On the learning power of equivalence relations

Luca San Mauro (University of Bari, Italy)

ASL North American Meeting 2024, Iowa State University

Joint with Bazhenov, Cipriani, Fokina, Jain, Marcone, and Stephan

Ranking countable isomorphism problems

Borel reducibility is a decisive tool for assessing whether familiar classes of countable structures may be nicely classified. However, isomorphism problems with countably many isomorphism types have all the same Borel complexity; thus, finer notions are required to investigate them properly.

This issue is well-known and fuels the study of, e.g., Turing computable embeddings or (more recently) countable reductions.

Here, we propose another framework for ranking countable isomorphism problems which is inspired by algorithmic learning theory.

Algorithmic Learning Theory

Algorithmic Learning Theory (**ALT**) dates back to the work of **Gold** and **Putnam** in the 1960's and comprises different formal models for the inductive inference.

Generally speaking, **ALT** deals with the question of how a learner, provided with more and more data about some environment, is eventually able to achieve systematic knowledge about it.

Most work in **ALT** deals with two main paradigms: learning total computable functions; learning formal languages. These paradigms model the data to be learned as an *unstructured flow* — what if one deals with data having some structural content?

Our framework

All our structures $\mathcal{A}$ have a relation signature, have domain $\mathbb{N}$, and are identified with their atomic diagrams (i.e., with reals in 2^ω). All our families of structures are at most countable.

For a family $\mathfrak{K}$ of nonisomorphic structures:

- The learning domain $\mathrm{LD}(\mathfrak{K})$ is the collection of all isomorphic copies of the structures from $\mathfrak{K}$.
- A learner **L** sees, by stages, all positive and negative data about any structure from $\mathrm{LD}(\mathfrak{K})$ and is required to output conjectures about the isomorphism type of the observed structure.
- The learning is successful if, for each structure $\mathcal{S} \in \mathrm{LD}(\mathfrak{K})$, the learner eventually stabilizes to a correct conjecture.

$\mathfrak{K}$ is **Ex**-learnable (for *explanatory*), if some learner **L** successfully learns $\mathfrak{K}$.

A few basic examples

- $\{\omega, \omega^*\}$ *is* Ex-*learnable*
  (although, at each finite stage, fragments of ω and ω^* are indistinguishable, any copy of either of the structures will eventually show the true least or greatest element)
- $\{\omega, \zeta\}$ *is not* Ex-*learnable*
  (to diagonalize against a learner L, one constructs by stages a copy of either ω or ζ that forces L to fail or to change mind infinitely often)
- It's easy to build a family $\mathfrak{K}$ so that all finite sub-families of $\mathfrak{K}$ are Ex-learnable but $\mathfrak{K}$ is not.

Characterizing Ex-learning, I

It turns that the **Ex**-learnability (or, lack thereof) of a family $\mathfrak{K}$ depends on the existence of a suitable collection of $\mathcal{L}_{\omega_1\omega}$-formulas and has a natural descriptive set theoretic interpretation.

Theorem (Bazhenov, Fokina, S.)

A family $\mathfrak{K} := \{\mathcal{A}_i : i \in \mathbb{N}\}$ *of structures is* ***Ex****-learnable iff there are* $\Sigma^{\inf}_2$ *formulas* $\{\psi_i : i \in \mathbb{N}\}$ *so that*

$$\mathcal{A}_j \models \psi_i \Leftrightarrow i = j,$$

for every i and j.

Characterizing Ex-learning, II

Recall that E_0 denotes the relation of eventual agreement on reals, i.e.,

$$p \, E_0 \, q \Leftrightarrow (\exists m)(\forall n \geq m)(p(n) = q(n)).$$

Theorem (Bazhenov, Cipriani, S.)

A family $\mathfrak{K}$ of structures is **Ex***-learnable iff* $\mathrm{LD}(\mathfrak{K})$ *is continuously reducible E_0, i.e., there is a continuous function* Γ *so that*

$$\mathcal{A} \cong \mathcal{B} \Leftrightarrow \Gamma(\mathcal{A}) \, E_0 \, \Gamma(\mathcal{B}),$$

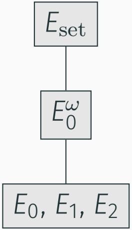
for all $\mathcal{A}, \mathcal{B} \in \mathrm{LD}(\mathfrak{K})$.

