

# On the computability of optimal Scott sentences

Rachael Alvir, Barbara Csima, and Matthew Harrison-Trainer

April 12, 2025

## Abstract

Given a countable mathematical structure, its *Scott sentence* is a sentence of the infinitary logic  $\mathcal{L}_{\omega_1\omega}$  that characterizes it among all countable structures. We can measure the complexity of a structure by the least complexity of a Scott sentence for that structure. It is known that there can be a difference between the least complexity of a Scott sentence and the least complexity of a computable Scott sentence; for example, Alvir, Knight, and McCoy showed that there is a computable structure with a  $\Pi_2$  Scott sentence but no computable  $\Pi_2$  Scott sentence. It is well known that a structure with a  $\Pi_2$  Scott sentence must have a computable  $\Pi_4$  Scott sentence. We show that this is best possible: there is a computable structure with a  $\Pi_2$  Scott sentence but no computable  $\Sigma_4$  Scott sentence. We also show that there is no reasonable characterization of the computable structures with a computable  $\Pi_n$  Scott sentence by showing that the index set of such structures is  $\Pi_1^1$ - $m$ -complete.

## 1 Introduction

Let  $\mathcal{A}$  be a countable mathematical structure, such a countable graph, group, or ring. Suppose that we want to characterize  $\mathcal{A}$  by writing down a sentence, or theory, which characterizes  $\mathcal{A}$  up to isomorphism. If we work in elementary first-order logic then, as a consequence of compactness, we cannot do this for most countable structures  $\mathcal{A}$ . However, suppose we strengthen our logic to the logic  $\mathcal{L}_{\omega_1\omega}$  which allows countably infinite conjunctions and disjunctions. Scott [Sco65] showed that for any countable structure  $\mathcal{A}$  there is a sentence  $\varphi$  of  $\mathcal{L}_{\omega_1\omega}$  that characterizes  $\mathcal{A}$  up to isomorphism among countable structures, i.e., for all countable  $\mathcal{B}$ ,

$$\mathcal{B} \models \varphi \iff \mathcal{A} \cong \mathcal{B}.$$

We call such a sentence a *Scott sentence* for  $\mathcal{A}$ . This fact implies, for example, that while isomorphism is analytic-complete if we fix a particular  $\mathcal{L}$ -structure  $\mathcal{A}$  then the set  $\{\mathcal{B} \in \text{Mod}(\mathcal{L}) : \mathcal{A} \cong \mathcal{B}\}$  is actually Borel.

The standard proof that every countable structure has a Scott sentence uses through the back-and-forth relations. First, one shows that they must stabilize at some countable ordinal, and then from this one can extract a Scott sentence, called the *canonical Scott sentence*. This Scott analysis of a structure has played an important role in the study of Vaught's conjecture, e.g. in Morley's theorem [Mor70] that the number of non-isomorphic countable models of a theory is either at most  $\aleph_1$  or is exactly  $2^{\aleph_0}$ .

For each particular structure  $\mathcal{A}$ , there is some ordinal  $\alpha$  at which the back-and-forth relations stabilize. This gives a way of assigning an ordinal rank to each countable

structure giving a measure of that structure's complexity. There are numerous non-equivalent definitions of Scott rank depending on the particular back-and-forth relations one chooses or exactly what one means by stabilizing. Within the last ten years the most commonly used has been the (*unparametrized*) *Scott rank* due to Montalbán [Mon15] which measures the least complexity of a Scott sentence for  $\mathcal{A}$ , namely, the Scott rank is the least  $\alpha$  such that  $\mathcal{A}$  has a  $\Pi_{\alpha+1}$  Scott sentence. The following theorem shows that this definition is particularly robust.

**Theorem 1.1** (Montalbán [Mon15]). *Let  $\mathcal{A}$  be a countable structure, and  $\alpha$  a countable ordinal. The following are equivalent:*

- (1)  $\mathcal{A}$  has a  $\Pi_{\alpha+1}$  Scott sentence.
- (2) Every automorphism orbit in  $\mathcal{A}$  is  $\Sigma_\alpha$ -definable.
- (3) The set  $\{\mathcal{B} \in \text{Mod}(\mathcal{L}) : \mathcal{B} \cong \mathcal{A}\}$  of isomorphic copies of  $\mathcal{A}$  is  $\Pi_{\alpha+1}^0$ .
- (4)  $\mathcal{A}$  is (boldface)  $\Delta_\alpha^0$ -categorical.

In computable structure theory, there has been a long history of studying the Scott analysis of computable structures, or more generally how the Scott analysis of a structure relates computationally to the structure. Nadel [Nad74] showed that every computable structure  $\mathcal{A}$  has Scott rank at most  $\omega_1^{\mathcal{A}} + 1$ . Harrison [Har68], Millar and Knight [KM10] (building on work of Makkai [Mak81]), Harrison-Trainor, Igusa, and Knight [HTIK18], Harrison-Trainor [HT18a], and Alvir, Greenberg, Harrison-Trainor, and Turetsky [AGHTT21] constructed various examples of computable structures with non-computable Scott ranks.

In particular, if  $\mathcal{A}$  is computable and has computable Scott rank  $< \omega_1^{CK}$ , then the canonical Scott sentence constructed by Scott will also be a computable sentence. However this construction is not optimal: Alvir, Knight, and McCoy [AKM20] showed that there is a computable structure with a  $\Pi_2$  Scott sentence but with no computable  $\Pi_2$  Scott sentence. It is well-known (see, e.g., Lemma VI.14 of [Mon]) that if a computable structure  $\mathcal{A}$  has a  $\Pi_\alpha$  Scott sentence, with  $\alpha$  computable, then it has a computable  $\Pi_{2\alpha}$  Scott sentence. This is obtained by noting that if  $\mathcal{A}$  has a  $\Pi_\alpha$  Scott sentence  $\varphi$ , then  $\mathcal{A}$  has the property that for countable structure  $\mathcal{B}$ ,

$$\mathcal{A} \leq_\alpha \mathcal{B} \iff \mathcal{A} \cong \mathcal{B}.$$

Writing out the definition of the back-and-forth relations, we can see that there is a computable  $\Pi_{2\alpha}$  sentence  $\psi$  such that

$$\mathcal{B} \models \psi \iff \mathcal{A} \leq_\alpha \mathcal{B}.$$

This  $\psi$  is a computable  $\Pi_{2\alpha}$  Scott sentence.

In this paper, we restrict to the case of structures with  $\Pi_2$  Scott sentences, which are called by Montalbán [Mon17] the  $\exists$ -atomic structures because every automorphism orbit is isolated by a (finitary) existential formula. Even in this seemingly simple class of structures we find that there is significant complexity. As described above, such structures have a computable  $\Pi_4$  Scott sentence; we improve the result of [AKM20] to show that this upper bound is optimal.

**Theorem 1.2.** *There is a computable structure with a  $\Pi_2$  Scott sentence but no computable  $\Sigma_4$  Scott sentence.*

By taking Marker extensions / jump inversions, one can also obtain that for any  $n$  there is a computable structure with a  $\Pi_n$  Scott sentence but no computable  $\Sigma_{n+2}$  Scott sentence. (One can also extend this to non-limit ordinals though we will remain at finite levels since all of the complexity appears there; see Corollary 5.8.) We leave open the question of improving this to an example with no  $\Sigma_{2n}$  Scott sentence.

In the second part of this paper, we consider the *effective Scott rank*, i.e., the least  $\alpha$  such that a given structure  $\mathcal{A}$  has a computable  $\Pi_{\alpha+1}$  Scott sentence (when a computable Scott sentence exists). Whether or not this effective Scott rank is as robust as the (non-effective) Scott rank was an open problem from [AKM20]. It is known that there are a few conditions equivalent to having a formally  $\Sigma_\alpha$  Scott family; a “boldface” version of this result led to many of the equivalences of Theorem 1.1, so it was not unreasonable to expect similar robustness in the effective setting. In fact, it was known that for a computable structure, having a formally  $\Sigma_\alpha$  Scott family implies that the structure has a computable  $\Pi_{\alpha+1}$  Scott sentence. Whether the reverse implication held was left open.

In Section 4, we show that the effective Scott rank is not robust and that there is no good way to characterize when a structure has a computable  $\Pi_n$  Scott sentence, even when the structure is computable.

**Theorem 1.3.** *With  $(\mathcal{A}_i)_{i \in \omega}$  an effective list of (possibly partial) structures in a sufficiently rich language, the set*

$$\{i \mid \mathcal{A}_i \text{ has a computable } \Pi_2 \text{ Scott sentence}\}$$

*is  $\Pi_1^1$ - $m$ -complete.*

By a sufficiently rich language we mean any language which is universal, e.g., those including at least a binary relation symbol. This result was also obtained independently and around the same time by Knight, Lange, and McCoy [KLMC].

Again by taking Marker extensions / jump inversions, one can replace  $\Pi_2$  with  $\Pi_n$ , or through the hyperarithmetic hierarchy, as in Corollary 5.9. One can obtain from this theorem several interesting corollaries which we give in Section 5, for example:

**Corollary 1.4.** *There is a computable structure  $\mathcal{A}$  with a  $\Pi_2$  Scott sentence but no computable  $\Pi_2$  Scott sentence, but with a computable  $\Pi_2$  sentence  $\varphi$  such that, for all hyperarithmetic  $\mathcal{B}$ ,*

$$\mathcal{B} \models \varphi \iff \mathcal{A} \cong \mathcal{B}.$$

We leave open several questions. Though there are further generalizations that one can ask (e.g., to infinite ordinals), we see the two main questions to be the following.

**Question 1.5.** *Is there, for every  $n$ , a computable structure  $\mathcal{A}$  with a  $\Pi_n$  Scott sentence but with no computable  $\Sigma_{2n}$  Scott sentence?*

**Question 1.6.** *Is the set of computable structures with a  $\Pi_2$  Scott sentence and a computable  $\Pi_3$  Scott sentence  $\Pi_1^1$ - $m$ -complete?*

For Question 1.5, recall that the upper bound of  $\Pi_{2n}$  was obtained as follows. Given  $\mathcal{A}$  with a  $\Pi_n$  Scott sentence,  $\mathcal{A} \leq_n \mathcal{B} \implies \mathcal{A} \cong \mathcal{B}$ . We find a  $\Pi_{2n}$  sentence  $\psi$  such that  $\mathcal{B} \models \psi \iff \mathcal{A} \leq_n \mathcal{B}$ . Chen, Gonzalez, and Harrison-Trainor [CGHT] have recently showed

that the set of pairs  $\{(\mathcal{A}, \mathcal{B}) : \mathcal{A} \leq_n \mathcal{B}\}$  is  $\Pi_{2n}^0$ -complete. This means that there is no way of defining the set  $\{\mathcal{B} : \mathcal{A} \leq_n \mathcal{B}\}$  in a way that is both better than  $\Pi_{2n}$  and which is “schematic” in  $\mathcal{A}$ . On the other hand, they showed that the set  $\{\mathcal{B} : \mathcal{A} \leq_n \mathcal{B}\}$  is  $\Pi_{n+2}^0$ . This relies on the fact that every  $\Pi_n$ -type in  $\mathcal{A}$  is  $\Pi_n$ -definable, and so is a non-effective argument. The case of Question 1.5 solved in this paper,  $n = 2$ , satisfies  $n + 2 = 2n = 4$  and the distinction does not yet show up at this level. Thus one might expect Question 1.5 for  $n \geq 3$  to involve some new insight.

