LEARNING ALGEBRAIC STRUCTURES
WITH THE HELP OF BOREL EQUIVALENCE RELATIONS

NIKOLAY BAZHENOV, VITTORIO CIPRIANI, AND LUCA SAN MAURO

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 $E_{\text{set}}$).

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 [Gol67] and Putnam [Put65] in the 1960s and it comprehends 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 accessing 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., [ZZ08, LZZ08]), 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 [HS07, SV01, MS04, GSWY12].

In previous works [FKSM19, BFSM20], 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 structure. Then, the learning is successful if the conjectures eventually stabilize to a correct guess. See also [Gly85, MO98] 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 [BFSM20], 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 [BSM21], 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 descriptive set theory. This is readily justified. Indeed, 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., [Gao09, Kan08, Hjo10]). 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 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 [FSS9, Mek81, CG01]).

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 \( \mathcal{R} \) of algebraic structures is, in a sense, rooted in the complexity of the isomorphism relation associated with \( \mathcal{R} \). 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 [BSM21]);

(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 $\mathcal{R}$ is learnable if and only the isomorphism relation associated with $\mathcal{R}$ 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 [HKL90]), 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 $\mathcal{R}$ is $E$-learnable, for a Borel equivalence relation $E$, if there is a continuous reduction from the isomorphism relation associated with $\mathcal{R}$ 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 4.2) 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 [Soa16]. In particular, by $(\varphi_e)_{e \in \omega}$, $(W_e)_{e \in \omega}$, and $(\Phi^X_e)_{e \in \omega}$ 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., \( \alpha, \beta \)). Functions on reals are denoted by uppercase Greek letters (e.g., \( \Gamma, \Psi \)). The \( m \)-th binary digit of a real \( \alpha \) is denoted by \( \alpha(m) \). By \( \alpha^{[m]} \), we denote the real representing the \( m \)-th column \( \alpha(\langle m, \cdot \rangle) \) of \( \alpha \). The symmetric difference \( \alpha \triangle \beta \) of two reals is defined in the usual way:

\[
(\alpha \triangle \beta)(i) = 1 \iff \alpha(i) \neq \beta(i).
\]

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

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

These cylinders form a basis of \( 2^\omega \). A subset \( X \) of Cantor space is Borel, if it can be constructed by 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 \( \Gamma \) of any open set is open. A Turing operator \( \Phi \) can be naturally regarded as a partial function \( \Phi : 2^\omega \to 2^\omega \), where \( \Phi(\alpha) \) is defined if and only \( \Phi \) 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 \( \alpha \), and a number \( s \in \omega \), the notation \( \Gamma(\alpha)(s) \) refers to the \( s \)-th 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 \( \Phi \) so that

\[
\Gamma(\alpha) = \Phi_{X^\alpha}
\]

for every real \( \alpha \).

**Proof.** The continuity of \( \Gamma \) 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 \( \sigma \leq \tau \), then \( h(\sigma) \leq h(\tau) \);
2. for all \( \alpha \in 2^\omega \), \( \Gamma(\alpha) = \bigcup_{\sigma \leq \alpha} h(\sigma) \).

Hence, it is straightforward to define the desired Turing operator by choosing an oracle \( X \) that computes \( h \). \[ \square \]

In this paper, it is convenient to call every Turing operator of the form \( \Phi_{X^\alpha} \) a Turing \( X \)-operator. Intuitively, Turing \( X \)-operators can be identified with a Turing machine which has three tapes: the input tape (on which
the machine is allowed to work), the output tape, and the oracle tape, where
the oracle tape always contains the characteristic function of $X$.

**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 $2^{\omega}$. A *reduction* from $E$ to $F$ is a function
$\Gamma : 2^{\omega} \rightarrow 2^{\omega}$ such that
\[
\alpha E \beta \iff \Gamma(\alpha) F \Gamma(\beta),
\]
for all reals $\alpha, \beta$. It is common to impose definability requirements on the
functions inducing a reduction. *Borel reductions*, introduced in \[FS89\], 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 con-
tinuous function $\Gamma : 2^{\omega} \rightarrow 2^{\omega}$ which reduces $E$ to $F$.

The following combinatorial equivalence relations on reals are widely con-
sidered in descriptive set theory as benchmarks to gauge the complexity of
natural classification problems (see, e.g., \[Kan08\]):

(a) $\alpha E_0 \beta$ if and only if
\[
(\exists m)(\forall n \geq m)((\alpha(n) = \beta(n))).
\]
(b) $\alpha E_1 \beta$ if and only if
\[
(\forall x m \in \omega)((\alpha^m = \beta^m)).
\]
(c) $\alpha E_2 \beta$ if and only if
\[
\sum_{k=0}^{\infty} \frac{(\alpha \triangle \beta)(k)}{k + 1} < \infty.
\]
(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.
\[
\lim_{k \rightarrow \infty} \frac{\text{card}(|\{i \leq k : \alpha \triangle \beta(i) = 1\}|}{k + 1} = 0.
\]
These benchmark equivalence relations lie at the base of the Borel hierarchy:
Figure 2, which is taken from \[CHM12\], shows all continuous reducibilities
between them (in fact, the diagram is the same even if we restrict to com-
putable reductions). For more background about Borel, continuous, and
computable reductions, see \[Gao09, Hjo10, Mil21, BMSMZ21\].
Computable structures. A signature $L$ lists all function symbols and relation symbols which characterize an algebraic structure. All our structures have domain $\omega$. 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 $A$ by identifying $A$ with its atomic diagram, i.e., the collection of atomic formulas which are true of $A$. Up to a suitable Gödel numbering of $L$-formulas, the atomic diagram of $A$ 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 $A$ in a relational signature (i.e., with no function symbols) can be presented as an increasing union of its finite substructures

$$A \models_0 \subseteq A \models_1 \subseteq \ldots \subseteq A \models_i \subseteq \ldots,$$

where $A \models_n$ denotes the restriction of $A$ to the domain $\{0,1,\ldots,n\}$ and $A = \bigcup A \models_i$. For more background about computable structures, see [AK00, EG00].

