for different reasons:

1. $\mathbb{Q}$-vector spaces—only $\aleph_0$ isomorphism types,
2. subfields of the algebraic numbers—isomorphism relation is Borel,
3. Abelian p-groups—subtler reason.

Turing computable embeddings

Kechris suggested that my students and I consider effective embeddings.

Definition (Calvert-Cummins-K-Quinn, 2004). For classes K, K', closed under isomorphism, a *Turing computable embedding* of K in K' is a Turing operator $\Phi : K \to K'$ s.t. for $\mathcal{A}, \mathcal{B} \in K$, $\mathcal{A} \cong \mathcal{B}$ iff $\Phi(\mathcal{A}) \cong \Phi(\mathcal{B})$. We write $K \leq_{tc} K'$.

Examples. The Borel embeddings of Friedman-Stanley, Lavrov, Mekler, Maltsev, Marker are actually Turing computable.

Pullback Theorem

Theorem (K-Quinn-Vanden Boom, 2007). Suppose $K \leq_{tc} K'$ via Φ. For each computable infinitary sentence φ in the language of K', we can find a sentence φ^* in the language of K s.t. $\mathcal{A} \models \varphi^*$ iff $\Phi(\mathcal{A}) \models \varphi$. Moreover, for $0 < \alpha < \omega_1^{CK}$, if φ is computable Σ_α, then so is φ^*.

Proof. Forcing, definability of forcing.

Example. For each $n \geq 1$, there is a computable Σ_2-sentence φ_n saying of a $\mathbb{Q}$-vector space that the dimension is at least n. If $K \leq_{tc} \mathbb{Q}$-vector spaces, then the pullbacks of the φ_n describe invariants for K.

Return to Abelian p-groups

Fokina-K-Melnikov-Quinn-Safranski. There is no Turing computable embedding of graphs in Abelian p-groups or in other classes of "Ulm Type."

Proof: Apply Pullback Theorem.

Results of Safranski let us generalize to Borel embeddings.

Invariants

Suppose $K \leq_B K'$ via Φ. The embedding Φ reduces the classification problem for K to that for K'.

Having the same Borel cardinality means essentially having the same invariants. Exactly what counts as "useful" invariants is vague.

1. $\mathbb{Q}$-**vector spaces**: dimension—universally accepted as useful.
2. **Abelian** p-**groups**: Ulm sequence plus dimension of the divisible part—complicated, but pretty much accepted as useful.
3. **Boolean algebras**: Ketonen invariants—according to Camerlo and Gao, these are "very complicated."

Torsion-free Abelian groups

Let TFA_n be the class of torsion-free Abelian groups of rank n. These are isomorphic to subgroups of $\mathbb{Q}^n$ with n $\mathbb{Z}$-linearly independent elements.

Baer invariants for TFA_1. To describe $G \in TFA_1$, take a non-zero element a and give the set of n s.t. $n|a$. For different choices of a, the sets we obtain differ only finitely. Baer gave invariants based on this. Baer's invariants are accepted as useful.

Invariants of Maltsev, Kurosh. For $n \geq 2$, we can describe $G \in TFA_n$ by choosing a $\mathbb{Z}$-independent n-tuple, and, for each $\mathbb{Z}$-linear combination, saying which n divide it. The sets depend on the chosen tuple. Maltsev and Kurosh gave invariants based on this. As Hjorth and Thomas both report, Fuchs dismissed these invariants as no better than the group itself.

More on torsion-free groups of finite rank

Theorem (Hjorth, 1999). $TFA_1 <_B TFA_2$.

Theorem (Thomas, 2001). For $n \geq 2$, $TFA_n <_B TFA_{n+1}$.

Theorem (Harrison-Trainor-Ho). The class of torsion-free groups of finite rank lies on top, under $\leq_B$, among classes of finitely generated structures.

Project of Ho-K-R. Miller

1. Hjorth, Thomas use deep results of descriptive set theory, measure. Try to simplify the proofs, using methods of computability (generic subgroups of $\mathbb{Q}^n$).

2. For the Friedman-Stanley embedding L of graphs in linear orderings, show that there is no Borel operator that, given a copy of $L(G)$, produces a copy of G.

3. Show that the Paolini-Shelah embedding of graphs in torsion-free Abelian groups, like the Friedman-Stanley embedding of graphs in linear orderings, does not correspond to a uniform interpretation by formulas of $L_{\omega_1\omega}$.