Theoretical Computer Science 951 (2023) 113762

Contents lists available at ScienceDirect

Theoretical Computer Science

journal homepage: www.elsevier.com/locate/tcs

Learning algebraic structures with the help of Borel equivalence relations ☆,☆☆

Nikolay Bazhenov [a], Vittorio Cipriani [b], Luca San Mauro [c,*]

[a] Sobolev Institute of Mathematics, 4 Acad. Koptyug Ave., Novosibirsk, 630090, Russia
[b] Department of Mathematics, Computer Science and Physics, University of Udine, Italy
[c] Institute of Discrete Mathematics and Geometry, Vienna University of Technology, Austria

ARTICLE INFO

Article history:
Received 27 October 2021
Received in revised form 24 July 2022
Accepted 6 February 2023
Available online 10 February 2023
Communicated by M.J. Golin

Keywords:
Inductive inference
Algorithmic learning theory
Computable structures
Borel equivalence relations
Continuous reducibility

ABSTRACT

We study algorithmic learning of algebraic structures. In our framework, a learner receives larger and larger pieces of an arbitrary copy of a computable structure and, at each stage, is required to output a conjecture about the isomorphism type of such a structure. The learning is successful if the conjectures eventually stabilize to a correct guess. We prove that a family of structures is learnable if and only if its learning domain is continuously reducible to the relation E_0 of eventual agreement on reals. This motivates a novel research program, that is, using descriptive set theoretic tools to calibrate the (learning) complexity of nonlearnable families. Here, we focus on the learning power of well-known benchmark Borel equivalence relations (i.e., E_1, E_2, E_3, Z_0, and E_{set}).

1. Introduction

This paper wishes to connect two seemingly distant areas of research: algorithmic learning theory and the theory of Borel equivalence relations.

Algorithmic learning theory dates back to the work of Gold [16] and Putnam [30] in the 1960s and it encompasses several formal frameworks for the inductive inference. Broadly construed, this research program models the ways in which a *learner* can achieve systematic knowledge about a given *environment*, by access to more and more data about it. Although in classical paradigms the objects to be inferred are either formal languages or recursive functions (see, e.g., [33,23]), in recent times there has been a growing interest in the learning of data embodied with a structural content, with special attention paid to familiar classes of algebraic structures, such as vector spaces, rings, trees, and matroids [20,32,28,17].

In previous works [12,3], relying on ideas and technology from computable structure theory, we introduced and explored our own framework. Intuitively (formal details will be given below), an agent receives larger and larger pieces of an arbitrary copy of a computable structure and, at each stage, is required to output a conjecture about the isomorphism type of such a

☆ This article belongs to Section A: Algorithms, automata, complexity and games, Edited by Paul Spirakis.

☆☆ *Acknowledgments.* Bazhenov was supported by the Mathematical Center in Akademgorodok under the agreement No. 075-15-2022-281 with the Ministry of Science and Higher Education of the Russian Federation. Cipriani's research was partially supported by the Italian PRIN 2017 Grant "Mathematical Logic: Models, Sets, Computability". San Mauro's research was funded in whole or in part by the Austrian Science Fund (FWF) P36304-N. The authors wish to thank two anonymous referees for valuable comments.

* Corresponding author.
E-mail addresses: bazhenov@math.nsc.ru (N. Bazhenov), cipriani.vittorio@spes.uniud.it (V. Cipriani), luca.san.mauro@tuwien.ac.at (L. San Mauro).

https://doi.org/10.1016/j.tcs.2023.113762

structure. Then, the learning is successful if the conjectures eventually stabilize to a correct guess. See also [15,27] for other frameworks which can be similarly applied to arbitrary structures.

As single countable structures can always be learned, the emphasis of our research is on the learnability (or lack thereof) of families of structures. In [3], by adopting infinitary logic, we obtained a complete model theoretic characterization of which families of algebraic structures are learnable. From such a characterization, it immediately follows that some seemingly innocent learning problems are out of reach: e.g., no agent can learn whether the observed structure is a copy of the isomorphism type of the natural numbers or of the integers, that is, the pair of linear orders $\{\omega, \zeta\}$ is nonlearnable. We have also addressed the question of how much computational power is needed to handle a given learning problem: in [5], we constructed a pair of structures which is learnable, but no Turing machine can learn it.

A defect of our framework has been that, until this day, we had no way of calibrating the complexity of nonlearnable families. The present paper aims at rectifying this situation, by offering a new hierarchy to classify the complexity of learning problems for algebraic structures. To this end, we borrow several ideas from topology and descriptive set theory. This is readily justified. On the one hand, it is known that there are many connections between topology and inductive inference (see, e.g., [29,10,9]). On the other hand, a primary theme of modern descriptive set theory is the study of the complexity of equivalence relations defined on suitable topological spaces, with a special focus on the so-called *Borel equivalence relations*, to be defined below (see, e.g., [14,21,18]). A popular way of evaluating the complexity of Borel equivalence relations is by defining an appropriate reducibility: in general, a reduction from an equivalence relation E on X to an equivalence relation F on Y is a (nice) function $f : X \to Y$ which induces an embedding on the equivalence classes, $X_{/E} \to Y_{/F}$.

A large body of literature, within the theory of Borel equivalence relations, concerns equivalence relations associated to *classification problems*, i.e., problems which ask to scaffold a given family of mathematical structures up to a certain notion of "similarity". Crucially to our interests, isomorphism problems form an important subclass of classification problems, and descriptive set theorists have put serious effort in ranking the complexity of isomorphism problems for various familiar classes of countable structures (such as groups, trees, linear orderings, and Boolean algebras [13,25,7]).

The above description, albeit necessarily brief and oversimplified, may resound with our learning framework. Indeed, in our paradigm the learner is required to guess the isomorphism type for each structure from the family to be learned. Hence, the nonlearnability of a certain family $\mathfrak{K}$ of algebraic structures is, in a sense, rooted in the complexity of the isomorphism relation associated with $\mathfrak{K}$. Yet, two aspects shall be stressed:

(1) The isomorphism relations customarily studied in descriptive set theory refer to *large* collections of countable structures (e.g., *all* graphs, abelian groups, or metric spaces). On the contrary, here we focus on learning *small* families (i.e., countable families, and in fact often finite ones as in [5]);
(2) At any finite stage, the learner sees only a finite fragment of the structure to be learned, and each conjecture must be formulated without knowing how the observed structure will be extended. In topological terms, this coincides with asking that the learning must be a *continuous* process.

These observations are clearly informal. But, in Section 3, we'll be able to make them precise, while offering a new characterization of learnability, this time from a descriptive set theoretic point of view. Namely, we'll show that *a family of structures $\mathfrak{K}$ is learnable if and only if the isomorphism relation associated with $\mathfrak{K}$ is continuously reducible to the relation E_0 of eventual agreement on reals* (Theorem 3.1). As the relation E_0 is a fundamental benchmark in the theory of Borel equivalence relations (e.g., the celebrated Glimm-Effros dichotomy states that E_0 is the successor of the identity on reals within the Borel hierarchy [19]), such a new characterization of learnability for structures may serve as a piece of evidence that our paradigm is a natural one.

Furthermore, by replacing E_0 with Borel equivalence relations of higher complexity, one immediately unlocks the promised hierarchy of learning problems. That is, we'll say that a family of structures $\mathfrak{K}$ is *E-learnable*, for a Borel equivalence relation E, if there is a continuous reduction from the isomorphism relation associated with $\mathfrak{K}$ to E. Then, Sections 4–6 are dedicated to an investigation of the learning power of several benchmark Borel equivalence relations, offering both examples of relations which do not enlarge the scope of E_0-learnability (Theorems 4.1 and 4.3) and equivalence relations which do so (Theorem 5.2 and 5.4). Interestingly, we'll show that the learning power of some equivalence relations is affected by whether we restrict the attention to families containing only finitely many isomorphism types, or we rather allow countably infinite families. The final section contains a brief description of our intended future research in this area.

2. Preliminaries

As this paper is at the crossroad of a number of areas – namely, computable structure theory, algorithmic learning theory, and descriptive set theory – it will be convenient to break down these preliminaries in multiples subsections. We assume that the reader has some basic knowledge of topology and computability, as it can be found in [31]. In particular, by $(\varphi_e)_{e\in\omega}$, $(W_e)_{e\in\omega}$, and $(\Phi_e^X)_{e\in\omega}$ we denote a uniformly computable list of, respectively, all partial computable functions, all computably enumerable (c.e.) sets, and all Turing operators with oracle X.

Reals. We adopt the common habit of calling infinite binary sequences *reals*. To distinguish them from the natural numbers, reals are denoted by lowercase Greek letters (e.g., α, β). Functions on reals are denoted by uppercase Greek letters (e.g.,

Γ, Ψ). The m-th binary digit of a real α is denoted by $\alpha(m)$. By $\alpha^{[m]}$, we denote the real representing the m-th column $\alpha(\langle m, \cdot\rangle)$ of α. The *symmetric difference* $\alpha \triangle \beta$ of two reals is defined in the usual way:

$$(\alpha \triangle \beta)(i) = 1 \Leftrightarrow \alpha(i) \neq \beta(i).$$

The Cantor space. In this paper, we focus on equivalence relations defined on the *Cantor space*. Such a space, written as 2^ω, can be represented as the collection of reals, equipped with the product topology of the discrete topology on $\{0, 1\}$. For a binary string σ, the *cylinder* $[\sigma]$ is defined as the collection of reals extending σ, i.e.,

$$[\sigma] := \{\alpha \in 2^\omega : \sigma \subseteq \alpha\}.$$

These cylinders form a basis of 2^ω. A subset X of Cantor space is *Borel*, if it can be constructed from open sets, taking countable unions, countable intersections, and complements. A function $\Gamma : 2^\omega \to 2^\omega$ is *Borel*, if the preimage of any Borel set is Borel; it is *continuous*, if the preimage Γ of any open set is open. A Turing operator Φ can be naturally regarded as a partial function $\Phi : 2^\omega \to 2^\omega$, where $\Phi(\alpha)$ is defined if and only if Φ^α is total. Note that this partial function is continuous, since converging oracle computations are always determined by a finite initial segment of the oracle. For a function $\Gamma : 2^\omega \to 2^\omega$, a real α, and a number $s \in \omega$, the notation $\Gamma(\alpha)(s)$ refers to the sth bit of $\Gamma(\alpha)$.

Throughout the paper, we will often rely on the following lemma which expresses that every continuous function is computable with respect to some powerful enough oracle.

Lemma 2.1 *(folklore). If $\Gamma : 2^\omega \to 2^\omega$ is continuous, then there are an oracle X and Turing operator Φ so that*

$$\Gamma(\alpha) = \Phi^{X \oplus \alpha},$$

for every real α.

Proof. The continuity of Γ guarantees that there is a function $h : 2^{<\omega} \to 2^{<\omega}$ which satisfies the following requirements:

(1) for $\sigma, \tau \in 2^{<\omega}$, if $\sigma \subseteq \tau$, then $h(\sigma) \subseteq h(\tau)$;
(2) for all $\alpha \in 2^\omega$, $\Gamma(\alpha) = \bigcup_{\sigma \subset \alpha} h(\sigma)$.

