Comparison of query learning and Gold-style learning in dependence of the hypothesis space

Steffen Lange[1] and Sandra Zilles[2]

[1] Fachhochschule Darmstadt,
FB Informatik, Haardtring 100, 64295 Darmstadt, Germany,
e-mail: s.lange@fbi.fh-darmstadt.de
[2] Technische Universität Kaiserslautern,
FB Informatik, Postfach 3049, 67653 Kaiserslautern, Germany,
e-mail: zilles@informatik.uni-kl.de

Abstract. Different formal learning models address different aspects of learning. Below we compare *learning via queries*—interpreting learning as a *one-shot process* in which the learner is required to identify the target concept with just one hypothesis—to *Gold-style learning*—interpreting learning as a *limiting process* in which the learner may change its mind arbitrarily often before converging to a correct hypothesis.

Although these two approaches seem rather unrelated, a previous study has provided characterisations of different models of Gold-style learning (learning in the limit, conservative inference, and behaviourally correct learning) in terms of query learning. Thus under certain circumstances it is possible to replace limit learners by equally powerful one-shot learners. Both this previous and the current analysis are valid in the general context of learning indexable classes of recursive languages.

The main purpose of this paper is to solve a challenging open problem from the previous study. The solution of this problem leads to an important observation, namely that there is a natural query learning type hierarchically in-between Gold-style learning in the limit and behaviourally correct learning. Astonishingly, this query learning type can then again be characterised in terms of Gold-style inference.

In connection with this new in-between inference type we have gained new insights into the basic model of conservative learning and the way conservative learners work. In addition to these results, we compare several further natural inference types in both models to one another.

1 Introduction

Undeniably, there is no formal scheme spanning all aspects of human learning. Thus each learning model analysed within the scope of learning theory addresses only special facets of our understanding of learning. For example, Angluin's [2, 3] model of *learning with queries* focusses learning as a finite process of interaction between a learner and a teacher. The learner asks questions of a specified type about the target concept and the teacher answers these questions truthfully. After finitely many steps of interaction the learner is supposed to return its sole

hypothesis—correctly describing the target concept. Here the crucial features of the learner are its ability to demand special information on the target concept and its restrictiveness in terms of mind changes. Since a query learner is required to identify the target concept with just a single hypothesis, we refer to this phenomenon as *one-shot learning*.[3]

In contrast to that, Gold's [7] model of *identification in the limit* is concerned with learning as a limiting process of creating, modifying, and improving hypotheses about a target concept. These hypotheses are based upon instances of the target concept offered as information. In the limit, the learner is supposed to stabilize on a correct guess, but during the learning process one will never know whether or not the current hypothesis is already correct. Here the ability to change its mind is a crucial feature of the learner.

[11] is concerned with a first systematic analysis of common features of these two seemingly unrelated approaches, thereby focussing on the identification of formal languages, ranging over indexable classes of recursive languages, as target concepts, see [1, 9, 14]. Characterising different types of Gold-style language learning in terms of query learning has pointed out interesting correspondences between the two models. In particular, the results in [11] demonstrate how learners identifying languages in the limit can be replaced by one-shot learners without loss of learning power. That means, under certain circumstances the capabilities of limit learners are equal to those of one-shot learners using queries.

The analysis summarized in this paper has initially been motivated by an open problem in [11], namely whether or not the capabilities of query learners using superset queries in Gödel numberings are equal to those of behaviourally correct Gold-style learners using Gödel numberings as their hypothesis spaces. Below we will answer this question to the negative, which will lead to the following astonishing observation: there is a natural inference type (learning via superset queries in Gödel numberings) which lies *in-between* Gold-style learning in the limit from text and behaviourally correct Gold-style learning from text in Gödel numberings.[4] Up to now, no such inference type has been known.

This observation immediately raises a second question, namely whether there is an analogue of this query learning type in terms of Gold-style learning and thus whether there is also a *Gold-style* inference type between learning in the limit and behaviourally correct learning. Indeed such a relation can be observed with conservative inference in Gödel numberings by learners using an oracle for the halting problem; see [13] for further results on learning with oracles.

Studying such relations between two different approaches to language learning allows for transferring theoretically approved insights from one model to the other. In particular, our characterisations may serve as 'interfaces' between an analysis of query learning and an analysis of Gold-style learning through which proofs on either model can be simplified using properties of the other.

---

[3] Most studies on query learning mainly deal with the efficiency of query learners, whereas below we are only interested in qualitative learnability results in this context.

[4] That means that the capabilities of the corresponding learners lie in-between. Concerning the notions of inference, see [7, 1, 14] and the preliminaries below.

Most of our proofs provided below make use of recursion-theoretic conceptions, thus in particular providing a second quite accessible example—after the one in [11]—for a class identifiable by a behaviourally correct learner in Gödel numberings but not identifiable in the limit. The interesting feature of these classes is that they are defined without any diagonal construction—very unlike the corresponding classes known before, see for instance [1].

Comparing our results to a result from [13] points out a related open problem in Gold-style learning: Note that, by [13], an indexable class is learnable conservatively using an oracle for the halting problem and a *uniformly recursive hypothesis space* if and only if it is learnable in the limit. In contrast to that, we show that conservative learners using an oracle for the halting problem and a *Gödel numbering as a hypothesis space* are more capable than limit learners. This implies that, in the context of conservative inference, oracle learners may benefit from using a Gödel numbering instead of uniformly recursive numbering as a hypothesis space. Now the related open problem is: do conservative learners (without the help of oracles) also benefit from Gödel numberings instead of uniformly recursive numberings? Though this question is quite natural, it has not been discussed in the literature so far. Unfortunately, we can provide an answer only for a special case: if a learner is required to work both conservatively and consistently on the relevant data, Gödel numberings do not increase the capabilities when compared to uniformly recursive hypothesis spaces. Additional results below relate several further natural inference types in both models to each other.

2 Preliminaries

Familiarity with standard recursion theoretic and language theoretic notions is assumed, see [12, 8]. Subsequently, let Σ be a finite alphabet with $\{a, b\} \subseteq \Sigma$. A *word* is any element from Σ^* and a *language* any subset of Σ^*. The *complement* $\overline{L}$ of a language L is the set $\Sigma^* \setminus L$. Any infinite sequence $t = (w_i)_{i \in \mathbb{N}}$ with $\{w_i \mid i \in \mathbb{N}\} = L$ is called a *text* for L. Then, for any $n \in \mathbb{N}$, t_n denotes the initial segment $(w_0, \ldots, w_n)$ and *content*(t_n) denotes the set $\{w_0, \ldots, w_n\}$.

A family $(A_i)_{i \in \mathbb{N}}$ of languages is *uniformly recursive (uniformly r. e.)* if there is a recursive (partial recursive) function f with $A_i = \{w \in \Sigma^* \mid f(i, w) = 1\}$ for all $i \in \mathbb{N}$. A family $(A_i)_{i \in \mathbb{N}}$ is *uniformly 2-r. e.*, if there is a recursive function g with $A_i = \{w \in \Sigma^* \mid g(i, w, n) = 1$ for all but finitely many $n\}$ for all $i \in \mathbb{N}$. Note that for uniformly recursive families membership is uniformly decidable.

Let $\mathcal{C}$ be a class of recursive languages. $\mathcal{C}$ is said to be an *indexable class of recursive languages* (in the sequel we will write *indexable class* for short), if there is a uniformly recursive family $(L_i)_{i \in \mathbb{N}}$ of all and only the languages in $\mathcal{C}$. Such a family will subsequently be called an *indexing* of $\mathcal{C}$.

A family $(T_i)_{i \in \mathbb{N}}$ of *finite* languages is *recursively generable*, if there is a recursive function that, given $i \in \mathbb{N}$, enumerates all elements of T_i and stops.

