TURING DEGREES AND THE ERSHOV HIERARCHY

FRANK STEPHAN, YUE YANG AND LIANG YU

ABSTRACT. An n-r.e. set can be defined as the symmetric difference of n recursively enumerable sets. The classes of these sets form a natural hierarchy which became a well-studied topic in recursion theory. In a series of ground-breaking papers, Ershov generalized this hierarchy to transfinite levels based on Kleene's notations of ordinals and this work lead to a fruitful study of these sets and their many-one and Turing degrees. The Ershov hierarchy is a natural measure of complexity of the sets below the halting problem. In this paper, we survey the early work by Ershov and others on this hierarchy and present the most fundamental results. We also provide some pointers to concurrent work in the field.

1. INTRODUCTION

This paper aims to achieve two goals. One is to give a survey on the Turing degree structure of the Ershov hierarchy, including some recent results on elementary equivalence and elementary substructure problems. The other is presenting short proofs to some selected results of transfinite levels of the Ershov hierarchy. In his three classic papers [10], [11] and [12], Ershov studied a hierarchy of sets generated by the recursively enumerable sets. At the finite levels, the hierarchy is also referred as difference hierarchy, which forms a proper subclass of the Δ^0_2-sets. The transfinite extension of the difference hierarchy exhausts all Δ^0_2-sets. Although the finite levels of the Ershov hierarchy have been studied by recursion theorists extensively, the same cannot be said on transfinite levels. Things started to change only recently. The transfinite levels of the Ershov hierarchy appeared naturally in many areas, for example, in recursion theory and recursive model theory, see Downey and Gale [14] and Khoussainov, Stephan and Yang [18], and in inductive inference, see Ambainis, Freivalds and Smith [2], Carlucci, Case and Jain [6] and Freivalds and Smith [13]. Yet we still feel that Ershov's theorems did not achieve the popularity which they deserved. It is a pity that such a treasure remains obscure due to partially the limited availability of Ershov's original papers and the notations used almost 40 years ago. We hope our selection of results and our presentation would make the transfinite levels of the Ershov hierarchy more accessible for beginners. We refer the interested reader to the excellent and detailed article of Arslanov [3] about the same topic which goes in depth much beyond our short survey.

The paper is organized as follows. In Section 2, we discuss the elementary equivalent problems among finite levels of the Ershov hierarchy; in Section 3, we remind the readers some basic facts about ordinal notations which is necessary for the later part; in Section 4, we present some basic results on transfinite levels of Ershov hierarchy; in the Section

This work was supported by NUS Academic Research Grants R146-000-114-112 (F. Stephan and Y. Yang) and R252-000-308-112 (F. Stephan); National Natural Science Fund of China, No. 10701041 (L. Yu) and Research Fund for the Doctoral Program of Higher Education, No. 20070284043 (L. Yu).

5 we discuss some applications to the area of inductive inference. We conclude with two interesting open problems.

2. The Finite Levels of the Ershov Hierarchy

The finite levels of the Ershov hierarchy cover those sets which are the symmetric difference of finitely many r.e. sets. This class coincides with the sets which are bounded truth-table reducible to the halting problem $K = \{e : \varphi_e(e)\downarrow\}$.

2.1. The basic definitions.

Let us begin with the well-known Limit Lemma which characterizes the class of sets Turing reducible to K as the class of the limit-recursive sets.

Theorem 2.1 (Shoenfield). *Let $A \subseteq \omega$. The following are equivalent:*

- *(a) $A \leq_T K$.*
- *(b) A is Δ^0_2.*
- *(c) There is a recursive function $f : \omega \times \omega \to \{0, 1\}$ such that*
  - *(1) for all $x \in \omega$, $f(x, 0) = 0$;*
  - *(2) for all $x \in \omega$, $\lim_{s\to\infty} f(x, s) = A(x)$, here we identified A with its characteristic function.*

Definition 2.2. *Let n be a fixed natural number. If a set $A \subseteq \omega$ satisfies condition (c) in Limit Lemma together with*

(3) for all x, $|\{s : f(x, s + 1) \neq f(x, s)\}| \leq n$

*then A is called an n-*r.e. *set.*

The classes of n-r.e. sets form the finite levels of Ershov hierarchy. Clearly a 1-r.e. set is just a r.e. set in the usual sense. A 2-r.e. set can be expressed as the difference of two r.e. sets, hence they are often called *d-r.e.* sets. In fact, any n-r.e. set can be expressed as the symmetric difference of n recursively enumerable sets. This is the reason why finite levels of the Ershov hierarchy is also called the *difference hierarchy*. An easy diagonalization argument shows that if $n < m$ then there is an m-r.e. set which is not n-r.e. Consequently the hierarchy is proper.

Traditionally when n-r.e. sets were studied, the so-called ω-r.e. sets were normally included also, even if strictly speaking they do not belong to the finite levels of the Ershov hierarchy. In section 4, we will look at the same class of sets again but with a different definition. Below is the definition of ω-r.e. sets used in this section.

Definition 2.3. *Let $g : \omega \to \omega$ be a fixed recursive function. If a set A satisfies condition (c) in Limit Lemma together with*

(3) for all x, $|\{s : f(x, s + 1) \neq f(x, s)\}| \leq g(x)$

*then A is called an ω-*r.e. *set.*

Historically the notion of n-r.e. sets was first introduced by Putnam [23] and Gold [15] in the middle of 1960's. It should be noted that we do not state the definitions as in the original literature.

The notions of n-r.e. and ω-r.e. sets can be naturally extended to degrees. Ershov had studied the m-degrees intensively in his three papers. We will study their Turing degrees instead. Let D_n and D_ω denote the class of n-r.e. and ω-r.e. sets respectively. What we are interested in is the partially ordered structures $\mathcal{D}_n = (D_n, \leq)$ and $\mathcal{D}_\omega = (D_\omega, \leq)$ where $\leq$ is the Turing reducibility. For historical reasons, the notation $\mathcal{R}$ is often used to denote $\mathcal{D}_1$. Clearly the class of n-r.e. and ω r.e. degrees are contained inside $\mathcal{D}(\leq \mathbf{0}')$ which are the Δ^0_2-degrees.

The study of n-r.e. Turing degrees can be traced back to Cooper [7] who showed that the degrees of difference hierarchy do not collapse.

Theorem 2.4 (Cooper)**.** *There is a d-r.e. set D such that for all r.e. set W, $D \not\equiv_T W$.*

The proof of Cooper's theorem can be generalized to show the proper containment of the following nested degree structures.

$$\mathcal{R} \subsetneq \mathcal{D}_2 \subsetneq \cdots \subsetneq \mathcal{D}_\omega \subsetneq \mathcal{D}(\leq \mathbf{0}'). \tag{1}$$

2.2. **Elementary equivalence problems.**

Thus when the number of changes allowed in the approximation increases, we get more and more degrees. One natural question to ask is whether the corresponding partial order structures also become more and more complicated. More precisely:

(a) Are degrees appeared in (1) elementarily equivalent?
(b) If not, at which level does the difference show up?