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

- $\Sigma_0^{\inf}$ and $\Pi_0^{\inf}$ formulas are quantifier-free first-order formulas.
- A $\Sigma_n^{\inf}$ formula $\psi(x)$ is a countably infinite disjunction

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

where each $\xi_i$ is a $\Pi_n^{\inf}$ formula.
- A $\Pi_n^{\inf}$ formula $\psi(x)$ is a countably infinite conjunction

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

FIGURE 1. Reductions up to continuous reducibility.
where each $\xi_i$ is a $\Sigma^{\text{inf}}_n$ formula.

Next, computable infinitary $\Sigma_n$ formulas (or $\Sigma^c_n$ 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 $\Sigma_n$ formulas is denoted by $\Sigma^c_n(X)$. For more background about infinitary formulas, see [Mar16].

2.1. Our framework. We shall now revisit the learning framework presented in [BFSM20]. Our exposition closely follows [BSM21]. 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 $'\mathcal{A}'$, to be understood as conveying the piece of information “this is $\mathcal{A}$”.

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

- The learning domain (LD) is the collection of all copies of the structures from $\mathcal{K}$. That is,
  $$\text{LD}(\mathcal{K}) := \bigcup_{\mathcal{A} \in \mathcal{K}} \{S : S \cong \mathcal{A}\}.$$  
  As we identify each countable structure with an element of Cantor space, we obtain that $\text{LD}(\mathcal{K}) \subseteq 2^\omega$.

- The hypothesis space (HS) contains, for each $\mathcal{A} \in \mathcal{K}$, a formal symbol $'\mathcal{A}'$ and a question mark symbol. That is,
  $$\text{HS}(\mathcal{K}) := \{'\mathcal{A}' : \mathcal{A} \in \mathcal{K}\} \cup \{?\}.$$  

- A learner $\mathcal{M}$ 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 $\mathcal{M}$ is a function
  $$\lim_{n \to \infty} \mathcal{M}(S \upharpoonright n) = '\mathcal{A}'$$  
  if and only if $S$ is a copy of $\mathcal{A}$.

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

Remark 2.3. In [BSM21], the domain of a learner was limited to
  $$X := \{S \upharpoonright n : S \in \text{LD}(\mathcal{K})\},$$
the collection of (finite) initial segments of structures from the learning domain. For our present purposes, it is more convenient to let $\mathcal{M}$ 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 $h(\sigma) = ?$, for all $\sigma \in 2^{<\omega} \setminus X$. 

In [BFSM20, Theorem 3], we obtained the following model theoretic characterization of which families of structures are learnable.

**Theorem 2.4** (Bazhenov, Fokina, San Mauro). Let $\mathcal{K} := (\mathcal{A}_i)_{i \in \omega}$ be a countable family of pairwise nonisomorphic structures. Then, $\mathcal{K}$ is learnable if and only if there are $\Sigma^\inf_2$ formulas $\varphi_0, \ldots, \varphi_n, \ldots$ such that

$$\mathcal{A}_i \models \varphi_j \iff i = j.$$

The interested reader is referred to [BFSM20] for motivating examples and a detailed discussion about our framework (there named Inf Ex$_{\omega}$-learning).

**Definition 2.5.** We say that a family of structures $\mathcal{K}$ is countable if it contains at most countably many isomorphism types. Similarly, $\mathcal{K}$ 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.** ([CCKM04, KMV07]). A Turing operator $\Phi$ is a Turing computable embedding of $\mathcal{K}_0$ into $\mathcal{K}_1$ (notation: $\mathcal{K}_0 \leq_{tc} \mathcal{K}_1$) if it induces an embedding $\mathcal{K}_0 \models_{\omega} \Phi$, that is, if $\Phi$ 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
  $$\mathcal{A} \cong \mathcal{B} \iff \Phi(\mathcal{A}) \cong \Phi(\mathcal{B}).$$

It is common to abbreviate the term “Turing computable embedding” as tc-embedding. One of the most powerful tools in the theory of tc-embeddings is the so-called Pullback Theorem [KMV07]. In this paper, we’ll adopt a natural relativization of this result, already employed in [BFSM20].

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

$$\mathcal{A} \models \psi^* \iff \Phi(\mathcal{A}) \models \psi.$$

Note that the Relativized Pullback Theorem can be applied to any continuous operator $\Phi$. Indeed, if $\Phi$ is a 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

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

**Theorem 3.1.** A family of structures $\mathcal{R}$ 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 $\mathcal{A}, \mathcal{B} \in \text{LD}(\mathcal{R})$.

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

$(\Rightarrow)$: Let $\Gamma$ be a function which induces a continuous reduction from $\text{LD}(\mathcal{R}) / \cong$ to $E_0$. We need to show that $\mathcal{R}$ is learnable. Certainly, $(\beta_i \equiv_{\mathcal{R}} \beta_j)$, for all $i \neq j$. Since $\Gamma$ is continuous, by Lemma 2.1 there exists an oracle $\Phi$ so that

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

Let $\alpha$ be a real. We define an auxiliary function $f_{\text{sim}}(\alpha; i, s)$. Informally speaking, $f_{\text{sim}}(\alpha; i, s)$ is a measure of similarity (at the stage $s$) between the reals $\Gamma(\alpha)$ and $\beta_i$.

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

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

$$f_{\text{sim}}(\alpha; i, s) := \max \left\{ k \leq \ell : \text{for every } j \leq k, \Phi_{(X \oplus \alpha)|s}(\ell - j) = \beta_i(\ell - j) \right\}, \text{ if } \Phi_{(X \oplus \alpha)|s}(\ell) = \beta_i(\ell);$$

$$-1, \text{ otherwise.}$$

Here by $\ell$ 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 $f_{\text{sim}}$ satisfies the following properties. Suppose that a real $\alpha$ 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 $x \geq m_0$, we have $\Gamma(\alpha)(x) = \beta_{i_0}(x)$. This implies that there exists a stage $s_0$ such that every $s \geq s_0$ satisfies $f_{\text{sim}}(\alpha; i_0, s + 1) \geq f_{\text{sim}}(\alpha; i_0, s) > -1$. In addition,

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

(b) Let $i \neq i_0$. Since $(\Gamma(\alpha) \equiv_{\mathcal{R}} \beta_i)$ and $\ell[s + 1] \leq \ell[s] + 1$ for all $s$, there are infinitely many stages $s$ such that $\Phi_{(X \oplus \alpha)|s}(\ell[s]) \neq \beta_i(\ell[s])$ and
\[
f_{\text{sim}}(\alpha; i, s) = -1. \text{ Therefore,} \\
\liminf_{s} f_{\text{sim}}(\alpha; i, s) = -1.
\]

**Construction.** We build our desired learner $M$. Let $\alpha$ be a real, and let $s$ be a natural number. We put $M(\text{empty string}) := ?$.