In the sequel, let φ be a Gödel numbering of all partial recursive functions and Φ the associated Blum complexity measure, see [5] for a definition. For $i, n \in \mathbb{N}$ we will write $\varphi_i[n]$ for the initial segment $(\varphi_i(0), \ldots, \varphi_i(n))$ and say that $\varphi_i[n]$ is

defined if all the values $\varphi_i(0), \ldots, \varphi_i(n)$ are defined. For convenience, $\varphi_i[-1]$ is always considered defined. Moreover, let $Tot = \{i \in \mathbb{N} \mid \varphi_i \text{ is a total function}\}$ and $K = \{i \in \mathbb{N} \mid \varphi_i(i) \text{ is defined}\}$. The family $(W_i)_{i \in \mathbb{N}}$ of languages is given by $W_i = \{w_j^* \mid \varphi_i(j) \text{ is defined}\}$ for all $i \in \mathbb{N}$, where $(w_j^*)_{j \in \mathbb{N}}$ is some fixed effective enumeration of Σ^* without repetitions. Moreover, we use a bijective recursive function coding a pair (x, y) with $x, y \in \mathbb{N}$ into a number $\langle x, y \rangle \in \mathbb{N}$.

2.1 Language learning via queries

In the query learning model, a learner has access to a teacher that truthfully answers queries of a specified kind. A *query learner* M is an algorithmic device that, depending on the reply on the previous queries, either computes a new query or returns a hypothesis and halts, see [2]. Its queries and hypotheses are coded as natural numbers; both will be interpreted with respect to an underlying *hypothesis space*. When learning an indexable class $\mathcal{C}$, any indexing $\mathcal{H} = (L_i)_{i \in \mathbb{N}}$ of $\mathcal{C}$ may form a hypothesis space. So, as in the original definition, see [2], when learning $\mathcal{C}$, M is only allowed to query languages belonging to $\mathcal{C}$.

More formally, let $\mathcal{C}$ be an indexable class, let $L \in \mathcal{C}$, let $\mathcal{H} = (L_i)_{i \in \mathbb{N}}$ be an indexing of $\mathcal{C}$, and let M be a query learner. *M learns L with respect to $\mathcal{H}$ using some type of queries* if it eventually halts and its only hypothesis, say i, correctly describes L, i. e., $L_i = L$. So M returns its unique and correct guess i after only finitely many queries. Moreover, *M learns $\mathcal{C}$ with respect to $\mathcal{H}$ using some type of queries*, if it learns every $L' \in \mathcal{C}$ with respect to $\mathcal{H}$ using queries of the specified type. In order to learn a language L, a query learner M may ask:

Membership queries. The input is a string w and the answer is 'yes' or 'no', depending on whether or not w belongs to L.

Restricted superset queries. The input is an index of a language $L' \in \mathcal{C}$. The answer is 'yes' or 'no', depending on whether or not L' is a superset of L.

Restricted disjointness queries. The input is an index of a language $L' \in \mathcal{C}$. The answer is 'yes' or 'no', depending on whether or not L' and L are disjoint.

$MemQ$, $rSupQ$, and $rDisQ$ denote the collections of all indexable classes $\mathcal{C}'$ for which there are a query learner M' and a hypothesis space $\mathcal{H}'$ such that M' learns $\mathcal{C}'$ with respect to $\mathcal{H}'$ using membership, restricted superset, and restricted disjointness queries, respectively. In the sequel we will omit the term 'restricted' for convenience and also neglect other types of queries analysed in the literature, see [2, 3]. Obviously, superset and disjointness queries are in general not decidable, i. e. the teacher may be non-computable.

Note that learning via queries focusses the aspect of one-shot learning, i. e., it is concerned with scenarios in which learning eventuates without mind changes.

Having a closer look at the different models of query learning, one easily finds negative learnability results. Some examples in [11] point to a drawback of Angluin's query model, namely the demand that a query learner is restricted to pose queries concerning languages contained in the class of possible target languages. That means there are very simple classes of languages, for which any learner must fail just because it is barred from asking the 'appropriate' queries.

To overcome this drawback, it seems reasonable to allow the query learner to formulate its queries with respect to any uniformly recursive family comprising the target class $\mathcal{C}$. An *extra query learner* (see also [10, 11]) for an indexable class $\mathcal{C}$ is permitted to query languages in any uniformly recursive family $(L_i')_{i \in \mathbb{N}}$ comprising $\mathcal{C}$. We say that $\mathcal{C}$ is learnable with extra superset (disjointness) queries respecting $(L_i')_{i \in \mathbb{N}}$ iff there is an extra query learner M learning $\mathcal{C}$ with respect to $(L_i')_{i \in \mathbb{N}}$ using superset (disjointness) queries concerning $(L_i')_{i \in \mathbb{N}}$. Then $rSupQ_{\mathrm{rec}}$ ($rDisQ_{\mathrm{rec}}$) denotes the collection of all indexable classes $\mathcal{C}$ learnable with extra superset (disjointness) queries respecting a uniformly recursive family.

It is conceivable to permit even more general hypothesis spaces, i.e., to demand a more potent teacher. Thus, let $rSupQ_{\mathrm{r.e.}}$ ($rDisQ_{\mathrm{r.e.}}$) denote the collection of all indexable classes which are learnable with superset (disjointness) queries respecting a uniformly r.e. family. Obviously, each class in $rSupQ_{\mathrm{r.e.}}$ ($rDisQ_{\mathrm{r.e.}}$) can be identified respecting our fixed numbering $(W_i)_{i \in \mathbb{N}}$. Similarly, replacing the subscript 'r.e.' by '2-r.e.', we consider learning in a uniformly 2-r.e. family.

Note that the capabilities of $rSupQ$-learners ($rDisQ$-learners) already increase with the additional permission to ask membership queries. Yet, as has been shown in [11], combining superset or disjointness queries with membership queries does not yield the same capability as extra queries do. For convenience, we denote the family of classes which are learnable with a combination of superset (disjointness) and membership queries by $rSupMemQ$ ($rDisMemQ$).

2.2 Gold-style language learning

Let $\mathcal{C}$ be an indexable class, $\mathcal{H} = (L_i)_{i \in \mathbb{N}}$ any uniformly recursive family (called *hypothesis space*), and $L \in \mathcal{C}$. An *inductive inference machine* (*IIM*) M is an algorithmic device that reads longer and longer initial segments t_n of a text and outputs numbers $M(t_n)$ as its hypotheses. An IIM M returning some i is construed to hypothesize the language L_i. Given a text t for L, M *identifies L from t with respect to $\mathcal{H}$ in the limit*, if the sequence of hypotheses output by M, when fed t, stabilizes on a number i (i.e., past some point M always outputs the hypothesis i) with $L_i = L$. M *identifies $\mathcal{C}$ in the limit from text* with respect to $\mathcal{H}$, if it identifies every $L' \in \mathcal{C}$ from every corresponding text. $Lim\,Txt_{\mathrm{rec}}$ denotes the collection of all indexable classes $\mathcal{C}'$ for which there are an IIM M' and a uniformly recursive family $\mathcal{H}'$ such that M' identifies $\mathcal{C}'$ in the limit from text with respect to $\mathcal{H}'$. A quite natural and often studied modification of $Lim\,Txt_{\mathrm{rec}}$ is defined by the model of *conservative inference*, see [1]. M is a *conservative* IIM for $\mathcal{C}$ with respect to $\mathcal{H}$, if M performs only justified mind changes, i.e., if M, on some text t for some $L \in \mathcal{C}$, outputs hypotheses i and later j, then M must have seen some element $w \notin L_i$ before returning j. An important property of conservative learners is that they never hypothesize proper supersets of the language currently to be learned. The collection of all indexable classes identifiable from text by a conservative IIM is denoted by $Consv\,Txt_{\mathrm{rec}}$. Note that $Consv\,Txt_{\mathrm{rec}} \subset Lim\,Txt_{\mathrm{rec}}$ [14]. Since we consider learning from text only, we will assume in the sequel that all languages to be learned are *non-empty*. One main aspect of human learning modelled in the approach of learning in the limit

is the ability to change one's mind during learning. Thus learning is a process in which the learner may change its hypothesis arbitrarily often before stabilizing on its final correct guess. In particular, it is undecidable whether or not the final hypothesis has been reached, i.e., whether or not a success in learning has already eventuated.