Recall that two structures $\mathfrak{A}$ and $\mathfrak{B}$ over the some language are called *elementarily equivalent* (written $\mathfrak{A} \equiv \mathfrak{B}$) if for any sentence σ in that language,

$$\mathfrak{A} \models \sigma \quad \text{iff} \quad \mathfrak{B} \models \sigma.$$

If the above holds only for Σ_k-sentences, then we say that $\mathfrak{A}$ is Σ_k-elementarily equivalent to $\mathfrak{B}$, written $\mathfrak{A} \equiv_{\Sigma_k} \mathfrak{B}$ or simply $\mathfrak{A} \equiv_k \mathfrak{B}$.

The elementary differences among structures in (1) would not occur at the Σ_1-level, which is a fact first observed by Sacks.

Theorem 2.5. *For any natural numbers $m, n \geq 1$, $\mathcal{D}_m \equiv_1 \mathcal{D}_n$, $\mathcal{D}_m \equiv_1 \mathcal{D}_\omega$ and $\mathcal{D}_m \equiv_1 \mathcal{D}(\leq \mathbf{0}')$.*

The outline of the proof is as follows: By modifying the proof of Friedberg-Muchnik Theorem, we can embed all countable partial orders into any of the above degrees structures, say $\mathcal{D}_*$. Therefore for any Σ_1-sentence σ, $\mathcal{D}_*$ satisfies σ if and only if σ is consistent with the theory of partial orderings. In other words, σ is true in one of the structures if and only if it is true in any other structures.

The same cannot be said at Σ_2-level. Indeed we are going to quote contrasting pairs of results about degree structures and conclude that most of the degree structures in (1) are not elementarily equivalent and the differences show up at level two.

We begin with separating $\mathcal{D}(\leq \mathbf{0}')$ with the rest.

Theorem 2.6 (Sacks [26], 1961)**.** *There is a minimal degree less than $\mathbf{0}'$.*

In fact the minimal degree built by Sacks is in $\mathcal{D}_\omega$.

Theorem 2.7 (Sacks [27], 1964)**.** *$\mathcal{R}$ is dense.*

Corollary 2.8. *$\mathcal{R} \not\equiv_2 \mathcal{D}_\omega$ and $\mathcal{R} \not\equiv_2 \mathcal{D}(\leq \mathbf{0}')$.*

One can also use the existence of minimal degrees to separate $\mathcal{D}_\omega$ from $\mathcal{D}_n$ by the following theorem.

Theorem 2.9 (Lachlan). *Let $n \geq 2$ be a natural number. For any $\mathbf{d} \in \mathcal{D}_n$, if $\mathbf{d} \neq \mathbf{0}$ then there is $\mathbf{c} \in \mathcal{D}_{n-1}$ such that $\mathbf{0} < \mathbf{c} \leq \mathbf{d}$. Moreover $\mathbf{d}$ is r.e. in $\mathbf{c}$. Consequently for all natural number n, $\mathcal{D}_n$ is downward dense.*

Proof. We prove the case when $n = 2$ to illustrate the idea. Let D be any nonrecursive d-r.e. set. Fix a recursive approximation D_s of D, such that $D_0 = \emptyset$, $D_s(x)$ can at most change twice and at any stage only one element can enter or leave D. Let

$$A = \{s : (\exists x)[x \text{ enters } D \text{ at stage } s \text{ and } \exists t > s(x \notin D_t)]\}.$$

It's easy to verify that A is r.e., $A \leq_T D$ and D is r.e. in A. If D is properly d-r.e., then the set A cannot be recursive, since otherwise D is r.e. in a recursive set, contradicting D being properly d-r.e. To finish the proof of downward density, if D is not properly d-r.e., then it follows from Sacks Density Theorem. □

Corollary 2.10. *For any natural number $n \geq 1$, $\mathcal{D}_n \not\equiv_2 \mathcal{D}_\omega$ and $\mathcal{D}_n \not\equiv_2 \mathcal{D}(\leq 0')$.*

Next we separate $\mathcal{R}$ from $\mathcal{D}_n$ ($n \geq 2$). The clearest distinction is the nondensity theorem of n-r.e. degrees:

Theorem 2.11 (Cooper, Harrington, Lachlan, Lempp and Soare [8], 1991). *For every natural number $n \geq 2$ there is a maximal element in $\mathcal{D}_n$.*

The technical theorem they showed is actually stronger:

Theorem 2.12. *There is a d-r.e. degree $\mathbf{d}$ such that no ω-r.e. degree $\mathbf{a}$ can have $\mathbf{d} < \mathbf{a} < \mathbf{0}'$.*

Corollary 2.13. *For every natural number $n \geq 2$,*

$$\mathcal{R} \not\equiv_2 \mathcal{D}_n.$$

It also gives another instance of

$$\mathcal{R} \not\equiv_2 \mathcal{D}_\omega.$$

Furthermore, observe that for any Δ_2-set $D <_T K$, K is r.e. in D. Thus by relativizing Sacks Density Theorem, we see that $\mathcal{D}(\leq 0')$ is upward dense, thus we can separate $\mathcal{D}_n$ and $\mathcal{D}_\omega$ from $\mathcal{D}(\leq \mathbf{0}')$.

Corollary 2.14. *For each natural number $n \geq 2$*

$$\mathcal{D}_n \not\equiv_2 \mathcal{D}(\leq \mathbf{0}') \text{ and } \mathcal{D}_\omega \not\equiv_2 \mathcal{D}(\leq \mathbf{0}').$$

Historically, it was Downey [9] who first demonstrated a Σ_2-difference between $\mathcal{R}$ and $\mathcal{D}_n$ by showing that diamond exists in $\mathcal{D}_n$, in contrast with Lachlan's Nondiamond Theorem [20], which says there is no diamond in $\mathcal{R}$. The result motivated Downey to propose the following Conjecture 2.15:

Conjecture 2.15 (Downey [9]). *For any natural number $n, m \geq 2$, the structures $\mathcal{D}_n$ and $\mathcal{D}_m$ are elementarily equivalent.*

It is said that Downey did not believe that $\mathcal{D}_n$ and $\mathcal{D}_m$ ($n \neq m$) are elementarily equivalent. He proposed his conjecture in this way to challenge people to explore the structural differences among $\mathcal{D}_n$ and $\mathcal{D}_m$ for $n, m > 1$. It was one of the major problems in the area of difference hierarchy.

Recently Arslanov, Kalimullin and Lempp [4] announced a negative solution to Conjecture 2.15.

Theorem 2.16 (Arslanov, Kalimullin and Lempp [4]). $\mathcal{D}_2 \not\equiv_2 \mathcal{D}_3$.

The technical statement of the difference goes as follows.

Theorem 2.17 (Arslanov, Kalimullin and Lempp [4]).

(a) There are 3-r.e. degrees $\mathbf{f} > \mathbf{e} > \mathbf{d} > \mathbf{0}$ *such that any 3-r.e. degree* $\mathbf{u} \leq \mathbf{e}$ *is comparable with* $\mathbf{d}$*, and any 3-r.e. degree* $\mathbf{u}$ *with* $\mathbf{d} \leq \mathbf{u} \leq \mathbf{f}$ *is comparable with* $\mathbf{e}$.

(b) There are no d-r.e. degrees $\mathbf{f} > \mathbf{e} > \mathbf{d} > \mathbf{0}$ *such that any d-r.e. degree* $\mathbf{u} \leq \mathbf{e}$ *is comparable with* $\mathbf{d}$*, and any d-r.e. degree* $\mathbf{u}$ *with* $\mathbf{d} \leq \mathbf{u} \leq \mathbf{f}$ *is comparable with* $\mathbf{e}$.

