Isomorphism and Classification for Countable Structures

Russell Miller

Queens College & CUNY Graduate Center

Continuity, Computability, Constructivity
26 June 2017
LORIA, Nancy, France

Where this begins: computable structures

Definition

In a computable signature, a structure $\mathcal{A}$ is *computable* if its domain is ω and its atomic diagram $\Delta(\mathcal{A})$ is decidable. (It follows that subsets of $\mathcal{A}$ definable by existential formulas must be computably enumerable, and those definable by $\exists\forall$ formulas must be arithmetically Σ_2, etc.)

This definition excludes:

- Finite structures. (Sometimes we allow the domain to be an initial segment of ω.)
- Infinite structures whose domain is a decidable subset of ω. However, all such structures have computable isomorphisms onto computable structures.
- Uncountable structures. Too bad. Deal with it.

Often we write $\mathcal{A}$ for the atomic diagram $\Delta(\mathcal{A})$.

The Isomorphism Problem

Let $\mathfrak{C}$ be a class of computable structures – often the computable models of a given theory. For example, for computable models of **ACF**$_0$, we use a Godel coding of the atomic formulas in the signature $(+, \cdot, 0, 1, \langle c_n \rangle_{n \in \omega})$, with the c_n representing elements in the domain ω:

$\mathfrak{C} = \{e \in \omega : \varphi_e$ decides the atomic diagram of some model of **ACF**$_0\}$.

Write F_e for the computable ACF determined by φ_e, for $e \in \mathfrak{C}$. The *Isomorphism Problem* for this class is:

$$I = \{(i, j) \in \mathfrak{C}^2 : F_i \cong F_j\}.$$

For computable models of **ACF**$_0$, this set is Π_3^0: $F_i \cong F_j$ iff

$$(\forall d \in \omega)[(\exists x_1, \ldots, x_d \text{ independent in } F_i)$$
$$\iff (\exists y_1, \ldots, y_d \text{ independent in } F_j)].$$

Complexity of the Isomorphism Problem

Usually, for natural classes $\mathfrak{C}$ of computable structures such as above, the Isomorphism Problem is complete at some level.

- For computable models of $\textbf{ACF}_0$, $\cong$ is Π_3^0-complete.
- For computable algebraic field extensions of $\mathbb{Q}$, $\cong$ is Π_2^0-complete, as also for computable finite-branching trees.

Complexity of the Isomorphism Problem

Usually, for natural classes $\mathfrak{C}$ of computable structures such as above, the Isomorphism Problem is complete at some level.

- For computable models of **ACF**$_0$, $\cong$ is Π_3^0-complete.
- For computable algebraic field extensions of $\mathbb{Q}$, $\cong$ is Π_2^0-complete, as also for computable finite-branching trees.
- For equivalence structures (i.e., equivalence relations on ω), $\cong$ is Π_4^0-complete.

Complexity of the Isomorphism Problem

Usually, for natural classes $\mathfrak{C}$ of computable structures such as above, the Isomorphism Problem is complete at some level.

- For computable models of **ACF**$_0$, $\cong$ is Π_3^0-complete.
- For computable algebraic field extensions of $\mathbb{Q}$, $\cong$ is Π_2^0-complete, as also for computable finite-branching trees.
- For equivalence structures (i.e., equivalence relations on ω), $\cong$ is Π_4^0-complete.
- For torsion-free abelian groups of rank 1, and also for finite-valence connected graphs, $\cong$ is Σ_3^0-complete.

Complexity of the Isomorphism Problem

Usually, for natural classes $\mathfrak{C}$ of computable structures such as above, the Isomorphism Problem is complete at some level.

- For computable models of **ACF**$_0$, $\cong$ is Π_3^0-complete.
- For computable algebraic field extensions of $\mathbb{Q}$, $\cong$ is Π_2^0-complete, as also for computable finite-branching trees.
- For equivalence structures (i.e., equivalence relations on ω), $\cong$ is Π_4^0-complete.
- For torsion-free abelian groups of rank 1, and also for finite-valence connected graphs, $\cong$ is Σ_3^0-complete.
- For graphs, for fields, for trees, and for many other broad classes of computable structures, $\cong$ is Σ_1^1-complete.
  (Cf. Hirschfeldt-Khoussainov-Shore-Slinko 2002.)