If only uniformly recursive families are used as hypothesis spaces, $Lim\,Txt_{\mathrm{rec}}$ coincides with the collection of indexable classes identifiable in a behaviourally correct manner, see [6]: If $\mathcal{C}$ is an indexable class, $\mathcal{H} = (L_i)_{i\in\mathbb{N}}$ a uniformly recursive family, M an IIM, then M is a *behaviourally correct* learner for $\mathcal{C}$ from text with respect to $\mathcal{H}$, if for each $L \in \mathcal{C}$ and each text t for $\mathcal{C}$, all but finitely many outputs i of M on t fulfil $L_i = L$. Here M may alternate different correct hypotheses instead of converging to a single hypothesis. Defining $Bc\,Txt_{\mathrm{rec}}$ correspondingly as usual yields $Bc\,Txt_{\mathrm{rec}} = Lim\,Txt_{\mathrm{rec}}$ (a folklore result).

This relation no longer holds, if more general types of hypothesis spaces are considered. Assume $\mathcal{C}$ is an indexable class and $\mathcal{H}^+ = (A_i)_{i\in\mathbb{N}}$ is any uniformly r.e. family of languages comprising $\mathcal{C}$. Then it is also conceivable to use $\mathcal{H}^+$ as a hypothesis space. For $I \in \{Lim, Consv, Bc\}$, $I\,Txt_{\mathrm{r.e.}}$ denotes the collection of all indexable classes learnable as in the definition of $I\,Txt_{\mathrm{rec}}$, if the demand for a uniformly recursive family $\mathcal{H}$ as a hypothesis space is loosened to demanding a uniformly r.e. family $\mathcal{H}^+$ as a hypothesis space. Note that each class in $I\,Txt_{\mathrm{r.e.}}$ can also be $I\,Txt$-identified in the hypothesis space $(W_i)_{i\in\mathbb{N}}$. Interestingly, $Lim\,Txt_{\mathrm{rec}} = Lim\,Txt_{\mathrm{r.e.}}$ (a folklore result), i.e., in learning in the limit, the capabilities of IIMs do not increase, if the constraints concerning the hypothesis space are weakened by allowing for arbitrary uniformly r.e. families. In contrast to that, for $Bc\,Txt$-identification, weakening these constraints yields an add-on in learning power, i.e., $Bc\,Txt_{\mathrm{rec}} \subset Bc\,Txt_{\mathrm{r.e.}}$. In particular, $Lim\,Txt_{\mathrm{rec}} \subset Bc\,Txt_{\mathrm{r.e.}}$ and so $Lim\,Txt$- and $Bc\,Txt$-learning no longer coincide for identification with respect to arbitrary uniformly r.e. families, see also [4, 1].

The main results of our analysis will be comparisons of these inference types with different query learning types. For that purpose we will make use of well-known characterizations based on so-called families of *telltales*, see [1].

Definition 1. *Let $(L_i)_{i\in\mathbb{N}}$ be a uniformly recursive family and $(T_i)_{i\in\mathbb{N}}$ a family of finite non-empty sets. $(T_i)_{i\in\mathbb{N}}$ is a telltale family for $(L_i)_{i\in\mathbb{N}}$ iff for all $i, j \in \mathbb{N}$:*

1. *$T_i \subseteq L_i$.*
2. *If $T_i \subseteq L_j \subseteq L_i$, then $L_j = L_i$.*

Telltale families are the best known concept to illustrate the specific differences between indexable classes in $Lim\,Txt_{\mathrm{rec}}$, $Consv\,Txt_{\mathrm{rec}}$, and $Bc\,Txt_{\mathrm{r.e.}}$. Their algorithmic structure has turned out to be crucial for learning, see [1, 9, 4]:

Theorem 1. *Let $\mathcal{C}$ be an indexable class of languages.*

1. *$\mathcal{C} \in Lim\,Txt_{\mathrm{rec}}$ iff there is an indexing of $\mathcal{C}$ possessing a uniformly r.e. family of telltales.*
2. *$\mathcal{C} \in Consv\,Txt_{\mathrm{rec}}$ iff there is a uniformly recursive family comprising $\mathcal{C}$ and possessing a recursively generable family of telltales.*
3. *$\mathcal{C} \in Bc\,Txt_{\mathrm{r.e.}}$ iff there is an indexing of $\mathcal{C}$ possessing a family of telltales.*

3 Hypothesis spaces in query learning

Concerning the influence of the query and hypothesis spaces in query learning, various interesting results have been established in [11]. These reveal a hierarchy of capabilities of query learners resulting from a growing generality of the hypothesis spaces. Interestingly, in some but not in all cases, the capabilities of superset query learners and disjointness query learners coincide:

Theorem 2. *[11]*
$rSupQ_{\mathrm{rec}} = rDisQ_{\mathrm{rec}} \subset rDisQ_{\mathrm{r.e.}} \subset rSupQ_{\mathrm{r.e.}} \subseteq rSupQ_{\text{2-r.e.}} = rDisQ_{\text{2-r.e.}}$.

In [11] it has remained open, whether or not there is an indexable class in $rSupQ_{\text{2-r.e.}} \setminus rSupQ_{\mathrm{r.e.}}$—a problem which will be solved below. Moreover, [11] studies original superset (disjointness) query learners which are additionally permitted to ask membership queries. Their capabilities are in-between those of the original learners and extra query learners.

Theorem 3. *[11] (a) $rSupQ \subset rSupMemQ \subset rSupQ_{\mathrm{rec}}$.*
(b) $rDisQ \subset rDisMemQ \subset rDisQ_{\mathrm{rec}}$.

Comparing these results, notice that Theorem 2 analyses the relationship between superset and disjointness query learning, whereas Theorem 3 avoids corresponding statements. That means, it has remained open, how the inference types $rSupQ$ and $rSupMemQ$ relate to $rDisQ$ and $rDisMemQ$.

As an answer to this question, we can state that $rSupQ$ and $rSupMemQ$ are incomparable to both $rDisQ$ and $rDisMemQ$, an immediate consequence of the following theorem.

Theorem 4. *(a) $rSupQ \nsubseteq rDisMemQ$.   (b) $rDisQ \nsubseteq rSupMemQ$.*

Proof. We provide the separating indexable classes without a proof.

(a) The class $\mathcal{C}_a$ containing $L = \{b\}$ and $L_k = \{a^k, b\}$ for $k \geq 0$ belongs to $rSupQ \setminus rDisMemQ$.

(b) A class $\mathcal{C}_b \in rDisQ \setminus rSupMemQ$ is defined as follows: for $k \in \mathbb{N}$, let $\mathcal{C}_b$ contain the languages $L_k = \{a^k b^z \mid z \geq 0\}$ and $L'_k = \{ba^{\langle k,j \rangle} \mid j \geq 0\}$. Additionally, if $k \notin K$, then $\mathcal{C}_b$ contains $L^1_{k,j} = \{ba^{\langle k,j \rangle}\}$ for all $j \in \mathbb{N}$; whereas, if $k \in K$, then $\mathcal{C}_b$ contains $L^2_{k,j} = \{a^k b^z \mid z \leq \Phi_k(k) \text{ or } z > \Phi_k(k) + j\}$ as well as $L^3_{k,j} = \{ba^{\langle k,j \rangle}\} \cup \{a^k b^{\Phi_k(k)+j}\}$ for all $j \in \mathbb{N}$. $\qquad\square$

The more challenging open question is whether or not the inference types $rSupQ_{\mathrm{r.e.}}$ and $rSupQ_{\text{2-r.e.}}$ coincide. Interestingly, this is not the case, that means, 2-r. e. numberings provide a further benefit for learning with superset queries.

Theorem 5. $rSupQ_{\mathrm{r.e.}} \subset rSupQ_{\text{2-r.e.}}$.