The question whether $\mathcal{D}_n \equiv_2 \mathcal{D}_m$ for $2 < n < m$ remains open, though Arslanov, Kalimullin and Lempp suggested that their techniques might be generalized to settle it negatively.

2.3. **Elementary substructure problems.**

What remains is whether one structure can be a Σ_1-elementary substructure of the other.

Recall: When $\mathfrak{A}$ is a substructure of $\mathfrak{B}$, there is a finer notion to gauge the structural differences by allowing parameters from the universe of $\mathfrak{A}$. More precisely, let $\mathcal{L}_A$ be the extended language $\mathcal{L} \cup \{\mathbf{a} : a \in A\}$ obtained by adding a constant symbol $\mathbf{a}$ for each element a in A (here A denotes the universe of $\mathfrak{A}$).

Definition 2.18. *Let n be a natural number. We say that $\mathfrak{A}$ is a Σ_n-*substructure *of $\mathfrak{B}$, written $\mathfrak{A} \preccurlyeq_n \mathfrak{B}$, if for all Σ_n-formulas $\varphi(x_1, x_2, ..., x_m)$ and all $a_1, a_2, ..., a_m \in A$,*

$$\mathfrak{A} \models \varphi(\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_m) \text{ if and only if } \mathfrak{B} \models \varphi(\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_m).$$

In Ershov hierarchy the natural question is: Can one structure in (1) be a Σ_1-elementary substructure of another?

The first breakthrough related to the Σ_1-elementary substructure problem was obtained by Slaman [30] in 1983.

Theorem 2.19 (Slaman).

(i) There are r.e. sets A, B and C and a Δ^0_2-set X such that
- *(a)* $\emptyset <_T X \leq_T A$;
- *(b)* $C \not\leq_T B \oplus X$;
- *(c) for all r.e. set W,* $\emptyset <_T W \leq_T A \Rightarrow C \leq_T B \oplus W$.

(ii) For each natural number $n \geq 1$, $\mathcal{D}_n \not\preccurlyeq_1 \mathcal{D}(\leq \mathbf{0}')$.

The r.e. sets A, B and C in Theorem 2.19 form the so-called *Slaman triple*. If one tries to use Slaman triples to separate $\mathcal{R}$ and $\mathcal{D}_2$, one immediately runs into the following difficulty. By Theorem 2.9 of Lachlan, every nonrecursive n-r.e. degree bounds a nonrecursive r.e. degree, one cannot hope that any n-r.e. degree D plays the role of X as in Theorem 2.19. In 2006, Yang and Yu [31] solved the difficulty by introducing another parameter to control the r.e. degrees below the degree of D. The technical statement is as follows:

Theorem 2.20 (Yang and Yu).

(i) There are r.e. sets A, B, C and E and a d.r.e. set D such that
- *(a)* $D \leq_T A$ *and* $D \not\leq_T E$;
- *(b)* $C \not\leq_T B \oplus D$;
- *(c) for all r.e. set W, if $W \leq_T A$ then either $C \leq_T B \oplus W$ or $W \leq_T E$.*

(ii) For each natural number $n \geq 2$, $\mathcal{R} \not\preccurlyeq_1 \mathcal{D}_n$.

Shore and Slaman (2007) have announced that $\mathcal{D}_m \not\preccurlyeq_1 \mathcal{D}_n$ for natural numbers $m < n$, thus completely settled the Σ_1-elementary equivalence problem for finite levels of the Ershov hierarchy. However, many questions on transfinite levels of the Ershov hierarchy are still open.

3. Ordinal Notations

In this section, we remind the readers some basic definitions and facts about ordinal notations, which are necessary for the development of the Ershov hierarchy to transfinite levels. We will only prove Markwald Theorem (Theorem 3.5) in detail, because its technique will be used in Section 4. Other materials can be found in standard reference books such as Rogers [24], Sacks [25] and Ash and Knight [5].

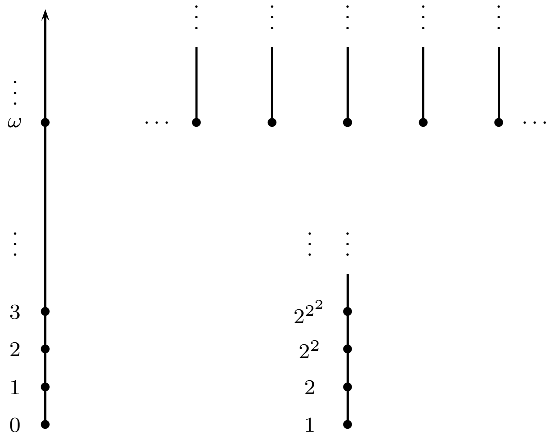
3.1. Kleene's $\mathcal{O}$.

We define a set of notations $\mathcal{O} \subset \omega$, a function $|\_\_|_{\mathcal{O}}$ taking each $a \in \mathcal{O}$ to an ordinal $\alpha = |a|_{\mathcal{O}}$ and a strict partial ordering $<_{\mathcal{O}}$ on $\mathcal{O}$ simultaneously. The elements of $\mathcal{O}$ are the notations and $|a|_{\mathcal{O}}$ is the ordinal represented by the notation a.

- We let 1 be the notation for 0; that is, $|1|_{\mathcal{O}} = 0$.
- If a is a notation for α, then 2^a is a notation for $\alpha + 1$, i.e., $|2^a|_{\mathcal{O}} = \alpha + 1$. Let $b <_{\mathcal{O}} 2^a$ if $b <_{\mathcal{O}} a$ or $b = a$.
- If φ_e is a total recursive function such that for each $n \in \omega$, we have defined $|\varphi_e(n)|_{\mathcal{O}} = \alpha_n$ and $\varphi_e(n) <_{\mathcal{O}} \varphi_e(n+1)$, then $3 \cdot 5^e$ is a notation for $\alpha = \lim_n \alpha_n$, i.e., $|3 \cdot 5^e|_{\mathcal{O}} = \alpha$. Let $b <_{\mathcal{O}} 3 \cdot 5^e$ if there exists some n such that $b <_{\mathcal{O}} \varphi_e(n)$.

That finishes the definition of our system of notations.

The initial part of $\mathcal{O}$ looks like the figure below, where the vertical line to the left indicates the ordinal line:



We will also need the following results on effective addition on $\mathcal{O}$.

Lemma 3.1. *There is a 2-place total recursive function $+_{\mathcal{O}}$ such that for all $a, b \in \mathcal{O}$,*

$$|a +_{\mathcal{O}} b|_{\mathcal{O}} = |a|_{\mathcal{O}} + |b|_{\mathcal{O}}.$$

3.2. **Constructive and recursive ordinals.**

Definition 3.2. *The ordinals having notations in $\mathcal{O}$ are called* constructive ordinals*.*

Since $\mathcal{O}$ is countable, there are countably many constructive ordinals. The first nonconstructive ordinal is denoted by ω_1^{CK} for Church-Kleene.