For the number $s$, we define two auxiliary parameters $b[s] \in \omega$ and $c[s] \in \{0, 1, 2\}$. Informally, these parameters control our learning strategy, i.e., what the learner $M$ should output. We will ensure the following property: $c[s] = 0$ if and only if $M(\alpha \upharpoonright s) = ?$. We put $b[0] := 0$ and $c[0] := 0$. If not specified otherwise, we assume that $b[s + 1] = b[s]$ and $c[s + 1] = c[s]$.

Suppose that $M(\alpha \upharpoonright s)$ is defined. We describe how to obtain the value $M(\alpha \upharpoonright s + 1)$. Consider the following three cases:

**Case 1.** Suppose that $c[s] = 0$. Then we search for the least $j \leq s$ such that $f_{\text{sim}}(\alpha; j, s) \neq -1$ and
\[
f_{\text{sim}}(\alpha; j, s) = \max\{f_{\text{sim}}(\alpha; m, s) : m \leq s\}.
\]
- If there is such $j$, then set $M(\alpha \upharpoonright s + 1) := 'A_j'$, $c[s + 1] := 1$, and $b[s + 1] := s$;
- Otherwise, put $M(\alpha \upharpoonright s + 1) := ?$.

**Case 2.** Suppose that $c[s] = 1$ and $M(\alpha \upharpoonright s) = 'A_i'$. 
- If $f_{\text{sim}}(\alpha; i, s + 1) \neq -1$, then define $M(\alpha \upharpoonright s + 1) := 'A_i'$;
- If $f_{\text{sim}}(\alpha; i, s + 1) = -1$, then put $M(\alpha \upharpoonright s + 1) := 'A_0'$ and $c[s] := 2$.

**Case 3.** Suppose that $c[s] = 2$ and $M(\alpha \upharpoonright s) = 'A_i'$. Our construction will ensure that in this case, we have $i \leq b[s]$.
- If $f_{\text{sim}}(\alpha; i, s + 1) \neq -1$, then $M(\alpha \upharpoonright s + 1) := 'A_i'$;
- If $f_{\text{sim}}(\alpha; i, s + 1) = -1$ and $i < b[s]$, then set $M(\alpha \upharpoonright s + 1) := 'A_{i+1}'$;
- If $f_{\text{sim}}(\alpha; i, s + 1) = -1$ and $i = b[s]$, then define $M(\alpha \upharpoonright s + 1) := ?$ and $c[s] := 0$.

**Verification.** We show that $M$ learns our family $\mathcal{R}$. Suppose that $\alpha$ is a real, which encodes a copy $S$ of some $A_{i_0}$.

Let $s_0$ be a stage such that $f_{\text{sim}}(\alpha; i_0, s) \neq -1$ for all $s \geq s_0$. Case 1 of the construction ensures that there are infinitely many stages $s \geq s_0$ such that $M(\alpha \upharpoonright s) = 'A_j'$ for some $j \in \omega$.

If there exists a stage $s_1 \geq s_0$ such that $M(\alpha \upharpoonright s_1) = 'A_{i_0}'$, then we have
\[
M(\alpha \upharpoonright s) = 'A_{i_0}' \text{ for all } s \geq s_1,
\]
and in the limit, $M$ outputs the correct conjecture. Thus, it is sufficient to establish the existence of this $s_1$.

Let $s'$ be an arbitrary stage such that $s' \geq s_0$ and $M(\alpha \upharpoonright s') = 'A_j'$ for some $j$. If $j = i_0$, then there is nothing to prove. Assume that $j \neq i_0$. There are again three cases:
Case I. If $c[s'] = 1$, then eventually we will witness a sequence of stages $s' = s'' < s''_0 < s''_1 < \cdots < s''_{i_0-1}$ such that

$$f_{\text{sim}}(\alpha; j, s'') = f_{\text{sim}}(\alpha; 0, s''_0) = f_{\text{sim}}(\alpha; 1, s''_1) = \cdots = f_{\text{sim}}(\alpha; i_0 - 1, s''_{i_0-1}) = -1.$$ 

Hence, the construction ensures that $M(\alpha \uparrow s''_{i_0-1}) = \mathcal{A}_{i_0}^\sim$.

Case II. If $c[s'] = 2$ and $j < i_0$, then we will find a sequence of stages $s' = s'' < s''_j < s''_{j+1} < \cdots < s''_{i_0-1}$ such that

$$f_{\text{sim}}(\alpha; k, s''_k) = -1$$
for every $j \leq k < i_0$.

Again, we have $M(\alpha \uparrow s''_{i_0-1}) = \mathcal{A}_{i_0}^\sim$.

Case III. If $c[s'] = 2$ and $j > i_0$, then there will be a sequence of stages $s' = s'' < s''_j < s''_{j+1} < \cdots < s''_{b[s']}$ such that

$$f_{\text{sim}}(\alpha; k, s''_k) = -1$$
for every $j \leq k \leq b[s']$.

Then for $s^* := s''_{b[s']} + 1$, we have $c[s^*] = 1$ and $M(\alpha \uparrow s^*) = \mathcal{A}_\ell$ for some $\ell$. After that, the argument proceeds similarly to the first case.

Therefore, we proved that there is $s_1 \geq s_0$ with $M(\alpha \uparrow s_1) = \mathcal{A}_{i_0}^\sim$. This implies that in the limit, $M$ says that “$S$ is a copy of $A_{i_0}$”, and the family $\mathcal{R}$ is learnable by $M$.

