Replacing limit learners with equally powerful one-shot query learners

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 human learning. Below we compare *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—to *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. Although these two approaches seem rather unrelated at first glance, we provide characterizations of different models of Gold-style learning (learning in the limit, conservative inference, and behaviourally correct learning) in terms of query learning. Thus we describe the circumstances which are necessary to replace limit learners by equally powerful one-shot learners. Our results are valid in the general context of learning indexable classes of recursive languages.
In order to achieve the learning capability of Gold-style learners, the crucial parameters of the query learning model are the type of queries (membership, restricted superset, or restricted disjointness queries) and the underlying hypothesis space (uniformly recursive, uniformly r. e., or uniformly 2-r. e. families). The characterizations of Gold-style language learning are formulated in dependence of these parameters.

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, Gold's [8] 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.

In contrast to that, 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—having the target concept in mind—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*.

Our analysis concerns common features and coincidences between these two seemingly unrelated approaches, thereby focussing our attention on the identification of formal languages, ranging over indexable classes of recursive languages, as target concepts, see [1, 10, 14]. In case such coincidences exist, their revelation might allow for transferring theoretically approved insights from one model to the other. In this context, our main focus will be on characterizations of Gold-style language learning in terms of learning via queries. Characterizing different types of Gold-style language learning in such a way, we will point out interesting correspondences between the two models. In particular, our results 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 capability of limit learners is equal to that of one-shot learners using queries.

The crucial question in this context is what abilities of the teacher are required to achieve the learning capability of Gold-style learners for query learners. In particular, it is of importance which types of queries the teacher is able to answer (and thus the learner is allowed to ask). This addresses two facets: first, the kind of information prompted by the queries (we consider membership, restricted superset, and restricted disjointness queries) and second, the hypothesis space used by the learner to formulate its queries and hypotheses (we consider uniformly recursive, uniformly r. e., and uniformly 2-r. e. families). Note that both aspects affect the demands on the teacher.

Our characterizations reveal the corresponding necessary requirements that have to be made on the teacher. Thereby we formulate coincidences of the learning capabilities assigned to Gold-style learners and query learners in a quite general context, considering three variants of Gold-style language learning. Moreover, we compare our results to several insights in Gold-style learning via oracles, see [13] for a formal background. As a byproduct of our analysis, we provide a special indexable class of recursive languages which can be learned in a behaviourally correct manner[3] in case a uniformly r. e. family is chosen as a hypothesis space, but which is not learnable in the limit, no matter which hypothesis space is chosen. Although such classes have already been offered in the literature, see [1], up to now all examples—to the authors' knowledge—are defined via diagonalisation

[3] Behaviourally correct learning is a variant of learning in the limit, see for example [7, 4, 13]. A definition is given later on.

in a rather involved manner. In contrast to that, the class we provide below is very simply and explicitly defined without any diagonal construction.

2 Preliminaries and basic results

2.1 Notations

Familiarity with standard mathematical, recursion theoretic, and language theoretic notions and notations is assumed, see [12, 9]. From now on, a fixed finite alphabet Σ with $\{a, b\} \subseteq \Sigma$ is given. 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.

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 such that $A_i = \{w \in \Sigma^* \mid f(i, w) = 1\}$ 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 such that $A_i = \{w \in \Sigma^* \mid g(i, w, n) = 1 \text{ for all but finitely many } n\}$ 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 over Σ^*. $\mathcal{C}$ 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].

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 σ of a text and outputs numbers $M(\sigma)$ 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. $\mathit{LimTxt}_{\mathrm{rec}}$ 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 $\mathit{LimTxt}_{\mathrm{rec}}$ 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. 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 is modelled in the approach of learning in the limit: the ability to change one's mind during learning. Thus learning is considered as a process in which the learner may change its hypothesis arbitrarily often until reaching its final correct guess. In particular, it is in general impossible to find out whether or not the final hypothesis has been reached, i. e., whether or not a success in learning has already eventuated.

Note that in the given context, where only uniformly recursive families are considered as hypothesis spaces for indexable classes, $Lim\,Txt_{\mathrm{rec}}$ coincides with the collection of all indexable classes identifiable from text in a behaviourally correct manner, see [7]: 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 when fed t fulfil $L_i = L$. Here M may alternate different correct hypotheses arbitrarily often instead of converging to a single hypothesis. Defining the notion $Bc\,Txt_{\mathrm{rec}}$ correspondingly as usual yields $Bc\,Txt_{\mathrm{rec}} = Lim\,Txt_{\mathrm{rec}}$ (a folklore result). In particular, each IIM $Bc\,Txt$-identifying an indexable class $\mathcal{C}'$ in some uniformly recursive family $\mathcal{H}'$ can be modified to an IIM $Lim\,Txt$-identifying $\mathcal{C}'$ in $\mathcal{H}'$.

This coincidence no longer holds, if more general types of hypothesis spaces are considered. Assume $\mathcal{C}$ is an indexable class and $\mathcal{H}^+ = (U_i)_{i\in\mathbb{N}}$ is any uniformly r. e. family of languages comprising $\mathcal{C}$. Then it is also conceivable to use $\mathcal{H}^+$ as a hypothesis space. $Lim\,Txt_{\mathrm{r.e.}}$ ($Bc\,Txt_{\mathrm{r.e.}}$) denotes the collection of all indexable classes learnable as in the definition of $Lim\,Txt_{\mathrm{rec}}$ ($Bc\,Txt_{\mathrm{rec}}$), 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. 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, in the context of $Bc\,Txt$-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].

Hence, in what follows, our analysis of Gold-style language learning will focus on the inference types $Lim\,Txt_{\mathrm{rec}}$, $Consv\,Txt_{\mathrm{rec}}$, and $Bc\,Txt_{\mathrm{r.e.}}$.