Definition 3.3. *An ordinal α is* recursive *if it is finite or it is isomorphic to some recursive well-ordering of ω.*

Using Recursion Theorem and transfinite induction up to ω_1^{CK}, one can show that

Theorem 3.4. *Fix an enumeration of all r.e. sets $\{W_e : e \in \omega\}$.*

(1) There exists a total recursive function p such that for all b in $\mathcal{O}$,

$$W_{p(b)} = \{a : a <_{\mathcal{O}} b\}.$$

(2) There exists a total recursive function q such that for all b in $\mathcal{O}$,

$$W_{q(b)} = \{\langle u, v\rangle : u <_{\mathcal{O}} v <_{\mathcal{O}} b\}.$$

Therefore all constructive ordinals are recursive because r.e. linear orderings are recursive. The converse is also true:

Theorem 3.5 (Markwald)**.** *Every recursive ordinal is constructive.*

The proof follows from the two Lemmas 3.6 and 3.7, which we give detailed proof in a moment. Assuming the two Lemmas, the argument goes as follows. For any recursive ordinal α, first embed it into a recursive well-ordering $\mathfrak{B}$ as specified in Lemma 3.6; then into $\mathcal{O}$ as in Lemma 3.7. Since $|\mathfrak{B}| \geq \alpha$ as ordinals, α is embedded too.

Lemma 3.6. *Given a recursive index for a linear ordering $\mathfrak{A}$, we can find a recursive index for a linear order $\mathfrak{B}$ of type $\omega \cdot (1 + \mathfrak{A}) + 1$ in which we can apply uniform effective procedure to*

(a) determine whether a given element b is first, last, a successor, or a limit point;
(b) for any successor element b, find all the immediate predecessor, and for any limit b, determine an increasing sequence $l(b, i)$ with limit b.

Furthermore, we can effectively determine a recursive index for the embedding that takes a in $\mathfrak{A}$ to the first element of the corresponding copy of ω in $\mathfrak{B}$.

Note if $\mathfrak{A}$ is a linear ordering then $\mathfrak{B}$ is a well-ordering if and only if $\mathfrak{A}$ is a well-ordering. Moreover if $\mathfrak{A}$ is well-ordered of order type α then the order type of $\mathfrak{B}$ is $\beta > \alpha$.

Proof. We may suppose that the universe of $\mathfrak{A}$ is $\omega \setminus \{0\}$ or a finite initial segment. Let $\mathfrak{A}^*$ be the result of adding 0 at the front of $\mathfrak{A}$. Let $\mathfrak{B}$ be the linear order which is the product of $\mathfrak{A}^*$ and ω with lexicographic order, with an extra last point ∞.

Now $\mathfrak{B}$ is recursive and its limit points are exactly $(x, 0)$ for $x > 1$ and ∞. Also its successors are exactly the newly add elements, (x, y) for $y \neq 0$. Hence the predecessor of (x, y) is $(x, y - 1)$.

It remains to show the second part of (b). If b is a limit point in $\mathfrak{B}$, define $b_{i+1} = \mu\langle x, y\rangle[b_i <_{\mathfrak{B}} \langle x, y\rangle <_{\mathfrak{B}} b]$, where the μ-operator searches through the normal order of ω. It's easy to see that for any $b' <_{\mathfrak{B}} b$ there is i such that $b' <_{\mathfrak{B}} b_i <_{\mathfrak{B}} b$. $\square$

Lemma 3.7. *Given an index for a recursive linear ordering $\mathfrak{B}$ with the special properties in Lemma 3.6 (there are first, last element, the sets of successor and limit elements are recursive with recursive witnessing functions and the index for $\mathfrak{B}$ tells us how to compute them), we can find a recursive index for a function g defined on all of $\mathfrak{B}$ such that*

(a) if $\operatorname{pred}(x)$, *which is the set $\{y : y <_{\mathfrak{B}} x\}$, is well-ordered, then $g(x)$ is a notation in $\mathcal{O}$ for the height x;*

(b) if $\operatorname{pred}(x)$ *is not well-ordered then $g(x) \notin \mathcal{O}$.*

Proof. We begin by defining a partial recursive function f such that if e is a recursive index for a function with the behavior we want for g on $\operatorname{pred}(x)$ (thinking of $\mathfrak{B}$ as well-ordered), then $f(e, x)$ has the value we want for $g(x)$. We consider three cases.

Case 1. If x is the first element of $\mathfrak{B}$, then $f(e, x) = 1$.

Case 2. If x is the successor of y in $\mathfrak{B}$, then $f(e, x) = 2^{\varphi_e(y)}$.

Case 3. If x is a limit element of $\mathfrak{B}$ and $l(x, i)$ is the recursive increasing sequence with limit b defined in the proof of Lemma 3.6, let $h(e, b)$ be the recursive function given by s-m-n Theorem such that

$$\varphi_{h(e,x)}(i) = \varphi_e(l(e, i)),$$

then $f(e, x) = 3 \cdot 5^{h(e,x)}$.

Note that $f(e, x)$ is defined for all e and all $x \in \mathfrak{B}$, even when $\mathfrak{B}$ is not well-ordered or the function with index e is badly behaved. Now $f(e, x) = \varphi_{k(e)}(x)$ for some k, by Recursion Theorem, there is an n such that for all $x \in \mathfrak{B}$,

$$\varphi_n(x) = \varphi_{k(n)}(x) = f(n, x).$$

Let $g(x) = \varphi_n(x)$. We show that g is what we wanted. Note that $g(x)$ is defined for all $x \in \mathfrak{B}$ since f is.

Claim 1. If $\operatorname{pred}(x)$ is well-ordered and x has height α, then $|g(x)|_{\mathcal{O}} = \alpha$.

Claim 1 can be proven by transfinite induction on α.

Claim 2. If $g(x) \in \mathcal{O}$, then $\operatorname{pred}(x)$ is well-ordered in $\mathfrak{B}$.

Claim 2 can be proven by showing that whenever $|a|_{\mathcal{O}} = \alpha$ and $g(x) = a$ then x has height α in $\mathfrak{B}$. This is done by transfinite induction on α, here we used the definition of f and the choice of n. $\square$

3.3. $\mathcal{O}$ is Π^1_1-complete.

Recall that a set $X \subseteq \omega$ is called Σ^1_1 if $X = \{z : \exists f \forall y R(f, y, z)\}$ where R is recursive and the quantifiers $\exists f$ and $\forall y$ are ranging over functions and numbers respectively. X is Π^1_1 if the complement of X is Σ^1_1. X is Δ^1_1 if X is both Σ^1_1 and Π^1_1. The connection between Π^1_1-sets and well-orderings can be seen below. Kleene [19] showed that $\mathcal{O}$ is Π^1_1-complete and hence $\mathcal{O}$ is not Σ^1_1.

Theorem 3.8 (Kleene [19])**.**

(1) Every Π^1_1-set is many-one reducible to $\mathcal{O}$.

(2) $\mathcal{O}$ is Π^1_1.

Theorem 3.9 (Kleene Boundedness Theorem)**.** *Suppose $X \subseteq \mathcal{O}$ and X is Σ^1_1. Then $X \subseteq \mathcal{O}_b$ for some $b \in \mathcal{O}$, where $\mathcal{O}_b = \{a : a \in \mathcal{O} \wedge |a|_{\mathcal{O}} < |b|_{\mathcal{O}}\}$.*

Note that in Kleene's $\mathcal{O}$, each finite ordinal has a unique notation, while for an infinite ordinal if there is one notation, then there are infinitely many. In other words, the notations of many ordinals are not unique. This is a consequence of the fact that $\mathcal{O}$ is some type of universal structure in which every well-ordering with possibilities to identify successors and limit-ordinals can be embedded. Harrison [17] showed that one can have all recursive ordinals without losing uniqueness.