Though our current tools allow for a verification of this theorem, the proof would be rather lengthy. Since a characterisation of $rSupQ_{\mathrm{r.e.}}$ in terms of Gold-style learning simplifies the proof considerably, we postpone the proof for now.

4 Query learning and Gold-style learning — relations

In [11], a couple of relations between query learning and Gold-style learning have been elaborated. The following theorem summarizes the corresponding results.

Theorem 6. *[11] (a)* $rSupQ_{\mathrm{rec}} = rDisQ_{\mathrm{rec}} = Consv\,Txt_{\mathrm{rec}}$.
(b) $rDisQ_{\mathrm{r.e.}} = Lim\,Txt_{\mathrm{rec}} = Bc\,Txt_{\mathrm{rec}}$.
(c) $rSupQ_{\text{2-r.e.}} = rDisQ_{\text{2-r.e.}} = Bc\,Txt_{\mathrm{r.e.}}$.

By Theorems 2 and 5 this implies $Lim\,Txt_{\mathrm{rec}} \subset rSupQ_{\mathrm{r.e.}} \subset Bc\,Txt_{\mathrm{r.e.}}$, i.e., we have found a natural type of learners the capabilities of which are strictly between those of $Lim\,Txt$-learners and those of $Bc\,Txt$-learners. This raises the question whether the learning type $rSupQ_{\mathrm{r.e.}}$ can also be characterised in terms of Gold-style learning. This is indeed possible if we consider learners which have access to some oracle. In the sequel the notion $Consv\,Txt_{\mathrm{r.e.}}[K]$ refers to the collection of indexable classes which are learnable in the sense of $Consv\,Txt_{\mathrm{r.e.}}$, if also K-recursive learners are admitted, see [13].

Theorem 7. $rSupQ_{\mathrm{r.e.}} = Consv\,Txt_{\mathrm{r.e.}}[K]$.

Proof. First, we prove $rSupQ_{\mathrm{r.e.}} \subseteq Consv\,Txt_{\mathrm{r.e.}}[K]$. For that purpose assume $\mathcal{C}$ is an indexable class in $rSupQ_{\mathrm{r.e.}}$. Let M be a query learner identifying $\mathcal{C}$ in $(W_i)_{i\in\mathbb{N}}$ and assume wlog that each hypothesis ever returned by M corresponds to the intersection of all queries answered with 'yes' in the preceding scenario.[5]

Let $L \in \mathcal{C}$, t a text for L. Let $M'(t_0)$ be an index of the language $content(t_0)$. Given $n \geq 1$, a learner M' works on input t_n as follows: M' simulates M for n steps of computation. Whenever M asks a superset query i, M' transmits the answer 'yes' to M, if $content(t_n) \subseteq W_i$, the answer 'no', otherwise (* this test is K-recursive *). If M returns a hypothesis i within n steps of computation, let M' return i on t_n; otherwise let $M'(t_n) = M'(t_{n-1})$.

Note that there must be some n, such that M' answers all queries of M truthfully respecting L. Thus it is not hard to verify that the K-recursive IIM M' learns L in the limit from text. Moreover, $W_{M'(t_n)} \not\supset L$ for all n: assuming $W_{M'(t_n)} \supset L$ implies, by normalisation of M, that all queries M' has answered with 'yes' in the simulation of M indeed represent supersets of L. Since all 'no'-answers are truthful respecting L by definition, this yields a truthful query-scenario for L. As M learns L from superset queries, the hypothesis i must correctly describe L—a contradiction. So M' learns $\mathcal{C}$ without ever returning an index of a proper superset of a language currently to be identified. Now it is not hard to modify M' into a K-recursive IIM which works conservatively for the class $\mathcal{C}$ (a hypothesis will only be changed if its inconsistency is verified with the help of a K-oracle). Thus $\mathcal{C} \in Consv\,Txt_{\mathrm{r.e.}}[K]$ and $rSupQ_{\mathrm{r.e.}} \subseteq Consv\,Txt_{\mathrm{r.e.}}[K]$.

---

[5] Think of M as a normalisation of a superset query learner M^-: M copies M^- until M^- returns the hypothesis i. Now M asks a query for the language W_i instead of returning a hypothesis. Then let M return a hypothesis j representing the intersection of all queries answered with 'yes' in its preceding scenario. Given a fair scenario for W_i and a successful learner M^-, this implies $W_i = W_j$ and thus M is successful.

Second, we show $Consv\,Txt_{\mathrm{r.e.}}[K] \subseteq rSupQ_{\mathrm{r.e.}}$. For that purpose assume $\mathcal{C}$ is an indexable class in $Consv\,Txt_{\mathrm{r.e.}}[K]$. Let M be a K-recursive IIM identifying $\mathcal{C}$ with respect to $(W_i)_{i\in\mathbb{N}}$. Suppose $L \in \mathcal{C}$ is the target language. An $rSupQ$-learner M' for L with respect to $(W_i)_{i\in\mathbb{N}}$ is defined by steps, starting in step 0. Note that representations in $(W_i)_{i\in\mathbb{N}}$ can be computed for all queries to be asked. In step 0, M' finds the minimal m, such that the query for $\Sigma^* \setminus \{w_m^*\}$ is answered 'no'. M' sets $t(0) = w_m^*$ and goes to step 1. In general, step $n+1$ reads as follows:

– Ask a superset query for $\Sigma^* \setminus \{w_{n+1}^*\}$. If the answer is 'no', let $t(n+1) = w_{n+1}^*$; if the answer is 'yes', let $t(n+1) = t(n)$.
(* Note that $content(t_{n+1}) = L \cap \{w_x^* \mid x \le n+1\}$. *)
– Simulate M on input t_{n+1}. Whenever M wants to access a K-oracle for the question whether $j \in K$, formulate a superset query for the language

$$W_j' = \begin{cases} \Sigma^*, & \text{if } \varphi_j(j) \text{ is defined}, \\ \emptyset, & \text{otherwise}. \end{cases}$$

and transmit the received answer to M.
(* Note that W_j' is uniformly r.e. in j and $W_j' \supseteq L$ iff $\varphi_j(j)$ is defined. *)
As soon as M returns $i = M(t_{n+1})$, pose a superset query for W_i. If the answer is 'yes', then return the hypothesis i and stop (* since M learns L conservatively, we have $W_i \not\supseteq L$ and thus $W_i = L$ *). If the answer is 'no', then go to step $n+2$.

Now it is not hard to verify that M' learns $\mathcal{C}$ with superset queries in $(W_i)_{i\in\mathbb{N}}$. Details are omitted. Thus $\mathcal{C} \in rSupQ_{\mathrm{r.e.}}$ and $Consv\,Txt_{\mathrm{r.e.}}[K] \subseteq rSupQ_{\mathrm{r.e.}}$. □

Using this characterisation, Theorem 5 translates as follows:

Theorem 5' $Consv\,Txt_{\mathrm{r.e.}}[K] \subset Bc\,Txt_{\mathrm{r.e.}}$.

Proof. By Theorems 6 and 7 it suffices to prove $Bc\,Txt_{\mathrm{r.e.}} \setminus Consv\,Txt_{\mathrm{r.e.}}[K] \ne \emptyset$. For that purpose we provide an indexable class $\mathcal{C}_{\mathrm{bc}} \in Bc\,Txt_{\mathrm{r.e.}} \setminus Consv\,Txt_{\mathrm{r.e.}}[K]$. For all $k \in \mathbb{N}$, $\mathcal{C}_{\mathrm{bc}}$ contains the language $L_k = \{a^k b^z \mid z \ge 0\}$. Moreover, for all $k, i, j \in \mathbb{N}$ for which $\varphi_k[i-1]$ is defined and $j \le i$, let $\mathcal{C}_{\mathrm{bc}}$ contain the language

$$L_{k,i,j} = \begin{cases} \{a^k b^z \mid z \le j\}, & \text{if } \varphi_k(i) \text{ is undefined}, \\ \{a^k b^z \mid z \le j\} \cup \{ba^{\Phi_k(i)}\}, & \text{if } \varphi_k(i) \text{ is defined}. \end{cases}$$

To show that $\mathcal{C}_{\mathrm{bc}} \in Bc\,Txt$, it suffices by Theorem 1 to prove the existence of telltales corresponding to some indexing of $\mathcal{C}_{\mathrm{bc}}$. This is quite simple: as each language $L_{k,i,j} \in \mathcal{C}_{\mathrm{bc}}$ is finite, it forms a telltale for itself. Moreover, as for all k there are only finitely many subsets of L_k in $\mathcal{C}_{\mathrm{bc}}$, telltales for L_k must exist, too.