Hence, it is straightforward to define the desired Turing operator by choosing an oracle X that computes h. □

In this paper, it is convenient to call every Turing operator of the form $\Phi^{X \oplus \alpha}$ a *Turing X-operator*. Intuitively, Turing X-operators can be identified with a Turing machine which has three tapes: the working tape (that we can assume contains the input as well), the oracle tape and the output tape.

Benchmark Borel equivalence relations. To evaluate the complexity of equivalence relations on reals, one defines a suitable reducibility. Let E and F be equivalence relations on 2^ω. A *reduction* from E to F is a function $\Gamma : 2^\omega \to 2^\omega$ such that

$$\alpha \; E \; \beta \Leftrightarrow \Gamma(\alpha) \; F \; \Gamma(\beta),$$

for all reals α, β. It is common to impose definability requirements on the functions inducing a reduction. *Borel reductions*, introduced in [13], are regarded as the most useful tools for calculating the relative complexity of equivalence relations. But in this paper, we'll concentrate on the following stronger reducibility: E is *continuously reducible* to F, if there is a continuous function $\Gamma : 2^\omega \to 2^\omega$ which reduces E to F.

The following combinatorial equivalence relations on reals are widely considered in descriptive set theory as benchmarks to gauge the complexity of natural classification problems (see, e.g., [21]):

(a) $\alpha \; E_0 \; \beta$ if and only if

$$(\exists m)(\forall n \geqslant m)(\alpha(n) = \beta(n)).$$

(b) $\alpha \; E_1 \; \beta$ if and only if

$$(\forall^\infty m \in \omega)(\alpha^{[m]} = \beta^{[m]}).$$

(c) $\alpha \; E_2 \; \beta$ if and only if

$$\sum_{k=0}^{\infty} \frac{(\alpha \triangle \beta)(k)}{k+1} < \infty.$$

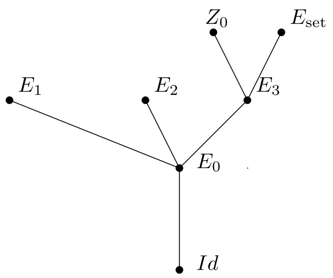

Fig. 1. Reductions up to continuous reducibility.

(d) $\alpha\ E_3\ \beta$ if and only if

$$(\forall m)(\alpha^{[m]}\ E_0\ \beta^{[m]}).$$

(e) $\alpha\ E_{set}\ \beta$ if and only if

$$\{\alpha^{[m]} : m \in \omega\} = \{\beta^{[m]} : m \in \omega\}.$$

(f) $\alpha\ Z_0\ \beta$ if and only if $\alpha \triangle \beta$ has (asymptotic) density zero, i.e.

$$\lim_{k\to\infty} \frac{\mathrm{card}(\{i \leqslant k : \alpha\triangle\beta(i) = 1\})}{k+1} = 0.$$

These benchmark equivalence relations lie at the base of the Borel hierarchy: Fig. 1, which is taken from [8], shows all continuous reducibilities between them (in fact, the diagram is the same even if we restrict to computable reductions). For more background about Borel, continuous, and computable reductions, see [14,18,26,4].

Computable structures. A *signature* L lists all function symbols and relation symbols which characterize an algebraic structure. All our structures have domain ω. We say that two structures are *copies* of each other if they are isomorphic.

In computable structure theory, one measures the complexity of an L-structure $\mathcal{A}$ by identifying $\mathcal{A}$ with its *atomic diagram*, i.e., the collection of atomic formulas which are true of $\mathcal{A}$. Up to a suitable Gödel numbering of L-formulas, the atomic diagram of $\mathcal{A}$ may be regarded as a real: this provides a natural way of assigning to each structure a Turing degree **d**, representing its algorithmic complexity. Any computable structure $\mathcal{A}$ in a relational signature (i.e., with no function symbols) can be presented as an increasing union of its finite substructures

$$\mathcal{A}\restriction_0 \subseteq \mathcal{A}\restriction_1 \subseteq \ldots \subseteq \mathcal{A}\restriction_i \subseteq \ldots,$$

where $\mathcal{A}\restriction_n$ denotes the restriction of $\mathcal{A}$ to the domain $\{0, 1, \ldots, n\}$ and $\mathcal{A} = \bigcup \mathcal{A}\restriction_i$. For more background about computable structures, see [1,11].

Infinitary formulas. To assess the model theoretic complexity of countable structures, it is common to work in the infinitary logic $\mathcal{L}_{\omega_1\omega}$, which allows to take the conjunctions or disjunctions of infinite sets of formulas. In particular, *infinitary* Σ_n *formulas* are defined as follows,

- Σ^{inf}_0 and Π^{inf}_0 formulas are quantifier-free first-order formulas.
- A $\Sigma^{\mathrm{inf}}_{n+1}$ formula $\psi(\bar{x})$ is a countably infinite disjunction

$$\bigvee_{i\in I} \exists \bar{y}_i \xi_i(\bar{x}, \bar{y}_i),$$

where each ξ_i is a Π^{inf}_n formula.
- A Π^{inf}_{n+1} formula $\psi(\bar{x})$ is a countably infinite conjunction

$$\bigwedge_{i\in I} \forall \bar{y}_i \xi_i(\bar{x}, \bar{y}_i),$$

where each ξ_i is a Σ^{inf}_n formula.

Next, *computable infinitary* Σ_n *formulas* (or Σ^c_n *formulas*, for short) are defined in the same way as above, but requiring infinite conjunctions and disjunctions to range over c.e. sets of (computable) formulas. Finally, computable infinitary formulas can be relativized to an arbitrary oracle X: the class of X-*computable infinitary* Σ_n *formulas* is denoted by $\Sigma^c_n(X)$. For more background about infinitary formulas, see [24].

2.1. Our framework

We shall now revisit the learning framework presented in [3]. Our exposition closely follows [5]. In particular, we ignore how a given family is enumerated and we just assume that any structure $\mathcal{A}$ gives rise to a corresponding *conjecture* $\ulcorner\mathcal{A}\urcorner$, to be understood as conveying the piece of information "this is $\mathcal{A}$".

Definition 2.2. Suppose that **P** is the learning problem associated to a countable family $\mathfrak{K}$ of nonisomorphic countable structures. The ingredients of our framework may be specified as follows. For **P**,

- The *learning domain* (LD) is the collection of all copies of the structures from $\mathfrak{K}$. That is,
$$\mathrm{LD}(\mathfrak{K}) := \bigcup_{\mathcal{A}\in\mathfrak{K}} \{\mathcal{S} : \mathcal{S} \cong \mathcal{A}\}.$$
As we identify each countable structure with an element of Cantor space, we obtain that $\mathrm{LD}(\mathfrak{K}) \subseteq 2^{\omega}$.
- The *hypothesis space* (HS) contains, for each $\mathcal{A} \in \mathfrak{K}$, a formal symbol $\ulcorner\mathcal{A}\urcorner$ and a question mark symbol. That is,
$$\mathrm{HS}(\mathfrak{K}) := \{\ulcorner\mathcal{A}\urcorner : \mathcal{A} \in \mathfrak{K}\} \cup \{?\}.$$
- A *learner* **M** sees, by stages, all positive and negative data about any given structure in the learning domain and is required to output conjectures. This is formalized by saying that **M** is a function
$$\text{from } 2^{<\omega} \text{ to } \mathrm{HS}(\mathfrak{K}).$$
- The learning is *successful* if, for each structure $\mathcal{S} \in \mathrm{LD}(\mathfrak{K})$, the learner eventually stabilizes to a correct conjecture about its isomorphism type. That is,
$$\lim_{n\to\infty} \mathbf{M}(\mathcal{S} \restriction_n) = \ulcorner\mathcal{A}\urcorner \text{ if and only if } \mathcal{S} \text{ is a copy of } \mathcal{A}.$$

In case $\mathcal{S} \notin \mathrm{LD}(\mathfrak{K})$ we make no assumption on **M**'s behavior, i.e., **M** either diverges or converges to some $x \in \mathrm{HS}(\mathfrak{K})$.

We say that $\mathfrak{K}$ is *learnable*, if some learner **M** successfully learns $\mathfrak{K}$.

Remark 2.3. In [5], the domain of a learner was limited to
$$X := \{\mathcal{S} \restriction_n \, : \mathcal{S} \in \mathrm{LD}(\mathfrak{K})\},$$
the collection of (finite) initial segments of structures from the learning domain. For our present purposes, it is more convenient to let **M** be defined on all binary strings. This change is not problematic. Indeed, any learner with domain X can be (non-effectively but continuously) transformed into a learner with domain $2^{<\omega}$ by simply let $\mathbf{M}(\sigma) =?$, for all $\sigma \in 2^{<\omega} \smallsetminus X$.

In [3, Theorem 3], we obtained the following model theoretic characterization of which families of structures are learnable.

Theorem 2.4 *(Bazhenov, Fokina, San Mauro). Let $\mathfrak{K} := (\mathcal{A}_i)_{i\in\omega}$ be a countable family of pairwise nonisomorphic structures. Then, $\mathfrak{K}$ is learnable if and only if there are Σ^{inf}_2 formulas $\varphi_0, \dots, \varphi_n, \dots$ such that*
$$\mathcal{A}_i \models \varphi_j \Leftrightarrow i = j.$$

The interested reader is referred to [3] for motivating examples and a detailed discussion about our framework (there named **InfEx**$_{\cong}$-learning).

Definition 2.5. We say that a family of structures $\mathfrak{K}$ is *countable* if it contains at most countably many isomorphism types. Similarly, $\mathfrak{K}$ is *finite*, if it contains only *finitely* many isomorphism types. In this paper, we won't consider uncountable families.

Turing computable embeddings. We conclude these preliminaries with a brief reminder about the technology of Turing computable embeddings, which was fundamental for proving Theorem 2.4 and will play a decisive role in Section 5.4.

Turing computable embeddings allow to compare the algorithmic complexity of different isomorphism problems.

Definition 2.6. ([6,22]). Let $\mathcal{K}_0$ and $\mathcal{K}_1$ be classes of structures. A Turing operator Φ is a *Turing computable embedding* of $\mathcal{K}_0$ into $\mathcal{K}_1$ (notation: $\mathcal{K}_0 \leqslant_{tc} \mathcal{K}_1$) if it induces an embedding $\mathcal{K}_{0/\cong} \to \mathcal{K}_{1/\cong}$, that is, if Φ satisfies the following:

- For any $\mathcal{A} \in \mathcal{K}_0$, the function $\Phi^{\mathcal{A}}$ is the characteristic function of the atomic diagram of a structure from $\mathcal{K}_1$. This structure is denoted by $\Phi(\mathcal{A})$.
- For any $\mathcal{A}, \mathcal{B} \in \mathcal{K}_0$, we have

$$\mathcal{A} \cong \mathcal{B} \Leftrightarrow \Phi(\mathcal{A}) \cong \Phi(\mathcal{B}).$$

It is common to abbreviate the term "Turing computable embedding" as *tc-embedding*. One of the most powerful tool in the theory of *tc*-embeddings is the so-called Pullback Theorem [22]. In this paper, we'll adopt a natural relativization of this result, already employed in [3].