Why just computable structures?

Working only with computable structures makes things difficult.
(Lange-M-Steiner) gave an "effective classification" of the computable algebraic fields: a kind of Friedberg construction of a computable list of algebraic fields such that every computable algebraic field extension of $\mathbb{Q}$ is isomorphic to exactly one field on the list.
As with the original Friedberg enumeration (of the c.e. sets, without repetition), this list is not useful.

Why just computable structures?

Working only with computable structures makes things difficult. (Lange-M-Steiner) gave an "effective classification" of the computable algebraic fields: a kind of Friedberg construction of a computable list of algebraic fields such that every computable algebraic field extension of $\mathbb{Q}$ is isomorphic to exactly one field on the list.

As with the original Friedberg enumeration (of the c.e. sets, without repetition), this list is not useful.

Knight et al. considered *Turing-computable embeddings*.

Defn.

A *tc-embedding* of $\mathfrak{C}$ into $\mathfrak{D}$ is a functional Φ such that, for every $\mathcal{C} \in \mathfrak{C}$, $\Phi^{\mathcal{C}} \in \mathfrak{D}$, and:

$$\mathcal{C}_0 \cong \mathcal{C}_1 \text{ (in } \mathfrak{C}) \iff \Phi^{\mathcal{C}_0} \cong \Phi^{\mathcal{C}_1} \text{ (in } \mathfrak{D}).$$

Now we are no longer restricted to computable structures: $\mathfrak{C}$ and $\mathfrak{D}$ can be any classes of structures with domain ω.

Classes of countable structures

A structure $\mathcal{A}$ with domain ω (in a fixed language) is identified with its atomic diagram $\Delta(\mathcal{A})$, making it an element of 2^ω. We now consider classes of such structures, e.g.:

$$Alg = \{D \in 2^\omega : D \text{ is an algebraic field of characteristic } 0\}.$$

$$ACF_0 = \{D \in 2^\omega : D \text{ is an ACF of characteristic } 0\}.$$

$$\mathcal{T} = \{D \in 2^\omega : D \text{ is an infinite finite-branching tree}\}.$$

On each class, we have the equivalence relation $\cong$ of isomorphism. The topology on the class is the quotient topology, modulo $\cong$:

$$\mathcal{V} \subseteq Alg/\cong \text{ is open} \iff \{D \in Alg : [D] \in \mathcal{V}\} \text{ is open in } Alg.$$

Thus a basic open set in $Alg/\cong$ is determined by a finite set of polynomials in $\mathbb{Q}[X]$ which must each have a root in the field.

Examining this topology

The quotient topology on $Alg/\cong$ is not readily recognizable. The isomorphism class of the algebraic closure $\overline{\mathbb{Q}}$ (which is universal for the class Alg) lies in *every* nonempty open set $\mathcal{U}$, since if $F \in \mathcal{U}$, then some finite piece of the atomic diagram of F suffices for membership in $\mathcal{U}$, and that finite piece can be extended to a copy of $\overline{\mathbb{Q}}$.

In contrast, the prime model $[\mathbb{Q}]$ lies in no open set $\mathcal{U}$ except the entire space $Alg/\cong$. If $\mathbb{Q} \in \mathcal{U}$, then some finite piece of the atomic diagram of $\mathbb{Q}$ suffices for membership in $\mathcal{U}$, and this piece can be extended to a copy of any algebraic field.

This does not noticeably illuminate the situation.

Expanding the language for *Alg*

Classifying *Alg*/ $\cong$ properly requires a jump, or at least a fraction of a jump. For each $d > 1$, add to the language of fields a predicate R_d:

$$\models_F R_d(a_0, \ldots, a_{d-1}) \iff X^d + a_{d-1}X^{d-1} + \cdots + a_0 \text{ has a root in } F.$$

Write *Alg** for the class of atomic diagrams of algebraic fields of characteristic 0 in this expanded language.