Finally, it remains to prove that $\mathcal{C}_{\mathrm{bc}} \notin Consv\,Txt_{\mathrm{r.e.}}[K]$. Assume the opposite, i.e., there is some K-recursive IIM M which $Consv\,Txt$-identifies $\mathcal{C}_{\mathrm{bc}}$ in $(W_i)_{i\in\mathbb{N}}$. The idea is to deduce a contradiction by concluding that Tot is K-recursive. For that purpose, define a K-recursive procedure on input k as follows:

– Let $t = a^k, a^k b, a^k b^2, \ldots$ be the 'canonical' text for L_k.
– Simulate M on input $t_0, t_1, t_2, \ldots$ until some n is found with $content(t_n) \subset W_{M(t_n)} \subseteq L_k$. (* n exists, as M learns L_k. Determining n is K-recursive. *)
– If $\varphi_k(i)$ is defined for all $i \le n$, then return '1'; otherwise return '0'.

Obviously, this procedure is K-recursive. Note that it returns '0' only in case φ_k is not total. So assume it returns '1'. Then there is some n such that $content(t_n) \subset W_{M(t_n)} \subseteq L_k$. If φ_k was not total, the minimal i for which $\varphi_k(i)$ is undefined would be greater than n. Thus $L = \{a^k b^z \mid z \leq n\} \in \mathcal{C}_{\mathrm{bc}}$. Now t_n is also a text segment for L, but $L = content(t_n) \subset W_{M(t_n)}$. Thus M hypothesizes a proper superset of L on input t_n and hence M fails to learn L conservatively. This contradicts the choice of M, so φ_k is total.

Consequently, our procedure decides Tot, i.e., Tot is K-recursive. As this is impossible, we have $\mathcal{C}_{\mathrm{bc}} \notin Consv\,Txt_{\mathrm{r.e.}}[K]$. □

In particular, the class $\mathcal{C}_{\mathrm{bc}}$ defined in this proof constitutes a second quite accessible example—after the one in [11]—for a class identifiable by a behaviourally correct learner in Gödel numberings but not identifiable in the limit. An interesting feature is that these classes are defined without any diagonal construction—very unlike the corresponding classes known before, see for instance [1].

Finally, thus Theorem 5 is proven, too. This is an example for the advantages of our characterisations; verifying Theorem 5 without Theorems 6 and 7 would have been possible, but more complicated. So features of Gold-style learning can be exploited in the context of query learning. Note that, with Theorems 6 and 7, we have characterised all types of extra query learning in terms of Gold-style learning. In order to better describe and understand the capabilities of the original query learners (the types $rSupQ$, $rSupMemQ$, $rDisQ$, $rDisMemQ$), let us have a closer look at the results established up to now.

We know that $rSupQ_{\mathrm{rec}} = rDisQ_{\mathrm{rec}} = Consv\,Txt_{\mathrm{rec}}$. Note that, according to [9], for successful conservative learning, it is decisive, whether or not the learner is allowed to hypothesize languages not belonging to the target class. That means, if we denote by $PConsv\,Txt$ the family of all indexable classes $\mathcal{C}$, for which there is an indexing $(L_i)_{i \in \mathbb{N}}$ *exactly* describing $\mathcal{C}$ and an IIM $Consv\,Txt$-learning $\mathcal{C}$ in $(L_i)_{i \in \mathbb{N}}$, then we obtain $PConsv\,Txt \subset Consv\,Txt_{\mathrm{rec}}$.

Similarly, the decisive difference between extra query learners and the original query learners is the ability to pose queries representing languages not belonging to the target class.

This raises the question how the original types of query learning can be compared to class-preserving conservative inference. As it turns out, $PConsv\,Txt$-learners have much in common with $rSupMemQ$-learners. In contrast, significant discrepancies between $PConsv\,Txt$ and the query learning types $rSupQ$, $rDisQ$, and $rDisMemQ$ can be observed.

Theorem 8. *(a) $PConsv\,Txt \subset rSupMemQ$.*
(b) $\mathcal{T} \# PConsv\,Txt$ for all $\mathcal{T} \in \{rSupQ, rDisQ, rDisMemQ\}$.

Proof. (a) The proof of $PConsv\,Txt \subseteq rSupMemQ$ results from a slight modification of the proof of $Consv\,Txt_{\mathrm{rec}} \subseteq rSupQ$ in [10]:

Fix $\mathcal{C} \in PConsv\,Txt$. Then there is an indexing $(L_i)_{i \in \mathbb{N}}$ of $\mathcal{C}$ and an IIM M, such that M is a $PConsv\,Txt$-learner for $\mathcal{C}$ in $(L_i)_{i \in \mathbb{N}}$. Note that, as in the general case of conservative inference, if $L \in \mathcal{C}$ and t is a text for L, then M never returns an index i with $L \subset L_i$ on any initial segment of t.

An *rSupMemQ*-learner M' identifying any $L \in \mathcal{C}$ may use membership queries to construct a text for L and then simulate M on this text until M returns an index of a superset of L. This index is then returned by M':

First, to effectively enumerate a text t for L, M' determines the set T of all words in Σ^*, for which a membership query is answered with 'yes'. Any recursive enumeration of T yields a text for L.

Second, to compute its hypothesis, M' executes steps 0, 1, 2, ... until it receives a stop signal. In general, step n, $n \in \mathbb{N}$, reads as follows:

– Determine $i := M(t_n)$, where t is a recursive enumeration of the set T. Pose a query referring to L_i. If the answer is 'yes', hypothesize the language L_i and stop. (* As M never hypothesizes a proper superset of L, L_i equals L. *) If the answer is 'no', then go to step $n+1$.

It is not hard to verify that M' is a successful *rSupMemQ*-learner for $\mathcal{C}$. Further details are omitted. So $P\,Consv\,Txt \subseteq rSupMemQ$.

To prove $rSupMemQ \setminus P\,Consv\,Txt \neq \emptyset$, we provide a separating class $\mathcal{C}_{\sup}$: for all $k \in \mathbb{N}$, let $\mathcal{C}_{\sup}$ contain the language $L_k = \{a^k b^z \mid z \geq 0\}$ and, if $k \in K$, additionally the languages $L_{k,j}^1 = \{a^k b^z \mid z \leq \Phi_k(k)$ or $[z > \Phi_k(k)+j$ and z is odd$]\}$ and $L_{k,j}^2 = \{a^k b^z \mid z \leq \Phi_k(k)$ or $[z > \Phi_k(k)+j$ and z is even$]\}$.

Using the indexing $(L'_{\langle k,j \rangle})_{k,j \in \mathbb{N}}$, given by $L'_{\langle k,0 \rangle} = L_k$, $L'_{\langle k,2j+y \rangle} = L^y_{\langle k,j \rangle}$ if $k \in K$, and $L'_{\langle k,2j+y \rangle} = L_k$ if $k \notin K$, one easily verifies $\mathcal{C}_{\sup} \in rSupQ$ and thus $\mathcal{C}_{\sup} \in rSupMemQ$. In contrast to that, $\mathcal{C}_{\sup} \notin P\,Consv\,Txt$, because otherwise K would be recursive. Details are omitted.