The main results of our analysis will be characterizations of these inference types in the query learning model. For that purpose we will make use of well-known characterizations concerning 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 called a family of telltales 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 $L_j = L_i$.*

The concept of telltale families is the best known notion to illustrate the specific differences between indexable classes in $\mathit{LimTxt}_{\mathrm{rec}}$, $\mathit{ConsvTxt}_{\mathrm{rec}}$, and $\mathit{BcTxt}_{\mathrm{r.e.}}$. Telltale families and their algorithmic structure have turned out to be characteristic for identifiability in our three models, see [1, 10, 14, 4]:

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

1. *$\mathcal{C} \in \mathit{LimTxt}_{\mathrm{rec}}$ iff there is an indexing of $\mathcal{C}$ possessing a uniformly r. e. family of telltales.*
2. *$\mathcal{C} \in \mathit{ConsvTxt}_{\mathrm{rec}}$ iff there is a uniformly recursive family comprising $\mathcal{C}$ and possessing a recursively generable family of telltales.*
3. *$\mathcal{C} \in \mathit{BcTxt}_{\mathrm{r.e.}}$ iff there is an indexing of $\mathcal{C}$ possessing a family of telltales.*

The notion of telltales is closely related to the notion of *locking sequences*, see [6]. If $\mathcal{H} = (U_i)_{i\in\mathbb{N}}$ is a hypothesis space, M an IIM, and L a language, then any finite text segment σ of L is called a *LimTxt*-locking sequence for M and L (a *BcTxt*-locking sequence for M, L and $\mathcal{H}$), if $M(\sigma) = M(\sigma\sigma')$ ($U_{M(\sigma)} = U_{M(\sigma\sigma')}$) for all finite text segments σ' of L. If L is *LimTxt*-learned by M (*BcTxt*-learned by M) respecting $\mathcal{H}$, then there exists a *LimTxt*-locking sequence σ for M and L (a *BcTxt*-locking sequence for M, L, and $\mathcal{H}$). Moreover, $U_{M(\sigma)} = L$ must be fulfilled for each such locking sequence.

2.3 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. Below we consider, for learning a target language L:

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.[4]

[4] The term "restricted" is used to distinguish these types of query learning from learning with superset (disjointness) queries, where, together with each negative answer the learner is provided a counterexample, i. e., a word in $L \setminus L_j$ (in $L \cap L_j$).

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. In the literature, see [2, 3], more types of queries such as (restricted) subset queries and equivalence queries have been analysed, but in what follows we concentrate on the three types explained above.

Note that, in contrast to the Gold-style models introduced above, learning via queries focusses the aspect of one-shot learning, i. e., it is concerned with learning scenarios in which learning may eventuate without mind changes.

Having a closer look at the different models of query learning, one easily finds negative learnability results. For instance, the class $\mathcal{C}_{\mathrm{sup}}$ consisting of the language $L^* = \{a\}^* \cup \{b\}$ and all languages $\{a^k \mid k \leq i\}$, $i \in \mathbb{N}$, is not learnable with superset queries. Assume a query learner M learns $\mathcal{C}_{\mathrm{sup}}$ with superset queries in an indexing $(L_i)_{i\in\mathbb{N}}$ of $\mathcal{C}$ and consider a scenario for M learning L^*. Obviously, a query j is answered 'yes', iff $L_j = L^*$. After finitely many queries, M hypothesizes L^*. Now let i be maximal, such that a query j with $L_j = \{a^k \mid k \leq i\}$ has been posed. The above scenario is also feasible for the language $\{a^k \mid k \leq i+1\}$. Given this language as a target, M will return a hypothesis representing L^* and thus fail. This yields a contradiction, so $\mathcal{C}_{\mathrm{sup}} \notin \mathit{rSupQ}$.

Moreover, as can be verified easily, the class $\mathcal{C}_{\mathrm{dis}}$ consisting only of the languages $\{a\}$ and $\{a, b\}$ is not learnable with disjointness queries.

Both examples 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. Note that the class $\mathcal{C}_{\mathrm{sup}}$ would be learnable with superset queries, if it was additionally permitted to query the language $\{a\}^*$, i. e., to ask whether or not this language is a superset of the target language. Similarly, $\mathcal{C}_{\mathrm{dis}}$ would be learnable with disjointness queries, if it was additionally permitted to query the language $\{b\}$. That means there are very simple classes of languages, for which any query 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 $\mathcal{C}$. So let $\mathcal{C}$ be an indexable class. An *extra query learner* for $\mathcal{C}$ 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 $\mathit{rSupQ}_{\mathrm{rec}}$ ($\mathit{rDisQ}_{\mathrm{rec}}$) denotes the collection of all indexable classes $\mathcal{C}$ learnable with extra superset (disjointness) queries respecting a uniformly recursive family.

Our classes $\mathcal{C}_{\mathrm{sup}}$ and $\mathcal{C}_{\mathrm{dis}}$ witness $\mathit{rSupQ} \subset \mathit{rSupQ}_{\mathrm{rec}}$ and $\mathit{rDisQ} \subset \mathit{rDisQ}_{\mathrm{rec}}$. Note that both classes would already be learnable, if in addition to the superset (disjointness) queries the learner was allowed to ask a membership query for the word b. So the capabilities of *rSupQ*-learners (*rDisQ*-learners) already increase with the additional permission to ask membership queries. Yet, as Theorem 2 shows, combining superset or disjointness queries with membership queries does not yield the same capability as extra queries do. For convenience, denote the family of classes which are learnable with a combination of superset (disjointness) queries and membership queries by *rSupMemQ* (*rDisMemQ*).

Theorem 2. *1. $rSupQ \subset rSupMemQ \subset rSupQ_{\mathrm{rec}}$.*
2. $rDisQ \subset rDisMemQ \subset rDisQ_{\mathrm{rec}}$.

Proof. ad 1. $rSupQ \subseteq rSupMemQ$ is evident; the class $\mathcal{C}_{\mathrm{sup}}$ yields the inequality.

In order to prove $rSupMemQ \subseteq rSupQ_{\mathrm{rec}}$, note that, for any word w and any language L, $w \in L$ iff $\Sigma^* \setminus \{w\} \not\supseteq L$. This helps to simulate membership queries with extra superset queries. Further details are omitted.

$rSupQ_{\mathrm{rec}} \setminus rSupMemQ \neq \emptyset$ is witnessed by the class $\mathcal{C}$ of all languages L_k and $L_{k,l}$ for $k, l \in \mathbb{N}$, where $L_k = \{a^k b^z \mid z \in \mathbb{N}\}$, $L_{k,l} = L_k$, if $\varphi_k(k)$ is undefined, and $L_{k,l} = \{a^k b^z \mid z \leq \Phi_k(k) \ \vee \ z > \Phi_k(k) + l\}$, if $\varphi_k(k)$ is defined, see [10].

To verify $\mathcal{C} \in rSupQ_{\mathrm{rec}}$ choose a uniformly recursive family comprising $\mathcal{C}$ and all languages $L_k^* = \{a^k b^z \mid z \leq \Phi_k(k)\}$, $k \in \mathbb{N}$. Note that $L_k^* \in \mathcal{C}$ iff $\varphi_k(k)$ is undefined. An $rSupQ_{\mathrm{rec}}$-learner M for $\mathcal{C}$ may act on the following instructions.

