Mitotic Classes

Sanjay Jain[⋆1] and Frank Stephan[∗2]

[1] Department of Computer Science,
National University of Singapore, Singapore 117543, Republic of Singapore.
sanjay@comp.nus.edu.sg
[2] Department of Computer Science and Department of Mathematics,
National University of Singapore, Singapore 117543, Republic of Singapore.
fstephan@comp.nus.edu.sg

Abstract. For the natural notion of splitting classes into two disjoint subclasses via a recursive classifier working on texts, the question is addressed how these splittings can look in the case of learnable classes. Here the strength of the classes is compared using the strong and weak reducibility from intrinsic complexity. It is shown that for explanatorily learnable classes, the complete classes are also mitotic with respect to weak and strong reducibility, respectively. But there is a weak complete class which cannot be split into two classes which are of the same complexity with respect to strong reducibility. Furthermore, it is shown that for complete classes for behaviourally correct learning, one half of each splitting is complete for this learning notion as well. Furthermore, it is shown that explanatorily learnable and recursively enumerable classes always have a splitting into two incomparable classes; this gives an inductive inference counterpart of Sacks Splitting Theorem from Recursion Theory.

1 Introduction

A well-known observation is that infinite sets can be split into two parts of the same cardinality as the original set, while finite sets cannot be split in such a way; for example, the integers can be split into the sets of the even and odd numbers while splitting a set of 5 elements would result in subsets of unequal sizes. In this sense, infinite sets are more perfect than finite ones. The corresponding question in Complexity and Recursion Theory is which sets are so perfect that they can be split into two sets of the same complexity $[1, 9, 10, 14]$.

Ambos-Spies [1] defined one of the variants of mitocity using many-one reducibilities. Here a set A is many-one reducible to a set B iff there is a recursive function f such that $A(x) = B(f(x))$. That is, one translates every input x for A into an input $f(x)$ for B and then takes the solution provided by B (in the set or out of the set) and copies this to obtain the solution for A. Similarly one considers also complexity-theoretic counterparts of many-one reductions; for example one can translate an instance (G_1, G_2) of the Graph-Isomorphism problem into an instance ϕ of the Satisfiability problem in polynomial time, where G_1 is isomorphic to G_2 iff ϕ is satisfiable. Indeed, NP-complete problems are characterized as those into which every NP problem can be translated. Here, one can choose the reduction such that one does not only test membership but can also translate a solution of ψ into an isomorphism between G_1 and G_2 whenever such a solution exists for ψ. This general method of reducing problems and translating solutions (although here the translation of the solution is just the identity) occurs quite frequently in

⋆ Supported in part by NUS grant number R252-000-212-112.

other fields of mathematics. In inductive inference, intrinsic complexity is based on the notion of reducing one learning problem $\mathcal{L}$ to another problem $\mathcal{H}$: so first an operator translates a text T for a set L in $\mathcal{L}$ into a text $\Theta(T)$ for a set H in $\mathcal{H}$ and then another operator translates a solution E which is a sequence converging to an index e of H into a solution for L given as a sequence converging to an index e' of L. Before explaining this in more detail, some terminology is necessary to make it precise.

- A language is a recursively enumerable subset of the natural numbers.
- A class $\mathcal{L}$ is a set of recursively enumerable sets.
- A text T for a set $L \in \mathcal{L}$ is a mapping from the set $\mathbb{N}$ of natural numbers to $\mathbb{N} \cup \{\#\}$ such that $L = \{T(n) \mid n \in \mathbb{N} \wedge T(n) \in \mathbb{N}\}$. The latter set on the right hand side of the equation is called the content of T, in short, content(T). $T[n]$ denotes the first n elements of sequence T, that is $T[n] = T(0), T(1) \ldots, T(n-1)$.
- A learner is a general recursive operator (see [19]) which translates T into another sequence E. The learner converges on T iff there is a single e such that $E(n) = e$ for almost all n — in this case we say that the learner converges on T to e. The learner identifies (see [11]) T if it converges to some e such that the e-th recursively enumerable set W_e coincides with the content of T: $W_e = \text{content}(T)$. A learner identifies L if it identifies every text for L and it identifies $\mathcal{L}$ iff it identifies every $L \in \mathcal{L}$.
- A classifier is a general recursive operator which translates texts to sequences over $\{0, 1\}$. A classifier C converges on a text T to a iff $C(T[n]) = a$ for almost all n.
- For learning criteria considered in this paper, one can assume without loss of generality that learner computes $E(n-1)$ recursively from input $T[n]$. Thus, for learner M we use $M(T[n])$ to denote $E(n-1)$. Similar convention holds for classifiers.

Freivalds, Kinber and Smith [6] consider reduction between learnability problems for function classes. Jain and Sharma [13] carried this idea over to the field of learning sets from positive data and formalized the following two reducibilities for learnability problems. The main difference between these two notions is that Θ can be one-to-many in the case of the weak reducibility as different texts for the same language can go to texts for different languages while for the strong reducibility this is not allowed, at least for texts of sets in the given class.

- A class $\mathcal{L}$ is weakly reducible to $\mathcal{H}$ iff there are general recursive operators Θ and Ψ such that
  - Whenever T is a text for a language in $\mathcal{L}$ then $\Theta(T)$ is a text for a language in $\mathcal{H}$;
  - Whenever E is a sequence converging to a single index e with $W_e = \text{content}(\Theta(T))$ for some text T of a language in $\mathcal{L}$ then $\Psi(E)$ is a sequence converging to a single e' with $W_{e'} = \text{content}(T)$.

  One writes $\mathcal{L} \leq_{weak} \mathcal{H}$ in this case.
- A class $\mathcal{L}$ is strongly reducible to $\mathcal{H}$ iff there are general recursive operators Θ, Ψ as above with the additional constraint that whenever T, T' are texts for the same language in $\mathcal{L}$ then $\Theta(T), \Theta(T')$ are texts for the same language in $\mathcal{H}$. One writes $\mathcal{L} \leq_{strong} \mathcal{H}$ in this case.