Theorem 2.7 *(Relativized Pullback Theorem). Suppose that $X \subseteq \omega$ and $\mathfrak{K}_0 \leqslant_{tc} \mathfrak{K}_1$ via a Turing X-operator Φ. Then, for any X-computable infinitary sentence ψ in the signature of $\mathfrak{K}_1$, one can find, effectively with respect to X, an X-computable infinitary sentence $\psi^\star$ in the signature of $\mathfrak{K}_0$ such that, for all $\mathcal{A} \in \mathfrak{K}_0$, we have*

$$\mathcal{A} \models \psi^\star \Leftrightarrow \Phi(\mathcal{A}) \models \psi.$$

Note that the Relativized Pullback Theorem can be applied to any continuous operator Φ. Indeed, if Φ is continuous, then, by Lemma 2.1, it is equivalent to a Turing X-operator for some suitable oracle X.

We have amassed enough formal ingredients. Let's start.

3. A new characterization of learnability

In this section, we offer the promised descriptive set theoretic interpretation of our learning framework. Remember that E_0 denotes the relation of eventual agreement of reals, i.e., $\alpha \; E_0 \; \beta$ holds if and only if

$$(\exists m)(\forall n \geqslant m)(\alpha(n) = \beta(n)).$$

Theorem 3.1. *A family of structures $\mathfrak{K}$ is learnable if and only if there is a continuous function $\Gamma : 2^\omega \to 2^\omega$ such that*

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

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

Proof. For the sake of exposition, we'll assume that $\mathfrak{K}$ is infinite (the other case being easier) and it coincides with $(\mathcal{A}_i)_{i \in \omega}$. Denote $\Gamma(\mathcal{A}_i)$ by β_i.

($\Leftarrow$): Let Γ be a function which induces a continuous reduction from $\mathrm{LD}(\mathfrak{K})_{/\cong}$ to E_0. We need to show that $\mathfrak{K}$ is learnable. Certainly, $(\beta_i \not{E_0} \; \beta_j)$, for all $i \neq j$. Since Γ is continuous, by Lemma 2.1 there exists an oracle $X \in 2^\omega$ and a Turing operator Φ so that

$$\Gamma(\alpha) = \Phi^{X \oplus \alpha}, \text{ for every } \alpha \in 2^\omega.$$

Let α be a real and consider the uniform join

$$Y := X \oplus \bigoplus_{i \in \omega} \beta_i.$$

We define a Y-computable auxiliary function $f_{sim}(\alpha; i, s)$. Informally speaking, $f_{sim}(\alpha; i, s)$ is a measure of similarity (at the stage s) between the reals $\Gamma(\alpha)$ and β_i.

Let $\ell[s]$ be the greatest number such that for every $x \leqslant \ell[s]$, the value $\Phi^{(X \oplus \alpha) \upharpoonright s}(x)[s]$ is defined. If there is no such $\ell[s]$, then set $f_{sim}(\alpha; i, s) := -1$ for all $i \in \omega$.

Otherwise, for an index $i \in \omega$, we put

$$f_{sim}(\alpha; i, s) := \begin{cases} \max\left\{k : \ell - k \geqslant i, \text{ and for every } j \leqslant k, \; \Phi^{(X\oplus\alpha)\upharpoonright s}(\ell - j) = \beta_i(\ell - j)\right\}, & \text{if } \Phi^{(X\oplus\alpha)\upharpoonright s}(\ell) = \beta_i(\ell) \\ & \text{and } \ell \geqslant i; \\ -1, & \text{otherwise.} \end{cases}$$

Here by ℓ we denote $\ell[s]$. Without loss of generality, we assume that $\ell[s+1] \in \{\ell[s], \ell[s]+1\}$.

It is not hard to show that the function f_{sim} satisfies the following properties. Suppose that a real α encodes a copy of the structure $\mathcal{A}_{i_0}$, for some $i_0 \in \omega$.

(a) Note that there is an index m_0 such that for all $x \geqslant m_0$, we have $\Gamma(\alpha)(x) = \beta_{i_0}(x)$. This implies that there exists a stage s_0 such that every $s \geqslant s_0$ satisfies $f_{sim}(\alpha; i_0, s+1) \geqslant f_{sim}(\alpha; i_0, s) > -1$. In addition,

$$\lim_s f_{sim}(\alpha; i_0, s) = \infty.$$

(b) Let $i \neq i_0$. Since $(\Gamma(\alpha) \not{E_0}\, \beta_i)$ and $\ell[s+1] \leqslant \ell[s]+1$ for all s, there are infinitely many stages s such that $\Phi^{(X \oplus \alpha)\restriction s}(\ell[s]) \neq \beta_i(\ell[s])$ and $f_{sim}(\alpha; i, s) = -1$. Therefore,

$$\liminf_s f_{sim}(\alpha; i, s) = -1.$$

Construction. We build our desired learner **M**. Put **M**(empty string) :=?.

Let α be a real, and let s be a non-zero natural number. If the set

$$B_s = \{j \leqslant s : f_{sim}(\alpha; j, s) > -1\}$$

is empty, then define $\mathbf{M}(\alpha \restriction s) := \mathbf{M}(\alpha \restriction s-1)$. Otherwise, take $\mathbf{M}(\alpha \restriction s) := \ulcorner \mathcal{A}_{j^*} \urcorner$, where

$$j^* = \min\left\{ j \in B_s : f_{sim}(\alpha; j, s) = \max\{f_{sim}(\alpha; m, s) : m \in B_s\}\right\}. \tag{1}$$

Verification. We show that **M** learns our family $\mathfrak{K}$. Suppose that α is a real, which encodes a copy $\mathcal{S}$ of some $\mathcal{A}_{i_0}$.

Let s_0 be a stage such that $f_{sim}(\alpha; i_0, s) \neq -1$ for all $s \geqslant s_0$. Set $\ell^* := \ell[s_0]$. Observe the following:

- By the definition of the function f_{sim}, we have $\ell^* \geqslant i_0$. In addition, at each stage $s \geqslant s_0$, the value $\mathbf{M}(\alpha \restriction s)$ is defined according to Eq. (1).
- Suppose that $j > \ell^*$ and $s > s_0$. If $j > \ell[s]$, then $f_{sim}(\alpha; j, s) = -1$. If $j \leqslant \ell[s]$, then

  $$f_{sim}(\alpha; j, s) \leqslant \ell[s] - j < \ell[s] - \ell^* \leqslant (\ell[s] - \ell^*) + f_{sim}(\alpha; i_0, s_0) = f_{sim}(\alpha; i_0, s)$$

  Hence, Eq. (1) implies that $\mathbf{M}(\alpha \restriction s) \neq \ulcorner \mathcal{A}_j \urcorner$.

We deduce that for all $s > s_0$, we have $\mathbf{M}(\alpha \restriction s) \in \{\ulcorner \mathcal{A}_i \urcorner : 0 \leqslant i \leqslant \ell^*\}$. Choose a stage $s_1 > s_0$ such that for each $j \in \{0, 1, \ldots, \ell^*\} \setminus \{i_0\}$, there is another stage t_j satisfying $\max(s_0, \ell^*) < t_j < s_1$ and $f_{sim}(\alpha; j, t_j) = -1$.

For every $j \in \{0, 1, \ldots, \ell^*\} \setminus \{i_0\}$ and $s > s_1$, observe the following: if $t_j \leqslant \ell[s]$, then we have

$$f_{sim}(\alpha; j, s) \leqslant \ell[s] - t_j < \ell[s] - \ell^* \leqslant f_{sim}(\alpha; i_0, s).$$

Choose $s_2 > s_1$ such that $\ell[s_2] \geqslant \max t_j$. Then Eq. (1) implies that $\mathbf{M}(\alpha \restriction s) = \ulcorner \mathcal{A}_{i_0} \urcorner$ for all $s > s_2$. Therefore, in the limit, **M** says that "$\mathcal{S}$ is a copy of $\mathcal{A}_{i_0}$", and the family $\mathfrak{K}$ is learnable by **M**.

($\Rightarrow$): For the converse direction, let **M** be a learner of $\mathfrak{K}$. We need to construct a continuous function $\Gamma : 2^\omega \to 2^\omega$ which induces a reduction from $\mathrm{LD}(\mathfrak{K})_{/\cong}$ to E_0. To this end, it suffices to fix a countably infinite transversal $(\alpha_i)_{i\in\omega}$ of E_0 (i.e., a set intersecting countably many equivalence classes of E_0 in exactly one point) and define Γ as follows,

$$\Gamma(\beta)(s) := \alpha_{\mathbf{M}(\beta\restriction_s)}(s).$$

Here we use the following convention:

- if $\mathbf{M}(\beta \restriction_s) = \ulcorner \mathcal{A}_i \urcorner$, then $\alpha_{\mathbf{M}(\beta\restriction_s)} = \alpha_i$;
- if $\mathbf{M}(\beta \restriction_s) =?$, then $\alpha_{\mathbf{M}(\beta\restriction_s)} = 0^\infty$.

To verify that this Γ works, it is enough to observe the following: if a real β encodes a copy of some $\mathcal{A}_i$ from $\mathrm{LD}(\mathfrak{K})$, then there must be a stage s_0 such that, for all $s \geqslant s_0$, $\mathbf{M}(\beta \restriction_s)$ outputs $\ulcorner \mathcal{A}_i \urcorner$, and thus $\Gamma(\beta)$ is E_0-equivalent to α_i. So, since the α_i's form a transversal for E_0, we deduce that, if β_0 and β_1 encode copies of $\mathcal{A}_i$ and $\mathcal{A}_j$ respectively, then

$$\Gamma(\beta_0)\ E_0\ \Gamma(\beta_1) \Leftrightarrow i = j.$$

This concludes the proof. □

The above theorem unlocks a natural way to stratify learning problems, by simply replacing E_0 with Borel equivalence relations of higher and higher complexity.

Definition 3.2. A family of structures $\mathfrak{K}$ is *E-learnable* if there is a function $\Gamma : 2^\omega \to 2^\omega$ which continuously reduce $\mathrm{LD}(\mathfrak{K})_{/\cong}$ to E.

Definition 3.3. Let E, F be Borel equivalence relations. E is Learn$^{\omega}$*-reducible* to F, if every countable E-learnable family is also F-learnable. E is Learn$^{<\omega}$*-reducible* to F, if every finite E-learnable family is also F-learnable.

The rest of the paper is devoted to the study of the learning of power of some benchmark Borel equivalence relations.

4. When oracle equivalence relations don't help

In this section, we analyze the learning power of E_1 and E_2. These equivalence relations are incomparable and strictly above E_0 with respect to continuous reductions; in fact, the same is true if one requires computable reductions. But, as is proven in Theorems 4.1 and 4.3, E_1 and E_2 coincide and collapse to E_0 with respect to their learning power.