## 2 A simplifying remark

Given a structure  $\mathcal{A}$ , we can consider the structure  $\mathcal{A} \cdot \omega$  which consists of an equivalence relation  $E$  with infinitely many equivalence classes, on each of which is a copy of  $\mathcal{A}$ . Then:

- (1) if  $\mathcal{A}$  has a (computable)  $\Pi_\alpha$  Scott sentence, then  $\mathcal{A} \cdot \omega$  will have a (computable)  $\Pi_\alpha$  Scott sentence, and
- (2) if  $\mathcal{A}$  has no (computable)  $\Pi_\alpha$  Scott sentence, then  $\mathcal{A} \cdot \omega$  will have no (computable)  $\Sigma_{\alpha+1}$  Scott sentence.

(1) is straightforward, and (2) uses the fact that  $\mathcal{B}$  has a  $\Sigma_{\alpha+1}$  if and only if there is  $\bar{b} \in \mathcal{B}$  such that  $(\mathcal{B}, \bar{b})$  has a  $\Pi_\alpha$  Scott sentence. (This fact was first stated by Montalbán in [Mon15] and proved in [Mon17, AGHTT21].)

Given the above facts, we do not need to consider  $\Sigma$  Scott sentences, as e.g., the fact that there is a computable structure with a  $\Pi_2$  Scott sentence but no computable  $\Pi_2$  Scott sentence yields that there is a structure with a  $\Pi_2$  Scott sentence but no computable  $\Sigma_3$  Scott sentence. When we prove Theorem 1.2 we will prove that there is a computable structure with a  $\Pi_2$  Scott sentence but no computable  $\Pi_3$  Scott sentence; it will follow that there is a computable structure with a  $\Pi_2$  Scott sentence but no computable  $\Sigma_4$  Scott sentence.

## 3 A computable structure with a $\Pi_2$ Scott sentence but no computable $\Pi_3$ Scott sentence

As a warmup we sketch a new construction of a computable structure with a  $\Pi_2$  Scott sentence but no computable  $\Pi_2$  Scott sentence. Theorems 1.2 and 1.3 will build on this technique and so we begin by presenting it in its simplest form.

**Theorem 3.1** (Alvir, Knight, and McCoy [AKM20]). *There is a computable structure with a  $\Pi_2$  Scott sentence but with no computable  $\Pi_2$  Scott sentence.*

*Proof sketch.* We list all of the computable  $\Pi_2$  sentences as  $(\theta_e)_{e \in \omega}$  where

$$\theta_e = \bigwedge_{i \in \omega} \forall \bar{x}_{e,i} \varphi_{e,i}(\bar{x}_{e,i})$$

where the  $\varphi_{e,i}(\bar{x}_{e,i})$  are computable  $\Sigma_1$  formulas uniformly in  $e, i$  and the arities of each of the  $\bar{x}_{e,i}$  are also computable in the indices.

We can take  $\mathcal{A}$  to be the “bouquet graph”  $\mathcal{G}_1(\mathcal{F})$  of a collection  $\mathcal{F}$  of subsets of  $\omega$ .  $\mathcal{A}$  will be  $\exists$ -atomic and hence have a  $\Pi_2$  Scott sentence. Recall that by Lemma 8.17 of [Mon21] the structure  $\mathcal{G}_1(\mathcal{F})$  is  $\exists$ -atomic exactly when  $\mathcal{F}$  is discrete.

The structure  $\mathcal{A}$  will consist of a number of elements each of which is given various c.e. labels. We can think of  $\mathcal{A}$  as consisting of the elements which are the centers of the flowers, labeled by labels  $\ell_n$ ; putting a label  $\ell_n$  on a flower means to add a loop of length  $n + 3$  to the flower.

We divide our label into *sort labels*  $(u_e)_{e \in \omega}$  such that only exactly one label holds of each element, dividing the domain into the disjoint sets  $U_e = \{x \in A : u_e(x)\}$ ; we think of these as different sorts of the structure, and call the elements of  $U_e$  the *eth sort*. Then we have two other sets of labels  $(\ell_i)_{i \in \omega}$  and  $(\ell_i^\dagger)_{i \in \omega}$  which we call simply *labels*. For example, we may use loops of length  $2e + 3$  for the labels  $u_e$ , and of length  $4i + 4$  for the  $\ell_i$ , and  $4i + 6$  for the labels  $\ell_i^\dagger$ .

We use the *eth sort* to diagonalize against  $\theta_e$ . We do this by simultaneously building a structure  $\mathcal{B}_e$  such that if  $\mathcal{A} \models \theta_e$  then  $\mathcal{B}_e \models \theta_e$  and  $\mathcal{A} \not\equiv \mathcal{B}_e$ .  $\mathcal{B}_e$  will differ from  $\mathcal{A}$  only on the *eth sort*  $U_e$ . We build  $\mathcal{A}$  by finite approximations  $\mathcal{A} = \bigcup_s \mathcal{A}_s$ .  $\mathcal{B}_e$  will also be built by approximations  $\mathcal{B}_e = \bigcup_s \mathcal{B}_{e,s}$  with each  $\mathcal{B}_{e,s} \cong \mathcal{A}_s$ .

When we describe the construction of  $\mathcal{A} = \bigcup_s \mathcal{A}_s$ , we will describe the construction of the *eth sort*. The constructions for the different sorts should be thought of as happening simultaneously.

During the construction certain stages will be *e-expansionary stages* where we get evidence that  $\mathcal{A}_s \models \theta_e$ . We use  $k = k[s]$  to keep track of the number of expansionary stages and  $n = n[s] = k[s] + 1$  to keep track of the number of flowers in  $\mathcal{A}_s$ . At stage  $s$ ,  $\mathcal{A}_s$  will have elements  $a_1, \dots, a_n$  and  $\mathcal{B}_{e,s}$  will have elements  $b_1, \dots, b_{n-1}, c$ .

*Construction.*

*Stage 0.* We begin with  $n = n[0] = 1$  and  $k = k[0] = 0$ .  $\mathcal{A}_0$  consists of a single element  $a_1$ , and  $\mathcal{B}_{e,0}$  of a single element  $c$ . We give  $a_1$  and  $c$  the same label  $\ell_0$ .

*Stage  $s + 1$ .* Suppose that we have constructed  $\mathcal{A}_s$  and  $\mathcal{B}_{e,s}$  with  $k = k[s]$  and  $n = n[s] = k + 1$ . The elements of  $\mathcal{A}_s$  will be  $a_1, \dots, a_n$ . For each  $i \leq n$ ,  $a_i$  will have the labels  $\ell_j$  for  $j \leq i$ ; and for  $i < n$ ,  $a_i$  will have the label  $\ell_i^\dagger$ . The elements of  $\mathcal{B}_{e,s}$  will be  $b_1, \dots, b_{n-1}, c$ . For each  $i \leq n - 1$ ,  $b_i$  will have the labels  $\ell_j$  for  $j \leq i$  and  $\ell_i^\dagger$ . The element  $c$  will have the labels  $\ell_j$  for  $j \leq n$ .

Let  $s_k$  be the last expansionary stage. Check whether

$$\mathcal{A}_s \models \bigwedge_{i \leq k} \forall \bar{x}_{e,i} \in \mathcal{A}_{s_k} \varphi_{e,i}(\bar{x}_{e,i}).$$

If not, then this is not an expansionary stage. Make no changes to  $\mathcal{A}$ ,  $\mathcal{B}_e$ ,  $k[s + 1] = k[s]$ , or  $n[s + 1] = n[s]$ . If so then this stage is an expansionary stage. In  $\mathcal{A}_{s+1}$  put the label  $\ell_i^\dagger$  on  $a_n$ , and add a new element  $a_{n+1}$  to with labels  $\ell_i$  for  $i \leq n + 1$ . In  $\mathcal{B}_{e,s+1}$  add a new element  $b_n$  with labels  $\ell_i$  for  $i \leq n$  and  $\ell_n^\dagger$ . Put the label  $\ell_{n+1}$  on  $c$ . Set  $k[s + 1] = k[s] + 1$  and  $n[s + 1] = n[s] + 1$ .

*End construction.*

We will not give the full details of the verification. However, the general idea of the argument is as follows. We argue that there are infinitely many expansionary stages if and only if  $\mathcal{A} \models \theta_e$ . On the one hand, if there are infinitely many expansionary stages  $s$  then for every  $k$  there is  $s$  such that

$$\mathcal{A}_s \models \bigwedge_{i \leq k} \forall \bar{x}_{e,i} \in \mathcal{A}_{s_k} \varphi_{e,i}(\bar{x}_{e,i}).$$

and so as  $\varphi_{e,i}$  is  $\Sigma_1$  we have  $\mathcal{A} \models \theta_e$ . On the other hand, if there are only finitely many expansionary stages, then there is  $i$  and  $\bar{x}$  such that for all  $s$  we have

$$\mathcal{A}_s \models \neg\varphi_{e,i}(\bar{x}).$$

Then  $\mathcal{A} \models \neg\theta_e$ .

If there are infinitely many expansionary stages, then  $\mathcal{B}_e \not\models \mathcal{A}$  as the element  $c \in \mathcal{B}_e$  has infinitely many labels  $\ell_i$  for each  $i$ , but  $\mathcal{A}$  has no such element. We can also argue that  $\mathcal{B}_{e,s} \models \theta_n$ ; the tricky part is to check tuples  $\bar{x}_{e,i}$  containing  $c$ . Here we use the fact that at the  $k+1$ st expansionary stage  $s > s_k$  we have

$$\mathcal{A}_s \models \bigwedge_{i \leq k} \forall \bar{x}_{e,i} \in \mathcal{A}_{s_k} \varphi_{e,i}(\bar{x}_{e,i})$$

and that the isomorphism  $\mathcal{A}_s \cong \mathcal{B}_{e,s}$  extends the isomorphism  $\mathcal{A}_{s_k} \cong \mathcal{B}_{e,s_k}$ .

Finally, we need to check that  $\mathcal{A}$  is always  $\exists$ -atomic and hence always has a  $\Pi_2$  Scott sentence. Any element with a label  $\ell_i^\dagger$  is isolated by that label in its sort; and in each sort, there is at most one element without such a label (and only if there were finitely many expansionary stages in that sort), and that element is isolated by a label  $\ell_i$  which no other element of the sort has.  $\square$

The construction of a computable structure with a  $\Pi_2$  Scott sentence but no computable  $\Pi_3$  Scott sentence is similar, but the guessing as to whether a  $\Pi_3$  sentence is true in  $\mathcal{A}$  is more involved. Otherwise, the general ideas in the argument remain the same. Recall that by proving that such a structure exists, we have also shown that there is a computable structure with a  $\Pi_2$  Scott sentence but no computable  $\Sigma_4$  Scott sentence.