Jain, Kinber, Sharma and Wiehagen investigated these concepts in several papers [12, 13]. They found that there are complete classes for $\leq_{weak}$ and $\leq_{strong}$. Here a class $\mathcal{H}$ is complete if $\mathcal{H}$ can be learned in the limit from text and for every learnable class $\mathcal{L}$ it holds that $\mathcal{L} \leq_{weak} \mathcal{H}$ and $\mathcal{L} \leq_{strong} \mathcal{H}$, respectively. If $\sqsubset$ is a recursive dense linear ordering on $\mathbb{N}$ (which makes $\mathbb{N}$ to an order-isomorphic copy of the rationals) then

$$\mathcal{Q} = \{\, \{y \in \mathbb{N} \mid y \sqsubseteq x\} \mid x \in \mathbb{N}\}$$

is a class which is complete for both, $\leq_{weak}$ and $\leq_{strong}$. The following classes are complete for $\leq_{weak}$ but not for $\leq_{strong}$:

$$\begin{aligned}\mathcal{I} &= \{\, \{0, 1, \ldots, x\} \mid x \in \mathbb{N}\};\\ \mathcal{CS} &= \{\, \mathbb{N} - \{x\} \mid x \in \mathbb{N}\}\end{aligned}$$

If one looks at $\mathcal{CS}$, one can easily see that it is the disjoint union of two classes of equivalent intrinsic complexity, namely the class $\{\, \mathbb{N} - \{x\} \mid x \text{ is even}\}$ and $\{\, \mathbb{N} - \{x\} \mid x \text{ is odd}\}$. All three classes can be translated into each other and a classifier can witness the splitting: if T is a text for a member of $\mathcal{CS}$ the classifier converges in the limit to the remainder of x divided by 2 for the unique $x \notin \text{content}(T)$. This type of splitting can be formalized to the notion of a mitotic class.

Definition 1. Two infinite classes $\mathcal{L}_0$ and $\mathcal{L}_1$ are called a *splitting* of a class $\mathcal{L}$ iff $\mathcal{L}_0 \cup \mathcal{L}_1 = \mathcal{L}$, $\mathcal{L}_0 \cap \mathcal{L}_1 = \emptyset$ and there exists a classifier C such that for all $a \in \{0, 1\}$ and for all texts T with $\text{content}(T) \in \mathcal{L}_a$, C converges on T to a.

A class $\mathcal{L}$ is *strong mitotic* (*weak mitotic*) iff there is a splitting $\mathcal{L}_0, \mathcal{L}_1$ of $\mathcal{L}$ such that $\mathcal{L} \equiv_{strong} \mathcal{L}_0 \equiv_{strong} \mathcal{L}_1$ ($\mathcal{L} \equiv_{weak} \mathcal{L}_0 \equiv_{weak} \mathcal{L}_1$).

The study of such notions is motivated from Recursion Theory [15, 19] where a recursively enumerable set is called mitotic iff it is the disjoint union of two other recursively enumerable sets which have the same Turing degree. The importance of this notion is reflected by Ladner's result that an r.e. set is mitotic iff it is autoreducible, that is, iff there is an oracle Turing machine M such that $A(x) = M^{A \cup \{x\}}(x)$ for all x [14]. Furthermore the notion had been carried over to complexity theory where it is still an important research topic [1, 9, 10].

Although intrinsic complexity is not the exact counterpart of Turing degrees in Recursion Theory, it is the only type of complexity which is defined via reducibilities and not via measures such as counting mind changes or the size of long term memory in inductive inference. Therefore, from the viewpoint of inductive inference, the above defined version of mitotic classes is reasonable. Indeed, there are some obvious parallels: in Recursion Theory, any r.e. cylinder is mitotic where a cylinder A is a set of the form $\{(x, y) \mid x \in B, y \in \mathbb{N}\}$ for some set $B \subseteq \mathbb{N}$. A corresponding cylindrificated version of a class $\mathcal{L}$ would be the class

$$\{\, \{(x, y) \mid y \in L\} \mid x \in \mathbb{N}, L \in \mathcal{L}\}$$

and it can easily be seen that this class is strong mitotic and thus also weak mitotic. Indeed, two constraints are placed there in order to be as near to the original definition of mitoticity as possible:

- If A is split into two r.e. sets A_0, A_1 with $A_0 \equiv_T A_1$ then $A \equiv_T A_0 \equiv_T A_1$. Thus all three classes involved are required to have the same intrinsic complexity degree.
- There is a partial-recursive function with domain A mapping the elements of A_a to a for all $a \in \{0, 1\}$. This is taken over by requiring the existence of a classifier which works correctly on all texts of the class. It is not required to converge on every text as then many naturally strong mitotic classes like $\mathcal{CS}$ would no longer be mitotic. This has a parallel in recursion theory: if one splits a maximal set into two r.e. sets A_0 and A_1 which are both not recursive then the sets A_0 and A_1 are recursively inseparable.

Besides the reducibilities $\leq_{weak}$ and $\leq_{strong}$ considered here, other reducibilities have also been considered [12, 13]. This paper deals only with $\leq_{weak}$ and $\leq_{strong}$, as these two are the most natural and representative.

One emphasis of the current work is on the search for natural classes which split or do not split. Therefore it is always required that the class under consideration is learnable in the limit. Furthermore, one tries to show properties for complete classes, recursively enumerable classes and indexed families. Angluin [2] defined that $\{L_0, L_1, L_2, \ldots\}$ is an indexed family iff the function $e, x \mapsto L_e(x)$ is recursive. For indexed families $\{L_0, L_1, L_2, \ldots\}$ one can without loss of generality assume that $L_n \neq L_m$ whenever $n \neq m$. A learner for this family is called exact iff it converges on every text for L_n to n. The following remark is important for several proofs.

Remark 2. Let $T[n]$ be the first n elements of a text T. One says that a learner M or a classifier C converges on T to a value a iff $M(T[n]) = a$ or $C(T[n]) = a$ for almost all n, respectively. But it does not matter – in the framework of inductive inference – how fast this convergence is and it can be slowed down by starting with an arbitrary guess and later repeating hypotheses. Similarly, if one translates one text of a language L into a text of a language H, it is not important how fast the symbols of H show up in the translated text, it is only important that they show up eventually. Therefore the translator can put into the translated text pause symbols until more data is available or certain simulated computations have terminated.

Thus, learners, operators translating texts and classifiers can be made primitive recursive by the just mentioned delaying techniques. Thus one can have recursive enumerations $\Theta_0, \Theta_1, \Theta_2, \ldots$ of translators from texts to texts, $M_0, M_1, M_2, \ldots$ of learners and $C_0, C_1, C_2, \ldots$ of classifiers such that for every given translator, learner or classifier this list contains an equivalent one. These lists can be used in proofs where diagonalizations are needed.

Given a text T and a number n, one denotes by $\Theta(T[n])$ the initial part $\Theta(T)[m]$ for the largest $m \leq n$ such that $\Theta(T)[m]$ is produced without accessing any datum in T beyond the n-th position. Note that for every m there is an n such that $\Theta(T[n])$ extends $\Theta(T)[m]$ and that $\Theta(T[n])$ can be computed from $T[n]$.

2 Complete Classes

The two main results are that those classes which are complete for $\leq_{strong}$ are strong mitotic and those which are complete for $\leq_{weak}$ are weak mitotic. This stands in contrast to the situation in Recursion Theory where some Turing-complete r.e. sets are not mitotic [14]. Note that certain classes which are complete only for $\leq_{weak}$ fail to be strong mitotic; thus the main results cannot be improved.