4.1. E_1-learning

Recall that the equivalence relation E_1 is given by

$$(\alpha\ E_1\ \beta) \Leftrightarrow (\forall^{\infty} m \in \omega)(\alpha^{[m]} = \beta^{[m]}).$$

Theorem 4.1. *A family $\mathfrak{K}$ is E_1-learnable if and only if $\mathfrak{K}$ is E_0-learnable. That is, E_1 and E_0 are* Learn$^{\omega}$*-equivalent.*

Proof. Since E_0 is continuously reducible to E_1 (see Fig. 1), every E_0-learnable family is also E_1-learnable.

On the other hand, let $\mathfrak{K} := (\mathcal{A}_i)_{i\in\omega}$ be an E_1-learnable family. Let $\Gamma : 2^{\omega} \to 2^{\omega}$ induce a continuous reduction from $\mathrm{LD}(\mathfrak{K})_{/\cong}$ to E_1. For each $i \in \omega$, we choose a real β_i such that Γ maps all copies of $\mathcal{A}_i$ into the class $[\beta_i]_{E_1}$. Fix a computable bijection ξ from the set $\{(i, j) \in \omega^2 : i \neq j\} \times \omega$ onto ω.

We build a set $X = \{m_s : s \in \omega\}$ as follows. Put $m_0 := 0$. Suppose that $s = \xi(i, j, t)$ and m_s is already defined. Since $(\beta_i \not\!\!E_1\ \beta_j)$, there exists the least $q > m_s$ such that $\beta_i^{[q]} \neq \beta_j^{[q]}$. We choose the least $\ell \in \omega$ with $\beta_i^{[q]}(\ell) \neq \beta_j^{[q]}(\ell)$, and put $m_{s+1} := \langle q, \ell\rangle$. It is not hard to see that for every $q \in \omega$, there is at most one ℓ such that $\langle q, \ell\rangle$ belongs to X.

We define an operator Ψ as follows: for every real α and $s \in \omega$, set

$$\Psi(\alpha)(s) := \alpha(m_s).$$

It is clear that the operator Ψ is X-computable — hence, Ψ is continuous.

We show that the operator $\Phi := \Psi \circ \Gamma$ provides a continuous reduction from $\mathrm{LD}(\mathfrak{K})_{/\cong}$ to E_0. Let α be a real which encodes a copy of some $\mathcal{A}_{i_0}$. Since $(\Gamma(\alpha)\ E_1\ \beta_{i_0})$, almost every $s \in \omega$ satisfies $\Gamma(\alpha)(m_s) = \beta_{i_0}(m_s)$. Thus, $\Phi(\alpha)$ is E_0-equivalent to $\Psi(\beta_{i_0})$.

Suppose that $i \neq i_0$. Then for almost all $t \in \omega$, we have

$$\Gamma(\alpha)(m_{\xi(i,i_0,t)}) = \beta_{i_0}(m_{\xi(i,i_0,t)}) \neq \beta_i(m_{\xi(i,i_0,t)}).$$

This implies that $(\Phi(\alpha)\ \not\!\!E_0\ \Psi(\beta_i))$. Therefore, we deduce that our family $\mathfrak{K}$ is E_0-learnable. The theorem is proved. □

Remark 4.2. The previous theorem can also be restated in purely descriptive set theoretic terms as follows. Let $(\beta_i)_{i\in\omega}$ be a sequence of reals such that $\beta_i E_1 \beta_j$ iff $i = j$. Then, there is a continuous function Γ such that for every $\gamma \in \bigcup_{i\in\omega}[\beta_i]_{E_1}$ we have that $\gamma E_1 \beta_i$ if and only if $\Gamma(\gamma) E_0 \Gamma(\beta_i)$, i.e. E_1, with domain restricted to $(\beta_i)_{i\in\omega}$, continuously reduces to E_0 via Γ. Theorems 4.3, 5.1 and 6.2 admit similar characterizations, but we will not make them explicit.

4.2. E_2-learning

Recall that the equivalence relation E_2 is given by

$$\alpha\ E_2\ \beta \Leftrightarrow \sum_{k=0}^{\infty} \frac{(\alpha \triangle \beta)(k)}{k+1} < \infty.$$

Theorem 4.3. *A countable family $\mathfrak{K}$ is E_2-learnable if and only if $\mathfrak{K}$ is E_0-learnable. That is, E_2 and E_0 are* Learn$^{\omega}$*-equivalent.*

Proof. Since E_0 is continuously reducible to E_2 (see Fig. 1), every E_0-learnable family is E_2-learnable.

Let $\mathfrak{K} := (\mathcal{A}_i)_{i\in\omega}$ be an E_2-learnable family. Let Γ be an operator, which induces a continuous reduction from $\mathrm{LD}(\mathfrak{K})_{/\cong}$ to E_2. For $i \in \omega$, we fix a real β_i such that Γ maps all copies of $\mathcal{A}_i$ into $[\beta_i]_{E_2}$.

By Lemma 2.1, there exist an oracle X and a Turing operator Φ such that $\Gamma(\alpha) = \Phi^{X\oplus\alpha}$ for all $\alpha \in 2^{\omega}$.

Construction. We define a $(X \oplus \bigoplus_{i\in\omega} \beta_i)$-computable operator Ψ. For a real α, we describe how to construct the real $\gamma_\alpha = \Psi(\alpha)$. For $s \in \omega$, by $\ell[s]$ we denote the greatest number such that for every $x \leqslant \ell[s]$, the value $\Phi^{(X\oplus\alpha)\upharpoonright s}(x)[s]$ is defined. Without loss of generality, one may assume that $\ell[s]$ is defined for every s.

For $i, s \in \omega$, we consider the partial sum

$$p(i,s) := \sum_{k=0}^{\ell[s]} \frac{(\beta_i \triangle \Phi^{X\oplus\alpha})(k)}{k+1}.$$

At a stage s, we define auxiliary values $i[s], b[s] \in \omega$ and $c[s] \in \{0,1\}$. Similarly to the proof of Theorem 3.1, these parameters control the flow of the construction. Moreover, at each stage s, we set $\gamma_\alpha(s) := \beta_{i[s]}(s)$. Our construction will ensure that $i[s] \leqslant b[s]$ for every s.

Stage 0. Set $i[0] = 0$, $b[0] = 1$, and $c[0] = 0$.

Stage s+1. We assume that the parameters $b[s]$, $c[s]$, and $i[s]$ are already defined. Consider the following four cases:

Case 1. If $p(i[s], s+1) \leqslant b[s]$, then do not change anything.

Case 2. If $p(i[s], s+1) > b[s]$ and $c[s] = 0$, then put $i[s+1] := 0$ and $c[s] := 1$.

Case 3. Suppose that $p(i[s], s+1) > b[s]$, $c[s] = 1$, and $i[s] < b[s]$. Define $i[s+1] := i[s] + 1$.

Case 4. Suppose that $p(i[s], s+1) > b[s]$, $c[s] = 1$, and $i[s] = b[s]$. Find the least $i_0 \leqslant b[s] + 1$ such that

$$p(i_0, s+1) = \min\{p(j, s+1) : j \leqslant b[s] + 1\}.$$

We put $i[s+1] := i_0$, $c[s+1] := 0$, and

$$b[s+1] := \max(b[s] + 1, \text{the integer part of } p(i_0, s+1) + 1).$$

This concludes the description of the construction. It is clear that the operator $\Psi\colon \alpha \mapsto \gamma_\alpha$ is Y-computable.

Verification. Suppose that a real α encodes a copy of the structure $\mathcal{A}_{i_0}$. We define:

$$N_0 := \sum_{k=0}^{\infty} \frac{(\beta_{i_0} \triangle \Gamma(\alpha))(k)}{k+1}.$$

Claim 4.1. *There exists a finite limit $b^* = \lim_s b[s]$. In addition, $b^* \geqslant i_0$.*

Proof. We distinguish two cases. First, assume that $b[s] < i_0$ for all s. Then we have $i[s] < i_0$ for every s. Furthermore, since the sequence $b[s]$ is non-decreasing, there exists $b^* = \lim_s b[s]$ with $b^* < i_0$.

Since $(\Gamma(\alpha) \not\!\!E_2\ \beta_j)$ for all $j \neq i_0$, there exists a stage s_0 such that $p(j, s_0) > i_0$ for all $j < i_0$, and $b[s] = b^*$ for all $s \geqslant s_0$. Then, our construction ensures that after the stage s_0, there will be a stage s_1 satisfying *Case 4*. This implies that $b[s_1] \geqslant b^* + 1$, which gives a contradiction. Thus, we deduce that there must exist a stage s_0' such that $b[s_0'] \geqslant i_0$.

Second, assume that $\lim_s b[s] = \infty$. This implies that there are infinitely many stages $s > s_0'$ satisfying *Case 4*. Choose a stage $s_1 > s_0'$ such that s_1 satisfies *Case 4* and $b[s_1] \geqslant N_0 + 1$. Consider the value $i^* := i[s_1]$.

- If $i^* = i_0$, then for every s, we have $p(i^*, s) < b[s_1]$. This implies that every stage $s > s_1$ satisfies *Case 1*, which gives a contradiction.
- If $i^* \neq i_0$, then find the least stage $s_2 > s_1$ with $p(i^*, s_2) > b[s_1]$. Then the stage s_2 satisfies *Case 2*, and we have $c[s_2] = 1$. Therefore, *Case 3* of the construction ensures that there is a sequence of stages

  $$s_2 = s_0'' < s_1'' < \cdots < s_{i_0}''$$

  such that $i[s_k''] = k$ for every $k \leqslant i_0$. Again, every stage $s > s_{i_0}''$ satisfies *Case 1*, which provides a contradiction.

Therefore, we proved that there is a finite limit $b^* = \lim_s b[s]$, and $b^* \geqslant i_0$. □

Now choose a stage s^* such that $b[s^*] = b^*$. There exists a stage $s_1 \geqslant s^*$ such that every $i \leqslant b^*$ satisfies the following: if $i \neq i_0$, then $p(i, s_1) > b^*$. Since after the stage s^*, there are no stages satisfying *Case 4*, it is not hard to deduce that for every $s \geqslant s_1 + b^* + 2$, we must have $i[s] = i_0$.

This implies that the real $\Psi(\alpha)$ is E_0-equivalent to β_{i_0}. For all $i \neq j$, we have $(\beta_i \not\!\!E_2\ \beta_j)$ — clearly, this implies $(\beta_i \not\!\!E_0\ \beta_j)$. Hence, we conclude that our operator Ψ provides a continuous reduction from $\mathrm{LD}(\mathfrak{K})_{/\cong}$ to E_0. In other words, the family $\mathfrak{K}$ is E_0-learnable, as desired. □

5. Characterizing the learning power of E_3

All equivalence relations considered so far (i.e., E_0, E_1, and E_2) are inseparable with respect to their learning power. In fact, by Theorem 3.1, they don't expand the boundaries of our original framework. The case of E_3, to be discussed in this section, is different. Namely, E_3 has strictly more learning power than E_0—but this fact is only witnessed by infinite families. Recall that the equivalence relation E_3 is given by

$$(\alpha\ E_3\ \beta) \Leftrightarrow (\forall m \in \omega)(\alpha^{[m]}\ E_0\ \beta^{[m]}).$$

Theorem 5.1. *A finite family $\mathfrak{K}$ is E_3-learnable if and only if $\mathfrak{K}$ is E_0-learnable. That is, E_3 and E_0 are* Learn$^{<\omega}$*-equivalent.*

Proof. One direction is again immediate: since E_0 is continuously reducible to E_3 (see Fig. 1), every E_0-learnable family is E_3-learnable.