($\Leftarrow$): For the converse direction, let $M$ be a learner of $\mathcal{R}$. We need to construct a continuous function $\Gamma : 2^\omega \to 2^\omega$ which induces a reduction from $\text{LD}(\mathcal{R})/\equiv$ 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 $\Gamma$ as follows,

$$\Gamma(\beta)(s) := \alpha_{M(\beta_{\uparrow s})}(s).$$

Here we use the following convention:

- if $M(\beta_{\uparrow s}) = \mathcal{A}_i$, then $\alpha_{M(\beta_{\uparrow s})} = \alpha_i$;
- if $M(\beta_{\uparrow s}) = ?$, then $\alpha_{M(\beta_{\uparrow s})} = 0^\omega$.

To verify that this $\Gamma$ works, it is enough to observe the following: if a real $\beta$ encodes a copy of some $\mathcal{A}_i$ from $\text{LD}(\mathcal{R})$, then there must be a stage $s_0$ such that, for all $s \geq s_0$, $M(\beta_{\uparrow s})$ outputs $\mathcal{A}_i^\sim$, and thus $\Gamma(\beta)$ is $E_0$-equivalent to $\alpha_i$. So, since the $\alpha_i$’s form a transversal for $E_0$, we deduce that, if $\beta_0$ and $\beta_1$ encode copies of $\mathcal{A}_i$ and $\mathcal{A}_j$ respectively, then

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

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 $\mathcal{R}$ is $E$-learnable if there is function $\Gamma : 2^\omega \rightarrow 2^\omega$ which continuously reduce $LD(\mathcal{R}) \cong$ 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.

The rest of the paper is devoted to the study of the learning 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 4.2, $E_1$ 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

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

Theorem 4.1. A family $\mathcal{R}$ is $E_1$-learnable if and only if $\mathcal{R}$ 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 Figure 1), every $E_0$-learnable family is also $E_1$-learnable.

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

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 $(\beta_i, \beta_j)$, there exists the least $q > m_s$ such that $\beta_i^{[q]} \not= \beta_j^{[q]}$. We choose the least $\ell \in \omega$ with $\beta_i^{[q]}(\ell) \not= \beta_j^{[q]}(\ell)$, 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 $\ell$ such that $\langle q, \ell \rangle$ belongs to $X$.

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

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

It is clear that the operator $\Psi$ is $X$-computable — hence, $\Psi$ is continuous.

We show that the operator $\Phi := \Psi \circ \Gamma$ provides a continuous reduction from $LD(\mathcal{R}) \cong$ to $E_0$. Let $\alpha$ be a real which encodes a copy of some $A_{i_0}$. 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

$$\Gamma(\alpha)(m_{\xi(i,i_0,t)}) = \beta_{i_0}(m_{\xi(i,i_0,t)}) \not= \beta_i(m_{\xi(i,i_0,t)}).$$
This implies that \((\Phi(\alpha) \not\equiv_0 \Psi(\beta))\). Therefore, we deduce that our family \(\mathcal{R}\) is \(E_0\)-learnable. The theorem is proved.



4.2. \(E_2\)-learning. Recall that the equivalence relation \(E_2\) is given by

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

**Theorem 4.2.** A countable family \(\mathcal{R}\) is \(E_2\)-learnable if and only if \(\mathcal{R}\) 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 Figure 1), every \(E_0\)-learnable family is \(E_2\)-learnable.

Let \(\mathcal{R} := (A_i)_{i \in \omega}\) be an \(E_2\)-learnable family. Let \(\Gamma\) be an operator, which induces a continuous reduction from \(LD(\mathcal{R})/\equiv_\mathcal{R}\) to \(E_2\). For \(i \in \omega\), we fix a real \(\beta_i\) such that \(\Gamma\) maps all copies of \(A_i\) into \([\beta_i]_{E_2}\).

By Lemma 2.1, there exist an oracle \(X\) and a Turing operator \(\Phi\) such that \(\Gamma(\alpha) = \Phi^X \otimes_\alpha\) for all \(\alpha \in 2^\omega\). Consider the uniform join

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

**Construction.** We define a \(Y\)-computable operator \(\Psi\). For a real \(\alpha\), 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 \(x \leq \ell[s]\), the value \(\Phi^X \otimes_\alpha|x|/s(x)|s\) 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

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

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 \(i[s] \leq b[s]\) for every \(s\).

**Stage 0.** Set \(i[0] = 0\), \(b[0] = 1\), and \(c[0] = 0\).

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

**Case 1.** If \(p(i[s], s+1) \leq b[s]\), 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 \( i_0 \leq b[s] + 1 \) such that
\[
p(i_0, s + 1) = \min\{p(j, s + 1) : j \leq b[s] + 1\}.
\]
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 \( \Psi : \alpha \mapsto \gamma_\alpha \) is \( Y \)-computable.

**Verification.** Suppose that a real \( \alpha \) encodes a copy of the structure \( A_{i_0}. \)

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, \( b^* \geq i_0. \)

**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 \( (\Gamma(\alpha) \notin \beta_j) \) 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 \( s \geq s^*. \) Then, our construction ensures that after the stage \( s_0, \) there will be a stage \( s_1 \) satisfying Case 4. This implies that \( b[s_1] = b^* + 1, \) which gives a contradiction. Thus, we deduce that there must exist a stage \( s'_0 \) such that \( b[s'_0] \geq i_0. \)

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 \( b[s_1] \geq N_0 + 1. \) Consider the value \( i^* := i[s_1]. \)

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.
2. 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 \( k \leq i_0. \) 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 \( b^* \geq i_0. \)

Now choose a stage \( s^* \) such that \( b[s^*] = b^*. \) There exists a stage \( s_1 \geq s^* \) such that every \( i \leq b^* \) 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 \( s \geq s_1 + b^* + 2, \) we must have \( i[s] = i_0. \)

This implies that the real \( \Psi(\alpha) \) is \( \mathcal{L}_0 \)-equivalent to \( \beta_{i_0}. \) For all \( i \neq j, \) we have \( (\beta_i \notin \beta_j) \) — clearly, this implies \( (\beta_i \notin \beta_j). \) Hence, we conclude that
our operator $\Psi$ provides a continuous reduction from $LD(\mathcal{R})/\equiv$ to $E_0$. In other words, the family $\mathcal{R}$ is $E_0$-learnable, as desired. \qed

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

$$(\alpha E_3 \beta) \iff (\forall m \in \omega)(\alpha^m E_0 \beta^m).$$

**Theorem 5.1.** A finite family $\mathcal{R}$ is $E_3$-learnable if and only if $\mathcal{R}$ 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 Figure 1), every $E_0$-learnable family is $E_3$-learnable.