Theorem 3.10 (Harrison [17])**.** *There is a recursive linear ordering $\sqsubset$ on the natural numbers such that for every recursive ordinal there is exactly one representative n such that $(\{m : m \sqsubset n\}, \sqsubset)$ is order isomorphic to this ordinal.*

It follows that an initial segment of the natural numbers with the ordering $\sqsubset$ is order isomorphic to ω_1^{CK} but this initial segment does not have a supremum with respect to $\sqsubset$. Note that one can improve this structure in order to get more operations on ordinals to be recursive.

Remark 3.11. *Given Harrison's recursive linear ordering $\sqsubset$, let*

$$L = \{(a_1, a_2, \ldots, a_n) : n \in \omega \wedge a_1 \sqsupset a_2 \sqsupset \ldots \sqsupset a_n\}$$

and

$$\begin{aligned} (a_1, a_2, \ldots, a_n) \sqsubset (b_1, b_2, \ldots, b_m) \ &\Leftrightarrow\ \exists k \leq 1 + \min\{m, n\}\ [\forall \ell \in \{1, 2, \ldots, k-1\} \\ &\qquad [a_\ell = b_\ell \wedge (n < k \leq m \vee (k \leq n, m \wedge a_k < b_k))]]; \\ (a_1, a_2, \ldots, a_n) + (b_1, b_2, \ldots, b_m) \ &=\ (a_1, a_2, \ldots, a_k, b_1, b_2, \ldots, b_m) \\ &\qquad \textit{for the maximal } k \leq n \textit{ with } a_k \leq b_1. \end{aligned}$$

The empty sequence λ is also an element of L and satisfies $\lambda \sqsubset z$ for all $z \in L - \{\lambda\}$ and $\lambda + z = z + \lambda = z$ for all $z \in L$.

One can show that for any $(a_1, a_2, \ldots, a_n) \in L$ where all a_k represent a recursive ordinal α_k in Harrison's given ordering, the set $\{z \in L : z \sqsubset (a_1, a_2, \ldots, a_n)\}$ is order isomorphic to the ordinal $\omega^{\alpha_1} + \omega^{\alpha_2} + \ldots + \omega^{\alpha_n}$. Hence $(L, \sqsubset)$ is a recursive linear ordering where the initial part represents all recursive ordinals and where one can test for the ordinals in the initial segment how they are composed of ω-powers and what the coefficients of the lowest ordinals $\omega^n, \omega^{n-1}, \ldots, \omega^1, \omega^0$ for a fixed constant n are.

This remark shows that Kleene's $\mathcal{O}$ is not the only system where one can represent all recursive ordinals adequately and $(L, \sqsubset)$ has the advantage that it is a linear ordering with unique representatives for the recursive ordinals where various natural operations on the ordinals are recursive.

Somehow, such a system of representatives would have the short-coming that not every notation in $\mathcal{O}$ could be translated into a notation in it effectively and so one could argue that it picks the representatives of the ordinals a bit arbitrarily out of $\mathcal{O}$. Furthermore, as already said, the set of the ordinals in a linear ordering can never by Π^1_1-complete; indeed, such a set is even not many-one hard for the halting problem.

4. Transfinite Levels of the Ershov Hierarchy

We are now ready to extend the Ershov hierarchy to transfinite levels. We begin with extending the notions of n-r.e. sets and degrees to the transfinite levels and then prove some important theorems of the Ershov.

4.1. a-r.e. Sets and Degrees.

In the following, we will define D_a for each $a \in \mathcal{O}$. As we shall see the definition is heavily notation dependent. For example, for ordinals $\alpha \geq \omega^2$ we may have a and b both are notations for α, yet D_a and D_b are different. This is the reason why we do not define α-r.e. sets for recursive ordinals α, instead we can only define a-r.e. sets for ordinal notations $a \in \mathcal{O}$. It should also be noted that our phrasing is different from the original definition by Ershov. In fact, there are several different versions of the same notion in the literature.

Definition 4.1. *For each $a \in \mathcal{O}$, a subset A of ω is a-r.e. if and only if there are recursive functions $f : \omega \times \omega \to \{0, 1\}$ and $o : \omega \times \omega \to \mathcal{O}$ such that*

(1) For all x, $f(x, 0) = 0$ and $o(x, 0) <_{\mathcal{O}} a$.
(2) For all x and s, $o(x, s + 1) \leq_{\mathcal{O}} o(x, s)$.
(3) For all x and s, if $f(x, s + 1) \neq f(x, s)$ then $o(x, s + 1) \neq o(x, s)$.
(4) For all x, $\lim_s f(x, s) = A(x)$.

we use D_a to denote the class of all a-r.e. sets.

It is easy to see that a set A is ω-r.e. as in Definition 2.3 if and only if A is a-r.e. for some a such that $|a|_{\mathcal{O}} = \omega$.

There is slight inconsistency of notations between the finite and transfinite levels. In this section, we only talk about a-r.e. sets where a is a notation of an infinite ordinal α.

It is easy to see that for all $a \in \mathcal{O}$ there is a a-r.e. set which is not b-r.e. for any $b <_{\mathcal{O}} a$. Selivanov [28] transferred this result to Turing degrees.

Theorem 4.2 (Selivanov [28]). *For all $a \in \mathcal{O}$, there is a set $X \in D_a$ such that for all $Y \in \bigcup\{D_b : b <_{\mathcal{O}} a\}$, $X \not\equiv_T Y$.*

The major steps of the proof can be outlined as follows: For each fixed notation a, by Theorem 3.4 we can effectively enumerate all $\{b : b <_{\mathcal{O}} a\}$. For each fixed b we can effectively enumerate all sets in D_b. Thus we can effectively enumerate all sets in $\bigcup\{D_b : b <_{\mathcal{O}} a\}$. Once we have the enumeration, we can apply Cooper's technique used in the proof of Theorem 2.4 to find an $X \in D_a$ such that $Y \not\equiv_T Y$.

Let $\mathcal{D}_a$ to denote class of a-r.e. Turing degrees. Selivanov's Theorem tells us that for any fixed notation $a \in \mathcal{O}$, the set $\bigcup\{\mathcal{D}_b : b <_{\mathcal{O}} a\}$ will not exhaust all Δ^0_2-degrees.

4.2. Ershov's Theorems.

We now present some interesting theorems by Ershov. We begin with the one which roughly says all Δ^0_2-sets have appeared at level ω^2.

Theorem 4.3 (Ershov). *Any Δ^0_2-set A is a-r.e. for some $a \in \mathcal{O}$ with $|a|_{\mathcal{O}} = \omega^2$.*

Proof. Let A be a Δ^0_2-set with recursive approximation $f(x, s)$ as stated in the Limit Lemma. We need to find a notation $a \in \mathcal{O}$ and the corresponding $o(x, s) : \omega \times \omega \to a$ for $f(x, s)$. We do it in two steps.