Introducing the learning hierarchy

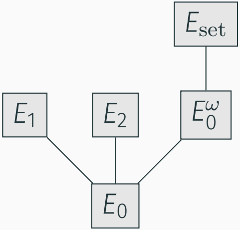
A natural hierarchy of learnability emerges by replacing E_0 with other Borel equivalence relations E, F:

- $\mathfrak{K}$ is E-learnable, if $\mathrm{LD}(\mathfrak{K})$ is continuously reducible to E;
- E is learning reducible to F ($E \leq_{\mathrm{learn}} F$), if every (countable) family $\mathfrak{K}$ which is E-learnable family is also F-learnable.

Clearly, if E is continuously reducible to F, then $E \leq_{\mathrm{learn}} F$. The converse (fortunately) fails.



Up to learning reducibility



Up to continuous reducibility

Characterizing the learning power of equivalence relations

Now, the syntactic characterization of E_0 suggests that the learning power of other E's can be similarly characterized by appealing to the logical complexity of the separating formulas of the E-learnable families.

A family $\mathfrak{K} := \{\mathcal{A}_i : i \in \mathbb{N}\}$ is a:

- $\Sigma_n^{\inf}$ poset, if $Th_{\Sigma_n^{\inf}}(\mathcal{A}_i) \neq Th_{\Sigma_n^{\inf}}(\mathcal{A}_j),$ for all $i \neq j$;
- $\Sigma_n^{\inf}$ antichain, if $Th_{\Sigma_n^{\inf}}(\mathcal{A}_i) \smallsetminus Th_{\Sigma_n^{\inf}}(\mathcal{A}_j) \neq \emptyset,$ for all $i \neq j$;
- $\Sigma_n^{\inf}$ strong antichain, if there are $\Sigma_n^{\inf}$ formulas $\{\psi_i : i \in \mathbb{N}\}$ so that $\mathcal{A}_j \models \psi_i \Leftrightarrow i = j$.

Hence, $\mathfrak{K}$ is E_0-learnable iff $\mathfrak{K}$ is a $\Sigma_2^{\inf}$ strong antichain.

E_{range} and E_{set}

The equivalence relations E_{range} and E_{set} are defined in the same way (stay tuned) but they live in different spaces:

- for $p, q \in \mathbb{N}^{\mathbb{N}}$,

$$p \, E_{\mathrm{range}} \, q \Leftrightarrow \{p(n) : n \in \mathbb{N}\} = \{q(n) : n \in \mathbb{N}\};$$

- for $p, q \in (\mathbb{N}^{\mathbb{N}})^{\mathbb{N}}$,

$$p \, E_{\mathrm{set}} \, q \Leftrightarrow \{p(n) : n \in \mathbb{N}\} = \{q(n) : n \in \mathbb{N}\}.$$

Proposition (Cipriani, Marcone, S.)

- *$\mathfrak{K}$ is E_{range}-learnable iff $\mathfrak{K}$ is a Σ_1^{inf} poset;*
- *$\mathfrak{K}$ is E_{set}-learnable iff $\mathfrak{K}$ is a Σ_2^{inf} poset.*

How to deal with all Σ_n^{inf} posets? Hint: Observe that E_{range} and E_{set} are Friedman-Stanley jumps.

The Friedman-Stanley jump operator

Let E be on X. The FS-jump E^+ of E is the equivalence relation on $X^{\mathbb{N}}$ given by

$$p\, E^+\, q \Leftrightarrow \{[p(n)]_E : n \in \mathbb{N}\} = \{[q(n)]_E : n \in \mathbb{N}\}.$$

So, E_{range} is $=^+_{\mathbb{N}}$ and E_{set} is $=^+_{\mathbb{N}^{\mathbb{N}}}$.