For the other direction, let $\mathcal{R} := (A_i)_{i \in \omega}$ be an $E_3$-learnable family and let $\Gamma$ induce a continuous reduction from $LD(\mathcal{R})/\equiv$ to $E_3$. For $i \leq n$, choose $\beta_i$ such that $\Gamma$ maps all copies of $A_i$ into $[\beta_i]_{E_3}$. For each pair of indices $i \neq j$, we choose a number $q(i,j)$ such that

$$(\beta_i[q(i,j)] E_0 \beta_j[q(i,j)]).$$

Then, we define a Turing operator $\Psi: 2^{\omega} \rightarrow 2^{\omega}$ as follows.

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

The operator $\Phi := \Psi \circ \Gamma$ provides a continuous reduction from $LD(\mathcal{R})/\equiv$ to $E_0$. Indeed, let $\alpha$ be a real which encodes a copy of $A_{i_0}$. 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

$$(\alpha[q(i,i_0)] E_0 \beta_{i_0}[q(i,i_0)] E_0 \beta_i[q(i,i_0)])$$

and $(\Phi(\alpha) E_0 \Psi(\beta_i)).$

Therefore, the family $\mathcal{R}$ is $E_0$-learnable. \qed

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 $\mathcal{R} := (A_i)_{i \in \omega}$ 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 $\mathcal{R}$ is allowed to be infinite. After that, we provide comments on how to build the desired $\mathcal{R}$ as a family of directed graphs.

Consider signature $L = \{R_j : j \in \omega\} \cup \{\leq\}$, where $R_j$ are unary predicates. Given a real $\alpha$, 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 $D(\alpha)$ 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 $\omega$ and $\omega^*$ are respectively the order types of the positive and negative integers. For a finite string $\sigma \in 2^{<\omega}$, let $A_\sigma$ be the structure $D(\sigma^{10^8})$. Our family $\mathcal{K}$ consists of all $A_\sigma$, $\sigma \in 2^{<\omega}$.

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

**Proof.** Recall that the family $\{\omega, \omega^*\}$ is learnable, as they are distinguishable by $\Sigma^1_2$ formulas [FSM20, Theorem 3]. By employing this fact, it is not hard to build a Turing operator $\Phi$, which acts as follows. Given a real $\alpha$, it treats $\alpha$ as a code for the atomic diagram of a countable partial order $L$. Then:

- If $L$ is a copy of $\omega$, then the output $\Phi(\alpha)$ is $E_0$-equivalent to $0^{10^8}$.
- If $L \cong \omega^*$, then we have $(\Phi(\alpha))_{E_0}^1$.\

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

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

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

Observe the following. Let $\beta$ be a real. If a real $\alpha$ encodes a copy of the structure $D(\beta)$, 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 $0^{10^8}$;
- if $\beta(j) = 1$, then $(\Theta(\alpha))_{[j]}^{[j]} E_0^{10^8}$.

This observation implies that the operator $\Theta$ witnesses the $E_3$-learnability of our family $\mathcal{K}$. Lemma 5.3 is proved.\[\square\]

Now, towards a contradiction, assume that the family $\mathcal{K}$ is $E_0$-learnable. Then $\mathcal{K}$ is $\text{InfEx}$-learnable, and by Theorem 2.4 one can choose an infinitary $\Sigma_2$ sentence $\theta$ such that $A_0 \models \theta$ and for every $\sigma \neq 0$, we have $A_\sigma \models \theta$.

Without loss of generality, one may assume that

$$\theta = 3\bar{x} \bigwedge_{i \in I} \forall \bar{y}_i \psi_i(\bar{x}, \bar{y}_i),$$

where every $\psi_i$ is a quantifier-free formula. Fix a tuple $\bar{c}$ from the structure $A_0$ such that

$$A_0 \models \bigwedge_{i \in I} \forall \bar{y}_i \psi_i(\bar{c}, \bar{y}_i).$$

Choose a natural number $N$ such that for every $j \geq N$, the $R_j$-box of $A_0$ does not contain elements from $\bar{c}$.

Consider a string $\tau := 010^N$ and the corresponding structure $A_\tau = D(\tau^{10^8})$. It is clear that for every $j < N$, the (contents of the) $R_j$-boxes...
inside $A_0$ and $A_\tau$ are isomorphic. Therefore, one can choose a tuple $\bar{d}$ inside $A_\tau$ as isomorphic copies of $\bar{c}$ (with respect to the isomorphism of the $R_j$-boxes, $j < N$).

**Claim 5.1.** The structures $(A_0, \bar{c})$ and $(A_\tau, \bar{d})$ satisfy the same $\exists$-sentences.

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

$$A_0 \models \exists \bar{y} \psi(\bar{c}, \bar{y}) \Rightarrow A_\tau \models \exists \bar{y} \psi(\bar{d}, \bar{y}).$$

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

Choose a tuple $\bar{b}$ from $A_0$ such that $A_0 \models \psi(\bar{c}, \bar{b})$. Suppose that $\bar{b} = b_0, b_1, \ldots, b_m$. We define a new tuple $\bar{b}' = b'_0, b'_1, \ldots, b'_m$ from $A_\tau$ 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 $A_\tau$, while preserving the ordering $\leq$. More formally, one needs to ensure the following: if $b_k \neq b_\ell$ both belong to this $R_j$-box, then we have:

$${A_0 \models b_k \leq b_\ell} \Leftrightarrow {A_\tau \models b'_k \leq b'_\ell}.$$

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 $A_\tau$ satisfies $\psi(\bar{d}, \bar{b}')$, and $A_\tau \models \exists \bar{y} \psi(\bar{d}, \bar{y})$. □

Claim 5.1 implies that

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

and hence, $A_\tau \models \theta$, which contradicts the choice of $\theta$. We deduce that the family $\mathcal{R}$ is not $E_0$-learnable.