Theorem 3. *Every class which is complete for $\leq_{strong}$ is also strong mitotic.*

Proof. Let $\mathcal{L}$ and $\mathcal{H}$ be any classes which are complete for $\leq_{strong}$. Then the class $\mathcal{K}$ consisting of the sets $I = \{1, 3, 5, 7, \ldots\}$, $J = \{0\} \cup I$, $\{2x+3 : x \in H\}$ and $J \cup \{2x+2 : x \in H\}$ for every $H \in \mathcal{H}$ is also complete. Since $\mathcal{L}$ is complete for $\leq_{strong}$, there is a translation Θ which maps languages in $\mathcal{K}$ to languages in $\mathcal{L}$ such that proper inclusion is preserved. Thus there is some element $e \in \Theta(J) - \Theta(I)$. As $\mathcal{H}$ is complete for $\leq_{strong}$, the subclasses $\{\Theta(\{2x + 3 : x \in H\}) : H \in \mathcal{H}\}$ and $\{\Theta(I \cup \{2x + 2 : x \in H\}) : H \in \mathcal{H}\}$ of $\mathcal{L}$ are also complete for $\leq_{strong}$. All members of the first class do not contain e while all members of the second class contain e as an element. It follows that the subclasses $\mathcal{L}_0 = \{L \in \mathcal{L} : e \notin L\}$ and $\mathcal{L}_1 = \{L \in \mathcal{L} : e \in L\}$ are complete and disjoint and can be classified by a C which conjectures 1 if e has shown up in the text so far and 0 otherwise. Therefore, $\mathcal{L}$ is strong mitotic. □

The following notion is used to formulate Proposition 6 which is a central ingredient of Theorem 7. Furthermore, learners with certain properties are needed.

Definition 4. For any sequence T of symbols, let all(T) be the length of the shortest prefix of T containing all symbols which show up in T, that is, let

$$\text{all}(T) = \sup\{n + 1 : \text{content}(T[n]) \subset \text{content}(T)\}.$$

Note that all$(T) < \infty$ iff content(T) is a finite set.

The following remark combines some ideas of Blum and Blum [4] and Fulk [8].

Remark 5. Let $\mathcal{L}$ be a learnable class. Then there is a learner M for $\mathcal{L}$ with the following properties:

- M is prudent, that is, whenever M outputs an index e on some input data then M learns W_e;
- if M learns L then there is an index e such that M converges on every text T for L to that index e;
- for every text T and index e, if $M(T[n]) = e$ for infinitely many n then actually $M(T[n]) = e$ for almost all n.

Proposition 6. *If $\mathcal{I} \leq_{weak} \mathcal{L}$ then the reduction to $\mathcal{L}$ and a learner M for $\mathcal{L}$ can be chosen such that for all texts T for a language in $\mathcal{I}$, M converges on the translation of T to an index $e \geq$ all(T).*

Proof. One assumes that a learner M satisfies the three conditions from Remark 5.

The key idea of the proof is the following: Given a reduction (Θ, Ψ) from $\mathcal{I}$ to $\{L_0, L_1, L_2, \ldots\}$, one constructs a further reduction (Θ', Ψ') from $\mathcal{I}$ to $\mathcal{I}$ such that, for every text T of a set in $\mathcal{I}$, M converges on $\Theta(\Theta'(T))$ to an index $e \geq$ all(T). By Remark 2, assume without loss of generality that Θ is primitive recursive. The idea is that in the first translation Θ' every set $I_n = \{0, 1, \ldots, n\}$ is translated into some set $I_{2^n(1+2m)}$ and Ψ' translates every index of every set $I_{2^n(1+2m)}$ into an index of I_n.

Given a sequence E of indices, $\Psi'(E)(s)$ is computed as follows. Let k be the least number such that $W_{E(s),s} \subseteq I_k$. Choose m, n such that $2^n(1 + 2m) = k$ and output the canonical index for I_n. It is easy to see that this translation works whenever E converges to an index of some set in $\mathcal{I}$.

The construction of Θ' is more involved. For the construction, the special properties of M from Remark 5 are important. The most adequate way to describe Θ' is as a method which is continuously extending the translations $\tau_0, \tau_1, \ldots$ of $\Theta'(T)$ which have not yet been built by taking the first of the below cases which applies. $\tau_0 = 0\#$ and in step s, τ_s is extended to τ_{s+1} according to the case which applies:

Case 1: $M(\Theta(\tau_s)) \leq$ all$(T[s])$. Then let τ_{s+1} be the first extension of τ_s found such that $M(\Theta(\tau_{s+1})) \neq M(\Theta(\tau_s))$;

Case 2: Case 1 does not hold but content$(\tau) \neq I_{2^n(1+2m)}$ for all m where n is the least number with content$(T[s]) \subseteq I_n$. Then let $\tau_{s+1} = \tau_s a$ for the least nonelement a of content(τ_s);

Case 3: Case 1 and Case 2 do not hold. Then $\tau_{s+1} = \tau_s \#$.

Here $\Theta(\tau_s)$ and $\Theta(\tau_{s+1})$ are defined as in Remark 2 and can be computed from τ_s and τ_{s+1}, respectively.

For the verification, assume now that a set $I_n = \{0, 1, 2, \ldots, n\} \in \mathcal{I}$ and a text T for I_n are given.

First it is necessary to note that the extension τ_{s+1} in the first condition of the construction of the τ_{s+1} from τ_s can always be found for any given parameter n. Indeed there are two texts T_1, T_2 extending σ_s for different sets in $\mathcal{I}$. It follows that $\Theta(T_1)$ and $\Theta(T_2)$ are texts of different sets and thus M converges on them to different indices. Thus one can take a sufficiently long prefix of one of them in order to get the desired τ_{s+1}.

Second, it can be shown by induction that $|\tau_s| > s$ in all stages s; this guarantees that Θ' is indeed a general recursive operator.

Third, one shows that M does not converge on $\Theta(\Theta'(T))$ to any index less than or equal to all(T). By Case 1 in the construction, M cannot converge on $\Theta(\Theta'(T))$ to an index $e \leq$ all(T). Thus, by Remark 5, there is a stage s_0 such that Case 1 of the construction is never taken after stage s_0.

Fourth, one shows that $\Theta'(T)$ is a text for some language in $\mathcal{I}$. There is a least m such that content$(\tau_{s_0}) \subseteq I_{2^n(1+2m)}$. For all stages $s > s_0$, if content$(\tau_s) \subset I_{2^n(1+2m)}$, then τ_{s+1} is chosen by Case 2, else τ_{s+1} is chosen by Case 3. One can easily see that the resulting text $\Theta'(T) = \lim_{s\to\infty} \tau_s$ is a text for $I_{2^n(1+2m)}$. Indeed, $\Theta'(T) = \tau_{s_1}\#^\infty$ for $s_1 = s_0 + 2^n(1 + 2m) + 2$.