For the other direction, let $\mathfrak{K} := (\mathcal{A}_i)_{i \leqslant n}$ be a finite E_3-learnable family and let Γ induce a continuous reduction from $\mathrm{LD}(\mathfrak{K})_{/\cong}$ to E_3. For $i \leqslant n$, choose β_i such that Γ maps all copies of $\mathcal{A}_i$ into $[\beta_i]_{E_3}$. For each pair of indices $i \neq j$, we choose a number $q(i,j)$ such that

$$\beta_i^{[q(i,j)]}\ \not{E_0}\ \beta_j^{[q(i,j)]}.$$

Then, we define a Turing operator $\Psi \colon 2^\omega \to 2^\omega$ as follows.

$$\Psi(\alpha) = \bigoplus_{i \neq j \leqslant n} \alpha^{[q(i,j)]}.$$

The operator $\Phi := \Psi \circ \Gamma$ provides a continuous reduction from $\mathrm{LD}(\mathfrak{K})_{/\cong}$ to E_0. Indeed, let α be a real which encodes a copy of $\mathcal{A}_{i_0}$. Then $(\Gamma(\alpha)\ E_3\ \beta_{i_0})$ and $(\Phi(\alpha)\ E_0\ \Psi(\beta_{i_0}))$. If $i \neq i_0$, then we have

$$(\alpha^{[q(i,i_0)]}\ E_0\ \beta_{i_0}^{[q(i,i_0)]}\ \not{E_0}\ \beta_i^{[q(i,i_0)]}) \text{ and } (\Phi(\alpha)\ \not{E_0}\ \Psi(\beta_i)).$$

Therefore, the family $\mathfrak{K}$ is E_0-learnable. □

Our next result separates E_3-learnability and E_0-learnability, thus proving that E_0 is strictly Learn$^{\omega}$-reducible to E_3

Theorem 5.2. *There exists an infinite family $\mathfrak{K} := (\mathcal{A}_i)_{i \in \omega}$ which is E_3-learnable, but not E_0-learnable.*

Proof. For the sake of exposition, first we give proof for the case when the signature of the class $\mathfrak{K}$ is allowed to be infinite. After that, we provide comments on how to build the desired $\mathfrak{K}$ as a family of directed graphs.

Consider signature $L = \{R_j : j \in \omega\} \cup \{\leqslant\}$, where R_j are unary predicates. Given a real α, we define an L-structure $\mathcal{D}(\alpha)$ as follows:

- Inside $\mathcal{D}(\alpha)$, the relations R_j, $j \in \omega$, are pairwise disjoint. We say that the set $R_j^{\mathcal{D}(\alpha)}$ is the *R_j-box* of $\mathcal{D}(\alpha)$.
- The R_j-box of $\mathcal{D}(\alpha)$ contains a linear order L_j such that

$$L_j \cong \begin{cases} \omega, & \text{if } \alpha(j) = 0, \\ \omega^*, & \text{if } \alpha(j) = 1, \end{cases}$$

where ω and ω^* are respectively the order types of the positive and negative integers. For a finite string $\sigma \in 2^{<\omega}$, let $\mathcal{A}_\sigma$ be the structure $\mathcal{D}(\sigma^\frown 10^\infty)$. Our family $\mathfrak{K}$ consists of all $\mathcal{A}_\sigma$, $\sigma \in 2^{<\omega}$. Notice that the relation $\leqslant$ in L provides an order between elements in L_j and does not provide any order between elements in different R_j-boxes.

Lemma 5.3. *The family $\mathfrak{K}$ is E_3-learnable.*

Proof. Recall that the family $\{\omega, \omega^*\}$ is learnable, as they are distinguishable by $\Sigma_2^{\inf}$ formulas [3, Theorem 3]. By employing this fact, it is not hard to build a Turing operator Φ, which acts as follows. Given a real α, it treats α as a code for the atomic diagram of a countable partial order $\mathcal{L}$. Then:

- If $\mathcal{L}$ is a copy of ω, then the output $\Phi(\alpha)$ is E_0-equivalent to 0^∞.
- If $\mathcal{L} \cong \omega^*$, then we have $(\Phi(\alpha)\ E_0\ 1^\infty)$.

For each index $j \in \omega$, we define a Turing operator Ψ_j. Given a real α, it treats α as a code of a countable L-structure $\mathcal{A}$. The output $\Psi_j(\alpha)$ encodes the partial order, which is contained inside the R_j-box of $\mathcal{A}$.

Finally, we define an operator Θ. For $\alpha \in 2^\omega$ and for $j, k \in \omega$, we set

$$\Theta(\alpha)(\langle j, k\rangle) := (\Phi \circ \Psi_j(\alpha))(k).$$

Observe the following. Let β be a real. If a real α encodes a copy of the structure $\mathcal{D}(\beta)$, then for every $j \in \omega$, we have:

- if $\beta(j) = 0$, then the j-th column $(\Theta(\alpha))^{[j]}$ is E_0-equivalent to 0^∞;
- if $\beta(j) = 1$, then $(\Theta(\alpha))^{[j]}\ E_0\ 1^\infty$.

This observation implies that the operator Θ witnesses the E_3-learnability of our family $\mathfrak{K}$. Lemma 5.3 is proved. □

Now, towards a contradiction, assume that the family $\mathfrak{K}$ is E_0-learnable. Then $\mathfrak{K}$ is **InfEx**-learnable, and by Theorem 2.4, one can choose an infinitary Σ_2 sentence θ such that $\mathcal{A}_0 \models \theta$ and for every $\sigma \neq 0$, we have $\mathcal{A}_\sigma \not\models \theta$.

Without loss of generality, one may assume that

$$\theta = \exists \bar{x} \bigwedge_{i\in I} \forall \bar{y}_i \psi_i(\bar{x}, \bar{y}_i),$$

where every ψ_i is a quantifier-free formula. Fix a tuple $\bar{c}$ from the structure $\mathcal{A}_0$ such that

$$\mathcal{A}_0 \models \bigwedge_{i\in I} \forall \bar{y}_i \psi_i(\bar{c}, \bar{y}_i).$$

Choose a natural number N such that for every $j \geqslant N$, the R_j-box of $\mathcal{A}_0$ does not contain elements from $\bar{c}$.

Consider a string $\tau := 010^N$ and the corresponding structure $\mathcal{A}_\tau = \mathcal{D}(\tau\widehat{\ }10^\infty)$. It is clear that for every $j < N$, the (contents of the) R_j-boxes inside $\mathcal{A}_0$ and $\mathcal{A}_\tau$ are isomorphic. Therefore, one can choose a tuple $\bar{d}$ inside $\mathcal{A}_\tau$ as isomorphic copies of $\bar{c}$ (with respect to the isomorphism of the R_j-boxes, $j < N$).

Claim 5.1. *The structures $(\mathcal{A}_0, \bar{c})$ and $(\mathcal{A}_\tau, \bar{d})$ satisfy the same $\exists$-sentences.*

Proof. It is sufficient to establish the following. Every quantifier-free formula $\psi(\bar{x}, \bar{y})$ satisfies

$$\mathcal{A}_0 \models \exists \bar{y} \psi(\bar{c}, \bar{y}) \ \Rightarrow\ \mathcal{A}_\tau \models \exists \bar{y} \psi(\bar{d}, \bar{y}).$$

The other direction ($\Leftarrow$) can be obtained via a similar argument.

Choose a tuple $\bar{b}$ from $\mathcal{A}_0$ such that $\mathcal{A}_0 \models \psi(\bar{c}, \bar{b})$. Suppose that $\bar{b} = b_0, b_1, \dots, b_m$. We define a new tuple $\bar{b}' = b'_0, b'_1, \dots, b'_m$ from $\mathcal{A}_\tau$ as follows:

- If b_k lies in an R_j-box, which contains elements from $\bar{c}$, then b'_k is defined as the copy of b_k with respect to the natural isomorphism of R_j-boxes, $j < N$.
- Suppose that b_k belongs to an R_j-box, which does not contain elements from $\bar{c}$. Then b'_k can be chosen as any element from the R_j-box of $\mathcal{A}_\tau$, while preserving the ordering $\leqslant$. More formally, one needs to ensure the following: if $b_k \neq b_\ell$ both belong to this R_j-box, then we have:

$$\mathcal{A}_0 \models b_k \leqslant b_\ell \ \Leftrightarrow\ \mathcal{A}_\tau \models b'_k \leqslant b'_\ell.$$

It is clear that the tuples $\bar{c}, \bar{b}$ and $\bar{d}, \bar{b}'$ satisfy the same atomic formulas. Therefore, we deduce that the structure $\mathcal{A}_\tau$ satisfies $\psi(\bar{d}, \bar{b}')$, and $\mathcal{A}_\tau \models \exists \bar{y} \psi(\bar{d}, \bar{y})$. □

Claim 5.1 implies that

$$\mathcal{A}_\tau \models \bigwedge_{i\in I} \forall \bar{y}_i \psi_i(\bar{d}, \bar{y}_i),$$

and hence, $\mathcal{A}_\tau \models \theta$, which contradicts the choice of θ. We deduce that the family $\mathfrak{K}$ is not E_0-learnable.

In order to obtain a family of directed graphs $\mathfrak{K}_{gr}$, which has the same properties as the family $\mathfrak{K}$, one can proceed as follows. Instead of distinguishing an R_j-box via the predicate R_j, one attaches to every element a of L_j of the corresponding R_j-box a cycle of length $(j+3)$. Indeed, for each $a \in L_j$ use fresh elements $c_{a,1}, c_{a,2} \dots, c_{a,j+3}$ and put the edges

(1) $(c_{a,i}, c_{a,i+1})$ for $i \leqslant j+2$,
(2) $(c_{a,j+3}, c_{a,1})$,

(3) and $(a, c_{a,1})$.

After that, the proof for the family $\mathfrak{K}_{gr}$ is essentially the same as the one provided above. Theorem 5.2 is proved. □

5.1. A syntactic characterization of E_3-learnability

As aforementioned, in the previous work we obtained a full syntactic characterization of which families of structures are learnable, by means of Σ_2^{inf} formulas (see Theorem 2.4). The next theorem offers an analogous characterization for E_3-learning.

Theorem 5.4. *Let $\mathfrak{K} := (\mathcal{A}_i)_{i\in\omega}$ be a countable family. The family $\mathfrak{K}$ is E_3-learnable if and only if there exists a countable family of Σ_2^{inf} sentences Θ with the following properties:*

(a) *if θ is a formula from Θ, then there is a formula $\psi \in \Theta$ such that for every $\mathcal{A} \in \mathfrak{K}$,*

$$\mathcal{A} \models \theta \;\Leftrightarrow\; \mathcal{A} \models \neg\psi;$$

(b) *if $\mathcal{A} \not\cong \mathcal{B}$ are structures from $\mathfrak{K}$, then there is a sentence $\theta \in \Theta$ such that*

$$\mathcal{A} \models \theta \text{ and } \mathcal{B} \models \neg\theta.$$

Proof. The proof of the theorem is inspired by ideas from [3]. In particular, we will adopt the technology of *tc*-embeddings and the Relativized Pullback Theorem reminded in the preliminaries.