**Theorem 1.2.** *There is a computable structure with a  $\Pi_2$  Scott sentence but no computable  $\Sigma_4$  Scott sentence.*

*Proof.* We list all of the computable  $\Pi_3$  sentences as  $(\theta_e)_{e \in \omega}$  where

$$\theta_e = \bigwedge_{i \in \omega} \forall \bar{x}_{e,i} \bigvee_{j \in \omega} \exists \bar{y}_{e,i,j} \varphi_{e,i,j}(\bar{x}_{e,i}, \bar{y}_{e,i,j})$$

where the  $\varphi_{e,i,j}(\bar{x}_{e,i}, \bar{y}_{e,i,j})$  are computable  $\Pi_1$  formulas uniformly in  $e, i, j$  and the arities of each of the  $\bar{x}_{e,i}$  and  $\bar{y}_{e,i,j}$  are also computable in the indices.

We will build a structure  $\mathcal{A}$  which is  $\exists$ -atomic. The structure  $\mathcal{A}$  consist of a number of elements each of which is given various c.e. labels. Formally,  $\mathcal{A}$  can be taken to be the “bouquet graph”  $\mathcal{G}_1(\mathcal{F})$  of a collection  $\mathcal{F}$  of subsets of  $\omega$ —see Section VIII of [Mon]. As we construct  $\mathcal{A}$ , we think of ourselves as giving the centers of various “flower graphs” and when we add labels this corresponds to adding loops to the flower graphs. Thus in a computable construction of  $\mathcal{A}$  we can add labels to elements in a c.e. way.

We introduce two sets of distinguished labels. First, we have *sort labels*  $(u_e)_{e \in \omega}$  such that only exactly one label holds of each element, dividing the domain into the disjoint sets  $U_e = \{x \in A : u_e(x)\}$ ; we think of these as different sorts of the structure, and call the elements of  $U_e$  the *eth sort*. The second set of labels  $(\ell_i)_{i \in \omega}$  we will just call *labels*, and they will be the most important labels for the construction.

Within  $U_e$  we will diagonalize against  $\theta_e$ . We will do this by constructing another countable structure  $\mathcal{B}_e$  such that if  $\mathcal{A} \models \theta_e$  then  $\mathcal{B}_e \models \theta_e$  and  $\mathcal{A} \not\models \mathcal{B}_e$ .  $\mathcal{B}_e$  will be the same as  $\mathcal{A}$  except for the *eth sort*  $U_e$  on which they will differ. At each stage  $s$  we will have a finite approximation  $\mathcal{A}_s$  to  $\mathcal{A}$ , with  $\mathcal{A} = \bigcup_s \mathcal{A}_s$ .  $\mathcal{B}_e$  will also be built by finite approximations

$\mathcal{B}_e = \bigcup_s \mathcal{B}_{e,s}$  with each  $\mathcal{B}_{e,s} \cong \mathcal{A}_s$ . Thus one can think of  $\mathcal{A}$  and  $\mathcal{B}_e$  as direct limits of direct systems of the same finite structures but with different embeddings

$$\mathcal{A}_0 \rightarrow \mathcal{A}_1 \rightarrow \mathcal{A}_2 \rightarrow \dots$$

At each stage  $s$ , each element of  $\mathcal{A}_s$  will satisfy some finite set of labels (in fact two, a domain label and another label) that is satisfied by no other elements except duplicates of itself which satisfy exactly the same labels. Thus the isomorphism  $\mathcal{A}_s \rightarrow \mathcal{B}_{e,s}$  will be easily and uniquely determined up to mapping duplicates.

For each  $e$ , we have requirements  $\mathcal{R}_{i,\bar{b}}^e$  for each  $i \in \omega$  and  $\bar{b} \in \mathcal{B}_e$  of the right arity; to make sense of this, we must fix ahead of time the domain of  $\mathcal{B}_e$  even though these elements will only be added to  $\mathcal{B}_e$  slowly over time. The elements  $\bar{b}$  may be in any sort of  $\mathcal{B}_e$ , not just the  $e$ th. Each requirement will have a stage  $t = t(\mathcal{R}_{i,\bar{b}}^e)$  when it began working; it will only start working at some stage when  $\bar{b} \in \mathcal{B}_{e,t}$ . At this stage, there will be an element  $\bar{a}$  of  $\mathcal{A}_t$  which corresponds to  $\bar{b} \in \mathcal{B}_{e,t}$  at this stage  $t$  via the isomorphism  $\mathcal{A}_t \cong \mathcal{B}_{e,t}$ . The requirement will be met if whenever  $\mathcal{A} \models \forall_j \exists \bar{y}_{e,i,j} \varphi_{e,i,j}(\bar{a}, \bar{y}_{e,i,j})$  then  $\mathcal{B}_e \models \forall_j \exists \bar{y}_{e,i,j} \varphi_{e,i,j}(\bar{b}, \bar{y}_{e,i,j})$ ; equivalently, if  $\mathcal{B}_e \models \forall_j \forall \bar{y}_{e,i,j} \neg \varphi_{e,i,j}(\bar{b}, \bar{y}_{e,i,j})$  then  $\mathcal{A} \models \forall_j \forall \bar{y}_{e,i,j} \neg \varphi_{e,i,j}(\bar{a}, \bar{y}_{e,i,j})$ . One can think of the requirements for a given  $e$  as splitting up, into the different possible outermost witnesses, the supreme requirement that if  $\mathcal{A} \models \theta_e$  then  $\mathcal{B}_e \models \theta_e$  (or equivalently if  $\mathcal{B} \models \neg \theta_e$  then  $\mathcal{A} \models \neg \theta_e$ ). The requirement will require attention when we believe that  $\mathcal{B}_e \models \forall \bar{y}_{e,i,j} \neg \varphi_{e,i,j}(\bar{b}, \bar{y}_{e,i,j})$ . This is  $\Pi_2^0$  and so if this is the case, this requirement will require attention infinitely often. Meeting a single requirement in this infinitary manner will be enough, as for  $\bar{b} \in \mathcal{B}_e$  we have that if  $\mathcal{B}_e \models \forall \bar{y}_{e,i,j} \neg \varphi_{e,i,j}(\bar{b}, \bar{y}_{e,i,j})$  then  $\mathcal{B}_e \not\models \theta_e$  and the requirement will have the outcome that  $\mathcal{B}_e \cong \mathcal{A}$  so that  $\mathcal{A} \not\models \theta_e$ . Thus even though the injury may be infinite, any guessing is simply finitary since if a requirement is injured infinitely often, then it does not need to be satisfied as a higher priority requirement will be satisfied. In addition to the value  $t = t(\mathcal{R}_{i,\bar{b}}^e)$ , each requirement will also have a value  $k = k(\mathcal{R}_{i,\bar{b}}^e)$  which denotes the number of times that it has previously received attention.

The general idea of the construction is that, over time, we work to build  $\mathcal{B}_e$  taking, at each stage, a step towards making  $\mathcal{B}_e \not\cong \mathcal{A}$ . In particular, there will be some special element of the  $e$ th sort of  $\mathcal{B}_e$  which we will work to make different from each element of  $\mathcal{A}$ . However, whenever we see evidence that  $\mathcal{B}_e \models \neg \theta_e$ , we will undo our previous work to make  $\mathcal{B}_e$  different from  $\mathcal{A}$ ; since this makes  $\mathcal{B}_e$  look similar to  $\mathcal{A}$ , we can at the same time transfer our evidence that  $\mathcal{B}_e \models \neg \theta_e$  to evidence that  $\mathcal{A} \models \neg \theta_e$ . If in fact  $\mathcal{B}_e \models \neg \theta_e$ , then we will continually find evidence of this, continually rolling back our construction, so that actually  $\mathcal{B}_e$  never becomes different from  $\mathcal{A}$ . Since  $\mathcal{A} \cong \mathcal{B}_e$ ,  $\mathcal{A} \models \neg \theta_e$ . On the other hand, if  $\mathcal{B}_e \models \theta_e$ , then we will have  $\mathcal{A} \not\cong \mathcal{B}_e$  so that  $\theta_e$  cannot be a Scott sentence for  $\mathcal{A}$ .

We describe the construction of  $\mathcal{A}$  and the  $\mathcal{B}_e$  stage-by-stage. We will describe for a fixed  $e$  the construction of the  $e$ th sort of  $\mathcal{A}$  together with the  $e$ th sort of  $\mathcal{B}_e$ . However all of these constructions for the different sorts should be thought of as happening simultaneously. All of the other sorts of  $\mathcal{B}_e$  will be exactly the same as  $\mathcal{A}$ , so when constructing the  $e$ th sort of  $\mathcal{A}$  we need only discuss  $\mathcal{B}_e$  and not any other  $\mathcal{B}_m$ . In the construction that follows, without explicitly saying so we put the label  $u_e$  on all elements that we add to  $\mathcal{A}$  and  $\mathcal{B}_e$  so that they are in  $U_e$ . We sometimes suppress  $e$  and the label  $u_e$  in what follows, e.g., we write for  $\theta = \theta_e$

$$\theta = \bigwedge_{i \in \omega} \forall \bar{x}_i \bigvee_{j \in \omega} \exists \bar{y}_{i,j} \varphi_{i,j}(\bar{x}_i, \bar{y}_{i,j}).$$

However, when we are guessing at whether some formula is true in  $\mathcal{A}$  or  $\mathcal{B}_e$ , we must

allow the quantifiers to quantify over all of the sorts; we will remind the reader of this later.

We begin at stage 0 with a single element  $a_0$  with no labels (other than  $u_e$ , which we no longer mention). At each stage  $s$  we will have a number  $n = n[s]$ . There will be  $n$  *active elements* designated  $a_1[s], \dots, a_n[s]$ . Each active element will have a label that no other active element has. The rest of the elements of  $\mathcal{A}_s$  will be *duplicate elements* and will have exactly the same labels as one of the active elements. When we declare a (formerly active) element to be a duplicate of  $a_i$ , we will give it exactly the same labels as  $a_i$ , and whenever  $a_i$  gets a new label, the element will as well; thus if later  $a_i$  becomes a duplicate of some other  $a_j$ , then any duplicates of  $a_i$  will become duplicates of  $a_j$ . When an active element  $a_i$  is declared to be a duplicate of another, becoming inactive, it gives up the designation  $a_i$ . Recall that  $\mathcal{B}_{e,s}$  will be isomorphic to  $\mathcal{A}_s$ ; however,  $\mathcal{B}_e$  will follow different embeddings from one stage to the next. The active elements of  $\mathcal{B}_{e,s}$  will be  $b_1[s], \dots, b_{n-1}[s], c$ . The isomorphism  $\mathcal{A}_s \rightarrow \mathcal{B}_{e,s}$  will map  $a_i \mapsto b_i$  and  $a_n \mapsto c$ . (The duplicate elements will also be mapped correspondingly.)