So it follows that Θ' maps every text of a set I_n to some text of some set $I_{2^n(1+2m)}$ as desired. So, for all texts T of sets in $\mathcal{I}$, M converges on $\Theta(\Theta'(T))$ to some index $e >$ all(T). This completes the proof. □

Theorem 7. *Let $\mathcal{L}$ be a learnable class which is complete for $\leq_{weak}$. Then $\mathcal{L}$ is weak mitotic.*

Proof. Let $I_n = \{0, 1, \ldots, n\}$. By Theorem 6 there is a reduction (Θ, Ψ) from $\mathcal{I}$ to $\mathcal{L}$ and a learner M such that for every text T of a member of $\mathcal{I}$, M converges on $\Theta(T)$ to an index $e >$ all(T). For this reason, using oracle K for the halting problem, one can check for every index e whether there is a text T for a language in $\mathcal{I}$ such that M on $\Theta(T)$ converges to e. One can assume without loss of generality, that, besides #, no data-item in a text is repeated. Also, among the texts for sets in $\mathcal{I}$, only the texts of the sets $\{0\}$, $\{0, 1\}$, $\{0, 1, 2\}$, …, $\{0, 1, 2, \ldots, e\}$ can satisfy all$(T) \leq e$. Thus, one has just to check the behaviour of the given learner M for the class $\mathcal{L}$ on the texts in the class

$$\mathcal{T}_e = \{\Theta(T') \mid T' \in \{0, 1, \ldots, e, \#\}^e \cdot \#^\infty \wedge \text{content}(T) \in \mathcal{I}\}.$$

Now, define a classifier C such that on a text T, the n-th guess of C is 1 iff there is an odd number $m \leq M(T[n])$ and a text $T'' \in \mathcal{T}_{M(T[n])}$ for I_m such that $M(\Theta(T'')[n]) = M(T[n])$.

For the verification that C is a classifier, assume that M converges on T to some index e. Then C converges on T to 1 iff there is a set $I_m \in \mathcal{I}$ with m odd and a text T' for I_m in $\mathcal{T}_e$ with content$(\Theta(T')) = W_e$. Otherwise C converges on T to 0. If M does not converge on T then T is not a text for a set in $\mathcal{L}$ and the behaviour of C on T is irrelevant. Thus C is a classifier which splits $\mathcal{L}$ into two classes $\mathcal{L}_0$ and $\mathcal{L}_1$. These classes $\mathcal{L}_0$ and $\mathcal{L}_1$ contain the images of repetition-free texts of sets in the classes $\{I_0, I_2, I_4, \ldots\}$ and $\{I_1, I_3, I_5, \ldots\}$, respectively. Thus both classes are complete for $\leq_{weak}$ and the splitting of $\mathcal{L}$ into $\mathcal{L}_0$ and $\mathcal{L}_1$ witnesses that $\mathcal{L}$ is weak mitotic. $\square$

As several proofs use known properties of maximal sets, the following remark summarizes some of these properties.

Remark 8. A set A is *maximal* iff (a) it is recursively enumerable, (b) has an infinite complement and (c) every recursively enumerable set B satisfies that either $B - A$ is finite or the complement of $A \cup B$ is finite.

For any partial-recursive function ψ and maximal set A, the following statements hold.

- Either $\psi(x)$ is defined for almost all $x \in \overline{A}$ or for only finitely many $x \in \overline{A}$.
- The set $\{x \notin A \mid \psi(x) \in A\}$ is either finite or contains almost all elements of $\overline{A}$.
- If for every x there is some $y > x$ such that $y \notin A$, $\psi(y)$ is defined, $\psi(y) > x$ and $\psi(y) \notin A$, then $\psi(z)$ is defined and $\psi(z) = z$ for almost all $z \in \overline{A}$.

Furthermore, A is dense simple. That is, for every recursive function f and for almost all $x \notin A$, $\{y : x < y \leq f(x)\} \subseteq A$.

These basic facts about maximal sets will be used in several proofs. Odifreddi [15, Pages 288–294]. provides more information on maximal sets including the proof of the existence by Friedberg [7].

Theorem 9. *There exists an indexed family $\{L_0, L_1, L_2, \ldots\}$ which is complete for $\leq_{weak}$ and weak mitotic but not strong mitotic.*

Proof. Let A be a maximal set with complement $\{a_0, a_1, \ldots\}$ where $a_n < a_{n+1}$ for all n. Now let $\mathcal{L}$ consist of the sets

- $\{x, x+1, x+2, \ldots, x+y\}$ for all $x \in A$ and $y \in \mathbb{N}$;
- $\{x, x+1, x+2, \ldots\}$ for all $x \notin A$.

As A is r.e., it is easy to see that the resulting family is indeed an indexed family. Learnability is also clear as the learner, on input σ, first determines $x = \min(\text{content}(\sigma))$ and then conjectures content(σ) if $x \in A_{|\sigma|}$ and conjectures $\{x, x+1, x+2, \ldots\}$ otherwise. Assuming that $a_0 > 0$, this class is a superclass of $\mathcal{I}$ and thus complete for $\leq_{weak}$. By Theorem 7, $\{L_0, L_1, L_2, \ldots\}$ is weak mitotic.

Let $\mathcal{L}_0$ and $\mathcal{L}_1$ be two disjoint classes with union $\mathcal{L}$. Without loss of generality, $\{a_0, a_0+1, a_0+2, \ldots\} \in \mathcal{L}_1$. Assume now by way of contradiction that $\mathcal{L} \equiv_{strong} \mathcal{L}_0$ as witnessed by a recursive operator Θ which translates the corresponding texts. As Θ has to preserve the proper subset relation on the content of the texts while translating, every text of a set of the form $\{a_n, a_n+1, a_n+2, \ldots\}$ has to be translated into a text for a set of the form $\{a_m, a_m+1, a_m+2, \ldots\}$ (to preserve the property that translation of $\{a_n, a_n+1, a_n+2, \ldots\}$ has infinitely many subsets in the class).