- For $k = 0, 1, 2, \ldots$ ask a superset query concerning L_k, until the answer 'yes' is received for the first time, i. e., until some k with $L_k \supseteq L$ is found.
- Pose a superset query concerning the language L_k^*. (* Note that L_k^* is a superset of the target language iff L_k^* is infinite iff $\varphi_k(k)$ is undefined. *)
  If the answer is 'yes', then output a hypothesis representing L_k and stop.
  If the answer is 'no' (* in this case $\varphi_k(k)$ is defined *), then compute $\Phi_k(k)$.
  Pose a superset query concerning $L_{k,1}$. (* Note that, for any target language $L \subseteq L_k$, this query will be answered with 'yes' iff $a^k b^{\Phi_k(k)+1} \notin L$. *)
    If the answer is 'no', then output a hypothesis representing L_k and stop.
    If the answer is 'yes', then, for any $l = 2, 3, 4, \ldots$, pose a superset query concerning $L_{k,l}$. As soon as such a query is answered with 'no', for some l, output a hypothesis representing $L_{k,l-1}$ and stop.

The details verifying that M learns $\mathcal{C}$ with extra superset queries are omitted.

In contrast to that one can show that $\mathcal{C} \notin rSupMemQ$. Otherwise the halting problem with respect to φ would be decidable. Details are omitted.

Hence $rSupMemQ \subset rSupQ_{\mathrm{rec}}$.

ad 2. $rDisQ \subseteq rDisMemQ$ is obvious; the class $\mathcal{C}_{\mathrm{dis}}$ yields the inequality.

In order to prove $rDisMemQ \subseteq rDisQ_{\mathrm{rec}}$, note that, for any word w and any language L, $w \in L$ iff $\{w\}$ and L are not disjoint. This helps to simulate membership queries with extra disjointness queries. Further details are omitted.

To prove the existence of a class in $rDisQ_{\mathrm{rec}} \setminus rDisMemQ$, define an indexable class $\mathcal{C}$ consisting of $L_0 = \{b\}$ and all languages $L_{i+1} = \{a^{i+1}, b\}$, $i \in \mathbb{N}$.

To show that $\mathcal{C} \in rDisQ_{\mathrm{rec}}$ choose a uniformly recursive family comprising $\mathcal{C}$ as well as $\{a\}^*$ and all languages $\{a^{i+1}\}$, $i \in \mathbb{N}$. A learner M identifying $\mathcal{C}$ with extra disjointness queries may work according to the following instructions.

Pose a disjointness query concerning $\{a\}^*$. (* Note that the only possible target language disjoint with $\{a\}^*$ is L_0. *)
If the answer is 'yes', then return a hypothesis representing L_0 and stop.

If the answer is ‘no’, then, for $i = 0, 1, 2, \ldots$ ask a disjointness query concerning $\{a^{i+1}\}$, until the answer ‘no’ is received for the first time. (* Note that this must eventually happen. *) As soon as such a query is answered with ‘no’, for some i, output a hypothesis representing L_{i+1} and stop.

The details verifying that M learns $\mathcal{C}$ with extra disjointness queries are omitted.

In contrast one can show that $\mathcal{C} \notin \mathit{rDisMemQ}$. For that purpose, to deduce a contradiction, assume that there is a query learner identifying $\mathcal{C}$ with disjointness and membership queries respecting an indexing $(L'_i)_{i\in\mathbb{N}}$ of $\mathcal{C}$. Consider a learning scenario of M for the target language L_0. Obviously, each disjointness query is answered with ‘no’; a membership query for a word w is answered with ‘no’ iff $w \neq b$. After finitely many queries, M must return a hypothesis representing L_0. Now let i be maximal, such that a membership query concerning a word a^i has been posed. The scenario described above is also feasible for the language $\{a^{i+1}, b\}$. If this language constitutes the target, then M will return a hypothesis representing L^* and thus fail. This yields the desired contradiction.

Hence $\mathit{rDisMemQ} \subset \mathit{rDisQ}_{\mathrm{rec}}$. □

3 Characterizations of Gold-style inference types

3.1 Characterizations in the query model

One main difference between Gold-style and query learning lies in the question whether or not a current hypothesis of a learner is already correct. A Gold-style learner is allowed to change its mind arbitrarily often (thus in general this question can not be answered), whereas a query learner has to find a correct representation of the target object already in the first guess, i. e., within “one shot” (and thus the question can always be answered in the affirmative). Another difference is certainly the kind of information provided during the learning process. So, at first glance, these models seem to focus on very different aspects of human learning and do not seem to have much in common.

Thus the question arises, whether there are any similarities in these models at all and whether there are aspects of learning both models capture. This requires a comparison of both models concerning the capabilities of the corresponding learners. In particular, one central question in this context is whether Gold-style (limit) learners can be replaced by equally powerful (one-shot) query learners. The rather trivial examples of classes not learnable with superset or disjointness queries already show that quite general hypothesis spaces—such as in learning with extra queries—are an important demand, if such a replacement shall be successful. In other words, we demand a more potent teacher, able to answer more general questions than in Angluin’s original model. Astonishingly, this demand is already sufficient to coincide with the capabilities of conservative limit learners: in [11] it is shown that the collection of indexable classes learnable with extra superset queries coincides with $\mathit{ConsvTxt}_{\mathrm{rec}}$. And, moreover, this also holds for the collection of indexable classes learnable with extra disjointness queries.

Theorem 3. $\mathit{rSupQ}_{\mathrm{rec}} = \mathit{rDisQ}_{\mathrm{rec}} = \mathit{ConsvTxt}_{\mathrm{rec}}$.

Proof. $rSupQ_{\mathrm{rec}} = ConsvTxt_{\mathrm{rec}}$ holds by [11]. Thus it remains to prove that $rSupQ_{\mathrm{rec}} = rDisQ_{\mathrm{rec}}$. For that purpose let $\mathcal{C}$ be any indexable class.

First assume $\mathcal{C} \in rDisQ_{\mathrm{rec}}$. Then there is a uniformly recursive family $(L_i)_{i\in\mathbb{N}}$ and a query learner M, such that M learns $\mathcal{C}$ with extra disjointness queries with respect to $(L_i)_{i\in\mathbb{N}}$. Now define $L'_{2i} = L_i$ and $L'_{2i+1} = \overline{L_i}$ for all $i \in \mathbb{N}$.