To illustrate the idea of the construction, we will describe a generic stage of the construction when no requirement requires attention. Suppose that we have defined  $\mathcal{A}_s$  and  $\mathcal{B}_{e,s}$  with  $n = n[s]$ . Define  $\mathcal{A}_{s+1}$  as follows. Add a new element  $a_{n+1}$ , and give  $a_{n+1}$  all of the labels that  $a_n$  had in  $\mathcal{A}_s$ . Give  $a_n$  and  $a_{n+1}$  each a new label unique only to it. For  $\mathcal{B}_{e,s+1}$ , add a new element  $b_n$  and let  $b_n$  have the same labels as  $a_n$ , and let  $c$  have the same labels as  $a_{n+1}$ . If we do this at every stage, then  $\mathcal{A}$  will have elements  $a_1, a_2, \dots$  and  $\mathcal{B}_e$  will have corresponding elements  $b_1, b_2, \dots$  with the same labels, but also an additional element  $c$  which does not correspond to any element of  $\mathcal{A}$ . Thus  $\mathcal{A} \not\cong \mathcal{B}_e$ . However we have not ensured that if  $\mathcal{A} \models \theta_e$  then  $\mathcal{B}_e \models \theta_e$ .

To ensure that if  $\mathcal{A} \models \theta_e$  then  $\mathcal{B}_e \models \theta_e$ , our strategy (for some  $i$  and  $\bar{b}$ ) is to wait for stages at which we think that  $\mathcal{B}_{e,s} \models \bigwedge_j \forall \bar{y}_{i,j} \neg \varphi_{i,j}^e(\bar{b}, \bar{y}_{i,j})$ . This is evidence that  $\mathcal{B}_e \not\models \theta_e$ . At stage  $s+1$  instead of taking our usual action towards making  $\mathcal{B}_e$  different from  $\mathcal{A}$ , we instead revert some of our previous actions to make  $\mathcal{B}_e$  look like  $\mathcal{A}$ . If  $\mathcal{B}_e \not\models \theta_e$  then we will get more and more evidence of this, and we will instead have  $\mathcal{B}_e \cong \mathcal{A}$  and thus  $\mathcal{A} \models \theta_e$ . How we do this is the heart of the construction.

*Construction.*

*Stage 0.* We begin with  $n = n[0] = 1$  and  $\mathcal{A}_0$  consisting of a single element  $a_1$ , and  $\mathcal{B}_{e,0}$  of a single element  $c$ . We give  $a_1$  a label and give  $c$  the same label.

*Stage  $s+1$ .* Suppose that we have constructed  $\mathcal{A}_s$  and  $\mathcal{B}_{e,s}$  with  $n = n[s]$  and active elements  $a_1, \dots, a_n$ . We first check whether any requirement  $\mathcal{R}_{i,\bar{b}}^e$  requires attention.

Given  $\mathcal{R}_{i,\bar{b}}^e$ , let  $t = t(\mathcal{R}_{i,\bar{b}}^e)$  be the corresponding stage and let  $k = k(\mathcal{R}_{i,\bar{b}}^e)$ . Then we say that  $\mathcal{R}_{i,\bar{b}}^e$  *requires attention* if

$$\mathcal{B}_{e,s} \models \bigwedge_{j \geq k} \forall \bar{y}_{i,j} \in \mathcal{B}_{e,t+k} \neg \varphi_{i,j}^e(\bar{b}, \bar{y}_{i,j}).$$

Note that this quantifier is not restricted to the  $e$ th sort of  $\mathcal{B}_{e,t+k}$  but is over the elements of all sorts.

If no requirement requires attention, then define  $\mathcal{A}_{s+1}$  as follows. Add a new element  $a_{n+1}$ , and give  $a_{n+1}$  all of the labels that  $a_n$  had in  $\mathcal{A}_s$ . Give each of  $a_n$  and  $a_{n+1}$  a new label unique only to it. Any duplicates of  $a_n$  receive the same labels that  $a_n$  did. For  $\mathcal{B}_{e,s+1}$ , add a new element  $b_n$  and let  $b_n$  have the same labels as  $a_n$ , and let  $c$  have the same labels as  $a_{n+1}$ . There may be elements of  $\mathcal{B}_{e,s}$  which look like duplicates of  $c$ ; these are



the elements which correspond to the duplicates of  $a_n$ . We call these *pending duplicates* and we now give them the same labels as  $b_n$ , making them duplicates of  $b_n$ . Let  $\mathcal{R}_{i,\bar{b}}^e$  be the next highest priority requirement; we may assume, by appropriately listing the requirements, that  $\bar{b} \in \mathcal{B}_{e,s+1}$ . We initialize  $\mathcal{R}_{i,\bar{b}}^e$  by setting  $t(\mathcal{R}_{i,\bar{b}}^e) = s + 1$  and  $k(\mathcal{R}_{i,\bar{b}}^e) = 0$ .

Otherwise, let  $\mathcal{R}_{i,\bar{b}}^e$  be the highest priority requirement which requires attention, and we will act on the requirement as follows. Let  $t = t(\mathcal{R}_{i,\bar{b}}^e)$  be the corresponding stage, and  $n^* = n[t]$  the corresponding value of  $n$ . Let  $n[s + 1] = n[t] = n^*$ . In  $\mathcal{A}_{s+1}$ , declare  $a_{n^*+1}[s], \dots, a_n[s]$  to be duplicates of  $a_{n^*}$ ; they give up these designations  $a_{n^*+1}, \dots, a_n$ . Moreover, any duplicates of these elements also become duplicates of  $a_{n^*}$ . Add labels to these elements and  $a_{n^*}$  so that they have the same labels, i.e., for each label of one of these elements, give it to each other element. So in  $\mathcal{A}_{s+1}$  we have active elements  $a_1[s + 1], \dots, a_{n^*}[s + 1]$  together with duplicates of these elements. In  $\mathcal{B}_{e,s+1}$  add each label of any of  $b_{n^*}, \dots, b_{n-1}, c$  to each of the others. In  $\mathcal{B}_{e,s+1}$  there are elements which correspond to the duplicates of  $a_n[s + 1]$ . One should think of these not as duplicates of  $c$ , but as duplicates of an element  $b_n$  which does not yet exist. We call these elements *pending duplicates*. Then put  $b_i[s + 1] := b_i[s]$  for  $i < n^*$ . Thus, at stage  $s + 1$ , the correspondence between  $\mathcal{A}_{s+1}$  and  $\mathcal{B}_{e,s+1}$  is  $a_i \mapsto b_i$  for  $i < n^*$  and  $a_{n^*} \mapsto c$  (together with the duplicates mapping to corresponding duplicates, with duplicates of  $a_n$  mapping to pending duplicates). Each lower priority  $e$ -requirement is injured, and increment  $k(\mathcal{R}_{i,\bar{a}}^e)$  by 1.

*End construction.*

We must now verify that the construction works. Standard arguments show that for each  $e$ , there are two possibilities. First, it might be that some requirement  $\mathcal{R}_{i,\bar{b}}^e$  is, after some point, never injured and acts infinitely often. In this case, let  $t = t(\mathcal{R}_{i,\bar{b}}^e)$  and let  $n = n[t]$ . Since no higher priority requirement acts, the elements  $a_1, \dots, a_n$  never become duplicates of any other elements, and every time  $\mathcal{R}_{i,\bar{b}}^e$  acts every other element which is not already becomes a duplicate of one of these elements. Thus the  $e$ th sort of  $\mathcal{A}$  consists of  $n$  elements  $a_1, \dots, a_n$  together with duplicates of them. We write  $a_1[\infty], \dots, a_n[\infty]$  for these elements.

The other possibility is that each requirement acts only finitely many times. If, after some stage  $s$  with  $n = n[s]$ , no requirement with  $t(\mathcal{R}_{i,\bar{b}}^e) < s$  ever acts, then after this stage  $a_1, \dots, a_n$  never become duplicates. In the limit, we build infinitely many elements  $a_1[\infty], a_2[\infty], a_3[\infty], \dots$  (Note that the first time we introduce some  $a_i$ , this element may become a duplicate; but there will be some point after which this stops happening.) All of the other elements will be duplicates of these.

**Lemma 3.2.**  *$\mathcal{A}$  has a  $\Pi_2$  Scott sentence.*

*Proof.* Recall that by Lemma 8.17 of [Mon21] the structure  $\mathcal{G}_1(\mathcal{F})$  is  $\exists$ -atomic exactly when  $\mathcal{F}$  is discrete. Therefore it suffices to show that for each element of  $\mathcal{A}$ , there is a finite set of labels on that element such that no other element has the same labels. Each element is in only one sort  $U_e$ , so it suffices to only compare elements within the same sort. Within each sort we will argue that each element has a label, that we call the distinguishing label, which it has and no other element has other than duplicates of that element.

Each element  $a_i[\infty]$ , when it is created, is given a label which is never given to  $a_1[\infty], \dots, a_{i-1}[\infty]$ . If  $a_{i+1}[\infty]$  exists then, when it is created,  $a_i[\infty]$  is given a label that

is never given to  $a_{i+1}[\infty]$ ; and no  $a_j[\infty]$ ,  $j > i + 1$ , is ever given that label. The other elements of  $\mathcal{A}$  are duplicates of the  $a_i[\infty]$  and hence share the same distinguishing label as the element they duplicate.  $\square$

**Lemma 3.3.** *Fix  $i$  and  $\bar{b}$ . Suppose that there is a stage  $s$  after which  $\mathcal{R}_{i,\bar{b}}^e$  is never injured. Let  $t = t(\mathcal{R}_{i,\bar{b}}^e)$ . Suppose that*

$$\mathcal{B}_e \models \bigwedge_j \forall \bar{y}_{i,j} \neg \varphi_{i,j}(\bar{b}, \bar{y}_{i,j}).$$

*Then there are infinitely many stages at which  $\mathcal{R}_{i,\bar{b}}^e$  requires attention.*

*Proof.* Given  $s_0 = s$ , we will argue that there are infinitely many stages  $s_u$  at which  $\mathcal{R}_{i,\bar{b}}^e$  requires attention. Given  $s_u$ , we have that

$$\mathcal{B}_e \models \bigwedge_j \forall \bar{y}_{i,j} \neg \varphi_{i,j}(\bar{b}, \bar{y}_{i,j}).$$

and so there must be some stage  $s_{u+1} > s_u$  such that

$$\mathcal{B}_{e,s_{u+1}} \models \bigwedge_{j \leq u+1} \forall \bar{y}_{i,j} \in B_{e,s_u} \neg \varphi_{i,j}(\bar{b}, \bar{y}_{i,j}).$$

At this stage  $s_{u+1}$ ,  $\mathcal{R}_{i,\bar{b}}^e$  requires attention. (Since  $\mathcal{R}_{i,\bar{b}}^e$  is not injured, no higher priority requirement can require attention at this stage.)  $\square$

**Lemma 3.4.** *Suppose that  $\mathcal{B}_e \models \neg \theta_e$ . Then  $\mathcal{A} \cong \mathcal{B}_e$  and  $\mathcal{A} \models \neg \theta_e$*

*Proof.* Then there is  $i$  and  $\bar{b} \in \mathcal{B}_e$  such that

$$\mathcal{B}_e \models \bigwedge_j \forall \bar{y}_{i,j} \neg \varphi_{i,j}(\bar{b}, \bar{y}_{i,j}).$$

By the previous lemma, for any such  $i$  and  $\bar{b}$ , if there is some stage after which it is not injured, the requirement  $\mathcal{R}_{i,\bar{b}}^e$  will require attention infinitely many times. Thus in particular there is some highest priority requirement  $\mathcal{R}_{i,\bar{b}}^e$  which requires attention infinite often.