Step 1: Obtain a recursive well-ordering of order type ω which "records" the changes of the approximation of f. The main idea informally is: When $f(x, s + 1) = f(x, s)$ we put the pair $(x, s + 1)$ after (x, s); when $f(x, s + 1) \neq f(x, s)$ we put $(x, s + 1)$ before all other pairs (x, y) where $y \leq s$. Fix an recursive ω-ordering $<$ on $\omega \times \omega$ as in the pairing function:

$$(0,0) < (0,1) < (1,0) < (0,2) < (1,1) < (2,0) < \dots$$

Define $\prec$ on $\omega \times \omega$ by recursion with respect to $<$: Suppose we have done all pairs before (x, y) with respect to the $<$-order. Assume the numbers between $(x-1, 0)$ and $(x, 0)$ are

$$(x-1,0) \prec (u_1, v_1) \prec \dots \prec (u_k, v_k) \prec (x, y_i) \prec (x, y_{i-1}) \prec \dots \prec (x, 0)$$

where $u_i < x$ for all $i \in \{1, 2, \dots, k\}$. Consider the pair (x, y). If $y = 0$ or ($y \neq 0$ and $f(x, y) = f(x, y-1)$), then put (x, y) in the list following the same $<$-order. If $y \neq 0$ and $f(x, y) \neq f(x, y-1)$ then insert (x, y) between (u_k, v_k) and (x, y_i).

Clearly $\prec$ is recursive. Since A is Δ^0_2, $\prec$ is a well-order of order type ω.

Step 2. We obtain a notation for $\prec$ by following the proof of Lemmas 3.6 and 3.7, the process of getting the $\mathfrak{B}$ as in Lemma 3.6 gives us the extra factor ω.

Let g be the embedding of $(\omega \times \omega, \prec)$ into $\mathcal{O}$. Then $g : \omega \times \omega \to a$ for some a with $|a|_{\mathcal{O}} = \omega^2$. Define $o : \omega \times \omega \to a$ by $o(x, 0) = g(x, 0)$ and $o(x, y+1) = o(x, y)$ if $f(x, y+1) = f(x, y)$; and $o(x, y+1) = g(x, y+1)$ if $f(x, y+1) \neq f(x, y)$. Then o is what we wanted. $\square$

The next theorem shows that ω^2 is the first level where one can exhaust all Δ^0_2-sets.

Theorem 4.4 (Ershov). *There is a Δ_2-set A such that $A \not\equiv_T B$ for any $B \in \bigcup\{D_b : b \in \mathcal{O} \wedge |b|_{\mathcal{O}} < \omega^2\}$.*

Proof. We first show two claims which tell us that D_α makes sense for $\alpha < \omega^2$. The point is that there are only finitely many limit points below α, thus we can code the information (nonuniformly).

Claim 1. If $a, b \in \mathcal{O}$ and $|a|_{\mathcal{O}} = |b|_{\mathcal{O}} = \alpha < \omega^2$, then there is a partial recursive function ϕ such that $\phi \restriction \{k : k <_{\mathcal{O}} a\}$ is an isomorphism from $\{k : k <_{\mathcal{O}} a\} \to \{k : k <_{\mathcal{O}} b\}$.

Claim 1 can be proven as follows: Suppose $\alpha = \omega \cdot i + j$ for some $i, j \in \omega$. Let $n_1, \dots, n_i$ and $m_1, \dots, m_i$ be the notations of $\omega, \dots, \omega \cdot i$ below a and b respectively. Define ϕ such that $\phi(n_l) = m_l$ and naturally extend to the notations of successor ordinals.

Claim 2. If $a, b \in \mathcal{O}$ and $|a|_{\mathcal{O}} = |b|_{\mathcal{O}} = \alpha < \omega^2$, then $D_a = D_b$.

For a proof of Claim 2, assume that f and o_1 are the witnessing functions for a set A being in D_a. Then f and $o_2 = \phi \circ o_1$ witness A being in D_b, where ϕ is the isomorphism in Claim 1.

Finally fix any notation a_0 such that $|a_0|_{\mathcal{O}} = \omega^2$. By Selivanov Theorem, there is $A \in D_{a_0}$ such that for any $B \in \bigcup\{D_a : a <_{\mathcal{O}} a_0\}$, $A \not\equiv_T B$. By Claim 2, diagonalizing against all $\{D_a : a <_{\mathcal{O}} a_0\}$ is as good as diagonalizing against all $\{D_a : |a|_{\mathcal{O}} < \omega^2\}$. Thus this A is what we wanted. $\square$

Since there are uncountably many paths of height $\geq \omega^2$ on $\mathcal{O}$, and $\mathcal{O}$ is countable, there are many paths of height ω^2. We now study the classes D_a when a belongs to some path T on $\mathcal{O}$.

Theorem 4.5 (Ershov).

(a) *There is a path $T \subset \mathcal{O}$ with $|T| = \omega^3$ such that $A \in \Delta^0_2$ if and only if $A \in D_a$ for some $a \in T$.*

(b) *For each path $T \subset \mathcal{O}$, if $|T| < \omega^3$, then there is some Δ^0_2-set A such that A is not in $\bigcup\{D_a : a \in T\}$.*

Proof. We just give a proof sketch. For (a), list all Δ_2-sets $\{X_i : i \in \omega\}$ and choose a notation a_i such that $|a_i|_{\mathcal{O}} = \omega^2$ and $X_i \in \mathcal{D}_i$. a_i exists by Theorem 4.3. Note that the listing and choice are done noneffectively. Then $T = \{a_0 +_{\mathcal{O}} \cdots +_{\mathcal{O}} a_n : n \in \omega\}$ is the path that we wanted.

For (b), first observe that for any given notation a_0, we can mimic the proof of Theorem 4.4 to get a notation a_1 with $|a_1|_{\mathcal{O}} = \omega^2$ and a set $A \in D_{a_0 +_{\mathcal{O}} a_1}$ which is not in

$$\bigcup\{D_a : a_0 <_{\mathcal{O}} a \text{ and } |a_0 +_{\mathcal{O}} a|_{\mathcal{O}} < |a_0|_{\mathcal{O}} + \omega^2\}.$$

Suppose that $|T| = \omega^2 \cdot i + \alpha$ for some $i \in \omega$, $\alpha < \omega^2$. If $\alpha \neq 0$, let $a_0 \in T$ be such that $|a_0|_{\mathcal{O}} = \omega^2 \cdot i$; if $\alpha = 0$, then $i \geq 1$ by Theorem 4.4, let $a_0 \in T$ be a notation such that $|a_0|_{\mathcal{O}} = \omega^2 \cdot (i-1)$. Apply the above observation to a_0, the result follows. □

One can also get the following result about Δ^0_2-sets and long paths through $\mathcal{O}$.

Theorem 4.6 (Ershov)**.**

(a) *There is an unbounded path T through $\mathcal{O}$ such that for any Δ^0_2-set A, there is an $a \in T$ such that $A \in D_a$. Furthermore, this T can be chosen recursive in $\mathcal{O}$.*

(b) *Such a path cannot be Π^1_1.*

Proof. Fix an enumeration of Δ^0_2-sets A_n, we build a path T recursively in $\mathcal{O}$ in stages as follows: At stage s, suppose we have had $T_s \in \mathcal{O}$. Let $a_s \in \mathcal{O}$ be such that $|a_s| = \omega^2$ and $A_s \in D_{a_s}$. Let

$$T_{s+1} = \begin{cases} T_s +_{\mathcal{O}} a_s +_{\mathcal{O}} s, & \text{if } s \in \mathcal{O}; \\ T_s +_{\mathcal{O}} a_s, & \text{otherwise.} \end{cases}$$

Clear $T = \bigcup_s \{a : a <_{\mathcal{O}} T_s\}$ is what we wanted.