Now consider the function f which outputs, on input x, the first element found to be in the range of the image $\Theta(\sigma)$ for some σ with $x = \min(\text{content}(\sigma))$. The function f is recursive, but as A is maximal, the relation $f(a_n) < a_{n+1}$ holds for almost all n. It follows that if n is sufficiently large, then some text of $\{a_n, a_n+1, a_n+2, \ldots\}$ is translated to a text of one of the sets $\{a_k, a_k+1, a_k+2, \ldots\}$ with $k \leq n$. Now fix a text T for $\{a_n, a_n+1, a_n+2, \ldots\}$. One can then inductively define a sequence of strings $\sigma_n, \sigma_{n-1}, \ldots, \sigma_0$ such that each sequence $\sigma_n \sigma_{n-1} \ldots \sigma_m T$ is a text for $\{a_m, a_m + 1, a_m + 2, \ldots\}$ and

$$\text{content}(\Theta(\sigma_n \sigma_{n-1} \ldots \sigma_m \sigma_{m-1})) \not\subseteq \text{content}(\Theta(\sigma_n \sigma_{n-1} \ldots \sigma_m T))$$

for each $m \leq n$. As Θ maps texts of infinite sets in $\mathcal{L}$ to texts of infinite sets in $\mathcal{L}$, one can conclude that

$$\text{content}(\Theta(\sigma_n \sigma_{n-1} \ldots \sigma_m T)) = \{a_m, a_m + 1, a_m + 2, \ldots\}.$$

Thus, for every m, some text of the set $\{a_m, a_m + 1, a_m + 2, \ldots\}$ is mapped to a text for the same set in contradiction to the assumption that Θ does not have $\{a_0, a_0 + 1, a_0 + 2, \ldots\}$ in its range. Therefore $\mathcal{L}$ is not strong mitotic. $\square$

3 Incomplete Learnable Classes

Finite classes are not mitotic and thus every nonempty class has a subclass which is not mitotic. For infinite classes, one can get that the corresponding subclass is also infinite. The proof is a standard application of Ramsey's Theorem: Given classifiers $C_0, C_1, C_2, \ldots$ one selects a subclass $\{H_0, H_1, H_2, \ldots\}$ of $\{L_0, L_1, L_2, \ldots\}$ such that each classifier C_n classifies $H_n, H_{n+1}, H_{n+2}, \ldots$ in the same way. The class $\{H_0, H_1, H_2, \ldots\}$ may not be an indexed family but a very thin class. Alternatively, one can also take the $H_0, H_1, H_2, \ldots$ such that for a given enumeration of primitive recursive operators, the text $\Theta_n(T_m)$ of the ascending text T_m of H_m is not a text for any H_k with $k > \max(\{n, m\})$. The latter method gives the following result.

Theorem 10. *Every infinite class $\mathcal{L}$ has an infinite subclass $\mathcal{H}$ such that $\mathcal{H}$ is not weakly isomorphic to any proper subclass of $\mathcal{H}$. In particular, $\mathcal{H}$ is not weak mitotic.*

There is an easier example of a class which is not weak mitotic. It is even an indexed family of finite sets, but such example cannot be build within any given indexed family.

Example 11. *Assume that $\{L_0, L_1, L_2, \ldots\}$ is given as $L_0 = \{0, 1\}$ and $L_n = \{n\}$ for all $n \in \mathbb{N} - \{0\}$. Then $\{L_0, L_1, L_2, \ldots\}$ is not weak mitotic.*

Proof. Given any splitting $\mathcal{L}_0, \mathcal{L}_1$ of $\{L_0, L_1, L_2, \ldots\}$, one of these classes, say $\mathcal{L}_0$, contains at most one of the sets L_0, L_1. Then, for any given reduction (Θ, Ψ) from $\{L_0, L_1, L_2, \ldots\}$ to $\mathcal{L}_0$, $\Theta(\sigma)$ produces some string of nonempty content for some $\sigma \in 1\#^*$ and thus there are texts T_0, T_1 extending σ for L_0 and L_1, respectively, such that $\Theta(L_0)$ and $\Theta(L_1)$ are texts for different sets in $\mathcal{L}_0$ with a nonempty intersection. But such sets do not exist, by choice of $\mathcal{L}_0$. $\square$

Note that the class

$$\{\{0, 1, 2\}, \{1, 2\}, \{2\}, \{3\}, \{4\}, \{5\}, \ldots, \{n\}, \ldots\}$$

compared with the class from Example 11 has the slight improvement that for any splitting $\mathcal{L}_0, \mathcal{L}_1$ of the class, one half of the splitting contains an ascending chain of two or three sets while the other half contains only disjoint sets so that the two halves are not equivalent with respect to $\leq_{weak}$.

As these two examples show, it is more adequate to study the splitting of more restrictive classes like the inclusion-free classes. A special case of such classes are the finitely learnable classes. Here a class is finitely learnable [11] iff there is a learner M which keeps outputting a special symbol denoting the absence of a reasonable conjecture until it has seen sufficiently much data and then outputs an index e forever, where W_e is the language to be learnt.

Theorem 12. $\{L_0, L_1, L_2, \ldots\} \equiv_{strong} \{H_0, H_1, H_2, \ldots\}$ *whenever both classes are infinite indexed families which are finitely learnable. In particular, every such class is strong mitotic.*

Proof. As $\{L_0, L_1, L_2, \ldots\}$ and $\{H_0, H_1, H_2, \ldots\}$ are infinite, one can without loss of generality assume that the underlying enumerations are one-to-one. Furthermore, they have exact learners M and N, respectively, which use the corresponding indexing. Now one translates M to N by mapping L_n to H_n; thus Ψ is the identity mapping each index to itself, where in the domain the n stands for H_n and in the range the n stands for L_n. $\Theta(T) = \#^k T_n$ where k is the least number such that M outputs a hypothesis n on input $T[k]$ (i.e., first position where M conjectures a hypothesis) and T_n is the ascending text of H_n. This completes the proof of the first statement.

Given now an infinite finitely learnable class $\{L_0, L_1, L_2, \ldots\}$, one can split it into $\{L_0, L_2, L_4, \ldots\}$ and $\{L_1, L_3, L_5, \ldots\}$ which are the subclasses of languages with even and odd index, respectively. Both classes are also infinite indexed families which are finitely learnable. Thus they are all equivalent by the above result. Furthermore, a classifier for splitting can be obtained by simulating the learner M on the input text, and then converging to 0 if the (only) grammar output by M on the input text is even, and to 1 if the (only) grammar output by M on the input text is odd. $\square$

4 Further Splitting Theorems

Another question is whether classes can be split into incomparable classes. So one would ask whether there is a parallel result to Sacks Splitting Theorem [18]: Every nonrecursive r.e. set A is the disjoint union of two r.e. sets A_0 and A_1 such that the Turing degrees of A_0 and A_1 are incomparable and strictly below the one of A. The next example shows that there are classes where every splitting is of this form; so these classes are not weak mitotic. Furthermore, splittings exist, so the result is not making use of a pathological diagonalization against all classifiers.