Let  $t = t(\mathcal{R}_{i,\bar{b}}^e)$ . Let  $n = n[t]$ . Let  $s_0$  be the stage at which  $\mathcal{R}_{i,\bar{b}}^e$  is initiated, and let  $s_1, s_2, \dots$  be the stages at which it requires attention. At each of these stages,  $\mathcal{A}_{s_i}$  consists of  $a_1, \dots, a_e$  and duplicates, and  $\mathcal{B}_{e,s_i}$  consists of  $b_1, \dots, b_{n-1}, c$  and duplicates; thus the isomorphisms  $a_1, \dots, a_e \mapsto b_1, \dots, b_{n-1}, c$  at these stages extend to an isomorphism  $\mathcal{A} \cong \mathcal{B}_e$ . (For this, we also need to know that corresponding elements have the same number of duplicates; this is maintained at each stage of the construction.) Thus  $\mathcal{A} \cong \mathcal{B}_e$ , and since  $\mathcal{B}_e \models \neg \theta_e$  we have  $\mathcal{A} \models \neg \theta_e$ .  $\square$

As a result, if  $\mathcal{A} \models \theta_e$  then  $\mathcal{B}_e \models \theta_e$ . Now we show that if  $\mathcal{B}_e \models \theta_e$  then  $\mathcal{A} \not\cong \mathcal{B}_e$ .

**Lemma 3.5.** *Suppose that  $\mathcal{B}_e \models \theta_e$ . Then  $\mathcal{A} \not\cong \mathcal{B}_e$ .*

*Proof.* We will argue that each requirement, after it is initialized and unless it is injured, requires attention only finitely many times. We argue by induction; given a requirement  $\mathcal{R}_{i,\bar{b}}^e$ , suppose that the higher priority requirements only require attention finitely many

times. Then there is a stage  $s_0$  after which  $\mathcal{R}_{i,\bar{b}}^e$  is never injured. Let  $t = t(\mathcal{R}_{i,\bar{b}}^e)$ . Whenever  $\mathcal{R}_{i,\bar{b}}^e$  acts for the  $k$ th time at stage  $s + 1$ , it is because

$$\mathcal{B}_{e,s} \models \bigwedge_{j \leq k} \forall \bar{y}_{i,j} \in \mathcal{B}_{e,t+k} \neg \varphi_{i,j}(\bar{b}, \bar{y}_{i,j}).$$

Then if  $\mathcal{R}_{i,\bar{b}}^e$  requires attention infinitely many times, we see that  $\mathcal{B}_e \models \bigwedge_j \forall \bar{y}_{i,j} \neg \varphi_{i,j}(\bar{b}, \bar{y}_{i,j})$  and so  $\mathcal{B}_e \models \neg \theta_e$ , contradicting the hypotheses of this lemma. Thus each requirement requires attention only finitely many times.

In particular, there are infinitely many “true expansionary stages” and  $\mathcal{A} \not\equiv \mathcal{B}_e$  because the element  $c$  in  $\mathcal{B}_e$  does not correspond to any element of  $\mathcal{A}$ .  $\square$

We have proved that if  $\mathcal{A} \models \theta_e$  then  $\mathcal{B}_e \models \theta_e$  and  $\mathcal{A} \not\equiv \mathcal{B}_e$ . Thus  $\theta_e$  cannot be a Scott sentence for  $\mathcal{A}$ . It follows that, since the  $\theta_e$  list all computable  $\Pi_3$  sentences,  $\mathcal{A}$  cannot have a computable Scott sentence.  $\square$

## 4 There is no characterization of the structures with a computable $\Pi_2$ Scott sentence

Given that we know that some computable structures with  $\Pi_2$  Scott sentences have no computable  $\Pi_2$  Scott sentences (or even computable  $\Pi_3$  Scott sentences), we might like to know how to decide whether or not some particular structure with a computable  $\Pi_2$  Scott sentence has a computable  $\Pi_2$  Scott sentence.

Having a  $\Pi_2$  Scott sentence is equivalent to being  $\exists$ -atomic, which means that it has a Scott family of existential formulas. This is equivalent (see [Mon17]) to the following statement: for each tuple  $\bar{a} \in \mathcal{A}$  there is an existential sentence  $\varphi(\bar{x})$  such that for any universal sentence  $\psi(\bar{x})$  with  $\mathcal{A} \models \psi(\bar{x})$ ,  $\mathcal{A} \models \forall \bar{x} \varphi(\bar{x}) \rightarrow \psi(\bar{x})$ . Thus it is not too difficult to tell whether a structure has a  $\Pi_2$  Scott sentence:  $\Pi_4^0$  if one counts the quantifiers (though we leave it as an open question to determine if this is best possible).

Since it is not too hard to decide whether a computable structure  $\mathcal{A}$  has a  $\Pi_2$  Scott sentence, we might as well assume that our given structure  $\mathcal{A}$  does. One possible attempt to characterize whether  $\mathcal{A}$  has a computable  $\Pi_2$  Scott sentence is to ask whether it is effectively  $\exists$ -atomic, that is, whether it has a c.e. Scott family of existential formulas. Alvir, Knight, and McCoy [AKMC20] proved that if a computable structure  $\mathcal{A}$  has a c.e. Scott family of existential formulas, then it has a computable  $\Pi_2$  Scott sentence. However the converse fails (and while our Theorem 1.3 gives a counterexample it is not too hard to construct a one directly).

In this section we will prove that there is no characterization of the structures with a computable  $\Pi_2$  Scott sentence. We will use the technique of index set complexity which has been used to great effect in computable structure theory, as suggested in [GK02], e.g., to show that there is no characterization of when a computable structure has a decidable or automatic presentation [HT18b, BHTK<sup>+</sup>19]. Given a listing of all (partial) computable structures in a given language, we consider the set of (indices for) structures with the particular property we are interested in. In this case, we consider the set

$$\{i \mid \mathcal{A}_i \text{ has a computable } \Pi_2 \text{ Scott sentence}\}.$$

This set is naively  $\Pi_1^1$ , because  $\mathcal{A}_i$  has a computable  $\Pi_2$  Scott sentence if and only if (a)  $\mathcal{A}$  has a  $\Pi_2$  Scott sentence and (b) there is a computable  $\Pi_2$  sentence  $\varphi$  such that for

all countable structures  $\mathcal{B}$ ,  $\mathcal{B} \models \varphi$  if and only if  $\mathcal{A} \equiv_2 \mathcal{B}$ . Note that give (a), if  $\mathcal{A} \equiv_2 \mathcal{B}$  then  $\mathcal{A} \cong \mathcal{B}$ . While (a) is  $\Pi_4^0$ , (b) is  $\Pi_1^1$ . If there was a better characterization of when a structure had a computable Scott sentence, then this index set would be simpler than  $\Pi_1^1$ , e.g., hyperarithmetic. However we show that the set is  $\Pi_1^1$ - $m$ -complete, and hence there is no simpler characterization.

**Theorem 1.3.** *With  $(\mathcal{A}_i)_{i \in \omega}$  an effective list of (possibly partial) structures in a sufficiently rich language, the set*

$$\{i \mid \mathcal{A}_i \text{ has a computable } \Pi_2 \text{ Scott sentence}\}$$

*is  $\Pi_1^1$ - $m$ -complete.*

*Proof.* By general results on the universality of certain languages up to effective bi-interpretability, we can construct examples in whatever language we wish to. We explained above why the class is  $\Pi_1^1$ , and it remains to show that it is  $\Pi_1^1$ - $m$ -complete. Let  $T \subseteq \omega^{<\omega}$  be a tree. We will define a computable structure  $\mathcal{A} = \mathcal{A}_T$  which is  $\exists$ -atomic and hence has a  $\Pi_2$  Scott sentence, and such that  $\mathcal{A}$  has a computable  $\Pi_2$  Scott sentence if and only if  $\mathcal{A}$  is well-founded.

We list all of the computable  $\Pi_2$  sentences as  $(\theta_e)_{e \in \omega}$  where

$$\theta_e = \bigwedge_{i \in \omega} \forall \bar{x}_{e,i} \varphi_{e,i}(\bar{x}_{e,i})$$

where the  $\varphi_{e,i}(\bar{x}_{e,i})$  are computable  $\Sigma_1$  formulas uniformly in  $e, i$  and the arities of each of the  $\bar{x}_{e,i}$  are also computable in the indices.

$\mathcal{A}$  will be  $\exists$ -atomic and hence have a  $\Pi_2$  Scott sentence. It will consist of a number of elements each of which is given various c.e. labels. As before, we can take  $\mathcal{A}$  to be the “bouquet graph”  $\mathcal{G}_1(\mathcal{F})$  of a collection  $\mathcal{F}$  of subsets of  $\omega$ . We introduce three sets of labels. First, we have *sort labels*  $(u_e)_{e \in \omega}$  such that only exactly one label holds of each element, dividing the domain into the disjoint sets  $U_e = \{x \in A : u_e(x)\}$ ; we think of these as different sorts of the structure, and call the elements of  $U_e$  the *eth sort*. Then we will have two set of labels  $(\ell_\sigma)_{\sigma \in \omega^{<\omega}}$  and  $(\ell_\sigma^\dagger)_{\sigma \in \omega^{<\omega}}$  we will just call *labels*.

Within  $U_e$  we will diagonalize against  $\theta_e$ , though this diagonalization may only be successful if  $T$  has an infinite path. Given a path  $\pi$  through  $T$ , we will diagonalize by constructing another countable structure  $\mathcal{B}_e = \mathcal{B}_{e,\pi}$  such that if  $\mathcal{A} \models \theta_e$  then  $\mathcal{B}_e \models \theta_e$  and  $\mathcal{A} \not\cong \mathcal{B}_e$ . At each stage  $s$  we will have an approximation  $\mathcal{A}_s$  to  $\mathcal{A}$ , with  $\mathcal{A} = \bigcup_s \mathcal{A}_s$ .  $\mathcal{B}_e$  will differ from  $\mathcal{A}$  only on the *eth sort*  $U_e$ .  $\mathcal{B}_e$  will also be built by approximations  $\mathcal{B}_e = \bigcup_s \mathcal{B}_{e,s}$  with each  $\mathcal{B}_{e,s} \cong \mathcal{A}_s$ , though the construction of  $\mathcal{B}_e$  is non-computable as it requires knowing a path through  $T$ .

When we describe the construction of  $\mathcal{A} = \bigcup_s \mathcal{A}_s$ , we will describe the construction of the *eth sort*. The constructions for the different sorts should be thought of as happening simultaneously.

During the construction certain stages will be *e-expansionary stages* where we get evidence that  $\mathcal{A} \models \theta_e$ . We use the variable  $k[s]$  to keep track of the number of expansionary stages. The elements of  $\mathcal{A}$  will all be of the form  $a_\sigma$  for some  $\sigma \in T$ . We write  $T_s$  for the set of  $\sigma$  such that  $a_\sigma \in \mathcal{A}_s$ , i.e., for the set of  $\sigma$  which correspond to elements of  $\mathcal{A}_s$  at stage  $s$ .  $T_s$  will be a subtree of  $T$ , with the property that if one child of  $\sigma \in T_s$  is in  $T_s$ , then all children of  $\sigma$  are in  $T_s$ .