In order to obtain a family of directed graphs $\mathcal{R}_{gr}$, which has the same properties as the family $\mathcal{R}$, one can proceed as follows. Instead of distinguishing an $R_j$-box via the predicate $R_j$, one attaches to each element $a$ of the (intended) $R_j$-box its own cycle of size $(j + 3)$. After that, the proof for the family $\mathcal{R}_{gr}$ 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 $\Sigma^\inf_2$ formulas (see Theorem 2.4). The next theorem offers an analogous characterization for $E_3$-learning.

**Theorem 5.4.** Let $\mathcal{R} := (A_i)_{i \in \omega}$ be a countable family. The family $\mathcal{R}$ is $E_3$-learnable if and only if there exists a countable family of $\Sigma^\inf_2$ sentences $\Theta$ with the following properties:
(a) if \( \theta \) is a formula from \( \Theta \), then there is a formula \( \psi \in \Theta \) such that for every \( \mathcal{A} \in \mathcal{R} \),
\[
\mathcal{A} \models \theta \iff \mathcal{A} \models \neg \psi;
\]
(b) if \( \mathcal{A} \not\models \mathcal{B} \) are structures from \( \mathcal{R} \), then there is a sentence \( \theta \in \Theta \) such that
\[
\mathcal{A} \models \theta \quad \text{and} \quad \mathcal{B} \models \neg \theta.
\]

**Proof.** The proof of the theorem is inspired by ideas from \cite{BFSM20}. In particular, we will adopt the technology of tc-embeddings and the Relativized Pullback Theorem reminded in the preliminaries.

Consider a signature \( L_{st} = \{ \leq \} \cup \{ P_i : i \in \omega \} \), where \( P_i \) are unary predicates. For an index \( i \in \omega \), an \( L \)-structure \( S_i \) satisfies the following properties:
- Inside \( S_i \), 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 \( \leq \)-incomparable. Let \( \eta \) be the order type of the rational numbers.
- The predicate \( P_1 \) contains an isomorphic copy of \( 1 + \eta \).
- Every \( P_j \), for \( j \neq i \), contains a copy of \( \eta \).

The class \( \mathcal{R}_{st} \) consists of all structures \( S_i \) with \( i \in \omega \).

In \cite{BFSM20}, it is shown \( \mathcal{R}_{st} \) is an archetypical \( E_0 \)-learnable family, in the sense that a countable family \( \mathcal{C} \) is learnable if and only if there is a continuous embedding from the class \( \mathcal{C} \) into \( \mathcal{R}_{st} \).

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

Consider a new signature \( L_1 = L_{st} \cup \{ Q_k : k \in \omega \} \), where \( Q_k \) are unary predicates. The class \( \mathcal{C}_{st} \) 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 \( \mathcal{R}_{st} \).

Note that our class \( \mathcal{C}_{st} \) has cardinality \( 2^{\aleph_0} \).

**Lemma 5.5.** The class \( \mathcal{C}_{st} \) has a computable family of \( \Sigma^2_{\text{inf}} \) sentences \( \Theta \), which satisfies properties (a) and (b) from the formulation of Theorem 5.4.

**Proof.** The desired family \( \Theta \) contains the following \( \Sigma^2_{\text{inf}} \) sentences:

1. For each \( i \) and \( j \), we add a finitary \( \Sigma_2 \) sentence \( \theta_{i,j} \), which states the following: “the \( P_j \)-predicate inside the \( Q_i \)-box has a \( \leq \)-least element”.
2. For each \( i \) and \( j \), we add a \( \Sigma^2_{\text{inf}} \) sentence \( \psi_{i,j} \), which is equivalent to the following formula:
\[
\bigvee_{k \neq j} \theta_{i,k}.
\]
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 $\mathcal{C}_{st}$. Since the $Q_i$-box of $\mathcal{M}$ is a structure from $\mathcal{K}_{st}$, it is not hard to show that

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

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

Suppose that $\mathcal{M} \neq \mathcal{N}$ are structures from $\mathcal{C}_{st}$. 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 $\eta$. Then, it is clear that

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

Therefore, $\mathcal{C}_{st}$ 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 $\mathcal{K}$ to $\mathcal{C}_{st}$. 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 $\alpha$, our operator $\Psi_{\vec{\gamma}}$ produces a new real $\delta_\alpha$, which encodes the atomic diagram of an $L$-structure $S(\alpha)$.

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

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

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

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

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

**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 $i \leq s + 1$ 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 $i_0 \leq s + 1$ and

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

  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 $i \leq s + 1$, 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 $(\bigoplus_{i \in \omega} \gamma_i)$-computable.
2. If $(\alpha E_0 \gamma_i)$ for some $i \in \omega$, then the structure $S(\alpha)$ is isomorphic to $S_i$.

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

We define a continuous operator $\Xi$ as follows. Let $\alpha$ 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 $\delta$, 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 $L_{\alpha}$-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 \( \mathcal{K} \) into a countable subclass of \( C_{st} \). So, by applying the Relativized Pullback Theorem (Theorem 2.7) to the continuous embedding \( \Xi \), we recover a countable family of formulas with the desired properties. Indeed, the following holds:

- by Lemma 2.1 \( \Phi \) is equivalent to a Turing \( X \)-operator, for a suitable oracle \( X \).

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

\[ p(q) \] This direction essentially follows from previous results. Assume that \( \mathcal{K} \):

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

Hence, we can apply Theorem 2.7 and deduce that \( \mathcal{K} \) has a family \( \Theta \) of \( \Sigma_2^{\inf} \) formulas which satisfies the properties (a) and (b) of the theorem, as desired.