(b) $rDisQ \setminus P\,Consv\,Txt \neq \emptyset$ and $rDisMemQ \setminus P\,Consv\,Txt \neq \emptyset$ follow from Theorems 8(a), 3(b), and 4(b). For the other claims we just provide the separating classes: $P\,Consv\,Txt \setminus rDisMemQ \neq \emptyset$ is witnessed by the class $\mathcal{C}_a$ from the proof of Theorem 4(a). The class $\mathcal{C}_{\sup}$ (see above) belongs to $rSupQ \setminus P\,Consv\,Txt$, whereas the class consisting of the language $L = \{a\}^* \cup \{b\}$ and all the languages $L_k = \{a, \ldots, a^k\}$, $k \geq 0$, belongs to $P\,Consv\,Txt \setminus rSupQ$. □

5 Discussion

New relations have been established between learning via queries and Gold-style language learning—depending on the hypothesis space. In particular, learning with superset queries in uniformly r.e. numberings has revealed a natural inference type in-between $Lim\,Txt_{\text{r.e.}}$ and $Bc\,Txt_{\text{r.e.}}$. In correspondence to other characterisations this inference type has an analogue in Gold-style learning.

As we have seen, the learning capabilities of query learners depend on the choice of the query and hypothesis space. A similar phenomenon may be observed also in the context of Gold-style language learning, where for instance $Bc\,Txt_{\text{r.e.}} \supset Bc\,Txt_{\text{rec}} = Lim\,Txt_{\text{rec}}$. In contrast to that, for some models of Gold-style learning, the choice of the hypothesis space is ineffectual: Recall that $Lim\,Txt_{\text{r.e.}} = Lim\,Txt_{\text{rec}}$. Now assume $(A_i)_{i \in \mathbb{N}}$ is any family (not necessarily uniformly r.e.) and $\mathcal{C}$ is any indexable class of recursive languages. It is not hard to prove that

- If $\mathcal{C}$ is $Lim\,Txt$-learnable wrt $(A_i)_{i\in\mathbb{N}}$, then $\mathcal{C} \in Lim\,Txt_{\mathrm{rec}}$.
- If $\mathcal{C}$ is $Bc\,Txt$-learnable wrt $(A_i)_{i\in\mathbb{N}}$, then $\mathcal{C} \in Bc\,Txt_{\mathrm{r.e.}}$.

But to what extent is the choice of the hypothesis space relevant in conservative learning in the limit? Whereas each class $\mathcal{C} \in Lim\,Txt_{\mathrm{rec}}$ can even be identified with respect to a uniformly recursive indexing exactly enumerating $\mathcal{C}$ (a folklore result), class-preserving conservative learners may be poor compared to unrestricted $Consv\,Txt_{\mathrm{rec}}$-learners, i.e., $PConsv\,Txt \subset Consv\,Txt_{\mathrm{rec}}$, see [9]. So it remains to analyse the relevance of uniformly r.e. hypothesis spaces in the context of conservative learning. It turns out that uniformly r.e. numberings are not sufficient for conservative IIMs to achieve the capabilities of $Lim\,Txt$-learners.

Theorem 9. $Consv\,Txt_{\mathrm{r.e.}} \subset Lim\,Txt_{\mathrm{rec}}$.

Proof. $Consv\,Txt_{\mathrm{r.e.}} \subseteq Lim\,Txt_{\mathrm{rec}}$ follows from $Lim\,Txt_{\mathrm{r.e.}} \subseteq Lim\,Txt_{\mathrm{rec}}$.

$Lim\,Txt_{\mathrm{rec}} \setminus Consv\,Txt_{\mathrm{r.e.}} \neq \emptyset$ is witnessed by an indexable class $\mathcal{C}$ from [9]: For each k, $\mathcal{C}$ contains the language $L_k = \{a^k b^z \mid z \geq 0\}$ and, if $k \in K$, additionally the languages $L_{k,j} = \{a^k b^z \mid z \leq j\}$ for all $j \leq \Phi_k(k)$. [9] shows that $\mathcal{C} \in Lim\,Txt_{\mathrm{rec}} \setminus Consv\,Txt_{\mathrm{rec}}$. Adopting the corresponding proof one can verify $\mathcal{C} \notin Consv\,Txt_{\mathrm{r.e.}}$ and thus $Consv\,Txt_{\mathrm{r.e.}} \subset Lim\,Txt_{\mathrm{rec}}$. $\square$

Whether or not each class in $Consv\,Txt_{\mathrm{r.e.}}$ can also be identified conservatively in some uniformly *recursive* numbering, remains unanswered. Interestingly, if a class $\mathcal{C} \in Consv\,Txt_{\mathrm{r.e.}}$ can be identified by a learner which is conservative *and consistent* for $\mathcal{C}$, then $\mathcal{C} \in Consv\,Txt_{\mathrm{rec}}$—see Theorem 10. If $(A_i)_{i\in\mathbb{N}}$ is any family of languages, M an IIM, and $\mathcal{C}$ some indexable class, then we say that M learns $\mathcal{C}$ consistently in $(A_i)_{i\in\mathbb{N}}$, if $content(t_n) \subseteq A_{M(t_n)}$ for all text segments t_n of languages in $\mathcal{C}$. For convenience, we denote by $Cons\text{-}ConsvTxt_{\mathrm{r.e.}}$ the family of all indexable classes $\mathcal{C}$, for which there is a uniformly r.e. family $(A_i)_{i\in\mathbb{N}}$ and an IIM M, such that M learns $\mathcal{C}$ both consistently and conservatively in $(A_i)_{i\in\mathbb{N}}$.

The proof of Theorem 10 will make use of the fact that $Consv\,Txt_{\mathrm{rec}} = Cons\text{-}ConsvTxt_{\mathrm{rec}}$ according to [9], where $Cons\text{-}ConsvTxt_{\mathrm{rec}}$ is defined as usual.

Theorem 10. $Consv\,Txt_{\mathrm{rec}} = Cons\text{-}ConsvTxt_{\mathrm{r.e.}}$.

Proof. By $Consv\,Txt_{\mathrm{rec}} \subseteq Cons\text{-}ConsvTxt_{\mathrm{rec}} \subseteq Cons\text{-}ConsvTxt_{\mathrm{r.e.}}$ it remains to prove $Cons\text{-}ConsvTxt_{\mathrm{r.e.}} \subseteq Consv\,Txt_{\mathrm{rec}}$. For that purpose suppose $\mathcal{C}$ is an indexable class in $Cons\text{-}ConsvTxt_{\mathrm{r.e.}}$. If $\mathcal{C}$ is finite, then $\mathcal{C}$ trivially belongs to $Consv\,Txt_{\mathrm{rec}}$. So suppose $\mathcal{C}$ is an infinite class.

By definition, there is an IIM M which learns $\mathcal{C}$ consistently and conservatively in the limit in $(W_i)_{i\in\mathbb{N}}$. Moreover, let $(L_i)_{i\in\mathbb{N}}$ be an indexing for $\mathcal{C}$.

Given $k \in \mathbb{N}$, define the *canonical text* t^k for L_k as follows: $t^k(0) = w_m^*$, where $m = \min\{i \mid w_i^* \in L_k\}$. For $n > 0$ let $t^k(n) = w_{m+n}^*$, if $w_{m+n}^* \in L_k$; $t^k(n) = t^k(n-1)$, if $w_{m+n}^* \notin L_k$. Now let $(k_i, n_i)_{i\in\mathbb{N}}$ be an effective enumeration of all pairs (k, n) of indices such that $n = 0$ or $M(t_n^k) \neq M(t_{n-1}^k)$.

The aim is to define an indexing $(L_i')_{i\in\mathbb{N}}$ comprising $\mathcal{C}$ and a recursively generable family $(T_i')_{i\in\mathbb{N}}$ of telltales for $(L_i')_{i\in\mathbb{N}}$. Using Theorem 1 this implies $\mathcal{C} \in Consv\,Txt_{\mathrm{rec}}$. For that purpose we will define several auxiliary indexings.