Now we have computable reductions in both directions between *Alg**/ $\cong$ and Cantor space 2^ω, and these reductions are inverses of each other. Hence *Alg**/ $\cong$ is homeomorphic to 2^ω.

2^ω is far more recognizable than the original topological space *Alg*/ $\cong$ (without the root predicates R_d). We consider this computable homeomorphism to be a legitimate classification of the class *Alg*, and therefore view the root predicates (or an equivalent) as essential for effective classification of *Alg*.

What do the R_d add?

We do *not* have the same reductions between $Alg/\cong$ and 2^ω: these are not homeomorphic. This seems strange: all R_d are definable in the smaller language, so how can they change the isomorphism relation?

The answer is that they do not change the underlying set: we have a bijection between Alg and Alg^* which respects $\cong$. However, the relations R_d change the topology on $Alg^*/\cong$ from that on $Alg/\cong$. (These are both the quotient topologies of the subspace topologies inherited from 2^ω.)

We do have a continuous map from $Alg^*/\cong$ onto $Alg/\cong$, by taking reducts, and so $Alg/\cong$ is also compact. This map is bijective, but its inverse is not continuous.

Too much information

Now suppose that, instead of merely adding the dependence relations R_d, we add *all* computable Σ_1^c predicates to the language. That is, instead of the algebraic field F, we now have its jump F'.

Fact

$$F \cong K \iff F' \cong K'.$$

However, the class Alg' of all (atomic diagrams of) jumps of algebraic extensions of $\mathbb{Q}$, modulo $\cong$, is no longer homeomorphic to 2^ω. In particular, the Σ_1^c property

$$(\exists p \in \mathbb{Q}[X])(\exists x \in F) \ [p \text{ irreducible of degree } > 1 \ \& \ p(x) = 0]$$

holds just in those fields $\ncong \mathbb{Q}$. Therefore, the isomorphism class of $\mathbb{Q}$ forms a singleton open set in the space $Alg'/\cong$.
(Additionally, $Alg'/\cong$ is not compact.)

Related spaces

From the preceding discussion, we infer that the root predicates are exactly the information needed for a nice classification of *Alg*.

(What does "nice" mean here? To be discussed....)

For another example, consider the class $\mathcal{T}$ of all finite-branching infinite trees, under the predecessor function *P*. As before, we get a topological space $\mathcal{T}/\cong$, which is not readily recognizable. (There is still a prime model, but no universal model.)

The obvious predicates to add are the *branching predicates B_n*:

$$\models_T B_n(x) \iff \exists^{=n} y \, (P(y) = x).$$

With these, the space becomes homeomorphic to Baire space ω^ω. So $\mathcal{T}^*/\cong$ and $Alg^*/\cong$ are *not homeomorphic*.

What constitutes a nice classification?

With both *Alg* and $\mathcal{T}$, we found very satisfactory classifications, by adding just the right predicates to the language. But it is not always so simple.

Let **TFAb**$_1$ be the class of torsion-free abelian groups G of rank exactly 1. We usually view these as being classified by tuples $(\alpha_0, \alpha_1, \ldots)$ from $(\omega + 1)^\omega$, saying that an arbitrary nonzero $x \in G$ is divisible by p_n exactly $f(n)$ times. To account for the arbitrariness of x, we must identify tuples $\vec{\alpha}$ and $\vec{\beta}$ with only finite differences:

$$\exists k[(\forall j > k \; \alpha_j = \beta_j) \; \& \; (\forall j \; |\alpha_j - \beta_j| < k)].$$

What constitutes a nice classification?

With both *Alg* and $\mathcal{T}$, we found very satisfactory classifications, by adding just the right predicates to the language. But it is not always so simple.

Let **TFAb**$_1$ be the class of torsion-free abelian groups G of rank exactly 1. We usually view these as being classified by tuples $(\alpha_0, \alpha_1, \ldots)$ from $(\omega + 1)^\omega$, saying that an arbitrary nonzero $x \in G$ is divisible by p_n exactly $f(n)$ times. To account for the arbitrariness of x, we must identify tuples $\vec{\alpha}$ and $\vec{\beta}$ with only finite differences:

$$\exists k[(\forall j > k \; \alpha_j = \beta_j) \; \& \; (\forall j \; |\alpha_j - \beta_j| < k)].$$

The space **TFAb**$_1/\cong$ has the indiscrete topology: no finite piece of an atomic diagram rules out any isomorphism type. More info needed!