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

**Lemma 5.6.** There is a collection of pairs of formulas \( (\rho_{i0}, \rho_{i1})_{i \in \omega} \) so that, for all structures \( A \) and \( B \) from \( \mathcal{K} \),

- (1) \( \bigcup_{i \in \omega} \{\rho_{i0}, \rho_{i1}\} = \Theta \);
- (2) for all \( i \in \omega \), \( A \) satisfies exactly one formula between \( \rho_{i0} \) and \( \rho_{i1} \);
- (3) if \( A \not= B \), then, for some \( j \in \omega \),
  \[ A \models \rho_{j0} \iff B \models \rho_{j1}. \]

The proof is similar to that of the direction (2) \( \Rightarrow \) (1) of [BFSM20, Theorem 3]. Let \( i \in \omega \). For \( k \in \{0, 1\} \), without loss of generality assume that

\[
\rho_{ik} := (\exists x) \bigwedge_{j \in J_k} \forall y \varphi_{ik,j}(\bar{x}, \bar{y}).
\]

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

**Construction.** Now, let \( \alpha \) be a real. Denote by \( \mathcal{F}_{\alpha_{\uparrow s}} \) the finite structure (in the signature of \( \mathcal{K} \)) encoded by the initial segment \( \alpha_{\uparrow s} \) of \( \alpha \). The continuous operator \( \Gamma_i \) 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 $\varphi_{i_0}$ is $\mathcal{F}_{\alpha_{1,s}}$-compatible via $\bar{c}$, and $\varphi_{i_1}$ is not $\mathcal{F}_{\alpha_{1,s}}$-compatible for all tuples $c < \bar{c}$. If so, let $\Gamma_i(\alpha)(2s + 1) := 0$; 
2. There is a tuple $\bar{c}$ so that $\varphi_{i_1}$ is $\mathcal{F}_{\alpha_{1,s}}$-compatible via $\bar{c}$, and $\varphi_{i_0}$ is not $\mathcal{F}_{\alpha_{1,s}}$-compatible for all tuples $c \leq \bar{c}$. 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 $\Gamma_i$ immediately follows from the construction. Next, suppose that $\beta \in 2^\omega$ encodes a copy of a structure $\mathcal{S} \in \mathcal{R}$. By Lemma 5.6, $\mathcal{S}$ satisfies exactly one formula between $\varphi_{i_0}$ and $\varphi_{i_1}$; 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 $\varphi_{i_1}$ is $\mathcal{F}_{\beta_{1,t}}$-compatible via $\bar{c}$, for all $t \geq t_0$. On the other hand, since $\mathcal{S} \models \varphi_{i_0}$, it must be the case that for all tuples $\bar{d}$ (and, in particular, all tuples $\leq \bar{c}$), there must be a stage $t_1$ so that, for all $t \geq t_1$, $\varphi_{i_0}$ is not $\mathcal{F}_{\beta_{1,t}}$-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 $1^\omega$, as desired. 

We can now construct a continuous reduction from $LD(\mathcal{R})/\equiv$ to $E_3$ by merging the operators $\Gamma_i$'s as follows:

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

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

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

Thus, $\Gamma(\beta_0) \leq^* E_3 \Gamma(\beta_1)$.

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 $\alpha$ and $\beta$ be reals, and let $s \in \omega$. We use the following notation:

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

Recall that the equivalence relation $Z_0$ is given by

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

**Lemma 6.1.** Suppose that $(\alpha Z_0 \beta)$ and $(\alpha Z_0 \gamma)$. 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 $\varepsilon$ such that $0 < \varepsilon < r$, we have

$$\limsup s dn(\alpha, \gamma; s) \geq r - \varepsilon.$$ 

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 $s \leq s_0$.

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

Note that every $s$ satisfies the following:

$$\text{card}\{i \leq s : \alpha \triangledown \gamma(i) = 1\} \geq \text{card}\{i \leq s : \beta \triangledown \gamma(i) = 1\} - \text{card}\{i \leq s : \beta \triangledown \alpha(i) = 1\}.$$ 

Hence, we have:

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

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

$$\limsup s dn(\alpha, \gamma; s) \geq q.$$ 

**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 $\mathcal{R}$ is $Z_0$-learnable if and only if $\mathcal{R}$ 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 Figure 1), every $E_0$-learnable family is also $Z_0$-learnable.

Suppose that $\mathcal{R} = (A_i)_{i \in \omega}$ is a $Z_0$-learnable family. Let $\Gamma$ be an operator which induces a continuous reduction from $LD(\mathcal{R})$ to $Z_0$. For $i \leq n$, we fix $\beta_i$ such that $\Gamma$ maps all copies of $A_i$ into $[\beta_i]_{Z_0}$. Notice that the reals $\beta_i$ are pairwise not $E_0$-equivalent.
We fix a positive rational $q_0$ such that

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

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

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

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

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

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

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

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

Suppose that a real $\alpha$ encodes a copy of a structure $A_{i_0}$ for some $i_0 \leq n$. 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 $t \geq t_0$, we have $dn(\Phi^{X \oplus \alpha}, \beta_{i_0}; t) \leq q_0$. Fix a stage $s_0$ with $i_0 \leq \ell[s_0]$. Then for all $s \geq s_0$, 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 $\beta_{i_0}$.

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

It is known that $E_3$ is continuously reducible to $Z_0$ (see Figure 1). So, $E_3$ is $\text{Learn}^\omega$ 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 $E_{\text{set}}$-learnable?

6.2. $E_{\text{set}}$-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 $\zeta$ 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 $\text{Inf Ex}$-learnable. Yet, we have observed that all Borel equivalence relations so far considered are $\text{Learn}^\omega$-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 $E_{\text{set}}$ suffices.

**Proposition 6.3.** The family $\{\omega, \zeta\}$ is $E_{\text{set}}$-learnable.
Proof. Given a real \( \alpha \), which encodes a graph 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 \( \alpha \).