*Construction of the eth sort of  $\mathcal{A}$ .*

*Stage 0.* We begin with  $\mathcal{A}_0$  consisting of a single element  $a_\emptyset$  with the single label  $\ell_\emptyset$ . Thus  $T_0 = \{\emptyset\}$ . Begin with  $k = 0$  as we have not yet had any expansionary stages.

*Stage  $s + 1$ .* Suppose that we have constructed  $\mathcal{A}_s$  with  $k = k[s]$ . We say that a tuple  $\bar{x}$  is  $k$ -small if it consists of elements  $a_\sigma \in \mathcal{A}$  of the  $e$ th sort with  $\sigma \in \{0, \dots, k\}^{\leq k}$  and elements not of the  $e$ th sort but among the first  $k$  elements of  $\mathcal{A}$ . First, we check whether this is an expansionary stage<sup>1</sup>: Check whether, for each  $i \leq k$  and each  $k$ -small tuple  $\bar{x}_i$  we have

$$\mathcal{A}_s \models \varphi_i(\bar{x}_i)$$

where the existential quantifier in  $\varphi_i(\bar{x}_i)$  is witnessed by one of the first  $s$  witnesses (being careful to use an appropriately dovetailed listing of possible witnesses). If this is the case, then stage  $s + 1$  is an expansionary stage, and set  $k[s + 1] = k[s] + 1$ . Otherwise, set  $k[s + 1] = k[s]$ .

If stage  $s + 1$  is not an expansionary stage, set  $\mathcal{A}_{s+1} = \mathcal{A}_s$ . If stage  $s + 1$  is an expansionary stage:

- (1) for each  $\sigma \in T_s \cap \{0, \dots, k\}^{\leq k}$  with a child  $\tau$  in  $T \cap \{0, \dots, k\}^{\leq k}$ , put the label  $\ell_\sigma^\dagger$  on  $a_\sigma$ .
- (2) for each  $\sigma \in T_s \cap \{0, \dots, k\}^{\leq k}$  which is a leaf of  $T_s$  and has a child  $\tau$  in  $T \cap \{0, \dots, k\}^{\leq k}$ , and each child  $\tau \in T$  of  $\sigma$ , add a new element  $a_\tau$  with the labels  $\ell_\rho$  for  $\rho \leq \tau$ . Let  $T_{s+1}$  be  $T_s$  together with all of these new  $\tau$ .

*End construction.*

What  $\mathcal{A}$  looks like depends on whether there are finitely many or infinitely many expansionary stages.

- (1) If there are finitely many expansionary stages, say  $k$ , then  $\mathcal{A}$  has finitely many elements  $a_\sigma$  with each  $|\sigma| \leq k$ . Write  $T_\infty$  for the set of such  $\sigma$ ;  $T_\infty = \bigcup T_s$  is a subtree of  $T$ . Each  $a_\sigma$  has the labels  $\ell_\rho$  for  $\rho \leq \sigma$ ; and if  $\sigma$  is not a leaf of  $T_\infty$ , then  $a_\sigma$  also has the label  $\ell_\sigma^\dagger$ .
- (2) If there are infinitely many expansionary stages, then  $\mathcal{A}$  has elements  $a_\sigma$  for  $\sigma \in T$ , and each  $a_\sigma$  has the labels  $\ell_\rho$  for  $\rho \leq \sigma$  and the label  $\ell_\sigma^\dagger$ . (We have  $T_\infty = \bigcup T_s = T$ .)

Thus it is easy to see that  $\mathcal{A}$  has a  $\Pi_2$  Scott sentence.

**Lemma 4.1.**  *$\mathcal{A}$  has a  $\Pi_2$  Scott sentence.*

*Proof.* For each  $a_\sigma$ , either  $a_\sigma$  has the label  $\ell_\sigma^\dagger$  and no other element has this label, or  $\sigma$  is a leaf of  $T_\infty$  and  $a_\sigma$  is the unique element with the label  $\ell_\sigma$ .  $\square$

**Lemma 4.2.** *Suppose that  $T$  is well-founded. Then  $\mathcal{A}$  has a computable  $\Pi_2$  Scott sentence.*

*Proof.* Note that  $T_\infty = \bigcup T_s$  is a c.e. set. Moreover, the set of non-leaves of  $T_\infty$  is also a c.e. set. Consider the computable  $\Pi_2$  sentence which says that:

---

<sup>1</sup>The conditions to have an expansionary stage, and exactly which elements we should add to  $\mathcal{A}_{s+1}$  at an expansionary stage, are quite subtle. We must ensure that we can perform the construction of  $\mathcal{B}_e$  given below and prove Lemmas 4.4 and Lemma 4.5.

- (1) if  $x$  has the label  $\ell_\sigma$  then it also has all of the labels  $\ell_\rho$  for  $\rho < \sigma$ .
- (2) if  $x$  has the label  $\ell_\sigma^\dagger$  then it has the label  $\ell_\sigma$  and does not have any label  $\ell_\tau$  for  $\tau > \sigma$ .
- (3) no  $x$  has labels  $\ell_{\sigma_1}$  and  $\ell_{\sigma_2}$  for incompatible  $\sigma_1$  and  $\sigma_2$ .
- (4) if  $x$  has the label  $\ell_\sigma$  then  $\sigma \in T_\infty$ .
- (5) every  $x$  has the label  $\ell_\emptyset$  and if  $x$  has the label  $\ell_\sigma$  and  $\sigma$  is not a leaf of  $T_\infty$  then  $x$  has either the label  $\ell_\sigma^\dagger$  or some label  $\ell_\tau$  where  $\tau$  is a child of  $\sigma$  on  $T$ .
- (6) if  $x$  has the label  $\ell_\sigma^\dagger$ , then  $\sigma$  is a non-leaf of  $T_\infty$ .
- (7) for each  $\sigma \in T_\infty$ , there is some  $x$  with the label  $\ell_\sigma$ .
- (8) for each  $\sigma \in T_\infty$  which is not a leaf, there is some  $x$  with the label  $\ell_\sigma^\dagger$ .
- (9) for every distinct  $x$  and  $y$ , either there are some incompatible  $\sigma, \tau$  such that one element has the label  $\ell_\sigma$  and the other has the label  $\ell_\tau$ , or there is some  $\sigma$  and a child  $\sigma^*$  of  $\sigma$  such that one element has the label  $\ell_\sigma^\dagger$  and the other has the label  $\ell_{\sigma^*}$ .

Since  $T$  is well-founded,  $T_\infty$  must also be well-founded. We can check that this sentence is true of  $\mathcal{A}$ .

Suppose that a structure  $\mathcal{C}$  satisfies this sentence. For each non-leaf  $\sigma$  of  $T_\infty$ , by (8) there is an element  $c_\sigma$  with the label  $\ell_\sigma^\dagger$  (and thus, by (1) and (2) also the labels  $\ell_\rho$  for  $\rho \leq \sigma$ , and by (2) and (3) these are the only labels). By (9) and (2) no other element of  $\mathcal{C}$  has the label  $\ell_\sigma^\dagger$ .

For each leaf  $\sigma$  of  $T_\infty$  by (7) there is an element  $c_\sigma$  with the label  $\ell_\sigma$ . By (1) it also has all of the labels  $\ell_\rho$  for  $\rho < \sigma$ , and by (2), (3), and (6) it has no other labels. Thus we have shown that  $\mathcal{A}$  embeds as a substructure of  $\mathcal{C}$ .

Suppose that  $d$  is some other element of  $\mathcal{C}$ , not included among the  $c_\sigma$  above. By (5)  $d$  has the label  $\ell_\emptyset$ . Let  $\sigma_0 = \emptyset$ . We argue inductively as follows. Given  $\sigma_i$  such that  $d$  has the label  $\ell_{\sigma_i}$ , by (5) either (a)  $\sigma_i$  is a leaf of  $T_\infty$ , (b)  $d$  has the label  $k_{\sigma_i}$ , or (c)  $d$  has some label  $\ell_{\sigma_{i+1}}$  for some child  $\sigma_{i+1} \in T_\infty$  of  $\sigma_i$ . Since  $T_\infty$  is well-founded, we can repeat this process to build a sequence of children  $\sigma_0 = \emptyset, \sigma_1, \dots, \sigma_m$  in  $T_\infty$  with either (a)  $\sigma_m$  is a leaf of  $T_\infty$ , or (b)  $d$  has the label  $k_{\sigma_m}$ . Recall that there is already an element  $c_{\sigma_m}$  with the label  $k_{\sigma_m}$ , and by (2) and (9) there are no other such elements. Thus  $\sigma_m$  is a leaf of  $T_\infty$  and  $d$  has the label  $\ell_{\sigma_m}$ . But we already have an element  $c_{\sigma_m}$  with the label  $\ell_{\sigma_m}$ , and by (9) we cannot have two such elements. So no such  $d$  exists, that is,  $\mathcal{C} \cong \mathcal{A}$ .  $\square$

**Lemma 4.3.** *Suppose that  $\mathcal{A} \models \theta_e$ . Then there are infinitely many expansionary stages.*

*Proof.* Let  $s_1 < s_2 < s_3 < \dots < s_k$  be expansionary stages. We must argue that there will later be another expansionary stage, the  $k+1$ st. For each  $i \leq k$  and  $k$ -small tuple  $\bar{x}_i$  we have

$$\mathcal{A} \models \varphi_i(\bar{x}_i).$$

Since  $\varphi_i$  is existential, and  $\mathcal{A} = \bigcup_s \mathcal{A}_s$ , there must be some stage  $s_{k+1} > s_k$  such that for each such  $i$  and  $\bar{x}_i$  we have

$$\mathcal{A}_{s_{k+1}} \models \varphi_i(\bar{x}_i)$$

with the existential quantifier in  $\varphi_i$  witnessed at this stage. This stage  $s_{k+1}$  is expansionary.  $\square$

Suppose that  $T$  has an infinite path  $\pi$ . Let  $\mathcal{B}_e$  be  $\mathcal{A}$  together with another element  $c$  of the  $e$ th sort satisfying  $\ell_\rho$  for each  $\rho < \pi$ . Clearly  $\mathcal{B}_e \not\cong \mathcal{A}$ . However we also want to know that if  $\mathcal{A} \models \theta_e$ , then  $\mathcal{B}_e \models \theta_e$ . To see this, we will show that  $\mathcal{B}_e$  is also a union  $\mathcal{B}_e = \bigcup \mathcal{B}_{e,s}$  and that  $\mathcal{B}_{e,s} \cong \mathcal{A}_s$ . Thus one can think of  $\mathcal{A}$  and  $\mathcal{B}_e$  as direct limits of direct systems of the same structures but with different embeddings

$$\mathcal{A}_0 \rightarrow \mathcal{A}_1 \rightarrow \mathcal{A}_2 \rightarrow \dots$$

However the definition of  $\mathcal{B}_{e,s}$  depends on the path  $\pi$  and thus is non-computable.