If, for all primes p, we add $D_p(x)$ and $D_{p^\infty}(x)$, saying that x is divisible by p and infinitely divisible by p, then we get the classification above. However, it is not homeomorphic to Baire space itself.

Reducibility on equivalence relations

To broaden our notion of classification, we apply descriptive set theory.

Definition

Let E and F be equivalence relations on 2^ω (or on ω^ω, or other spaces). A *reduction* of E to F is a function $g : 2^\omega \to 2^\omega$ satisfying:

$$(\forall x_0, x_1 \in 2^\omega)\ [x_0\ E\ x_1 \iff g(x_0)\ F\ g(x_1)].$$

Original context: $E \leq_B F$ if there is a reduction which is a Borel function on 2^ω.

Definition

A *continuous reduction g* is given by an oracle Turing functional Φ^S:

$$(\forall A \in 2^\omega)(\forall x \in \omega)\ \Phi^{A \oplus S}(x) = \chi_{g(A)}(x).$$

If $S = \emptyset$, then the reduction is *computable*.

Borel reducibility for 2^ω: the basic ERs

Standard Borel ERs are defined using the *columns A^k* of $A \in 2^\omega$:

- $A \; E_0 \; B \iff |A \Delta B| < \infty$.
- $A \; E_1 \; B \iff \forall^\infty k \; (A^k = B^k)$.
- $A \; E_3 \; B \iff \forall k \; (A^k \; E_0 \; B^k)$.
- $A \; E_{\text{set}} \; B \iff (\forall j \exists k) \; A^j = B^k \; \& \; (\forall j \exists k) \; B^j = A^k$.
- $A \; E_{\text{perm}} \; B \iff (\exists \text{ permutation } h \text{ of } \omega)(\forall j) A^j = B^{h(j)}$.

Additional ER's on 2^ω:

$$A \; E_{\text{card}} \; B \iff |A| = |B|.$$
$$A =^e B \iff \pi_1(A) = \pi_1(B), \text{ where } \pi_1(A) = \{x : \langle x, y \rangle \in A\}.$$
$$A =^f B \iff (\forall x) \; |\{y : \langle x, y \rangle \in A\}| = |\{y : \langle x, y \rangle \in B\}|.$$

$=^e$ and $=^f$ are both Borel-equivalent to $=$, but not homeomorphic.

Classifying other classes of structures

$2^\omega/=^f$ effectively classifies the class of (countable or finite) equivalence structures in the language with unary predicates $C_1, C_2, \ldots, C_\infty$ for the size of the equivalence class of an element. Just count the number of classes of each size $\leq \infty$ in the structure. (Equivalence structures can be classified by elements of ω^ω, but this requires Π^0_4 predicates in the language, much stronger than our C_i's and C_∞.)

$2^\omega/=^e$, which is homeomorphic to the *Scott topology* on 2^ω, effectively classifies the subrings of $\mathbb{Q}$: given a subring, just enumerate the set of those n such that the n-th prime p_n has a multiplicative inverse in the subring. Thus the subring $\mathbb{Z}\left[\frac{1}{p_{i_0}}, \frac{1}{p_{i_1}}, \ldots\right]$ gives an enumeration of the set $\{i_0, i_1, \ldots\}$.

Hence these two isomorphism spaces are *not* homeomorphic.

Project: Build useful new ER's on 2^ω and ω^ω

More ER's can be built from these. For example, let $A \, E_{\text{card}}^\forall \, B$ iff

$$|\{x : \forall y \, \langle x, y \rangle \in A\}| = |\{x : \forall y \, \langle x, y \rangle \in B\}|.$$

Then $2^\omega / E_{\text{card}}^\forall$ is homeomorphic to the isomorphism space for algebraically closed fields of characteristic 0.

If we adjoin independence predicates to the language of **ACF**$_0$, then this isomorphism space becomes homeomorphic to $2^\omega / E_{\text{card}}$.