Consider a signature $L_{st} = \{\leqslant\} \cup \{P_i : i \in \omega\}$, where P_i are unary predicates. For an index $i \in \omega$, an L-structure $\mathcal{S}_i$ satisfies the following properties:

- Inside $\mathcal{S}_i$, the relations P_j are pairwise disjoint. In addition, if $x \in P_j$ and $y \in P_k$ for some $j \neq k$, then x and y are $\leqslant$-incomparable. Let η be the order type of the rational numbers.
- The predicate P_i contains an isomorphic copy of $1 + \eta$.
- Every P_j, for $j \neq i$, contains a copy of η.

The class $\mathfrak{K}_{st}$ consists of all structures $\mathcal{S}_i$ with $i \in \omega$.

In [3], it is shown $\mathfrak{K}_{st}$ is an archetypical E_0-learnable family, in the sense that a countable family $\mathfrak{C}$ is learnable if and only if there is a continuous embedding from the class $\mathfrak{C}$ into $\mathfrak{K}_{st}$.

For dealing with E_3-learnability, we have to introduce a new, and more complicated, class $\mathfrak{C}_{st}$. But the informal idea behind $\mathfrak{C}_{st}$ is pretty simple: roughly speaking, this class contains all countable disjoint sums of the structures from $\mathfrak{K}_{st}$.

Consider a new signature $L_1 = L_{st} \cup \{Q_k : k \in \omega\}$, where Q_k are unary predicates. The class $\mathfrak{C}_{st}$ contains all L-structures $\mathcal{M}$, which satisfy the following properties:

- Their relations Q_k, $k \in \omega$, are pairwise disjoint. We say that the L_{st}-substructure with domain $\mathcal{M} \upharpoonright Q_k$ is the *Q_k-box* of $\mathcal{M}$.
- Every Q_k-box of $\mathcal{M}$ is isomorphic to a structure from the class $\mathfrak{K}_{st}$.

Note that our class $\mathfrak{C}_{st}$ has cardinality $2^{\aleph_0}$.

Lemma 5.5. *The class $\mathfrak{C}_{st}$ has a computable family of Σ_2^{inf} sentences Θ, which satisfies properties (a) and (b) from the formulation of Theorem 5.4.*

Proof. The desired family Θ contains the following Σ_2^{inf} sentences:

(1) For each i and j, we add a finitary Σ_2 sentence $\theta_{i,j}$, which states the following: "the P_j-predicate inside the Q_i-box has a $\leqslant$-least element".
(2) For each i and j, we add a Σ_2^{inf} sentence $\psi_{i,j}$, which is equivalent to the following formula:

$$\bigvee\!\!\!\bigvee_{k\neq j} \theta_{i,k}.$$

In other words, there is some $k \neq j$ such that the P_k-predicate inside the Q_i-box possesses the least element.

Let $\mathcal{M}$ be an arbitrary structure from $\mathfrak{C}_{st}$. Since the Q_i-box of $\mathcal{M}$ is a structure from $\mathfrak{K}_{st}$, it is not hard to show that

$$\mathcal{M} \models \theta_{i,j} \;\Leftrightarrow\; \mathcal{M} \models \neg\psi_{i,j}.$$

Hence, we deduce that the class $\mathfrak{C}_{st}$ satisfies property (a) of Theorem 5.4.

Suppose that $\mathcal{M} \not\cong \mathcal{N}$ are structures from $\mathfrak{C}_{st}$. Then there exist indices i and j such that for the structures $\mathcal{M}$ and $\mathcal{N}$, their P_j-predicates inside Q_i-boxes are not isomorphic. Without loss of generality, one may assume that in this P_j-place, $\mathcal{M}$ has order-type $1+\eta$, and $\mathcal{N}$ has order-type η. Then, it is clear that

$$\mathcal{M} \models \theta_{i,j} \& \neg\psi_{i,j} \text{ and } \mathcal{N} \models \neg\theta_{i,j} \& \psi_{i,j}.$$

Therefore, $\mathfrak{C}_{st}$ satisfies property (b) of the theorem. □

The rest of the proof for the direction ($\Rightarrow$) is devoted to building a continuous embedding from the given class $\mathfrak{K}$ to $\mathfrak{C}_{st}$. This embedding allows us to apply the Relativized Pullback Theorem (Theorem 2.7) for finishing our argument.

Consider a countable sequence of reals $\vec{\gamma} = (\gamma_i)_{i\in\omega}$. We define an auxiliary continuous operator $\Psi_{\vec{\gamma}}$ as follows. Given a real α, our operator $\Psi_{\vec{\gamma}}$ produces a new real δ_α, which encodes the atomic diagram of an L-structure $\mathcal{S}(\alpha)$.

We always assume that inside $\mathcal{S}(\alpha)$:

- all predicates P_i are disjoint;
- every predicate P_i contains at least one element;
- the domain of $\mathcal{S}(\alpha)$ equals ω.

Construction. The construction of $\mathcal{S}(\alpha)$ proceeds in stages. At a stage s, for each $i \in \omega$, we define the following auxiliary value:

$$v(i,s) = \begin{cases} \min\{t \leqslant s : (\forall x)[t \leqslant x \leqslant s \to \alpha(x) = \beta_i(x)]\}, & \text{if } \alpha(s) = \beta_i(s), \\ \infty, & \text{otherwise.} \end{cases}$$

We also define two parameters $p(s)$ and $b(s)$. Roughly speaking, at a stage s, our current "guess" is that the input real α is E_0-equivalent to $\beta_{p(s)}$, where $p(s) \leqslant b(s) \leqslant s$.

Stage 0. Put $p(0) = 0$ and $b(s) = 0$.

Stage $s+1$. Consider the following two cases:

Case 1. Suppose that there is an index $i \leqslant s+1$ such that $\alpha(s+1) = \beta_i(s+1)$.

If $v(p(s), s+1) \neq \infty$, then set $i_0 := p(s)$. Otherwise, i_0 is defined as follows.

- If $p(s) < b(s)$, then $p(s+1) := p(s)+1$ and $i_0 := p(s)+1$;
- If $p(s) = b(s)$, then we define i_0 as the least index such that $i_0 \leqslant s+1$ and

  $$v(i_0, s+1) = \min\{v(j, s+1) : j \leqslant s+1\}.$$

  We set $b(s+1) := s+1$ and $p(s+1) := i_0$.

Suppose that the relation P_{i_0} (at this particular moment) contains the following linear order: $a_0 < a_1 < \cdots < a_k$. We choose fresh elements $b_0, b_1, \ldots, b_k$, add them into P_{i_0}, and set:

$$a_0 < b_0 < a_1 < b_1 < \cdots < a_k < b_k.$$

Consider an index $j \neq i_0$, and suppose that the relation P_j contains the ordering $c_0 < c_1 < \cdots < c_\ell$. Choose fresh elements $d_{-1}, d_0, d_1, \ldots, d_\ell$, put them into P_j, and define:

$$d_{-1} < c_0 < d_0 < c_1 < d_1 < \cdots < c_\ell < d_\ell.$$

Case 2. If $\alpha(s) \neq \beta_i(s)$ for all $i \leqslant s+1$, then for every $j \in \omega$, the relation P_j is arranged in the same way as described in *Case 1*.

This concludes the description of the operator $\Psi_{\vec{\gamma}}$.

Verification. Similarly to the previous proofs, it is not hard to verify the following properties of $\Psi_{\vec{\gamma}}$:

(1) The operator $\Psi_{\vec{\gamma}}$ is $\left(\bigoplus_{i\in\omega} \gamma_i\right)$-computable.
(2) If $(\alpha \; E_0 \; \gamma_i)$ for some $i \in \omega$, then the structure $\mathcal{S}(\alpha)$ is isomorphic to $\mathcal{S}_i$.

Now, let Γ be a continuous operator which induces a reduction from $\mathrm{LD}(\mathfrak{K})_{/\cong}$ to E_3. For a structure $\mathcal{A}_i$ from $\mathfrak{K}$, fix a real β_i such that Γ maps all copies of $\mathcal{A}_i$ into the class $[\beta_i]_{E_3}$.

We define a continuous operator Ξ as follows. Let α be a real.

(1) First, we produce the real $\Gamma(\alpha)$.
(2) Second, for each $j \in \omega$, we consider the sequence $\vec{\beta}^{[j]} := (\beta_i^{[j]})_{i\in\omega}$. We compute the reals

$$\delta_{\alpha,j} := \Psi_{\vec{\beta}^{[j]}}((\Gamma(\alpha))^{[j]}).$$

(3) Finally, by using the reals $\delta_{\alpha,j}$, $j \in \omega$, we recover a new real δ, which encodes the atomic diagram of an L_1-structure $\mathcal{M}$. This structure $\mathcal{M}$ is defined as follows. For each j, the Q_j-box of $\mathcal{M}$ is an isomorphic copy of the L_{st}-structure encoded by $\delta_{\alpha,j}$, and this copy has domain $\{\langle j,k\rangle : k \in \omega\}$. We set $\Xi(\alpha) := \delta$.

It is straightforward to establish the following: the operator Ξ is a continuous embedding from the class $\mathfrak{K}$ into a countable subclass of $\mathfrak{C}_{st}$. So, by applying the Relativized Pullback Theorem (Theorem 2.7) to the continuous embedding Ξ, we recover a countable family of formulas with the desired properties. Indeed, the following holds:

- by Lemma 5.5, $\mathfrak{C}_{st}$ has a family of $\Sigma_2^{\inf}$ sentences Θ, which satisfies properties (a) and (b) of Theorem 5.4;
- by Lemma 2.1, Φ is equivalent to a Turing X-operator, for a suitable oracle X.

Hence, we can apply Theorem 2.7, and deduce that $\mathfrak{K}$ has a family $\Theta^\star$ of $\Sigma_2^{\inf}$ sentences Θ, which satisfies (a) and (b) of Theorem 5.4, as desired.

($\Leftarrow$). This direction essentially follows from previous results. Assume that $\mathfrak{K} := (\mathcal{A}_i)_{i\in\omega}$ has a family Θ of $\Sigma_2^{\inf}$ formulas which satisfies the properties (a) and (b) of the theorem. Then it's easy to check that the formulas of Θ can be arranged to satisfy the following lemma:

Lemma 5.6. *There is a collection of pairs of formulas $(\rho_{i_0}, \rho_{i_1})_{i\in\omega}$ so that, for all structures $\mathcal{A}$ and $\mathcal{B}$ from $\mathfrak{K}$,*

(1) $\bigcup_{i\in\omega}\{\rho_{i_0}, \rho_{i_1}\} = \Theta$;
(2) for all $i \in \omega$, $\mathcal{A}$ satisfies exactly one formula between ρ_{i_0} and ρ_{i_1};
(3) if $\mathcal{A} \ncong \mathcal{B}$, then, for some $j \in \omega$,

$$\mathcal{A} \models \rho_{j_0} \Leftrightarrow \mathcal{B} \models \rho_{j_1}.$$

The next lemma combines (a limited case of) Theorem 2.4 with Theorem 3.1.