Suppose L is a target language. A query learner M' identifying L with extra superset queries respecting $(L'_i)_{i\in\mathbb{N}}$ is defined via the following instructions:

- Simulate M when learning L.
- If M poses a disjointness query concerning L_i, then pose a superset query concerning L'_{2i+1} to your teacher. If the answer is 'yes', then transmit the answer 'yes' to M. If the answer is 'no', then transmit the answer 'no' to M. (* Note that $L_i \cap L = \emptyset$ iff $L \subseteq \overline{L_i}$ iff $L'_{2i+1} \supseteq L$. *)
- If M hypothesizes L_i, then output a representation for L'_{2i}.

It is not hard to verify that M' learns $\mathcal{C}$ with extra superset queries with respect to $(L'_i)_{i\in\mathbb{N}}$. Hence $\mathcal{C} \in rSupQ_{\mathrm{rec}}$. This implies $rDisQ_{\mathrm{rec}} \subseteq rSupQ_{\mathrm{rec}}$.

The opposite inclusion $rSupQ_{\mathrm{rec}} \subseteq rDisQ_{\mathrm{rec}}$ is verified analogously. □

As initially in Gold-style learning, we have only considered uniformly recursive families as hypothesis spaces for query learners. Similarly to the notion of $BcTxt_{\mathrm{r.e.}}$, it is conceivable to permit more general hypothesis spaces also in the query model, i. e., to demand an even more potent teacher. Thus, by $rSupQ_{\mathrm{r.e.}}$ ($rDisQ_{\mathrm{r.e.}}$) we denote the collection of all indexable classes which are learnable with superset (disjointness) queries respecting a uniformly r. e. family. Interestingly, this relaxation helps to characterize learning in the limit in terms of query learning.

Theorem 4. $rDisQ_{\mathrm{r.e.}} = LimTxt_{\mathrm{rec}}$.

Proof. First we show $rDisQ_{\mathrm{r.e.}} \subseteq LimTxt_{\mathrm{rec}}$. For that purpose, let $\mathcal{C} \in rDisQ_{\mathrm{r.e.}}$ be an indexable class. Fix a uniformly r. e. family $(U_i)_{i\in\mathbb{N}}$ and a query learner M identifying $\mathcal{C}$ with disjointness queries with respect to $(U_i)_{i\in\mathbb{N}}$.

The following IIM M' $LimTxt$-identifies $\mathcal{C}$ with respect to $(U_i)_{i\in\mathbb{N}}$. Given a text segment σ of length n, M' interacts with M simulating a learning process for n steps. In step k, $k \le n$, depending on how M' has replied to the previous queries posed by M, the learner M computes either (i) a new query i or (ii) a hypothesis i. In case (ii), M' returns the hypothesis i and stops simulating M. In case (i), M' checks whether there is a word in σ, which is found in U_i within n steps. If such a word exists, M' transmits the answer 'no' to M; else M' transmits the answer 'yes' to M. If $k < n$, M executes step $k+1$, else M' returns any auxiliary hypothesis and stops simulating M. Given segments σ of a text for some target language, if their length n is large enough, M' answers all queries of M correctly and M returns its sole hypothesis within n steps. So, the hypotheses returned by M' stabilize on this correct guess.

Hence $\mathcal{C} \in LimTxt_{\mathrm{r.e.}} (= LimTxt_{\mathrm{rec}})$ and therefore $rDisQ_{\mathrm{r.e.}} \subseteq LimTxt_{\mathrm{rec}}$.

Second we show that $LimTxt_{\mathrm{rec}} \subseteq rDisQ_{\mathrm{r.e.}}$. So let $\mathcal{C} \in LimTxt_{\mathrm{rec}}$ be an indexable class. Fix an indexing $\mathcal{H} = (L_i)_{i\in\mathbb{N}}$ of $\mathcal{C}$ and an IIM M, such that M $LimTxt$-identifies $\mathcal{C}$ with respect to $\mathcal{H}$.

Let $(U_i)_{i\in\mathbb{N}}$ be any Gödel numbering of all r. e. languages and $(w_x)_{x\in\mathbb{N}}$ an effective enumeration of Σ^*. Suppose $L \in \mathcal{C}$ is the target language. An *rDisQ*-learner M' for L with respect to $(U_i)_{i\in\mathbb{N}}$ is defined to act on the following instructions, starting in step 0. Note that Gödel numbers (representations in $(U_i)_{i\in\mathbb{N}}$) can be computed for all queries to be asked. Step n reads as follows:

- Ask disjointness queries for $\{w_0\}, \ldots, \{w_n\}$. Let $L_{[n]}$ be the set of words w_x, $x \leq n$, for which the corresponding query is answered with 'no'. (* Note that $L_{[n]} = L \cap \{w_x \mid x \leq n\}$. *)
- Let $(\sigma_x^n)_{x\in\mathbb{N}}$ be an effective enumeration of all finite text segments for $L_{[n]}$. For all $x, y \leq n$ pose a disjointness query for $\overline{L_{M(\sigma_x^y)}}$ and thus build $\mathrm{Cand}_n = \{\sigma_x^y \mid x, y \leq n \text{ and } \overline{L_{M(\sigma_x^y)}} \cap L = \emptyset\}$ from the queries answered with 'yes'. (* Note that $\mathrm{Cand}_n = \{\sigma_x^y \mid x, y \leq n \text{ and } L \subseteq L_{M(\sigma_x^y)}\}$. *)
- For all $\sigma \in \mathrm{Cand}_n$, pose a disjointness query for the language

  $$U'_\sigma = \begin{cases} \Sigma^*, & \text{if } M(\sigma\sigma') \neq M(\sigma) \text{ for some text segment } \sigma' \text{ of } L_{M(\sigma)}, \\ \emptyset, & \text{otherwise}. \end{cases}$$

  (* Note that U'_σ is uniformly r. e. in σ and $U'_\sigma \cap L = \emptyset$ iff σ is a *LimTxt*-locking sequence for M and $L_{M(\sigma)}$. *)

  If all these disjointness queries are answered with 'no', then go to step $n+1$. Otherwise, if $\sigma \in \mathrm{Cand}_n$ is minimal fulfilling $U'_\sigma \cap L = \emptyset$, then return a hypothesis representing $L_{M(\sigma)}$ and stop.