Likewise, **TFAb**$_1$ is homeomorphic to ω^ω / E_0 in the right signature. In its own signature, it appears to be $(2^\omega / =^f) / E_0$.

Back to *Alg*

Since *Alg*$^*/\cong$ is homeomorphic to 2^ω it seems natural to transfer the Lebesgue measure from 2^ω to *Alg*$/\cong$. But this requires care.

Fix a computable $\overline{\mathbb{Q}}$, and enumerate $\overline{\mathbb{Q}}[X] = \{f_0, f_1, \dots\}$. Let $F_\lambda = \mathbb{Q}$. Given $F_\sigma \subset \overline{\mathbb{Q}}$, we find the least i, with f_i irreducible in $F_\sigma[X]$ of prime degree, for which it is not yet determined whether f_i has a root in F_σ. Adjoin such a root to $F_{\sigma^\wedge 1}$, but not to $F_{\sigma^\wedge 0}$. This gives a homeomorphism from 2^ω onto *Alg*$^*/\cong$, via $h \mapsto \cup_n F_{h\restriction n}$.

If we transfer standard Lebesgue measure to *Alg*$^*/\cong$, we get a measure in which the odds of 2 having a 24-th root are $\frac{1}{16}$, but the odds of 2 having a 23-rd root are $\frac{1}{2}$.

Back to *Alg*[∗]

Since $Alg^*/\cong$ is homeomorphic to 2^ω it seems natural to transfer the Lebesgue measure from 2^ω to $Alg/\cong$. But this requires care.

Fix a computable $\overline{\mathbb{Q}}$, and enumerate $\overline{\mathbb{Q}}[X] = \{f_0, f_1, \ldots\}$. Let $F_\lambda = \mathbb{Q}$. Given $F_\sigma \subset \overline{\mathbb{Q}}$, we find the least i, with f_i irreducible in $F_\sigma[X]$ of prime degree, for which it is not yet determined whether f_i has a root in F_σ. Adjoin such a root to $F_{\sigma^\frown 1}$, but not to $F_{\sigma^\frown 0}$. This gives a homeomorphism from 2^ω onto $Alg^*/\cong$, via $h \mapsto \cup_n F_{h \restriction n}$.

If we transfer standard Lebesgue measure to $Alg^*/\cong$, we get a measure in which the odds of 2 having a 24-th root are $\frac{1}{16}$, but the odds of 2 having a 23-rd root are $\frac{1}{2}$.

Better: the odds of F_σ having a root of the next polynomial f_i (of prime degree d) should be $\frac{1}{d}$. This gives the measure on $Alg^*/\cong$ corresponding to the Haar measure on $\mathrm{Aut}(\overline{\mathbb{Q}})$.

Measuring properties of algebraic fields

Using either of these measures, for (the isomorphism type of) an algebraic field, the property of being normal has measure 0. So does the property of having relatively intrinsically computable predicates R_d.

In Alg^*, the property of being relatively computably categorical has measure 1: given two roots x_1, x_2 of the same irreducible polynomial, one can wait for them to become distinct, since with probability 1 there will be an f for which $f(x_1, Y)$ has a root in the field but $f(x_2, Y)$ does not. This allows computation of isomorphisms between copies of the field. The process works uniformly except on a measure-0 set of fields.

Surprisingly, measure-1-many fields (and all random fields) in Alg remain relatively computably categorical even when the root predicates are removed from the language. However, the procedures for computing isomorphisms are not uniform. A single procedure can succeed only for measure-$(1 - \epsilon)$-many fields.

Things to consider

Question

Use projection and complementation to create more Borel ER's out of the current ones. Look for regular structure among these under computable reducibility on 2^ω.

Question

For Alg^* and $\mathcal{T}^*$, the homeomorphisms onto 2^ω and ω^ω allow one to transfer notions of randomness to structures in these classes: an isomorphism type is random if and only if it maps to a random real in 2^ω or ω^ω. Do these correspond to other notions of random structures?

Question

Are there computable reductions in either direction between classes with Π^0_4 isomorphism problems? E.g., the classes of equivalence structures and of trees which are finite-branching except at the root?