For (b), suppose that T is Π^1_1. Define a function $h : \omega \to T$ by $h(n) =$ the $<_{\mathcal{O}}$-least $a \in T$ such that A_n appears in D_a. Then h is a total Π^1_1-function, therefore Δ^1_1. By Theorem 3.9 of Kleene, the range of h is bounded, contradiction. □

5. Applications to Inductive Inference

In inductive inference, transfinite ordinals were initially used to generalize the notion of counters. The general setting is that a learner M — which in the following discussion is always assumed to be recursive — reads from an infinite tape more and more data and eventually outputs a finite sequence of indices of r.e. sets such that for the last index e in this sequence it holds that W_e enumerates exactly all the elements the learner finds on the tape [16]. A class of r.e. sets is explanatorily learnable iff there is a single learner M which learns every set in the class from every possible way in which its elements can be written down on the tape; such tape contents are called texts in learning theory. Not every class is learnable.

Theorem 5.1 (Gold [16])**.** *The class of the set of all natural numbers plus all of its finite subsets is not explanatorily learnable.*

In his paper, Gold introduced the initial definitions and results of inductive inference which were then the basis for work which spans now already more than four decades. Freivalds and Smith [13] investigated restrictions on the number of mind changes or hypotheses by adding to the learner a counter [13]. This counter takes initially a fixed value $a \in \mathcal{O}$ before starting the learning process and outputting any hypothesis; then, whenever the learner outputs a hypothesis, the current value in the counter, say b, is replaced by a new value c with $c <_{\mathcal{O}} b$; if such a c does not exist, the learner does not output a new conjecture. Note that $b, c \leq_{\mathcal{O}} a$ by induction. Such a learner can be called an a-shot learner, as a bounds the number of attempts the learner can make to output a hypothesis. Note that it is enforced that an a-shot learner outputs only finitely many conjectures. There are learnable classes which cannot be learnt by any a-shot learner for any $a \in \mathcal{O}$. An example of such a class is the class of all finite sets. A characterization for such classes is based on the following notion: Osherson, Stob and Weinstein [22] called a learner M confident iff M converges on every text, even on those texts which are for languages which do not belong to the class to be learnt or which are even not recursively enumerable.

Theorem 5.2 (Ambainis, Freivalds and Smith [2]). *A class $\mathcal{L}$ of r.e. sets is confidently learnable iff there is an $a \in \mathcal{O}$ such that $\mathcal{L}$ is learnable by an a-shot learner.*

Sharma, Stephan and Ventsov [29] considered generalizations of this result and also characterized when a class is learnable by a learner converging on all recursive texts and so on. Freivalds and Smith [13] showed that the hierarchy of learning from ordinal notations does not collapse.

Theorem 5.3 (Freivalds and Smith [13]). *For every notation $a \in \mathcal{O}$ there is a class which can be learnt by an a-shot learner but not by a b-shot learner whenever $b \in \mathcal{O}$ and $|b|_{\mathcal{O}} < |a|_{\mathcal{O}}$.*

Proof. This can be proven by using the class of all sets $L_c = \{x : c \leq_{\mathcal{O}} x <_{\mathcal{O}} a\}$ with $c <_{\mathcal{O}} a$. The learner for this class is quite straight-forward: whenever it has seen a finite set D of data, it computes the minimum c of D with respect to $<_{\mathcal{O}}$ and it conjectures L_c. The ordinal counter is set to c. As any new hypothesis d after c would satisfy that $L_d \supset L_c$ and $d = \min(L_d) <_{\mathcal{O}} c = \min(L_c)$, the ordinal counter is counted down whenever the hypothesis is revised. On the other hand, if the ordinal counter would start with $c = b$ instead of a, one could feed the learner always with data for the L_c until it outputs a conjecture for L_c and counts down the ordinal from c to d; then one starts to feed more data into the learner which are then the data for the set L_d. By this method, one always feeds data for a set the learner is not conjecturing and when it revises the conjecture, the set of the data fed into the learner gets increased. This process can go on only finitely often and will get stuck in the learner conjecturing a different set than from which the data are; hence the learner would not succeed. It follows that the class of all L_c with $c <_{\mathcal{O}} a$ is learnable by an a-shot learner but not by a b-shot learner. $\square$

Note that the class $\{L_c : c <_{\mathcal{O}} a\}$ is even uniformly recursively enumerable. Ordinal counters serve as a measure of complexity in various areas of inductive inference.

There is also a connection to a-r.e. sets. Given an a-r.e. set A with $|a|_{\mathcal{O}} \geq \omega$ and an approximation f as defined in Definition 4.1, let

$$\begin{aligned} L_{x,y} &= \{\langle x,0\rangle, \langle x,1\rangle, \ldots, \langle x,y\rangle\} \text{ and} \\ \mathcal{L} &= \{L_{x,y} : x, y \in \omega \wedge (y = 0 \vee f(x, y-1) \neq f(x,y))\}. \end{aligned}$$

Now $\mathcal{L}$ is learnable by an a-shot learner. Furthermore, whenever $\mathcal{L}$ is learnable by a b-shot learner then A is a b-r.e. set. The converse does not hold, as one could make the approximation of $A = \emptyset$ having so many changes that it does not even witness that $\emptyset$ is ω-r.e.; the corresponding class has then no ω-shot learner although A is an ω-r.e. set.

Carlucci, Case and Jain [6] called approximations as the f in Definition 4.1 an a-correction-grammar. They considered the learning of classes of r.e. languages where the learner is permitted to output conjectures which are a-correction-grammars for the set to be learnt. In their work, they considered more general learners than before, namely explanatory learners (as already defined above) and behaviourally correct learners. Behaviourally correct learners are more general than explanatory learners: given a text for a set L to be learnt, the behaviourally correct learner outputs an infinite sequence $e_0, e_1, e_2, \ldots$ of a-correction-grammars such that almost all e_n are an a-correction grammar for L; but, in contrast to explanatory learning, no syntactic convergence is required and it is permtitted that there are infinitely many n with $e_{n+1} \neq e_n$. A behaviourally correct learner learns a whole class $\mathcal{L}$ iff it learns every $L \in \mathcal{L}$ this way. Note that behaviourally correct learners are more powerful than explanatory learners. Given an undecidable r.e. set K, for example the halting problem, the class $\{K \cup D : D \text{ is finite}\}$ is behaviourally correctly but not explanatorily learnable. Now Carlucci, Case and Jain [6] investigate how the usage of a-correction-grammars relates to explanatory and behaviourally correct learning.

For explanatory learning, they showed that there is no optimal hypothesis space; more precisely, given ordinal notations $a, b \in \mathcal{O}$ with $a <_{\mathcal{O}} b$ there is a class which is explanatorily learnable using b-correction grammars but not using a-correction grammars.

In the case of behaviourally correct learning, the situation is different and all hypothesis spaces which permit at least "ω-many" corrections are equivalent and optimal: They showed that for all notations $a, b \in \mathcal{O}$ with $|b|_{\mathcal{O}} \geq \omega$ and every class $\mathcal{L}$ of r.e. sets, if the class $\mathcal{L}$ is learnable using a-correction-grammars then $\mathcal{L}$ is also learnable using b-correction grammars.

Other applications of ordinals in learning theory refer to measuring how far the hypothesis can go off the minimal index of the target language [1]. Some papers also abstract from the requirement that learners have to be recursive and can then use ordinals directly in place of their notations [21].