M' identifies L with disjointness queries respecting $(U_i)_{i\in\mathbb{N}}$, because *(i)* M' eventually returns a hypothesis and *(ii)* this hypothesis is correct for L. To prove *(i)*, note that M is a *LimTxt*-learner for L respecting $(L_i)_{i\in\mathbb{N}}$. So there are i, x, y such that $M(\sigma_x^y) = i$, $L_i = L$, and σ_x^y is a *LimTxt*-locking sequence for M and L. Then $U'_{\sigma_x^y} = \emptyset$ and the corresponding disjointness query is answered with 'yes'. Thus M' returns a hypothesis. To prove *(ii)*, assume M' returns a hypothesis representing $L_{M(\sigma)}$ for some text segment σ of L. Then, by definition of M', $L \subseteq L_{M(\sigma)}$ and σ is a *LimTxt*-locking sequence for M and $L_{M(\sigma)}$. In particular, σ is a *LimTxt*-locking sequence for M and L. Since M learns L in the limit from text, this implies $L = L_{M(\sigma)}$. Hence the hypothesis M' returns is correct for L.

Therefore $\mathcal{C} \in rDisQ_{\text{r.e.}}$ and $LimTxt_{\text{rec}} \subseteq rDisQ_{\text{r.e.}}$. □

Reducing the constraints concerning the hypothesis spaces even more, let $rSupQ_{\text{2-r.e.}}$ ($rDisQ_{\text{2-r.e.}}$) denote the collection of all indexable classes which are learnable using superset (disjointness) queries with respect to a uniformly 2-r. e. family.[5] This finally allows for a characterization of the classes in $BcTxt_{\text{r.e.}}$.

Theorem 5. $rSupQ_{\text{2-r.e.}} = rDisQ_{\text{2-r.e.}} = BcTxt_{\text{r.e.}}$.

Proof. First we show $rSupQ_{\text{2-r.e.}} \subseteq BcTxt_{\text{r.e.}}$ and $rDisQ_{\text{2-r.e.}} \subseteq BcTxt_{\text{r.e.}}$. For that purpose, let $\mathcal{C} \in rSupQ_{\text{2-r.e.}}$ ($\mathcal{C} \in rDisQ_{\text{2-r.e.}}$) be an indexable class, $(L_i)_{i\in\mathbb{N}}$ an indexing of $\mathcal{C}$. Fix a uniformly 2-r. e. family $(V_i)_{i\in\mathbb{N}}$ and a query learner M identifying $\mathcal{C}$ with superset (disjointness) queries respecting $(V_i)_{i\in\mathbb{N}}$.

---

[5] With analogous definitions for Gold-style learning one easily obtains $LimTxt_{\text{2-r.e.}} = LimTxt_{\text{r.e.}} = LimTxt_{\text{rec}}$ and $BcTxt_{\text{2-r.e.}} = BcTxt_{\text{r.e.}}$.

To obtain a contradiction, assume that $\mathcal{C} \notin BcTxt_{\text{r.e.}}$. By Theorem 1, $(L_i)_{i\in\mathbb{N}}$ does not possess a telltale family. In other words, there is some $i \in \mathbb{N}$, such that for any finite set $W \subseteq L_i$ there exists some $j \in \mathbb{N}$ satisfying $W \subseteq L_j \subset L_i$. $(*)$

Consider M when learning L_i. In the corresponding learning scenario S

- M poses queries representing $V_{i_1^-}, \ldots, V_{i_k^-}, V_{i_1^+}, \ldots, V_{i_m^+}$ (in some order);
- the answers are 'no' for $V_{i_1^-}, \ldots, V_{i_k^-}$ and 'yes' for $V_{i_1^+}, \ldots, V_{i_m^+}$;
- afterwards M returns a hypothesis representing L_i.

That means, for all $z \in \{1, \ldots, k\}$, we have $V_{i_z^-} \not\supseteq L_i$ ($V_{i_z^-} \cap L_i \neq \emptyset$). In particular, for all $z \in \{1, \ldots, k\}$, there is a word $w_z \in L_i \setminus V_{i_z^-}$ ($w_z \in V_{i_z^-} \cap L_i$). Let $W = \{w_1, \ldots, w_k\} (\subseteq L_i)$. By $(*)$ there is some $j \in \mathbb{N}$ satisfying $W \subseteq L_j \subset L_i$.

Now note that the above scenario S is also feasible for L_j: $w_z \in L_j$ implies $V_{i_z^-} \not\supseteq L_j$ ($V_{i_z^-} \cap L_j \neq \emptyset$) for all $z \in \{1, \ldots, k\}$. $V_{i_z^+} \supseteq L_i$ ($V_{i_z^+} \cap L_i = \emptyset$) implies $V_{i_z^+} \supseteq L_j$ ($V_{i_z^+} \cap L_j = \emptyset$) for all $z \in \{1, \ldots, m\}$. Thus all queries in S are answered truthfully for L_j. Since M hypothesizes L_i in the scenario S, and $L_i \neq L_j$, M fails to identify L_j. This is the desired contradiction.

Hence $\mathcal{C} \in BcTxt_{\text{r.e.}}$, so $rSupQ_{\text{2-r.e.}} \subseteq BcTxt_{\text{r.e.}}$, $rDisQ_{\text{2-r.e.}} \subseteq BcTxt_{\text{r.e.}}$.

Second we show that $BcTxt_{\text{r.e.}} \subseteq rSupQ_{\text{2-r.e.}}$ and $BcTxt_{\text{r.e.}} \subseteq rDisQ_{\text{2-r.e.}}$. So let $\mathcal{C} \in BcTxt_{\text{r.e.}}$ be an indexable class. Fix a uniformly r. e. family $(U_i)_{i\in\mathbb{N}}$ and an IIM M, such that M $BcTxt_{\text{r.e.}}$-identifies $\mathcal{C}$ with respect to $(U_i)_{i\in\mathbb{N}}$.

Let $(V_i)_{i\in\mathbb{N}}$ be a uniformly 2-r. e. family such that indices can be computed for all queries to be asked below. Let $(w_x)_{x\in\mathbb{N}}$ an effective enumeration of Σ^*.