Example 13. *Let A be a maximal set and let $L_a = \{a\}$ if $a \notin A$ and $L_a = A$ if $a \in A$. Then $\{L_0, L_1, L_2, \ldots\}$ is recursively enumerable and finitely learnable but any splitting $\mathcal{L}_0, \mathcal{L}_1$ of $\{L_0, L_1, L_2, \ldots\}$ satisfies $\mathcal{L}_0 \not\leq_{weak} \mathcal{L}_1$ and $\mathcal{L}_1 \not\leq_{weak} \mathcal{L}_0$.*

Proof. Let $T_0, T_1, T_2, \ldots$ be a recursive enumeration of recursive texts for $L_0, L_1, L_2, \ldots$, respectively. Let $F(a)$ be the cardinality of $\{b < a \mid b \notin A\}$. It is easy to see that one can split $\{L_0, L_1, L_2, \ldots\}$ into $\{L_a \mid a \in A \vee F(a) \text{ is even}\}$ and $\{L_a \mid a \notin A \wedge F(a) \text{ is odd}\}$. Thus this class has a splitting; in fact there are infinitely many of them. Furthermore, $\{L_0, L_1, L_2, \ldots\}$ is finitely learnable by outputting an index for L_a for the first a occurring in a given text.

Assume now by way of contradiction that there is a splitting $\mathcal{L}_0, \mathcal{L}_1$ with $\mathcal{L}_0 \leq_{weak} \mathcal{L}_1$ via a reduction (Θ, Ψ). Now one defines the partial-recursive function f which outputs on input a the first number occurring in $\Theta(T_a)$; if there occurs no number then $f(a)$ is undefined. As $\mathcal{L}_0$ is infinite, there are infinitely many $a \notin A$ with $L_a \in \mathcal{L}_0$. For all but one of these, $\Theta(T_a)$ has to be a text for some set $L_b \neq A$ in $\mathcal{L}_1$. Then $L_b = \{b\}$ and $f(a) = b \notin A$ for these a. It follows that for every x there is an $a > x$ with $a \notin A \wedge f(a) \notin A \wedge f(a) > x$. Then, by Remark 8, $f(a) = a$ for almost all $a \notin A$. As infinitely many of these a belong to an $L_a \in \mathcal{L}_0$, one has that $\Theta(T_a)$ is a text for L_a and Θ translates some text for a set in $\mathcal{L}_0$ into a text for a set in $\mathcal{L}_0$ and not into a text for a set in $\mathcal{L}_1$. Thus $\mathcal{L}_0 \not\leq_{weak} \mathcal{L}_1$. By symmetry of the argument, $\mathcal{L}_1 \not\leq_{weak} \mathcal{L}_0$. $\square$

While the previous example showed that there are classes for which every splitting is a Sacks splitting, the next result shows that every learnable recursively enumerable class has a Sacks splitting; but it might also have other splittings.

Theorem 14. *Every infinite recursively enumerable and learnable class* $\{L_0, L_1, L_2, \ldots\}$ *has a splitting into two infinite subclasses* $\mathcal{L}_0, \mathcal{L}_1$ *such that* $\mathcal{L}_0 \not\leq_{weak} \mathcal{L}_1$ *and* $\mathcal{L}_1 \not\leq_{weak} \mathcal{L}_0$.

Proof. Let M be a learner for $\{L_0, L_1, L_2, \ldots\}$ which satisfies the three conditions from Remark 5. Now one defines inductively relative to the oracle K the following function F from $\mathbb{N}$ to $\{0, 1\}$. The definition uses parameters a, b to count the number of languages which have already gone to either side and which are initialized as 0. Furthermore, let $\Theta_0, \Theta_1, \Theta_2, \ldots$ be the enumeration of operators as given in Remark 2. Also, there is for every L_n a recursive text T_n which can be generated from the fact that the class is recursively enumerable. Furthermore, let U be the set of all minimal indices of languages L: $n \in U$ iff for all $m < n$, $L_m \neq L_n$. Note that U can be computed relative to the halting problem K since $L_m = L_n$ iff M converges on T_m to the same value as on T_n. Furthermore, let the auxiliary function $F^*(n, a)$ be the number of $k \in U$ with $k < n$ and $F(k) = a$. Also this value can be computed with oracle K.

The value of $F(n)$ is defined by applying the case with highest priority (reflected by least number) which qualifies:

- Priority 0: $n \notin U$. Then there is an $m < n$ such that $L_m = L_n$ and let $F(n) = F(m)$ for the least such m.
- Priority $4e + 1$: There is an $m < n$ such that M converges on $\Theta_e(T_m)$ and T_n to the same value and $F(m) = 0$ and there are no $i, j < n$ with $F(i) = 0$, $F(j) = 0$ and M converges on $\Theta_e(T_i)$ to the same value as on T_j. Then let $F(n) = 0$.
- Priority $4e + 2$: There is an $m < n$ such that M converges on $\Theta_e(T_m)$ and L_n to the same value and $F(m) = 1$ and there are no $i, j < n$ with $F(i) = 1$, $F(j) = 1$ and M converges on $\Theta_e(T_i)$ to the same value as on T_j. Then let $F(n) = 1$.
- Priority $4e + 3$: $F^*(m, 0) < F^*(m, 1) + e$ for all $m \leq n$. Then let $F(n) = 0$.
- Priority $4e + 4$: $F^*(m, 1) < F^*(m, 0) + e$ for all $m \leq n$. Then let $F(n) = 1$.

More precisely, when defining $F(n)$ one searches for the least number k such that the entry for priority k applies and defines F as described in this case. Note that the priorities $4e + 3$ and $4e + 4$ apply for all $e > n$ and thus there is always some entry which applies. Next it is shown that $F \leq_T K$. As M converges on every of the texts T_n, the test whether priority 0 applies can be done relative to K. For the test whether priority $4e+1$ applies it needs to be checked whether M on $\Theta_e(T_m)$ converges to the same value as on T_n. This is be done by first computing the value d to which M converges on T_n. Now, as M satisfies the constraints as in Remark 5, M on $\Theta_e(T_m)$ either converges to d or outputs d only finitely often. Thus one can check in the limit whether M converges on $\Theta_e(T_m)$ to d. Similarly one can check for any two numbers $i, j < n$ whether M converges on $\Theta_e(T_i)$ to that value which M converges on T_j. So one can test whether priority $4e + 1$ applies. Similarly one can test whether priority $4e + 2$ applies. The tests for priorities $4e + 3$ and $4e + 4$ are obviously doable as $U \leq_T K$ and the priorities only refer to statistics of previous values of F at places where the argument is in U.

So F can be computed in the limit. Having an approximation F_s to F, one defines a classifier C as $C(\sigma) = F_{|\sigma|}(m)$ for the least m with $m = |\sigma| \vee M(T_m[|\sigma|]) = M(\sigma)$. Assume now a text T of a language in $\{L_0, L_1, L_2, \ldots\}$ be given and n being the least index such that $L_n = \text{content}(T)$.