Define an indexing $(A_i)_{i \in \mathbb{N}}$ as follows: for $j \in \mathbb{N}$ let

$$
A'_j = \begin{cases} L_{k_j} & \text{if } M(t^{k_j}_{n_j+x}) = M(t^{k_j}_{n_j}) \text{ for all } x \geq 0\,, \\ content(t^{k_j}_{n_j+x-1}) & \text{if } x \text{ is minimal with } M(t^{k_j}_{n_j+x}) \neq M(t^{k_j}_{n_j})\,. \end{cases}
$$

Claim 1. For each $L \in \mathcal{C}$ there is some $i \in \mathbb{N}$ with $A'_i = L_{k_i} = L$.

The proof of Claim 1 is omitted.

Fix an indexing $(A_i)_{i \in \mathbb{N}}$ enumerating all the languages A'_j without repetitions. This is possible, since, by Claim 1, $(A'_i)_{i \in \mathbb{N}}$ comprises the infinite class $\mathcal{C}$. Accordingly, let $(y_i, z_i)_{i \in \mathbb{N}}$ be an effective enumeration with $t^{y_i}_{z_i} = t^{k_j}_{n_j}$ if $A_i = A'_j$.

Claim 2. For each $L \in \mathcal{C}$ there is some $i \in \mathbb{N}$ with $A_i = L_{y_i} = L$.

Claim 3. Let $i \in \mathbb{N}$. Then $A_i \subseteq W_h$ for $h = M(t^{y_i}_{z_i})$.

Claim 4. Let $i \in \mathbb{N}$. $A_i = L_{y_i}$ iff $W_h = L_{y_i}$ for $h = M(t^{y_i}_{z_i})$.

The proof of Claims 2–4 is left to the reader. Now we define indexings $(B_i)_{i \in \mathbb{N}}$, $(T_i)_{i \in \mathbb{N}}$: for $i \in \mathbb{N}$, construct B_i and T_i according to the following procedure:

(* The construction will yield $B_i = A_i$ or $B_i = \emptyset$. Finally, it will be uniformly decidable whether or not $B_i = \emptyset$. *)

– If $t^{y_j}_{z_j} \neq t^{y_i}_{z_i}$ for all $j < i$, then let $B_i = A_i$ and $T_i = content(t^{y_i}_{z_i})$.

– If some $j < i$ fulfils $t^{y_j}_{z_j} = t^{y_i}_{z_i}$, then act according to the following instructions:

  - Let $h = M(t^{y_i}_{z_i})$ and $\{j_1, \ldots, j_s\} = \{j < i \mid t^{y_j}_{z_j} = t^{y_i}_{z_i}\}$.
  - Execute Searches (a) and (b) until one of them terminates:

    (a) Search for some x with $M(t^{y_i}_{z_i+x}) \neq h$.
    (b) Search for $x_{j_1}, \ldots, x_{j_s}$ with $M(t^{y_j}_{z_j+x_j}) \neq h$ for all $j \in \{j_1, \ldots, j_s\}$.

    (* Note that, for all $j \in \{j_1, \ldots, j_s\}$, there must be some x with $M(t^{y_i}_{z_i+x}) \neq h$ or $M(t^{y_j}_{z_j+x}) \neq h \; (= M(t^{y_j}_{z_j}))$. Otherwise one would obtain $A_i = L_{y_i}$ and $A_j = L_{y_j}$. Since $h = M(t^{y_i}_{z_i}) = M(t^{y_j}_{z_j})$, Claim 4 would imply $W_h = L_{y_i} = L_{y_j}$ and thus $A_i = A_j$. This is impossible since $(A_i)_{i \in \mathbb{N}}$ avoids repetitions. *)
  - If Search (a) terminates first, let $B_i = T_i = \emptyset$. If Search (b) terminates first, then execute Searches (b.1) and (b.2) until one of them terminates:

    (b.1) Search for some x with $M(t^{y_i}_{z_i+x}) \neq h$.
    (b.2) Search for $w_{j_1}, \ldots, w_{j_s}$ with $w_j \in A_i \setminus B_j$ for all $j \in \{j_1, \ldots, j_s\}$.

    (* Suppose $j \in \{j_1, \ldots, j_s\}$. Note that there is some x with $M(t^{y_i}_{z_i+x}) \neq h$ or some $w \in A_i \setminus B_j$. Otherwise we would have $A_i \subseteq B_j$ and $M(t^{y_i}_{z_i+x}) = h$ for all $x \in \mathbb{N}$. This yields $A_i = L_{y_i}$ and, with Claim 4, $A_i = W_h$. Therefore $W_h \subseteq B_j$. Claim 3 then implies $B_j \subseteq W_h$ and hence $A_i = W_h = B_j = A_j$, which is a contradiction to the injectivity of the indexing $(A_i)_{i \in \mathbb{N}}$. *)
  - If Search (b.1) terminates first, let $B_i = T_i = \emptyset$. If Search (b.2) terminates first, then let $B_i = A_i$ and $T_i = content(t^{y_i}_{z_i}) \cup \{w_{j_1}, \ldots, w_{j_s}\}$.

This procedure yields a uniformly recursive indexing $(B_i)_{i \in \mathbb{N}}$ and a recursively generable family $(T_i)_{i \in \mathbb{N}}$ of (possibly empty) finite sets with $T_i \subseteq B_i$ for all i. The proof is left to the reader. Note that $T_i = \emptyset$ iff $B_i = \emptyset$. Moreover, $B_i = A_i$ iff $B_i \neq \emptyset$. In particular, it is uniformly decidable whether or not $B_i = \emptyset$.

Claim 5. $(B_i)_{i\in\mathbb{N}}$ comprises $\mathcal{C}$.

Proof of Claim 5. Let $L \in \mathcal{C}$. By Claim 2 there is some $i \in \mathbb{N}$ with $A_i = L_{y_i} = L$. So, by definition, $M(t_{z_i+x}^{y_i}) = M(t_{z_i}^{y_i}) = h$ for all $x \in \mathbb{N}$. Thus either $t_{z_j}^{y_j} \neq t_{z_i}^{y_i}$ for all $j < i$ or Searches (b) and (b.2) terminate first in the construction of B_i. This yields $B_i = A_i = L$. Hence $(B_i)_{i\in\mathbb{N}}$ comprises $\mathcal{C}$. *qed Claim 5.*

Finally, define an indexing $(L_i')_{i\in\mathbb{N}}$ by removing the empty language from $(B_i)_{i\in\mathbb{N}}$. For i and j with $L_i' = B_j$ let $T_i' = T_j$. Thus $(T_i')_{i\in\mathbb{N}}$ is a recursively generable family of non-empty finite sets with $T_i' \subseteq L_i'$ for all i. It remains to show that $\mathcal{C}$, $(L_i')_{i\in\mathbb{N}}$, and $(T_i')_{i\in\mathbb{N}}$ fulfil the conditions of Theorem 1.2, i.e.,

(i) $(L_i')_{i\in\mathbb{N}}$ comprises $\mathcal{C}$,
(ii) $(T_i')_{i\in\mathbb{N}}$ is a telltale family for $(L_i')_{i\in\mathbb{N}}$.

ad (i). This is an immediate consequence of Claim 5 and the definition of $(L_i')_{i\in\mathbb{N}}$.

ad (ii). To show that $T_i' \subseteq L_j' \subseteq L_i'$ implies $L_j' = L_i'$, suppose $T_{i'}' \subseteq L_{j'}' \subset L_{i'}'$ holds for some $i', j' \in \mathbb{N}$. Let $i, j \in \mathbb{N}$ with $L_{i'}' = B_i$ and $L_{j'}' = B_j$. This yields $L_{i'}' = A_i$ and $L_{j'}' = A_j$. By the properties of canonical texts, $t_{z_i}^{y_i}$ is an initial segment of t^{y_j}. Therefore one of the segments $t_{z_j}^{y_j}$, $t_{z_i}^{y_i}$ is an initial segment of the other. Let $h_i = M(t_{z_i}^{y_i})$, $h_j = M(t_{z_j}^{y_j})$ and consider three cases.