Assume $L \in \mathcal{C}$ is the target language. A query learner M' identifying L with superset (disjointness) queries respecting $(V_i)_{i\in\mathbb{N}}$ is defined according to the following instructions, starting in step 0. Step n reads as follows:

- Ask superset queries for $\Sigma^* \setminus \{w_i\}$ (disjointness queries for $\{w_i\}$) for all $i \leq n$. Let $L_{[n]}$ be the set of words w_x, $x \leq n$, for which the corresponding query is answered with 'no'. (* Note that $L_{[n]} = L \cap \{w_x \mid x \leq n\}$. *)
- Let $(\sigma_x^n)_{x\in\mathbb{N}}$ be an effective enumeration of all finite text segments for $L_{[n]}$. For all $x, y \leq n$ pose a superset query for $U_{M(\sigma_x^y)}$ (a disjointness query for $\overline{U_{M(\sigma_x^y)}}$) and thus build $\text{Cand}_n = \{\sigma_x^y \mid x, y \leq n \text{ and } U_{M(\sigma_x^y)} \supseteq L\} = \{\sigma_x^y \mid x, y \leq n \text{ and } \overline{U_{M(\sigma_x^y)}} \cap L = \emptyset\}$ from the queries answered with 'yes'.
- For all $\sigma \in \text{Cand}_n$, pose a superset (disjointness) query for the language

$$V'_\sigma = \begin{cases} \Sigma^*, & \text{if } U_{M(\sigma)} \neq U_{M(\sigma\sigma')} \text{ for some text segment } \sigma' \text{ of } U_{M(\sigma)}, \\ \emptyset, & \text{otherwise}. \end{cases}$$

  (* Note that V'_σ is uniformly 2-r. e. in σ and $V'_\sigma \not\supseteq L$ iff $V'_\sigma \cap L = \emptyset$ iff σ is a $BcTxt$-locking sequence for M and $U_{M(\sigma)}$. *)

  If all these superset queries are answered with 'yes' (all these disjointness queries are answered with 'no'), then go to step $n+1$. Otherwise, if $\sigma \in \text{Cand}_n$ is minimal fulfilling $V'_\sigma \not\supseteq L$ and thus $V'_\sigma \cap L = \emptyset$, then return a hypothesis representing $U_{M(\sigma)}$ and stop.

M' learns L with superset (disjointness) queries in $(V_i)_{i\in\mathbb{N}}$, because *(i)* M' eventually returns a hypothesis and *(ii)* this hypothesis is correct for L. To prove *(i)*, note that M is a $BcTxt$-learner for L in $(U_i)_{i\in\mathbb{N}}$. So there are x, y such that

$U_{M(\sigma_x^y)} = L$ and σ_x^y is a *BcTxt*-locking sequence for M, L, and $(U_i)_{i\in\mathbb{N}}$. Then $V'_{\sigma_x^y} = \emptyset$ and the corresponding superset query is answered with 'no' (the disjointness query with 'yes'). Thus M' returns a hypothesis. To prove *(ii)*, suppose M' returns a hypothesis representing $U_{M(\sigma)}$ for a text segment σ of L. Then, by definition of M', σ is a *BcTxt*-locking sequence for M, $U_{M(\sigma)}$, and $(U_i)_{i\in\mathbb{N}}$. In particular, σ is a *BcTxt*-locking sequence for M, L, and $(U_i)_{i\in\mathbb{N}}$. As M *BcTxt*-learns L, this implies $L = U_{M(\sigma)}$ and the hypothesis of M' is correct for L.

Therefore $\mathcal{C} \in rSupQ_{\text{2-r.e.}} \cap rDisQ_{\text{2-r.e.}}$, and thus $BcTxt_{\text{r.e.}} \subseteq rSupQ_{\text{2-r.e.}}$ and $BcTxt_{\text{r.e.}} \subseteq rDisQ_{\text{2-r.e.}}$. □

3.2 Characterizations in the model of learning with oracles—a comparison

In our characterizations we have seen that the capability of query learners strongly depends on the hypothesis space and thus on the demands concerning the abilities of the teacher. Of course a teacher might have to be more potent to answer questions with respect to some uniformly r. e. family than to work in some uniformly recursive family. For instance, teachers of the first kind might have to be able to solve the halting problem with respect to some Gödel numbering. In other words, the learner might use such a teacher as an *oracle* for the halting problem. The problem we consider in the following is to specify nonrecursive sets $A \subseteq \mathbb{N}$ such that A-recursive[6] query learners using uniformly recursive families as hypothesis spaces are as powerful as recursive learners using uniformly r. e. or uniformly 2-r. e. families. For instance, we know that $rDisQ_{\text{rec}} \subset rDisQ_{\text{r.e.}} = LimTxt_{\text{rec}}$. So we would like to specify a set A, such that $LimTxt_{\text{rec}}$ equals the collection of all indexable classes which can be identified with A-recursive $rDisQ_{\text{rec}}$-learners. The latter collection will be denoted by $rDisQ_{\text{rec}}[A]$. Subsequently, similar notions are used correspondingly.

In the Gold-style model, the use of oracles has been analysed for example in [13]. Most of the claims below use K-recursive or *Tot*-recursive learners, where $K = \{i \mid \varphi_i(i) \text{ is defined}\}$ and $Tot = \{i \mid \varphi_i \text{ is a total function}\}$. Concerning coincidences in Gold-style learning, the use of oracles is illustrated by Lemma 1.

Lemma 1. *1.* [13] $ConsvTxt_{\text{rec}}[K] = LimTxt_{\text{rec}}$.
2. $ConsvTxt_{\text{rec}}[Tot] = LimTxt_{\text{rec}}[K] = BcTxt_{\text{r.e.}}$.
3. $BcTxt_{\text{r.e.}}[A] = BcTxt_{\text{r.e.}}$ *for all* $A \subseteq \mathbb{N}$.

Proof. ad 3. Let $A \subseteq \mathbb{N}$. By definition $BcTxt_{\text{r.e.}} \subseteq BcTxt_{\text{r.e.}}[A]$. Thus it remains to prove the opposite inclusion, namely $BcTxt_{\text{r.e.}}[A] \subseteq BcTxt_{\text{r.e.}}$. For that purpose let $\mathcal{C} \in BcTxt_{\text{r.e.}}[A]$ be an indexable class. Fix an A-recursive IIM M such that $\mathcal{C}$ is $BcTxt_{\text{r.e.}}$-learned by M. Moreover, let $(L_i)_{i\in\mathbb{N}}$ be an indexing of $\mathcal{C}$.

Striving for a contradiction, assume $\mathcal{C} \notin BcTxt_{\text{r.e.}}$. By Theorem 1, $(L_i)_{i\in\mathbb{N}}$ does not possess a telltale family. In other words, there is some $i \in \mathbb{N}$, such that for any finite set $W \subseteq L_i$ there exists some $j \in \mathbb{N}$ satisfying $W \subseteq L_j \subset L_i$.

[6] A-recursive means recursive with the help of an oracle for the set A.