Then, for all sufficiently large s, $C(T[s]) = F_s(n)$. The reason is that M converges on T and T_n to the same index of L_n but, for $m < n$, M converges on T_m to an index of the language L_m which is not equal to L_n. Thus C converges on T to $F(n)$.

It is clear that the choice of priority 0 is applied iff $n \notin U$ and this is the only priority level which applies infinitely often. This is shown by induction. So assume that for some ℓ all priorities k strictly between 0 and $4e + c$ with $c \in \{1, 2, 3, 4\}$ do not apply at any $n \geq \ell$ and assume that priority $4e + c$ would apply at ℓ. Now one makes a case-distinction depending on which of the priorities $4e + c$ is applied at ℓ.

In the case of priority $4e + 1$ there is an m such that for $n = \ell$ it holds that $F(m) = 0$, $F(n) = 0$ and M converges on $\Theta_e(T_m)$ and T_n to the same value. Now let $i = m$ and $j = \ell$. Then, for $n > \ell$, these values i, j are below n and avoid that this priority qualifies again. So $n = \ell$ is the maximal n where $F(n)$ is taken according to this priority.

In the case of priority $4e + 3$, consider the set $\{n_0, n_1, n_2, \ldots, n_k\}$ of the least $k + 1$ elements of $U \cap \{\ell, \ell + 1, \ell + 2, \ldots\}$ where $k = e + 1 + F^*(\ell, 1) - F^*(\ell, 0)$. One can now prove by induction for $u = 0, 1, 2, \ldots$ that

- for n_u with $u < k$, Priority $4e + 3$ applies and $F(n_u) = 0$;
- for n_u with $u \leq k$, $F^*(n_u, 0) = F^*(\ell, 0) + u$ and $F^*(n_u, 1) = F^*(\ell, 1)$.

Thus, $F^*(n_k, 0) = F^*(n_k, 1) + e$ and there is no $n \geq n_k$ where priority $4e + 3$ applies.

The other two cases of priority $4e + 2$ and $4e + 4$ are symmetric to the two previous cases and so one can conclude that there are also only finitely many n where priority $4e + c$ applies. This completes the inductive step.

Now assume by way of contradiction that there is a reduction (Θ_e, Ψ) witnessing that $\mathcal{L}_0 \leq_{weak} \mathcal{L}_1$. Let ℓ be so large that all priorities $1, 2, \ldots, 4e$ are not used to define any $F(n)$ with $n \geq \ell$. Due to Priority $4\ell + 3$ there is an $\ell' \in U$ with $F^*(\ell', 0) \geq F^*(\ell', 1) + \ell + 1$; note that $\ell' > \ell$. So more sets in $\mathcal{L}_0$ than in $\mathcal{L}_1$ have an index below ℓ' and therefore there is an $m \leq \ell'$ such that $L_m \in \mathcal{L}_0$ and $\Theta_e(T_m)$ is not the text of any of the sets $L_0, L_1, \ldots, L_{\ell'}$. So let n be the minimal index of content$(\Theta_e(T_m))$; this index exists as Θ_e maps texts of languages in $\mathcal{L}_0$ to texts of languages in $\mathcal{L}_1$. It follows from the construction that either $F(n) = F(m) = 0$ and $L_n \in \mathcal{L}_0$ or there are $i, j < n$ with $F(i) = F(j) = 0$, $L_i, L_j \in \mathcal{L}_0$ and $\Theta_e(T_i)$ being a text for L_j. This contradicts the assumption that (Θ_e, Ψ) reduces $\mathcal{L}_0$ to $\mathcal{L}_1$. Hence $\mathcal{L}_0 \not\leq_{weak} \mathcal{L}_1$. Similarly one can show that $\mathcal{L}_1 \not\leq_{weak} \mathcal{L}_0$. $\square$

For this reason, one cannot give a recursively enumerable class where all splittings $\mathcal{L}_0, \mathcal{L}_1$ satisfy either $\mathcal{L}_0 \leq_{strong} \mathcal{L}_1$ or $\mathcal{L}_1 \leq_{strong} \mathcal{L}_0$. Furthermore, complete classes have comparable splittings like before as they are mitotic and have even equivalent splittings. The next example gives a class where some splittings are comparable but where they are never equivalent.

Example 15. *Let A be a maximal set. For all $a \in \mathbb{N}$ and $b \in \{0, 1, 2\}$, let $L_{3a+b} = \{3a + b\}$ if $a \notin A$ and $L_{3a+b} = \{3c + b \mid c \in A\}$ if $a \in A$. Then $\{L_0, L_1, L_2, \ldots\}$ is not weak mitotic but has a splitting $\mathcal{L}_0, \mathcal{L}_1$ with $\mathcal{L}_0 \leq_{strong} \mathcal{L}_1$.*

Proof. If one takes the splitting $\mathcal{L}_0 = \{L_0, L_3, L_6, \ldots\}$ and $\mathcal{L}_1 = \{L_1, L_2, L_4, L_5, L_7, L_8, \ldots\}$ then it is easy to see that $\mathcal{L}_0 \leq_{strong} \mathcal{L}_1$ via (Θ, Ψ) such that Θ is based on translating in every text every datum $3x$ to $3x + 1$ and Ψ is based on transforming every index e into an index for $\{3x \mid 3x + 1 \in W_e\}$. The details are left to the reader.

Given now a further splitting $\mathcal{L}_2, \mathcal{L}_3$ of $\{L_0, L_1, L_2, \ldots\}$, one of these two classes, say $\mathcal{L}_2$, must contain at least two of the sets $L_{3a}, L_{3a+1}, L_{3a+2}$ for infinitely many $a \notin A$. Assume by way of contradiction that (Θ, Ψ) would witness $\mathcal{L}_2 \leq_{weak} \mathcal{L}_3$. Now one defines the following functions f_b for $b = 0, 1, 2$ by letting $f_b(a)$ to be the first number x found such that $3x$ or $3x + 1$ or $3x + 2$ occurs in the text $\Theta((3a + b)^\infty)$. Now choose two different $b, b' \in \{0, 1, 2\}$ such that there are infinitely many $a \in \mathbb{N} - A$ with $L_{3a+b}, L_{3a+b'} \in \mathcal{L}_2$. Then one knows that for every bound c there are infinitely many $a \in \mathbb{N} - A$ such that $L_{3a+b} \in \mathcal{L}_2$ and $\Theta((3a + b)^\infty)$ is a text for some language in $\mathcal{L}_3 - \{L_0, L_1, L_2, \ldots, L_c\}$. It follows that $f_b(a) = a$ for almost all $a \notin A$. The same applies to $f_{b'}$. So there is an $a \notin A$ such that $L_{3a+b}, L_{3a+b'}$ are both in $\mathcal{L}_2$ and that Θ maps texts of both languages to texts of the sets $L_{3a}, L_{3a+1}, L_{3a+2}$. As only one of these sets can be in $\mathcal{L}_3$, Θ has to map texts of different languages to texts of the same language, a contradiction. Thus $\mathcal{L}_2 \not\leq_{weak} \mathcal{L}_3$ and the class cannot be weak mitotic. $\square$