At stage  $s$ , let  $\pi_s$  be the longest initial segment of  $\pi$  which is on  $T_s$ . Then  $\pi_s$  is a leaf of  $T_s$ —this is because  $T_s$  has the property that if there is any child of  $\pi_s$  in  $T_s$ , then all children of  $\pi_s$  from  $T$  are in  $T_s$ .  $\mathcal{B}_{e,s}$  will have domain consisting of elements  $\{b_\sigma \mid \sigma \in T_s, \sigma \neq \pi_s\} \cup \{c\}$ . We put the same labels on  $b_\sigma$  as on  $a_\sigma$ , and the same labels on  $c$  as on  $a_{\pi_s}$ . Note that since  $\pi_s$  is on the infinite path  $\pi$ , there cannot be a label  $k_{\pi_s}$  on  $c$  and  $a_{\pi_s}$ . We must check that  $\mathcal{B}_e$  is in fact the union of these  $\mathcal{B}_{e,s}$ .

**Lemma 4.4.** *If there are infinitely many expansionary stages, then  $\lim_s \pi_s = \pi$ , and so  $\mathcal{B}_e = \bigcup_s \mathcal{B}_{e,s}$ .*

*Proof.* Suppose not, so that for some  $t$  for all  $s \geq t$  we have  $\pi_s = \pi_t$ . But then there is  $k$  such that  $\pi \upharpoonright_{|\pi_t|+1} \in \{0, \dots, k\}^{\leq k}$ , and at some stage  $s \geq t$  there is a  $k'$ th expansionary stage for some  $k' \geq k$ . At this stage, we add the children of  $\pi_t$  on  $T$  to  $T_{s+1}$ , so that  $\pi_{t+1}$  is one of these children and strictly extends  $\pi_t$ .  $\square$

Now we can show that  $\mathcal{B}_e \models \theta_e$ .

**Lemma 4.5.** *Suppose that  $T$  has an infinite path. If there are infinitely many expansionary stages, then  $\mathcal{B}_e \models \theta_e$ .*

*Proof.* Fix  $i$  and  $\bar{y}_i \in \mathcal{B}_e$ . Fix some stage  $s$  sufficiently large such that (a) by stage  $s$  there have been  $k$  expansionary stages, (b) for each  $b_\sigma$  in  $\bar{y}_i$  of the  $e$ th sort,  $\sigma \in \{0, \dots, k\}^{\leq k}$  and  $|\sigma| < |\pi_s|$ , and (c) each element of  $\bar{y}_i$  not in the  $e$ th sort is among the first  $k$  elements of  $\mathcal{A}$  (recalling that  $\mathcal{B}_e$  is the same as  $\mathcal{A}$  on these other sorts). Let  $\bar{x}_i$  be  $\bar{y}_i$  except that each  $b_\sigma$  in  $\bar{y}_i$  is replaced by  $a_\sigma$ , and  $c$  is replaced by  $a_{\pi_s}$ , and elements not of the  $e$  sort are kept the same. Then at some stage  $s' \geq s$  there is for the first time a  $k'$ th expansionary stage with  $\pi_{s'} \in \{0, \dots, k'\}^{\leq k'}$ . Since  $\bar{x}_i$  is  $k$ -small we have  $\mathcal{A}_{s'} \models \varphi_i(\bar{x}_i)$ . Since  $\mathcal{B}_{e,s'} \cong \mathcal{A}_{s'}$  via an isomorphism mapping  $b_\sigma \mapsto a_\sigma$  and  $c \mapsto a_{\pi_s}$  (and hence  $\bar{y}_i \mapsto \bar{x}_i$ ), we have  $\mathcal{B}_{e,s'} \models \varphi_i(\bar{y}_i)$ . Since  $\varphi_i$  is  $\Sigma_1$ ,  $\mathcal{B}_e \models \varphi_i(\bar{y}_i)$ . Thus, since this is true for all  $i$  and  $\bar{y}_i \in \mathcal{B}_e$ ,  $\mathcal{B}_e \models \theta_e$ .  $\square$

We have shown in Lemma 4.1 that  $\mathcal{A}$  has a  $\Pi_2$  Scott sentence. If  $T$  is well-founded, then Lemma 4.2 says that  $\mathcal{A}$  has a computable  $\Pi_2$  Scott sentence. Otherwise, if  $T$  has an infinite path, we argue that  $\mathcal{A}$  has no computable  $\Pi_2$  Scott sentence. If it did, say  $\theta_e$ , then we have  $\mathcal{A} \models \theta_e$ . Then by Lemma 4.3 there are infinitely many  $e$ -expansionary stages. We construct the structure  $\mathcal{B}_{e,s} \not\cong \mathcal{A}$  using a path through  $T$ , and by Lemma 4.5 we have that  $\mathcal{B}_e \models \theta_e$ . Thus in fact  $\theta_e$  cannot have been a Scott sentence for  $\mathcal{A}$ . Thus if  $T$  has an infinite path, we argue that  $\mathcal{A}$  has no computable  $\Pi_2$  Scott sentence.  $\square$

## 5 Corollaries and other commentary

### 5.1 Complexity of Scott families

In [AKM20] it was shown that if a computable structure has a c.e. Scott family of computable  $\Sigma_\alpha$  formulas then the structure has a computable  $\Pi_{\alpha+1}$  Scott sentence. Note that it is easy to construct uncountably many structures with a c.e. Scott family of  $\Sigma_1$  formulas, so it is necessary here that the structure is computable.

The following corollary shows that this theorem does not, in general, reverse.

**Corollary 5.1.** *There is a computable structure with a computable  $\Pi_2$  Scott sentence but with no c.e. Scott family of computable  $\Sigma_1$  formulas.*

*Proof.* Any computable structure with a c.e. Scott family of computable  $\Sigma_1$  formulas has a computable  $\Pi_2$  Scott sentence. However the index set of computable structures with a c.e. Scott family of computable  $\Sigma_1$  formulas is  $\Sigma_5^0$  while the index set of computable structures with a computable  $\Pi_2$  Scott sentence is  $\Pi_1^1$ - $m$ -complete. Thus there must be a computable structure with a computable  $\Pi_2$  Scott sentence but with no c.e. Scott family of computable  $\Sigma_1$  formulas.  $\square$

Despite this, one can ask how bad the Scott family for such a structure must be. Note that by Theorem 1.2 these give necessary but not sufficient conditions to have a computable  $\Pi_2$  Scott sentence.

**Proposition 5.2.** *Suppose  $A$  is computable and has a computable  $\Pi_{\alpha+1}$  Scott sentence. Then  $A$  has a c.e. Scott family of computable  $\Sigma_{\alpha+1}$  formulas.*

*Proof.* Note it is enough to show there is a  $\Sigma_{\alpha+1}$  Scott family of computable  $\Sigma_{\alpha+1}$  formulas. We will show a uniform procedure that given  $\bar{a} \in A$  produces a computable  $\Sigma_{\alpha+1}$  formula defining the automorphism orbit of  $\bar{a}$ . It is enough to find a formula that supports the computable  $\Pi_\alpha$  type of  $\bar{a}$ . It is  $\Pi_\alpha$  to list all the computable  $\Pi_\alpha$  formulas true of  $\bar{a}$ . Taking the conjunction of all such formulas we get a formula that supports the type and thus defines the automorphism orbit of  $\bar{a}$ . However, the  $\Pi_\alpha$  conjunction of computable  $\Pi_\alpha$  formulas is equivalent to a computable  $\Pi_\alpha$  formula. This computable  $\Pi_\alpha$  formula is at worst  $\Pi_\alpha$  to find given  $\bar{a}$ .  $\square$

**Proposition 5.3.** *Suppose  $A$  is a computable structure with a computable  $\Pi_{\alpha+1}$  Scott sentence. Then  $A$  has a  $\Delta_{\alpha+2}$  Scott family of computable  $\Sigma_\alpha$  formulas.*

*Proof.* We will show that there is a  $\Delta_{\alpha+2}$  map that, given  $\bar{a}$ , can find a  $\Sigma_\alpha$  formula defining the automorphism orbit of  $\bar{a}$ .

Since  $A$  has a computable  $\Pi_{\alpha+1}$  Scott sentence, there is a computable  $\Sigma_\alpha$  formula defining the automorphism orbit of  $\bar{a}$ . Thus, we know there is some computable  $\Pi_{<\alpha}$  formula  $\theta$  such that if  $\theta(\bar{c}\bar{b})$  ever is true, then  $\bar{b} \sim \bar{a}$ . Moreover, this is true of  $\bar{a}$  for some  $\bar{c}$ . Note now if  $\psi$  is any such formula then  $\exists \bar{x}\psi(\bar{x}\bar{y})$  defines the automorphism orbit of  $\bar{a}$ .

We now claim the set  $S$  of all such formulas  $\exists \bar{x}\psi(\bar{x}\bar{y})$  is  $\Pi_{\alpha+1}$ . To construct the set, we list all of the tuples  $\bar{c}, \bar{b} \in A$  and all of the computable  $\Pi_{<\alpha}$  formulas  $\psi(\bar{x}\bar{y})$ . As we go, we check if  $\exists \bar{x}\psi(\bar{x}, \bar{a})$  is true. This is  $\Sigma_\alpha$  to check. If so, we continue; if not, we throw out the formula  $\exists \bar{x}\psi(\bar{x}\bar{y})$  from the set. Now, we check as we go along if  $\psi(\bar{c}, \bar{b})$  holds which is  $\Pi_{<\alpha}$  to check. If so, we check if  $\bar{a} \not\sim \bar{b}$ , which is  $\Pi_\alpha$  to check. If  $\psi(\bar{c}, \bar{b})$  ever holds for some  $\bar{c}$  and  $\bar{a} \not\sim \bar{b}$ , we throw out the formula. What remains are those computable  $\Pi_\alpha$  formulas  $\psi$  such that  $\exists \bar{x}\psi(\bar{x}\bar{y})$  defines the automorphism orbit of  $\bar{a}$ .



To find one such formula, we list all candidate formulas and check if they are in the set, stopping when we get a “yes.” This process will halt since we know by assumption there is at least one such formula. Therefore the desired map can be computed via a  $\Pi_{\alpha+1}$  oracle, so the map is  $\Delta_{\alpha+2}$ .  $\square$

**Question 5.4.** Let  $\mathcal{A}$  be a computable structure with a computable  $\Pi_{\alpha+1}$  Scott sentence. Are the bounds on the Scott family obtained in Proposition 5.2 and Proposition 5.3 best possible?

## 5.2 Pseudo-Scott sentences

Recall that a pseudo-Scott sentence for a computable structure  $\mathcal{A}$  is a computable sentence  $\varphi$  such that for all *computable* structure  $\mathcal{B}$ ,

$$\mathcal{B} \models \varphi \iff \mathcal{A} \cong \mathcal{B}.$$

There are several examples of structures with a computable pseudo-Scott sentence of a certain complexity but no computable Scott sentence of that complexity [Ho17, Qui20]. As a corollary we get a much stronger example.