Since M is a *BcTxt*-learner for L_i in some hypothesis space $\mathcal{H}$, there must be a *BcTxt*-locking sequence σ for M, L_i, and $\mathcal{H}$. If W denotes the set of words occurring in σ, there is some language $L_j \in \mathcal{C}$ with $W \subseteq L_j \subset L_i$. Thus σ is a *BcTxt*-locking sequence for M, L_j, and $\mathcal{H}$. In particular, M fails to $BcTxt_{\text{r.e.}}$-identify L_j. This yields the contradiction. Hence $BcTxt_{\text{r.e.}}[A] = BcTxt_{\text{r.e.}}$.

ad 2. The proofs of $ConsvTxt_{\text{rec}}[Tot] \subseteq BcTxt_{\text{r.e.}}$, $LimTxt_{\text{rec}}[K] \subseteq BcTxt_{\text{r.e.}}$ are obtained by similar means as the proof of 3. It suffices to use Theorem 1 for $ConsvTxt_{\text{rec}}$ and $LimTxt_{\text{rec}}$ instead of the accordant statement for $BcTxt_{\text{r.e.}}$. Note that $LimTxt_{\text{rec}}[K] = BcTxt_{\text{r.e.}}$ is already verified in [4].

Next we prove $BcTxt_{\text{r.e.}} \subseteq ConsvTxt_{\text{rec}}[Tot]$ and $BcTxt_{\text{r.e.}} \subseteq LimTxt_{\text{rec}}[K]$. For that purpose, let $\mathcal{C}$ be an indexable class in $BcTxt_{\text{r.e.}}$. By Theorem 1 there is an indexing $(L_i)_{i\in\mathbb{N}}$ of $\mathcal{C}$ which possesses a family of telltales. Next we show:

(i) $(L_i)_{i\in\mathbb{N}}$ possesses a *Tot*-recursively generable (uniformly K-r. e.) family of telltales.

(ii) A $ConsvTxt_{\text{rec}}$-learner ($LimTxt_{\text{rec}}$-learner) for $\mathcal{C}$ can be computed from any recursively generable (uniformly r. e.) family of telltales for $(L_i)_{i\in\mathbb{N}}$.

To prove *(i)*, Let $(w_x)_{x\in\mathbb{N}}$ be an effective enumeration of all words in Σ^*. Given $i \in \mathbb{N}$, let a function f_i enumerate a set T_i as follows.

- $f_i(0) = w_z$ for $z = \min\{x \mid w_x \in L_i\}$.
- If $f_i(0), \dots, f_i(n)$ are computed, then test whether or not there is some $j \in \mathbb{N}$ (some $j \le n$), such that $\{f_i(0), \dots, f_i(n)\} \subseteq L_j \subset L_i$. (* Note that this test is *Tot*-recursive (K-recursive). *)
- If such a number j exists, then $f_i(n+1) = w_z$ for $z = \min\{x \mid w_x \in L_i \setminus \{f_i(0), \dots, f_i(n)\}\}$. If no such number j exists, then $f_i(n+1) = f_i(n)$.

With $T_i = \{f_i(x) \mid x \in \mathbb{N}\}$, it is not hard to verify that $(T_i)_{i\in\mathbb{N}}$ is a *Tot*-recursively generable (uniformly K-r. e.) family of telltales for $(L_i)_{i\in\mathbb{N}}$. Here note that, in the case of using a *Tot*-oracle, $T_i = \{f_i(x) \mid f_i(y+1) \neq f_i(y) \text{ for all } y < x\}$.

Finally, *(ii)* holds since Theorem 1.1/1.2 has a constructive proof, see [1, 10].

Claims *(i)* and *(ii)* imply $\mathcal{C} \in ConsvTxt_{\text{rec}}[Tot]$ and $\mathcal{C} \in LimTxt_{\text{rec}}[K]$. So $BcTxt_{\text{r.e.}} \subseteq ConsvTxt_{\text{rec}}[Tot]$ and $BcTxt_{\text{r.e.}} \subseteq LimTxt_{\text{rec}}[K]$. □

Since this proof is constructive as are the proofs of our characterizations above, we can deduce results like for example $rDisQ_{\text{rec}}[K] = LimTxt_{\text{rec}}$: Given $\mathcal{C} \in LimTxt_{\text{rec}}$, a K-recursive conservative IIM for $\mathcal{C}$ can be constructed from a $LimTxt_{\text{rec}}$-learner for $\mathcal{C}$. Moreover, a $rDisQ_{\text{rec}}$-learner for $\mathcal{C}$ can be constructed from a conservative IIM for $\mathcal{C}$. Thus, a K-recursive $rDisQ_{\text{rec}}$-learner for $\mathcal{C}$ can be constructed from a $LimTxt_{\text{rec}}$-learner. Similar results are obtained by combining Lemma 1 with our characterizations above. This proves the following theorem.

Theorem 6. *1.* $rSupQ_{\text{rec}}[K] = rDisQ_{\text{rec}}[K] = LimTxt_{\text{rec}}$.
2. $rSupQ_{\text{rec}}[Tot] = rDisQ_{\text{rec}}[Tot] = BcTxt_{\text{r.e.}}$.
3. $rSupQ_{\text{2-r.e.}}[A] = rDisQ_{\text{2-r.e.}}[A] = BcTxt_{\text{r.e.}}$ *for all* $A \subseteq \mathbb{N}$.

4 Discussion

Our characterizations have revealed a correspondence between Gold-style learning and learning via queries—between limiting and one-shot learning processes.

Crucial in this context is that the learner may ask the "appropriate" queries. Thus the choice of hypothesis spaces and, correspondingly, the ability of the teacher is decisive. If the teacher is potent enough to answer disjointness queries in some uniformly r. e. family of languages, then, by Theorem 4, learning with disjointness queries coincides with learning in the limit. Interestingly, given uniformly recursive or uniformly 2-r. e. families as hypothesis spaces, disjointness and superset queries coincide respecting the learning capabilities. As it turns out, this coincidence is not valid, if the hypothesis space may be any uniformly r. e. family. That means, $rDisQ_{\text{r.e.}}$ (and $LimTxt_{\text{rec}}$) is not equal to the collection of all indexable classes learnable with superset queries in uniformly r. e. families.

Theorem 7. $LimTxt_{\text{rec}} \subset rSupQ_{\text{r.e.}}$.

Proof. To verify $LimTxt_{\text{rec}} \subseteq rSupQ_{\text{r.e.}}$, the proof of $LimTxt_{\text{rec}} \subseteq rDisQ_{\text{r.e.}}$ can be adapted. It remains to quote a class in $rSupQ_{\text{r.e.}} \setminus LimTxt_{\text{rec}}$.