6. Conclusion

In this paper we reviewed the usage of recursive ordinals in recursion and learning theory. We first reviewed the finite levels of the Ershov hierarchy and the degree structures of these sets. Then we turned to the concept of recursive ordinals and ways to represent them, with a particular emphasis on Kleene's $\mathcal{O}$ and notations of ordinals. The notation a used to represent a given recursive ordinal has a strong impact on which limit-recursive sets are a-r.e. and hence the world of a-r.e. sets is not only a study of recursive ordinals but also of notations of recursive ordinals. We summarize the ground-breaking results of Ershov and show some selected results on transfinite levels of the Ershov hierarchy. At the end we survey the usage of ordinals in inductive inference, where the notion of a recursive ordinal turned out to be a natural method to measure convergence speed and learning complexity.

We end our survey with two open problems. The first one can be viewed as a generalized version of Downey's Conjecture:

Conjecture 6.1. *For all $a \neq b \in \mathcal{O}$, if $\omega \leq |a|_{\mathcal{O}}, |b|_{\mathcal{O}} < \omega^2$ then $\mathcal{D}_a \equiv \mathcal{D}_b$.*

We know that $\bigcup\{\mathcal{D}_a : a \in \mathcal{O} \wedge |a|_{\mathcal{O}} < \omega^2\}$ is a proper subset of all Δ^0_2-degrees by Theorem 4.4, however, is there any degree theoretic difference?

Conjecture 6.2. $\bigcup\{\mathcal{D}_a : a \in \mathcal{O} \wedge |a|_{\mathcal{O}} < \omega^2\} \equiv \mathcal{D}(\leq \mathbf{0}')$.

Acknowledgments. We would like to thank John Case, Sanjay Jain and Theodore A. Slaman for fruitful discussions and information about the topics of this paper.

References

[1] Andris Ambainis, John Case, Sanjay Jain and Mandayam Suraj. Parsimony hierarchies for inductive inference. *The Journal of Symbolic Logic*, 69(1):287–327, 2004.

[2] Andris Ambainis, Rūsiņš Freivalds and Carl H. Smith. Inductive inference with procrastination: back to definitions. *Fundamenta Informaticae*, 40(1):1–16, 1999.

[3] Marat M. Arslanov. The Ershov hierarchy. *Computation and Logic in the Real World*, S.B. Cooper and A. Sorbi, eds., to appear.

[4] Marat M. Arslanov, Iskander Sh. Kalimullin and Steffen Lempp. On Downey's conjecture. Preprint, 2005.

[5] Christopher J. Ash and Julia Knight. *Computable structures and the hyperarithmetical hierarchy*, volume 144 of *Studies in Logic and the Foundations of Mathematics*. North-Holland Publishing Co., Amsterdam, 2000.

[6] Lorenzo Carlucci, John Case and Sanjay Jain. Learning Correction Grammars. *Proceedings of the 20th Annual Conference on Computational Learning Theory*, COLT 2007, *Springer Lecture Notes in Artificial Intelligence*, 4539:203–217, 2007.

[7] S. Barry Cooper. *Degrees of Unsolvability.* PhD thesis, Leicester University, 1971.

[8] S. Barry Cooper, Leo Harrington, Alistair H. Lachlan, Steffen Lempp and Robert I. Soare. The d.r.e. degrees are not dense. *Annals of Pure and Applied Logic*, 55(2):125–151, 1991.

[9] Rodney G. Downey. D-recursively enumerable degrees and the nondiamond theorem. *Bulletin of the London Mathematical Society*, 21:43–50, 1989.

[10] Yuri L. Ershov. A certain hierarchy of sets. I. *Algebra i Logika*, 7(1):47–74, 1968.

[11] Yuri L. Ershov. A certain hierarchy of sets. II. *Algebra i Logika*, 7(4):15–47, 1968.

[12] Yuri L. Ershov. A certain hierarchy of sets. III. *Algebra i Logika*, 9:34–51, 1970.

[13] Rūsiņš Freivalds and Carl H. Smith. On the role of procrastination in machine learning. *Information and Computation*, 107(2):237–271, 1993.

[14] Amy Gale and Rod Downey. On genericity and Ershov's hierarchy. *Mathematical Logic Quarterly*, 47(2):161–182, 2001.

[15] E. Mark Gold. Limiting recursion. *The Journal of Symbolic Logic*, 30(1):28–48, March 1965.

[16] E. Mark Gold. Language identification in the limit. *Information and Control*, 10:447–474, 1967.

[17] Joseph Harrison. Recursive pseudo well-orderings. *Transactions of the American Mathematical Society*, 131:526–543, 1968.

[18] Bakhadyr Khoussainov, Frank Stephan and Yue Yang. Computable categoricity and the Ershov hierarchy. to appear in *Annals of Pure and Applied Logic*.

[19] Stephen C. Kleene. On the forms of the predicates in the theory of constructive ordinals. II. *The American Journal of Mathematics*, 77:405–428, 1955.

[20] Alistair H. Lachlan. Lower bounds for pairs of recursively enumerable degrees. *Proceedings of the London Mathematical Society (3)*, 16:537–569, 1966.

[21] Eric Martin, Arun Sharma and Frank Stephan. On ordinal VC-dimension and some notions of complexity. *Theoretical Computer Science*, 364:62-76, 2006.

[22] Daniel Osherson, Michael Stob and Scott Weinstein. *Systems That Learn, An Introduction to Learning Theory for Cognitive and Computer Scientists.* Bradford — The MIT Press, Cambridge, Massachusetts, 1986.

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

[24] Hartley Rogers, Jr. *Theory of Recursive Functions and Effective Computability.* McGraw-Hill, New York, 1967.

[25] Gerald E. Sacks. *Higher Recursion Theory.* Perspectives in Mathematical Logic. Springer–Verlag, Heidelberg, 1990.

[26] Gerald E. Sacks. A minimal degree below $\mathbf{0}'$. *Bulletin of the American Mathematical Society*, 67:416–419, 1961.

[27] Gerald E. Sacks. The recursively enumerable degrees are dense. *Annals of Mathematics*, 80:300–312, 1964.

[28] Victor L. Selivanov. Ershov's hierarchy and T-jumps. *Algebra i Logika*, 27(4):464–478, 498–499, 1988.

[29] Arun Sharma, Frank Stephan and Yuri Ventsov. Generalized notions of mind change complexity. *Information and Computation*, 189:235-262, 2004.

[30] Theodore A. Slaman. The recursively enumerable degrees as a substructure of the Δ^0_2-degrees. Hand written notes, 1983.

[31] Yue Yang and Liang Yu. On Σ_1-structural differences among finite levels of the Ershov hierarchy. *The Journal of Symbolic Logic*, 71(4):1223–1236, 2006.

Frank Stephan, Department of Mathematics, Faculty of Science, National University of Singapore, 2 Science Drive 2, Singapore 117543.
E-mail address: fstephan@comp.nus.edu.sg

Yue Yang, Department of Mathematics, Faculty of Science, National University of Singapore, 2 Science Drive 2, Singapore 117543.
E-mail address: matyangy@nus.edu.sg

Liang Yu, Institute of Mathematical Science, Nanjing University, Nanjing, Jiangsu province 210093, P.R. of China.
E-mail address: yuliang.nju@gmail.com