Lemma 5.7. *For all i, there is a continuous operator $\Gamma_i : 2^\omega \to 2^\omega$ such that, for all structures $\mathcal{S} \in \mathrm{LD}(\mathfrak{K})$,*

- *if $\mathcal{S} \models \rho_{i_0}$, then $\Gamma_i(\mathcal{S})\ E_0\ 0^\infty$;*
- *if $\mathcal{S} \models \rho_{i_1}$, then $\Gamma_i(\mathcal{S})\ E_0\ 1^\infty$.*

Proof. The proof is similar to that of the direction (2) $\Rightarrow$ (1) of [3, Theorem 3]. Let $i \in \omega$. For $k \in \{0, 1\}$, without loss of generality assume that

$$\rho_{i_k} := (\exists \bar{x}) \bigwedge\!\!\!\bigwedge_{j\in J_{i_k}} \forall \bar{y} \varphi_{i_k,j}(\bar{x}, \bar{y}).$$

For a finite structure $\mathcal{F}$, say that φ_{i_k} is *$\mathcal{F}$-compatible* via tuple $\bar{a}$ if within the domain of $\mathcal{F}$ there is no pair $(j, \bar{b})$ with $j \in J_{i_k}$ such that $\mathcal{F} \models \neg\varphi_{i_k,j}(\bar{a}, \bar{b})$.

Construction. Now, let α be a real. Denote by $\mathcal{F}_{\alpha\restriction_s}$ the finite structure (in the signature of $\mathfrak{K}$) encoded by the initial segment $\alpha \restriction_s$ of α. The continuous operator Γ_i is defined by stages.

Stage 0. Let $\Gamma_i(\alpha)(0) := 0$ and $\Gamma_i(\alpha)(1) := 1$.

Stage s+1. At this stage, we define $\Gamma_i(\alpha)(2s)$ and $\Gamma_i(\alpha)(2s + 1)$. To this end, we distinguish three cases:

(1) There is a tuple $\bar{c}$ so that φ_{i_0} is $\mathcal{F}_{\alpha\restriction_s}$-compatible via $\bar{c}$, and φ_{i_1} is not $\mathcal{F}_{\alpha\restriction_s}$-compatible for all tuples $< \bar{c}$. If so, let $\Gamma_i(\alpha)(2s) = \Gamma_i(\alpha)(2s + 1) := 0$;
(2) There is a tuple $\bar{c}$ so that φ_{i_1} is $\mathcal{F}_{\alpha\restriction_s}$-compatible via $\bar{c}$, and φ_{i_0} is not $\mathcal{F}_{\alpha\restriction_s}$-compatible for all tuples $\leqslant \bar{c}$. If so, let $\Gamma_i(\alpha)(2s) = \Gamma_i(\alpha)(2s + 1) := 1$;
(3) If neither of the above cases hold, then let $\Gamma_i(\alpha)(2s) := 0$ and $\Gamma_i(\alpha)(2s + 1) := 1$.

Verification. The continuity of Γ_i immediately follows from the construction. Next, suppose that $\beta \in 2^\omega$ encodes a copy of a structure $\mathcal{S} \in \mathfrak{K}$. By Lemma 5.6, $\mathcal{S}$ satisfies exactly one formula between φ_{i_0} and φ_{i_1}; without loss of generality, assume that $\mathcal{S} \models \varphi_{i_1}$. This means that there is a tuple $\bar{c}$ and a stage t_0 so that φ_{i_1} is $\mathcal{F}_{\beta\upharpoonright t}$-compatible via $\bar{c}$, for all $t \geqslant t_0$. On the other hand, since $\mathcal{S} \not\models \varphi_{i_0}$, it must be the case that for all tuples $\bar{d}$ (and, in particular, all tuples $\leqslant \bar{c}$), there must be a stage t_1 so that, for all $t \geqslant t_1$, φ_{i_0} is not $\mathcal{F}_{\beta\upharpoonright t}$-compatible. So, for all sufficiently large x, $\Gamma_i(\beta)(x)$ is defined by performing action (2) above. Thus, $\Gamma_i(\beta)$ is E_0-equivalent to 1^∞, as desired. □

We can now construct a continuous reduction from $\mathrm{LD}(\mathfrak{K})_{/\cong}$ to E_3 by merging the operators Γ_i's as follows:

$$\Gamma(\alpha)(\langle i, x\rangle) := \Gamma_i(\alpha)(x).$$

It is an easy consequence of Lemma 5.7 that, if β_0 and β_1 are copies of the same structure $\mathcal{S} \in \mathfrak{K}$, then $\Gamma(\beta_0)\ E_3\ \Gamma(\beta_1)$. To deduce that Γ is the desired reduction, suppose that β_0 and β_1 are copies of nonisomorphic structures $\mathcal{A}$ and $\mathcal{B}$ from $\mathfrak{K}$. By Lemma 5.6, there are $j \in \omega$ and $k \in \{0, 1\}$ so that $\mathcal{A} \models \varphi_{i_k}$ and $\mathcal{B} \models \varphi_{i_{1-k}}$. But then, by Lemma 5.7, it follows that $\Gamma(\beta_0)$ and $\Gamma(\beta_1)$ differ on the jth column, that is,

$$\Gamma(\beta_0)^{[j]}\ E_0\ k^\infty \text{ but } \Gamma(\beta_1)^{[j]}\ E_0\ (1-k)^\infty.$$

Thus, $\Gamma(\beta_0) \not{E_3}\, \Gamma(\beta_1)$.

This concludes the proof of Theorem 5.4. □

6. Learning with the help of Z_0 and E_{set}

We conclude our examination of the learning power of combinatorial Borel equivalence relations by briefly focusing on two further examples: Z_0 and E_{set}. Here, the main goal is to finally individuate a Borel equivalence relation which is able to learn a finite family beyond the reach of our original framework.

6.1. Z_0-learning

Before proceeding to a new result, we give a simple useful fact. Let α and β be reals, and let $s \in \omega$. We use the following notation:

$$dn(\alpha, \beta; s) = \frac{\mathrm{card}(\{i \leqslant s : \alpha \triangle \beta(i) = 1\})}{s+1}.$$

Recall that the equivalence relation Z_0 is given by

$$(\alpha\ Z_0\ \beta) \Leftrightarrow \lim_{k\to\infty} dn(\alpha, \beta; k) = 0$$

Lemma 6.1. *Suppose that $(\alpha\ Z_0\ \beta)$. Then*

$$\limsup_s dn(\alpha, \gamma; s) = \limsup_s dn(\beta, \gamma; s).$$

Proof. Let $r := \limsup_s dn(\beta, \gamma; s)$. It is sufficient to show that for any ε such that $0 < \varepsilon < r$, we have

$$\limsup_s dn(\alpha, \gamma; s) \geqslant r - \varepsilon.$$

Define $q := r - \varepsilon$.

Let N be a non-zero natural number. Fix a number s_0 such that $dn(\alpha, \beta; s) < \frac{q}{N}$ for all $s \geqslant s_0$.

There exists a sequence $(s_j)_{j\in\omega}$, where $s_0 < s_1 < s_2 < \dots$, such that $dn(\beta, \gamma; s_j) > q$ for all j.

Note that every s satisfies the following:

$$\mathrm{card}(\{i \leqslant s : \alpha \triangle \gamma(i) = 1\}) \geqslant \mathrm{card}(\{i \leqslant s : \beta \triangle \gamma(i) = 1\}) - \mathrm{card}(\{i \leqslant s : \beta \triangle \alpha(i) = 1\}).$$

Hence, we have:

$$dn(\alpha, \gamma; s_j) \geqslant dn(\beta, \gamma; s_j) - dn(\alpha, \beta; s_j) > q - \frac{q}{N} = q \cdot \frac{N-1}{N}.$$

Since N was chosen as an arbitrary natural number, we deduce that for any $\delta > 0$, we have $\limsup_s dn(\alpha, \gamma; s) > q - \delta$. This implies

$$\limsup_s dn(\alpha, \gamma; s) \geqslant q.$$

Lemma 6.1 is proved. □

We show that learnability by finite families cannot distinguish between E_0 and Z_0:

Theorem 6.2. *A finite family $\mathfrak{K}$ is Z_0-learnable if and only if $\mathfrak{K}$ is E_0-learnable. That is, Z_0 and E_0 are* Learn$^{<\omega}$*-equivalent.*

Proof. Since E_0 is computably reducible to Z_0 (see Fig. 1), every E_0-learnable family is also Z_0-learnable.

Suppose that $\mathfrak{K} = (\mathcal{A}_i)_{i\in\omega}$ is a Z_0-learnable family. Let Γ be an operator which induces a continuous reduction from LD($\mathfrak{K}$) to Z_0. For $i \leqslant n$, we fix β_i such that Γ maps all copies of $\mathcal{A}_i$ into $[\beta_i]_{Z_0}$. Notice that the reals β_i are pairwise not E_0-equivalent.

We fix a positive rational q_0 such that

$$q_0 < \min\{\limsup_s dn(\beta_i, \beta_j; s) : i < j \leqslant n\}.$$

There exist an oracle X and a Turing operator Φ such that $\Gamma(\alpha) = \Phi^{X\oplus\alpha}$ for all $\alpha \in 2^\omega$.

We define an $(X \oplus \bigoplus_{i\leqslant n} \beta_i)$-computable operator Ψ. Let α be a real. For $s \in \omega$, by $\ell[s]$ we denote the greatest number such that for every $x \leqslant \ell[s]$, the value $\Phi^{(X\oplus\alpha)\upharpoonright s}(x)[s]$ is defined.

At a stage s, for each $i \leqslant n$, we compute the value

$$m_i[s] := \mathrm{card}(\{t \leqslant \ell[s] : dn(\Phi^{X\oplus\alpha}, \beta_i; t) > q_0\}).$$

We find the least $j \leqslant n$ such that

$$m_j[s] = \min\{m_i[s] : i \leqslant n\},$$

and set $\Psi(\alpha)(s) := \beta_j(s)$. This concludes the description of the operator Ψ.

Suppose that a real α encodes a copy of a structure $\mathcal{A}_{i_0}$ for some $i_0 \leqslant n$. Then by Lemma 6.1, we have:

$$\lim_s dn(\Gamma(\alpha), \beta_{i_0}; s) = 0 \text{ and } \limsup_s dn(\Gamma(\alpha), \beta_i; s) > q_0$$

for all $i \neq i_0$.

Choose a number t_0 such that for all $t \geqslant t_0$, we have $dn(\Phi^{X\oplus\alpha}, \beta_{i_0}; t) \leqslant q_0$. Fix a stage s_0 with $t_0 \leqslant \ell[s_0]$. Then for all $s \geqslant s_0$, we have $m_{i_0}[s] = m_{i_0}[s_0]$.