Case 1. $z_j < z_i$. Then $t_{z_j}^{y_j}$ is a proper initial segment of $t_{z_i}^{y_i}$. In particular, $t_{z_j}^{y_i} = t_{z_j}^{y_j}$ and $M(t_{z_j}^{y_i}) = h_j$. Note that $content(t_{z_i}^{y_i}) \subseteq T_{i'}' \subseteq L_{j'}' = A_j$. Moreover, by Claim 3, $A_j \subseteq W_{h_j}$ and thus $content(t_{z_i}^{y_i}) \subseteq W_{h_j}$. Since M is conservative on any text for $L_{y_i} \supseteq A_i$, this yields $h_j = M(t_{z_j}^{y_i}) = M(t_{z_j+1}^{y_i}) = \cdots = M(t_{z_i}^{y_i})$. By definition of the family $(k_i, n_i)_{i\in\mathbb{N}}$, we have $M(t_{z_i}^{y_i}) \neq M(t_{z_i-1}^{y_i})$. This results in $z_i = z_j$ and thus in a contradiction.

Case 2. $z_i < z_j$. Then $t_{z_i}^{y_i}$ is a proper initial segment of $t_{z_j}^{y_j}$. In particular, $t_{z_i}^{y_j} = t_{z_i}^{y_i}$ and $M(t_{z_i}^{y_j}) = h_i$. Note that $content(t_{z_j}^{y_j}) \subseteq L_{j'}' \subseteq L_{i'}' = A_i$. Moreover, by Claim 3, $A_i \subseteq W_{h_i}$ and thus $content(t_{z_j}^{y_j}) \subseteq W_{h_i}$. Similarly as above, this results in $z_i = z_j$ and thus in a contradiction.

Case 3. $z_i = z_j$. Then $t_{z_i}^{y_i} = t_{z_j}^{y_j}$. First, assume $i > j$. Since $B_i \neq \emptyset$, in the construction of B_i some $w \in A_i \setminus B_j$ is included in T_i. This yields $w \in T_{i'}' \subseteq L_{j'}'$. So $w \in L_{j'}' \setminus B_j$ in contradiction to $L_{j'}' = B_j$. Second, assume $i < j$. Since $B_j \neq \emptyset$, during the construction of B_j some $w \in A_j \setminus B_i$ is found. This yields $w \in L_{j'}' \setminus L_{i'}'$ in contradiction to $L_{j'}' \subseteq L_{i'}'$.
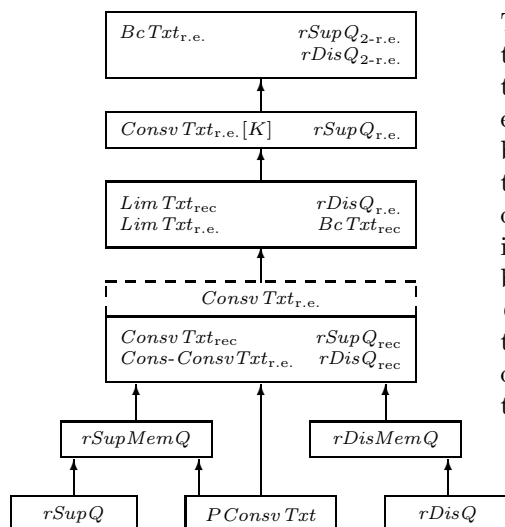
Since each case yields a contradiction, our assumption has been wrong, i.e., there are no indices $i', j' \in \mathbb{N}$ such that $T_{i'}' \subseteq L_{j'}' \subset L_{i'}'$. This finally proves (ii).

By Theorem 1.2, the families $(L_i')_{i\in\mathbb{N}}$ and $(T_i')_{i\in\mathbb{N}}$ witness $\mathcal{C} \in ConsvTxt_{\mathrm{rec}}$. Since $\mathcal{C} \in Cons\text{-}ConsvTxt_{\mathrm{r.e.}}$ was chosen arbitrarily, our argument directly proves $Cons\text{-}ConsvTxt_{\mathrm{r.e.}} \subseteq ConsvTxt_{\mathrm{rec}}$ and so $ConsvTxt_{\mathrm{rec}} = Cons\text{-}ConsvTxt_{\mathrm{r.e.}}$. □

As it turns out, throughout the whole proof we never use the fact that the hypothesis space for *Cons-ConsvTxt*-identification is uniformly r.e. This implies the following corollary: Assume $(A_i)_{i\in\mathbb{N}}$ is any family of languages (not necessarily uniformly r.e.) and $\mathcal{C}$ is any indexable class of recursive languages.

- If $\mathcal{C}$ is *Cons-ConsvTxt*-learnable wrt $(A_i)_{i\in\mathbb{N}}$, then $\mathcal{C} \in ConsvTxt_{\mathrm{rec}}$.

The following figure summarizes our main results.

$BcTxt_{\text{r.e.}}$ $rSupQ_{\text{2-r.e.}}$
$rDisQ_{\text{2-r.e.}}$

$ConsvTxt_{\text{r.e.}}[K]$ $rSupQ_{\text{r.e.}}$

$LimTxt_{\text{rec}}$ $rDisQ_{\text{r.e.}}$
$LimTxt_{\text{r.e.}}$ $BcTxt_{\text{rec}}$

$ConsvTxt_{\text{r.e.}}$

$ConsvTxt_{\text{rec}}$ $rSupQ_{\text{rec}}$
$Cons\text{-}ConsvTxt_{\text{r.e.}}$ $rDisQ_{\text{rec}}$

$rSupMemQ$ $rDisMemQ$

$rSupQ$ $PConsvTxt$ $rDisQ$

This graph illustrates the relations between different inference types studied above. If two inference types are contained in one box, they are equal. Each vector indicates a proper inclusion of inference types, whereas missing links symbolize incomparability. The dashed box around $ConsvTxt_{\text{r.e.}}$ is used to indicate that it is not yet known, whether or not $ConsvTxt_{\text{r.e.}}$ belongs to the adjacent box below.

References

1. D. Angluin. Inductive inference of formal languages from positive data. *Information and Control*, 45:117–135, 1980.
2. D. Angluin. Queries and concept learning. *Machine Learning*, 2:319–342, 1988.
3. D. Angluin. Queries revisited. *Theoretical Computer Science*, 313:175–194, 2004.
4. G. Baliga, J. Case, S. Jain. The synthesis of language learners. *Information and Computation*, 152:16–43, 1999.
5. M. Blum. A machine-independent theory of the complexity of recursive functions. *Journal of the ACM*, 14:322–336, 1967.
6. J. Case, C. Lynes. Machine inductive inference and language identification. In: *Proc. ICALP 1982*, LNCS 140, 107–115, Springer, 1982.
7. E. M. Gold. Language identification in the limit. *Information and Control*, 10:447–474, 1967.
8. J. E. Hopcroft, J. D. Ullman. *Introduction to Automata Theory, Languages, and Computation*. Addison-Wesley Publishing Company, 1979.
9. S. Lange, T. Zeugmann. Language learning in dependence on the space of hypotheses. In: *Proc. COLT 1993*, 127–136, ACM Press, 1993.
10. S. Lange, S. Zilles. On the learnability of erasing pattern languages in the query model. In: *Proc. ALT 2003*, LNAI 2842, 129–143, Springer, 2003.
11. S. Lange, S. Zilles. Replacing limit learners with equally powerful one-shot query learners. In: *Proc. COLT 2004*, LNAI 3120, 155–169, Springer, 2004.
12. H. Rogers. *Theory of Recursive Functions and Effective Computability*, MIT Press, 1987.
13. F. Stephan. *Degrees of Computing and Learning*. Habilitationsschrift, Ruprecht-Karls-Universität, Heidelberg, 1999.
14. T. Zeugmann, S. Lange. A guided tour across the boundaries of learning recursive languages. In: *Algorithmic Learning for Knowledge-Based Systems*, LNAI 961, 190–258, Springer, 1995.