Remark: It's known that the nth back-and-forth relation is continuously reducible to $(=_{\mathbb{N}^{\mathbb{N}}})^{n+}$, so FS-jumps are natural candidates for the task of learning Σ_n^{inf} posets. However, in the learning setting these bounds are not sharp (observe that E_{range} is smooth but doesn't continuously reduce to $=_{\mathbb{N}^{\mathbb{N}}}$).

A tower of benchmark relations

Iterating the *FS*-jump operator on $=^+_{\mathbb{N}}$ and $=^+_{\mathbb{N}^{\mathbb{N}}}$, we reach the learning power needed for capturing all Σ^{inf}_n posets:

Theorem (Cipriani, Marcone, S.)

- $\mathfrak{K}$ *is* $(=_{\mathbb{N}})^{(n+1)+}$*-learnable iff* $\mathfrak{K}$ *is a* $\Sigma^{\mathrm{inf}}_{2n+1}$ *poset;*
- $\mathfrak{K}$ *is* $(=_{\mathbb{N}^{\mathbb{N}}})^{(n+1)+}$*-learnable iff* $\mathfrak{K}$ *is a* $\Sigma^{\mathrm{inf}}_{2n+2}$ *poset.*

For both items, the left-to-right direction relies on computable structure theoretic machinary (e.g., the Pullback lemma). The converse direction relies on first proving, by a fairly combinatorial inductive argument, that the result holds for all families of size 2, and then showing that every *FS*-jump E is compact-for-learning (i.e., $\mathfrak{K}$ is E-learnable iff every pair of structures from $\mathfrak{K}$ is so).

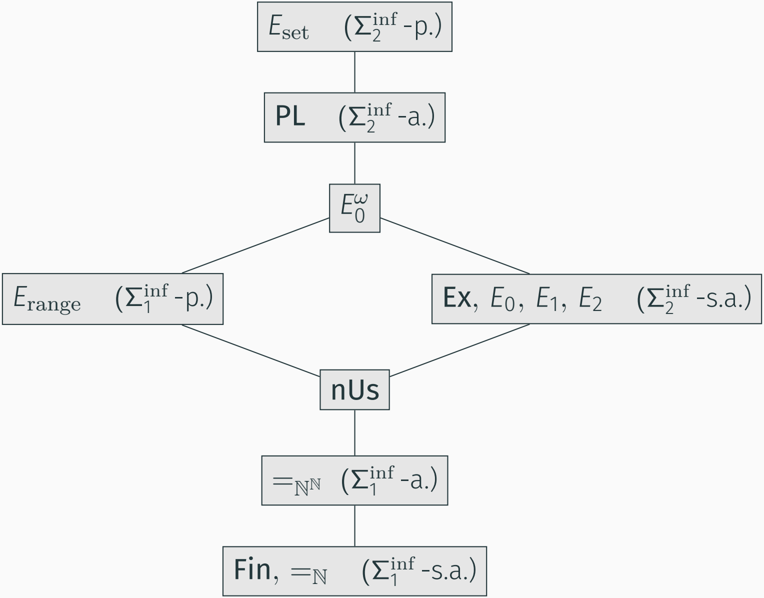
A quick glance beyond the *FS*-jumps

To gain further insights into the learning hierarchy, it's natural to investigate principles that are not equivalent to any *FS*-jump. This is the case of **Ex**.

Such a study is facilitated by analyzing principles that are well-established in classic **ALT**:

- In **Fin**-learning, the learner is not allowed to make any mind-change;
- In **PL**-learning, the learner can produce infinitely many mistakes as long as the only conjecture that is formulated infinitely often is the correct one.

We collect some of our results about the first few levels of the learning hierarchy into a concluding diagram.



E_{set} (Σ_2^{inf}-p.)
PL (Σ_2^{inf}-a.)
E_0^ω
E_{range} (Σ_1^{inf}-p.)
Ex, E_0, E_1, E_2 (Σ_2^{inf}-s.a.)
nUs
$=_{\mathbb{N}^\mathbb{N}}$ (Σ_1^{inf}-a.)
Fin, $=_\mathbb{N}$ (Σ_1^{inf}-s.a.)

Thanks!