On the other hand, it is not hard to show that for every $i \neq i_0$, we have $\lim_s m_i[s] = \infty$. This implies that the real $\Psi(\alpha)$ is E_0-equivalent to β_{i_0}.

We deduce that the operator Ψ provides a continuous reduction from LD($\mathfrak{K}$) to E_0. Theorem 6.2 is proved. □

It is known that E_3 is continuously reducible to Z_0 (see Fig. 1). So, E_3 is Learn$^\omega$ reducible to Z_0. The next question, which is left open, asks if the converse hold.

Question 1. *Is there a countable Z_0-learnable family, which is not E_{set}-learnable?*

6.2. E_{set}-learning

A distinctive feature of our learning framework is that there are finite families of structures which are not learnable. This is the case, most notably, of the pair of linear orders $\{\omega, \zeta\}$, where ζ is the order type of the integers. Such a feature is in sharp contrast with classical paradigms, since, e.g., any finite collection of recursive functions is **InfEx**-learnable. Yet, we have observed that all Borel equivalence relations so far considered are Learn$^{<\omega}$-equivalent to E_0. So, a question comes naturally: how high in the Borel hierarchy one needs to climb to reach an equivalence relation E which is able to learn a nonlearnable finite family? The next proposition shows that E_{set} suffices.

Proposition 6.3. *The family $\{\omega, \zeta\}$ is E_{set}-learnable.*

Proof. Given a real α, which encodes a linear order with infinite domain $A \subseteq \omega$, one can effectively recover a list $(a_i)_{i\in\omega}$, which enumerates the set A without repetitions. In addition, the recovery procedure is uniform in α.

We define a Turing operator Ψ. For a real α, the output $\Psi(\alpha)$ is constructed as follows. For all i and s, we put

$$\Psi(\alpha)(\langle 2i, s\rangle) := \begin{cases} 0, & \text{if } s < i, \\ 1, & \text{if } s \geqslant i. \end{cases}$$

Let B_s be the finite linear order, which is encoded by the finite string $\alpha \upharpoonright s$ (note that B_s can be empty). For $i \in \omega$, consider the element a_i (from the list discussed above). If $a_i \notin B_s$ or a_i is the $\leqslant_{B_s}$-least element inside B_s, then we set $\Psi(\alpha)(\langle 2i+1, s\rangle) := 0$. Otherwise, set $\Psi(\alpha)(\langle 2i+1, s\rangle) := 1$.

Suppose that a real α encodes a copy of $\mathcal{A} \in \{\omega, \zeta\}$. If $\mathcal{A}$ is isomorphic to ζ, then it is clear that

$$\{(\Psi(\alpha))^{[m]} : m \in \omega\} = \{0^i 1^\infty : i \in \omega\}.$$

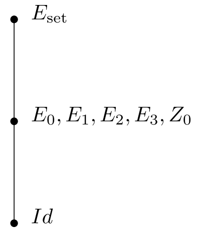

Fig. 2. Reductions up to Learn$^{<\omega}$-reducibility.

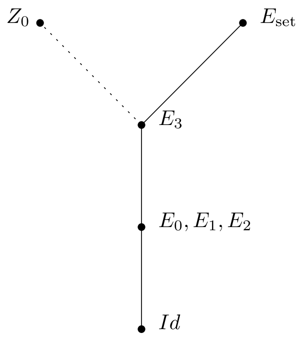

Fig. 3. Reductions up to Learn$^{\omega}$-reducibility.

If $\mathcal{A} \cong \omega$, then there is an element a_{i_0}, which is $\leqslant_{\mathcal{A}}$-least. This implies

$$\{(\Psi(\alpha))^{[m]} : m \in \omega\} = \{0^i 1^\infty : i \in \omega\} \cup \{0^\infty\}.$$

Therefore, we deduce that the family $\{\omega, \zeta\}$ is E_{set}-learnable. $\square$

7. Conclusions

The investigation conducted in this paper has been fueled by the discovery of a connection between algorithmic learning theory and descriptive set theory. Namely, we proved that the task of learning a given family of algebraic structures (up to isomorphism) is equivalent to the task of constructing a suitable continuous reduction to E_0. Then, we carefully analyzed the learning power of a number of well-known benchmark Borel equivalence relations. Our results are collected in Figs. 2 and 3.

We wish to conclude by mentioning three research directions that originate from the above results and which look promising:

(1) First, it seems natural to discuss the learning power of other Borel equivalence relations. There is a wide choice—even if one restricts to a small fragment of the Borel hierarchy, such as the $\mathbf{\Pi}^0_3$ equivalence relations (see [14]);
(2) Secondly, it would be nice to obtain learning theoretic or purely syntactic characterizations for E_{set}- and Z_0-learnabilities of countable families, along the lines of Theorems 5.4 and 2.4.
(3) Thirdly, observe that our original framework was inherently limited to the countable case, since the learner had to provide a conjecture (i.e., a finite object) for each isomorphism type of the observed family. But now the concept of E-learnability can be naturally applied to families of continuum size. This offers a new research opportunity, probably worth considering.
(4) Finally, as suggested by an anonymous reviewer, in a future work we plan to locate analogues of classic learning criteria (such as partial learning or non-U-shape learning) in the learning hierarchy introduced in this paper. See [2] for a similar discussion about **Fin**-learning of structures and learning with a bounded number of mistakes.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Chris J. Ash, Julia F. Knight, Computable Structures and the Hyperarithmetical Hierarchy, Studies in Logic and the Foundations of Mathematics, vol. 144, Elsevier Science B.V., Amsterdam, 2000.

[2] Nikolay Bazhenov, Vittorio Cipriani, Luca San Mauro, Calculating the mind change complexity of learning algebraic structures, in: Ulrich Berger, Johanna N.Y. Franklin, Florin Manea, Arno Pauly (Eds.), Revolutions and Revelations in Computability, Springer International Publishing, Cham, 2022, pp. 1–12.
[3] Nikolay Bazhenov, Ekaterina Fokina, Luca San Mauro, Learning families of algebraic structures from informant, Inf. Comput. 275 (2020) 104590.
[4] Nikolay Bazhenov, Benoit Monin, Luca San Mauro, Rafael Zamora, On the computational content of the theory of Borel equivalence relations, Oberwolfach Preprint OWP-2021-06 (2021).
[5] Nikolay Bazhenov, Luca San Mauro, On the Turing complexity of learning finite families of algebraic structures, J. Log. Comput. 31 (7) (2021) 1891–1900.
[6] W. Calvert, D. Cummins, J.F. Knight, S. Miller, Comparing classes of finite structures, Algebra Log. 43 (6) (2004) 374–392.
[7] Riccardo Camerlo, Su Gao, The completeness of the isomorphism relation for countable Boolean algebras, Trans. Am. Math. Soc. 353 (2) (2001) 491–518.
[8] Samuel Coskey, Joel David Hamkins, Russell Miller, The hierarchy of equivalence relations on the natural numbers under computable reducibility, Computability 1 (1) (2012) 15–38.
[9] Matthew de Brecht, Akihiro Yamamoto, Mind change complexity of inferring unbounded unions of restricted pattern languages from positive data, Theor. Comput. Sci. 411 (7–9) (2010) 976–985.
[10] Matthew de Brecht, Akihiro Yamamoto, Topological properties of concept spaces (full version), Inf. Comput. 208 (4) (2010) 327–340.
[11] Yu.L. Ershov, S.S. Goncharov, Constructive Models, Kluwer Academic/Plenum Publishers, New York, 2000.
[12] Ekaterina Fokina, Timo Kötzing, Luca San Mauro, Limit learning equivalence structures, in: Aurélien Garivier, Satyen Kale (Eds.), Proceedings of the 30th International Conference on Algorithmic Learning Theory, Chicago, Illinois, 22–24 Mar 2019, in: Proceedings of Machine Learning Research, vol. 98, 2019, pp. 383–403, PMLR.
[13] Harvey Friedman, Lee Stanley, A Borel reducibility theory for classes of countable structures, J. Symb. Log. 54 (3) (1989) 894–914.
[14] Su Gao, Invariant Descriptive Set Theory, CRC Press, Boca Raton, FL, 2009.
[15] Clark Glymour, Inductive inference in the limit, Erkenntnis 22 (1985) 23–31.
[16] E. Mark Gold, Language identification in the limit, Inf. Control 10 (5) (1967) 447–474.
[17] Ziyuan Gao, Frank Stephan, Guohua Wu, Akihiro Yamamoto, Learning families of closed sets in matroids, in: Michael J. Dinneen, Bakhadyr Khoussainov, André Nies (Eds.), Computation, Physics and Beyond - International Workshop on Theoretical Computer Science, WTCS 2012, in: Lecture Notes in Computer Science, vol. 7160, Springer, Berlin, 2012, pp. 120–139.
[18] Greg Hjorth, Borel equivalence relations, in: Handbook of Set Theory, Springer, 2010, pp. 297–332.
[19] L.A. Harrington, A.S. Kechris, A. Louveau, A Glimm-Effros dichotomy for Borel equivalence relations, J. Am. Math. Soc. 3 (4) (1990) 903–928.
[20] Valentina S. Harizanov, Frank Stephan, On the learnability of vector spaces, J. Comput. Syst. Sci. 73 (1) (2007) 109–122.
[21] Vladimir Grigor'evich Kanoveĭ, Borel Equivalence Relations: Structure and Classification, vol. 44, American Mathematical Soc., 2008.
[22] Julia F. Knight, Sara Miller, Michael Vanden Boom, Turing computable embeddings, J. Symb. Log. 72 (3) (2007) 901–918.
[23] Steffen Lange, Thomas Zeugmann, Sandra Zilles, Learning indexed families of recursive languages from positive data: a survey, Theor. Comput. Sci. 397 (1–3) (2008) 194–232.
[24] David Marker, Lectures on Infinitary Model Theory, Lecture Notes in Logic., vol. 46, Cambridge University Press, Cambridge, 2016.
[25] Alan H. Mekler, Stability of nilpotent groups of class 2 and prime exponent, J. Symb. Log. 46 (4) (1981) 781–788.
[26] Russell Miller, Computable reducibility for Cantor space, in: Structure and Randomness in Computability and Set Theory, World Scientific, Singapore, 2021, pp. 155–196.
[27] Eric Martin, Daniel Osherson, Elements of Scientific Inquiry, MIT Press, 1998.
[28] Wolfgang Merkle, Frank Stephan, Trees and learning, J. Comput. Syst. Sci. 68 (1) (2004) 134–156.
[29] Eric Martin, Arun Sharma, Frank Stephan, Unifying logic, topology and learning in parametric logic, Theor. Comput. Sci. 350 (1) (2006) 103–124.
[30] Hilary Putnam, Trial and error predicates and the solution to a problem of Mostowski, J. Symb. Log. 30 (1) (1965) 49–57.
[31] Robert I. Soare, Turing Computability. Theory and Applications, Springer, Berlin, 2016.
[32] Frank Stephan, Yuri Ventsov, Learning algebraic structures from text, Theor. Comput. Sci. 268 (2) (2001) 221–273.
[33] Thomas Zeugmann, Sandra Zilles, Learning recursive functions: a survey, Theor. Comput. Sci. 397 (1–3) (2008) 4–56.