**Corollary 5.5.** *There is a computable structure  $\mathcal{A}$  with a  $\Pi_2^0$  Scott sentence but no computable  $\Pi_2^0$  Scott sentence, but with a computable  $\Pi_2^0$  sentence  $\varphi$  such that, for all hyperarithmetic  $\mathcal{B}$ ,*

$$\mathcal{B} \models \varphi \iff \mathcal{A} \cong \mathcal{B}.$$

*Proof.* Consider the index set of computable structures  $\mathcal{A}$  with a  $\Pi_2$  Scott sentence and such that there is a computable  $\Pi_2^0$  sentence  $\varphi$  such that for all hyperarithmetic structures  $\mathcal{B}$ ,  $\mathcal{B} \models \varphi$  if and only if  $\mathcal{A} \cong_2 \mathcal{B}$  (so that  $\mathcal{A} \cong \mathcal{B}$ ). Since we are quantifying universally over hyperarithmetic structures, this index set is  $\Sigma_1^1$ . Moreover, it is a superset of the  $\Pi_1^1$ - $m$ -complete set of computable structures with a computable  $\Pi_2^0$  sentence. Thus it must be a proper superset, proving the corollary.  $\square$

## 5.3 Working in $\text{Mod}(\mathcal{L})$

Sometimes by working in  $\text{Mod}(\mathcal{L})$  rather than with index sets we get a stronger theorem. In this case, it depends on how we relativize the statement and whether we work with boldface or lightface Borel classes.

**Proposition 5.6.** *The set of structures with a computable  $\Pi_2^0$  Scott sentence is a Borel—in fact (boldface)  $\Sigma_3^0$ —set in  $\text{Mod}(\mathcal{L})$ . It is (lightface)  $\Pi_1^1$  but, in a sufficiently rich language, not (lightface)  $\Sigma_1^1$ .*

*Proof.* A similar argument as above (in the first paragraph of Theorem 1.3) shows that it is (lightface)  $\Pi_1^1$ . It is also (boldface)  $\Sigma_1^1$ : listing out all of the countably many structures with a computable  $\Pi_2^0$  Scott sentence, together with their Scott sentences, a structure has a computable  $\Pi_2^0$  Scott sentence if and only if it is isomorphic to one of these. This is (boldface)  $\Sigma_3^0$ : given  $\mathcal{A}$ , we ask whether there exists a structure in this list such that  $\mathcal{A}$  satisfies the corresponding  $\Pi_2^0$  Scott sentence. If this set was (lightface)  $\Sigma_1^1$ , then the index set of computable such structures would be  $\Sigma_1^1$ ; but we know that this index set is  $\Pi_1^1$ - $m$ -complete.  $\square$

**Corollary 5.7.** *The set of structures  $\mathcal{A}$  with an  $\mathcal{A}$ -computable  $\Pi_2^0$  Scott sentence is  $\Pi_1^1$ -Wadge-complete set in  $\text{Mod}(\mathcal{L})$ .*

*Proof.* The proof is the same as that of Theorem 1.3, working relative to the tree  $T$ .  $\square$

## 5.4 Jump inversion

We can use the method of jump inversion or Marker extensions to generalize our results from  $\Pi_2$  sentences to  $\Pi_n$  sentences for arbitrary  $n$ . We get the following corollaries of our main theorem:

**Corollary 5.8.** *For each computable ordinal  $\alpha$  there is a computable structure with a  $\Pi_{\alpha+2}$  Scott sentence but with no computable  $\Sigma_{\alpha+4}$  Scott sentence.*

**Corollary 5.9.** *With  $(\mathcal{A}_i)_{i \in \omega}$  an effective list of (possibly partial) structures in a sufficiently rich language, for each computable ordinal  $\alpha$ , the set*

$$\{i \mid \mathcal{A}_i \text{ has a computable } \Pi_{\alpha+2} \text{ Scott sentence}\}$$

*is  $\Pi_1^1$ - $m$ -complete.*

Note that Marker extensions are only additive, so that we do not answer whether every computable structure with a  $\Pi_n$  Scott sentence has a computable  $\Sigma_{2n}$  Scott sentence for  $n \geq 3$ .

The particular flavour of jump inversion that we use is due to Goncharov, Harizanov, Knight, McCoy, R. Miller, and Solomon [GHK<sup>+</sup>05] and also given in Chapter X.3 of [Mon]. In Chapter X.3 of [Mon] Montalbán shows that given a computable ordinal  $\alpha$  and a structure  $\mathcal{A}$  there is a structure  $\Phi_\alpha(\mathcal{A})$  such that  $\mathcal{A}$  is effectively bi-interpretable with the  $\alpha$ -canonical structural jump of the image. There is a uniform effective construction from  $\mathcal{A}$  of  $\Phi_\alpha(\mathcal{A})$  as discussed in [CGHT]. We leave to [Mon] the definitions and properties of effective bi-interpretations and jumps of structures (see also [Mon09]).

If two structures are effectively bi-interpretable, then one has a  $\Pi_\beta$  Scott sentence if and only if the other does. Thus  $\mathcal{A}$  has a  $\Pi_\beta$  Scott sentence if and only if the  $\alpha$ -canonical structural jump of  $\Phi_\alpha(\mathcal{A})$  does. And the  $\alpha$ -canonical structural jump of a structure has a  $\Pi_\beta$  Scott sentence if and only if the original structure has a  $\Pi_{\alpha+\beta}$  Scott sentence. The same is true for computable Scott sentences. Thus:

**Proposition 5.10.** *Let  $\mathcal{A}$  be a computable structure and let  $\alpha$  and  $\beta$  be computable ordinals.*

- *$\mathcal{A}$  has a  $\Pi_\beta$  Scott sentence if and only if  $\Phi_\alpha(\mathcal{A})$  has a  $\Pi_{\alpha+\beta}$  Scott sentence.*
- *$\mathcal{A}$  has a computable  $\Pi_\beta$  Scott sentence if and only if  $\Phi_\alpha(\mathcal{A})$  has a computable  $\Pi_{\alpha+\beta}$  Scott sentence.*

*The same is true for  $\Sigma$  Scott sentences.*

From this we can prove the corollaries above.

*Proof of Corollary 5.8.* Let  $\mathcal{A}$  be a computable structure with a  $\Pi_2$  Scott sentence but no computable  $\Sigma_4$  Scott sentence. Then  $\Phi_\alpha(\mathcal{A})$  is a computable structure with a  $\Pi_{\alpha+2}$  Scott sentence but no computable  $\Sigma_{\alpha+4}$  Scott sentence.  $\square$

*Proof of Corollary 5.9.* Given a  $\Pi_1^1$  set  $X$ , we can construct for  $i \in X$  a computable structure  $\mathcal{A}_i$  such that  $i \in X$  if and only if  $\mathcal{A}_i$  has a computable  $\Pi_2$  Scott sentence. Then  $i \in X$  if and only if  $\Phi_\alpha(\mathcal{A}_i)$  has a computable  $\Pi_{\alpha+2}$  Scott sentence.  $\square$

## References

- [AGHTT21] Rachael Alvir, Noam Greenberg, Matthew Harrison-Trainor, and Dan Turetsky. Scott complexity of countable structures. *J. Symb. Log.*, 86(4):1706–1720, 2021.
- [AKM20] Rachael Alvir, Julia F. Knight, and Charles McCoy. Complexity of Scott sentences. *Fund. Math.*, 251(2):109–129, 2020.
- [AKMC20] Rachael Alvir, Julia F. Knight, and Charles McCoy CSC. Complexity of Scott sentences. *Fund. Math.*, 251(2):109–129, 2020.
- [BHTK<sup>+</sup>19] Nikolay Bazhenov, Matthew Harrison-Trainor, Iskander Kalimullin, Alexander Melnikov, and Keng Meng Ng. Automatic and polynomial-time algebraic structures. *Journal of Symbolic Logic*, 84(4):1630–1669, 2019.
- [CGHT] Ruiyuan Chen, David Gonzalez, and Matthew Harrison-Trainor. Optimal syntactic definitions of back-and-forth types. preprint.
- [GHK<sup>+</sup>05] Sergey Goncharov, Valentina Harizanov, Julia Knight, Charles McCoy, Russell Miller, and Reed Solomon. Enumerations in computable structure theory. *Ann. Pure Appl. Logic*, 136(3):219–246, 2005.
- [GK02] Sergey Goncharov and Julia A. Knight. Computable structure and non-structure theorems. *Algebra and Logic*, 41:351–373, 2002.
- [Har68] Joseph Harrison. Recursive pseudo-well-orderings. *Trans. Amer. Math. Soc.*, 131:526–543, 1968.
- [Ho17] Meng-Che Ho. Describing groups. *Proc. Amer. Math. Soc.*, 145(5):2223–2239, 2017.
- [HT18a] Matthew Harrison-Trainor. Scott ranks of models of a theory. *Advances in Mathematics*, 330:109–147, 2018.
- [HT18b] Matthew Harrison-Trainor. There is no classification of the decidably presentable structures. *Journal of Mathematical Logic*, 18(02):1850010, 2018.
- [HTIK18] Matthew Harrison-Trainor, Gregory Igusa, and Julia F. Knight. Some new computable structures of high rank. *Proc. Amer. Math. Soc.*, 146(7):3097–3109, 2018.
- [KLMC] Julia F. Knight, Karen Lange, and Charles McCoy CSC. Computable  $\Pi_2$  scott sentences. *Preprint*.
- [KM10] J. F. Knight and J. Millar. Computable structures of rank  $\omega_1^{\text{CK}}$ . *J. Math. Log.*, 10(1-2):31–43, 2010.
- [Mak81] M. Makkai. An example concerning Scott heights. *J. Symbolic Logic*, 46(2):301–318, 1981.
- [Mon] Antonio Montalbán. Computable structure theory: Beyond the arithmetic. In preparation.

- [Mon09] Antonio Montalbán. Notes on the jump of a structure. In Klaus Ambos-Spies, Benedikt Löwe, and Wolfgang Merkle, editors, *Mathematical Theory and Computational Practice*, pages 372–378, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.
- [Mon15] A. Montalbán. A robust scott rank. *Proc. Amer. Math. Soc.*, 143(12):5427–5436, 2015.
- [Mon17] Antonio Montalbán. Effectively existentially-atomic structures. In *Computability and complexity*, volume 10010 of *Lecture Notes in Comput. Sci.*, pages 221–237. Springer, Cham, 2017.
- [Mon21] Antonio Montalbán. *Computable Structure Theory: Within the Arithmetic*. Perspectives in Logic. Cambridge University Press, 2021.
- [Mor70] Michael Morley. The number of countable models. *J. Symbolic Logic*, 35:14–18, 1970.
- [Nad74] Mark Nadel. Scott sentences and admissible sets. *Ann. Math. Logic*, 7:267–294, 1974.
- [Qui20] Sara B. Quinn. Scott sentences for equivalence structures. *Arch. Math. Logic*, 59(3-4):453–460, 2020.
- [Sco65] Dana Scott. Logic with denumerably long formulas and finite strings of quantifiers. In *Theory of Models (Proc. 1963 Internat. Sympos. Berkeley)*, pages 329–341. North-Holland, Amsterdam, 1965.