Let, for all $k, n \in \mathbb{N}$, $\mathcal{C}_{\text{lim}}$ contain the languages $L_k = \{a^k b^z \mid z \geq 0\}$ and

$$L_{k,n} = \begin{cases} \{a^k b^z \mid z \leq m\}\,, & \text{if } m \leq n \text{ is minimal such} \\ & \quad \text{that } \varphi_k(m) \text{ is undefined}\,, \\ \{a^k b^z \mid z \leq n\} \cup \ \{b^{n+1} a^{y+1}\}\,, & \text{if } \varphi_k(x) \text{ is defined for all } x \leq n \\ & \quad \text{and } y = \max\{\Phi_k(x) \mid x \leq n\}\,. \end{cases}$$

$\mathcal{C}_{\text{lim}}$ is an indexable class; the proof is omitted due to the space constraints.

To show $\mathcal{C}_{\text{lim}} \in rSupQ_{\text{r.e.}}$, let $(U_i)_{i \in \mathbb{N}}$ be a Gödel numbering of all r. e. languages. Assume $L \in \mathcal{C}$ is the target language. A learner M identifying L with superset queries respecting $(U_i)_{i \in \mathbb{N}}$ is defined to act on the following instructions:

- For $k = 0, 1, 2, \ldots$ ask a superset query concerning $L_k \cup \{b^r a^s \mid r, s \in \mathbb{N}\}$, until the answer 'yes' is received for the first time.
- Pose a superset query concerning the language L_k.
  If the answer is 'no', then, for $r, s = 0, 1, 2, \ldots$ ask a superset query concerning $L_k \cup \{b^{r+1} a^{s+1}\}$, until the answer 'yes' is received for the first time. Output a hypothesis representing $L_{k,r}$ and stop.
  If the answer is 'yes', then pose a superset query for the language

$$U'_k = \begin{cases} \{a^k b^z \mid z \leq n\}\,, & \text{if } n \text{ is minimal, such that } \varphi_k(n) \text{ is undefined}\,, \\ \{a^k b^z \mid z \geq 0\}\,, & \text{if } \varphi_k \text{ is a total function}\,. \end{cases}$$

  (* Note that U'_k is uniformly r. e. in k. U'_k is a superset of L iff $U'_k = L$. *)
  If the answer is 'yes', then return a hypothesis representing U'_k and stop.
  If the answer is 'no', then return a hypothesis representing L_k and stop.

The details proving that M $rSupQ$-identifies $\mathcal{C}_{\text{lim}}$ respecting $(U_i)_{i \in \mathbb{N}}$ are omitted.

Finally, $\mathcal{C}_{\text{lim}} \notin LimTxt_{\text{rec}}$ holds, since otherwise *Tot* would be K-recursive. To verify this, assume M is an IIM learning $\mathcal{C}_{\text{lim}}$ in the limit from text. Let $k \geq 0$. To decide whether or not φ_k is a total function, proceed as follows:

Let σ be a *LimTxt*-locking sequence for M and L_k. (* Note that σ exists by assumption and thus can be found by a K-recursive procedure. *) If there is some $x \leq \max\{z \mid a^k b^z \text{ occurs in } \sigma\}$, such that $\varphi_k(x)$ is undefined (* also a K-recursive test *), then return '0'. Otherwise return '1'.

It remains to show that φ_k is total, if this procedure returns '1'. So let the procedure return '1'. Assume φ_k is not total and n is minimal, such that $\varphi_k(n)$ is undefined. By definition, the language $L = \{a^k b^z \mid z \leq n\}$ belongs to $\mathcal{C}_{\mathrm{lim}}$. Then the sequence σ found in the procedure is also a text segment for L and by choice—since $L \subset L_k$—a LimTxt-locking sequence for M and L. As $M(\sigma)$ is correct for L_k, M fails to identify L. This is a contradiction; hence φ_k is total.

Thus the set *Tot* is K-recursive—a contradiction. So $\mathcal{C}_{\mathrm{lim}} \notin \mathit{LimTxt}_{\mathrm{rec}}$. □

Since $\mathit{rSupQ}_{\mathrm{r.e.}} \subseteq \mathit{rSupQ}_{\mathrm{2\text{-}r.e.}}$, one easily obtains $\mathit{rSupQ}_{\mathrm{r.e.}} \subseteq \mathit{BcTxt}_{\mathrm{r.e.}}$. Whether or not these two collections are equal, remains an open question. Still it is possible to prove that any indexable class containing just infinite languages is in $\mathit{rSupQ}_{\mathrm{r.e.}}$ iff it is in $\mathit{BcTxt}_{\mathrm{r.e.}}$. We omit the proof. In contrast to that there are classes of only infinite languages in $\mathit{BcTxt}_{\mathrm{r.e.}} \setminus \mathit{LimTxt}_{\mathrm{rec}}$.

Moreover, note that the indexable class $\mathcal{C}_{\mathrm{lim}}$ defined in the proof of Theorem 7 belongs to $\mathit{BcTxt}_{\mathrm{r.e.}} \setminus \mathit{LimTxt}_{\mathrm{rec}}$. Up to now, the literature has not offered many such classes. The first example can be found in [1], but its definition is quite involved and uses a diagonalisation. In contrast to that, $\mathcal{C}_{\mathrm{lim}}$ is defined compactly and explicitly without a diagonal construction and is—to the authors' knowledge—the first such class known in $\mathit{BcTxt}_{\mathrm{r.e.}} \setminus \mathit{LimTxt}_{\mathrm{rec}}$.

References

1. D. Angluin. Inductive inference of formal languages from positive data. *Inform. Control*, 45:117–135, 1980.
2. D. Angluin. Queries and concept learning. *Machine Learning*, 2:319–342, 1988.
3. D. Angluin. Queries revisited. *Theoret. Comput. Sci.*, 313:175–194, 2004.
4. G. Baliga, J. Case, S. Jain. The synthesis of language learners. *Inform. Comput.*, 152:16–43, 1999.
5. M. Blum. A machine-independent theory of the complexity of recursive functions. *J. ACM*, 14:322–336, 1967.
6. L. Blum, M. Blum. Toward a mathematical theory of inductive inference. *Inform. Control*, 28:125–155, 1975.
7. J. Case, C. Lynes. Machine inductive inference and language identification. In: *Proc. ICALP 1982*, LNCS 140, 107–115, Springer, 1982.
8. E. M. Gold. Language identification in the limit. *Inform. Control*, 10:447–474, 1967.
9. J. E. Hopcroft, J. D. Ullman. *Introduction to Automata Theory, Languages, and Computation.* Addison-Wesley Publishing Company, 1979.
10. S. Lange, T. Zeugmann. Language learning in dependence on the space of hypotheses. In: *Proc. COLT 1993*, 127–136, ACM Press, 1993.
11. 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.
12. H. Rogers, Jr. *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.