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

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

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 \( \leq B_s \)-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 \( \alpha \) encodes a copy of \( \mathcal{A} \in \{ \omega, \zeta \} \). If \( \mathcal{A} \) is isomorphic to \( \zeta \), then it is clear that

\[
\{ (\Psi(\alpha))^{[m]} : m \in \omega \} = \{ 0^i1^\omega : i \in \omega \}.
\]

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

\[
\{ (\Psi(\alpha))^{[m]} : m \in \omega \} = \{ 0^i1^\omega : i \in \omega \} \cup \{ 0^\omega \}.
\]

Therefore, we deduce that the family \( \{ \omega, \zeta \} \) is \( E_{\textrm{set}} \)-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 Figures 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 \( \Pi_3^0 \) equivalence relations (see [Gao09]);

2. Secondly, it would be nice to obtain learning theoretic or purely syntactic characterizations for \( E_{\textrm{set}} \)- 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.
Figure 2. Reductions up to \( \text{Learn}^{<\omega} \)-reducibility.

Figure 3. Reductions up to \( \text{Learn}^\omega \)-reducibility.

References

[AK00] Chris J. Ash and Julia F. Knight. *Computable Structures and the Hyperarithmetical Hierarchy*, volume 144 of *Studies in Logic and the Foundations of Mathematics*. Elsevier Science B.V., Amsterdam, 2000.

[BFSM20] Nikolay Bazhenov, Ekaterina Fokina, and Luca San Mauro. Learning families of algebraic structures from informant. *Information and Computation*, 275:104590, 2020.

[BMSMZ21] Nikolay Bazhenov, Benoit Monin, Luca San Mauro, and Rafael Zamora. On the computational content of the theory of Borel equivalence relations. *Oberwolfach Preprint OWP-2021-06*, 2021.

[BSM21] Nikolay Bazhenov and Luca San Mauro. On the Turing complexity of learning finite families of algebraic structures. *Journal of Logic and Computation*, 2021. Published online. arXiv preprint arXiv:2106.14515.

[CCKM04] W. Calvert, D. Cummins, J. F. Knight, and S. Miller. Comparing classes of finite structures. *Algebra and Logic*, 43(6):374–392, 2004.

[CG01] Riccardo Camerlo and Su Gao. The completeness of the isomorphism relation for countable Boolean algebras. *Transactions of the American Mathematical Society*, 353(2):491–518, 2001.

[CHM12] Samuel Coskey, Joel David Hamkins, and Russell Miller. The hierarchy of equivalence relations on the natural numbers under computable reducibility. *Computability*, 1(1):15–38, 2012.
Yu. L. Ershov and S. S. Goncharov. *Constructive models*. Kluwer Academic/Plenum Publishers, New York, 2000.

Ekaterina Fokina, Timo Kötzing, and Luca San Mauro. Limit learning equivalence structures. In Aurélien Garivier and Satyen Kale, editors, *Proceedings of the 30th International Conference on Algorithmic Learning Theory*, volume 98 of *Proceedings of Machine Learning Research*, pages 383–403, Chicago, Illinois, 22–24 Mar 2019. PMLR.

Harvey Friedman and Lee Stanley. A Borel reducibility theory for classes of countable structures. *J. Symbolic Logic*, 54(3):894–914, 1989.

Su Gao. *Invariant descriptive set theory*. CRC Press, Boca Raton, FL, 2009.

Clark Glymour. Inductive inference in the limit. *Erkenntnis*, 22:23–31, 1985.

E. Mark Gold. Language identification in the limit. *Information and Control*, 10(5):447–474, 1967.

Ziyuan Gao, Frank Stephan, Guohua Wu, and Akihiro Yamamoto. Learning families of closed sets in matroids. In Michael J. Dinneen, Bakhadyr Khoussainov, and André Nies, editors, *Computation, Physics and Beyond - International Workshop on Theoretical Computer Science, WTCS 2012*, volume 7160 of *Lecture Notes in Computer Science*, pages 120–139, Berlin, 2012. Springer.

Greg Hjorth. Borel equivalence relations. In *Handbook of set theory*, pages 297–332. Springer, 2010.

L. A. Harrington, A. S. Kechris, and A. Louveau. A Glimm-Effros dichotomy for Borel equivalence relations. *Journal of the American Mathematical Society*, 3(4):903–928, 1990.

Valentina S. Harizanov and Frank Stephan. On the learnability of vector spaces. *Journal of Computer and System Sciences*, 73(1):109–122, 2007.

Vladimir Grigor’evich Kanovei. *Borel equivalence relations: Structure and classification*, volume 44. American Mathematical Soc., 2008.

Julia F. Knight, Sara Miller, and Michael Vanden Boom. Turing computable embeddings. *J. Symb. Log.*, 72(3):901–918, 2007.

Steffen Lange, Thomas Zeugmann, and Sandra Zilles. Learning indexed families of recursive languages from positive data: A survey. *Theoretical Computer Science*, 397(1–3):194–232, 2008.

David Marker. *Lectures on infinitary model theory*, volume 46 of *Lecture Notes in Logic*. Cambridge University Press, Cambridge, 2016.

Alan H Meckler. Stability of nilpotent groups of class 2 and prime exponent. *The Journal of Symbolic Logic*, 46(4):781–788, 1981.

Russell Miller. Computable reducibility for Cantor space. In *Structure and Randomness in Computability and Set Theory*, pages 155–196. World Scientific, Singapore, 2021.

Eric Martin and Daniel Osherson. *Elements of scientific inquiry*. MIT Press, 1998.

Wolfgang Merkle and Frank Stephan. Trees and learning. *Journal of Computer and System Sciences*, 68(1):134–156, 2004.

Hilary Putnam. Trial and error predicates and the solution to a problem of Mostowski. *The Journal of Symbolic Logic*, 30(1):49–57, 1965.

Robert I. Soare. *Turing Computability. Theory and Applications*. Springer, Berlin, 2016.

Frank Stephan and Yuri Ventsov. Learning algebraic structures from text. *Theoretical Computer Science*, 268(2):221–273, 2001.

Thomas Zeugmann and Sandra Zilles. Learning recursive functions: A survey. *Theoretical Computer Science*, 397(1–3):4–56, 2008.
Sobolev Institute of Mathematics, 4 Acad. Koptyug Ave., Novosibirsk, 630090, Russia
Email address: bazhenov@math.nsc.ru

Department of Mathematics, Computer Science and Physics, University of Udine, Italy
Email address: cipriani.vittorio@spes.uniud.it

Department of Mathematics, Sapienza University of Rome, Italy
Email address: luca.sanmauro@uniroma1.it