5 Beyond Explanatory Learning

One could besides classes which are complete for (explanatorily) learning also consider classes which are complete for behaviourally correct learning [3, 5, 16] with respect to $\leq_{strong}$. Note that such a class $\mathcal{L}$ is no longer explanatorily learnable. But $\mathcal{L}$ satisfies the following two properties:

- The class $\mathcal{L}$ is behaviourally correct learnable, that is, there is a learner which outputs on every text T for a language in $\mathcal{L}$ an infinite sequence $e_0, e_1, e_2, \ldots$ of hypotheses such that $W_{e_n} = \text{content}(T)$ for almost all n;
- Every behaviourally correct learnable class $\mathcal{H}$ satisfies $\mathcal{H} \leq_{strong} \mathcal{L}$.

Note that the reduction $\leq_{strong}$ considered in this paper is always the same as defined for explanatory learning; reducibilities more adapted to behaviourally correct learning had also been studied [12, 13]. Completeness with respect to $\leq_{weak}$ is not considered in this section, so "complete" means "complete for $\leq_{strong}$" in this section.

It is easy to show that such complete classes exist, an example is the class of all sets $\{x\} \cup \{x + y + 1 \mid y \in L\}$ where the x-th behaviourally correct learner learns the set L. So given any behaviourally correct learnable class and an index x of its learner, the translation $L \mapsto \{x\} \cup \{x + y + 1 \mid y \in L\}$ would translate all the sets learnt by this learner into sets in the complete class.

Note that methods similar to those in Theorem 3 show that $\mathcal{L}$ is strong mitotic. The next result shows that for any splitting $\mathcal{L}_0, \mathcal{L}_1$ of $\mathcal{L}$, one of these two classes is complete for behaviourally correct learning as well and therefore this class cannot be split into two incomparable subclasses.

Theorem 16. *If $\mathcal{L}_0, \mathcal{L}_1$ are a splitting of a class which is complete for behaviourally correct learning with respect to $\leq_{strong}$ then either $\mathcal{L}_0 \equiv_{strong} \mathcal{L}_0 \cup \mathcal{L}_1$ or $\mathcal{L}_1 \equiv_{strong} \mathcal{L}_0 \cup \mathcal{L}_1$.*

Proof. Let $\mathcal{H}$ be a class which is complete for behaviourally correct learning with respect to $\leq_{strong}$. Furthermore, let $C_0, C_1, \ldots$ be a list of all primitive recursive classifiers. One can build, for each x, a sequence $\tau_{x,0}, \tau_{x,1}, \ldots$ starting with $\tau_{x,0} = x$. If $\tau_{x,y}$ has been defined, then one takes $\tau_{x,y+1}$ to be the first extension of $\tau_{x,y}$ found, if any, such that

$$\{x, x + 1, x + 2, \ldots, x + y\} \subseteq \text{content}(\tau_{x,y+1}) \subseteq \{x, x + 1, x + 2, \ldots\}$$

and $C_x(\tau_{x,y+1}) \neq C_x(\tau_{x,y})$. In the case that this process terminates at some y, that is, if $\tau_{x,y+1}$ does not exist, let $z = \max(\text{content}(\tau_{x,y}))$ and put for each $H \in \mathcal{H}$ the set

$$\{x, x+1, x+2, \ldots, z\} \cup \{z+u+1 \mid u \in H\}$$

into the class $\mathcal{L}$. If the process produces an infinite sequence $T = \lim \tau_{x,y}$ then T is a text for $\{x, x+1, x+2, \ldots\}$ on which C_x does not converge. Thus one puts the set

$$\{x, x+1, x+2, \ldots\}$$

into $\mathcal{L}$ and obtains that C_x does not split this class into two subclasses.

Now it is shown that $\mathcal{L}$ is behaviourally correct learnable. Given a behaviourally correct learner M for $\mathcal{H}$, the new learner N for $\mathcal{L}$ tries to establish in the limit the minimum x of the content of the text (which always succeeds) and the maximal y for this x such that $\tau_{x,y}$ is defined (which fails for the case that all $\tau_{x,u}$ are defined). Using current approximations x', y', z' to these parameters where $z' = \max(\text{content}(\tau_{x',y'}))$, the learner N constructs from the current input σ a new string η where it replaces every u by $u - z - 1$ if $u > z$ and by # otherwise. Then N conjectures the following set:

$$W_{N(\sigma)} = \begin{cases} \{x', x'+1, x'+2, \ldots, z'\} & \\ \quad \cup \ \{u+z'+1 \mid u \in W_{M(\eta)}\} & \text{if } \tau_{x',y'+1} \text{ does not exist;} \\ \{x', x'+1, x'+2, \ldots\} & \text{if } \tau_{x',y'+1} \text{ exists.} \end{cases}$$

Note that the language generated can always go from the first language to the second language when $\tau_{x',y'+1}$ turns out to exist as

$$\{x', x'+1, x'+2, \ldots, z'\} \cup \{u+z'+1 \mid u \in W_{M(\eta)}\} \subseteq \{x', x'+1, x'+2, \ldots\}.$$

Thus the above case-distinction can be coded into $N(\sigma)$ as indicated. The verification that N indeed behaviourally correct learns $\mathcal{L}$ is straightforward and thus skipped.

Now consider any classifier C_x which converges on every text for a language in $\mathcal{L}$ either to 0 or 1. Then there is a maximal y such that $\tau_{x,y}$ is defined since otherwise the above defined text T would exist on which C_x does not converge. Therefore, the class $\mathcal{H}$ is strongly reducible to the subclass

$$\{\, \{x, x+1, x+2, \ldots, z\} \cup \{z+u+1 \mid u \in H\} \mid H \in \mathcal{H}\}$$

of $\mathcal{L}$. Furthermore, every set in this subclass has a text starting with $\tau_{x,y}$ and C_x converges on all such texts to $C_x(\tau_{x,y})$. Therefore this complete class is contained in one member of the splitting of $\mathcal{L}$ defined by C_x and so one of these members is complete for behavioural correct learning with respect to $\leq_{strong}$.

After dealing with this special class $\mathcal{L}$, consider any splitting $\mathcal{L}_0, \mathcal{L}_1$ of a class $\mathcal{L}_0 \cup \mathcal{L}_1$ which is complete for behaviourally correct learning with respect to $\leq_{strong}$. There is a reduction (Θ, Γ) from $\mathcal{L}$ to $\mathcal{L}_0 \cup \mathcal{L}_1$ due to completeness and the classifier C doing the splitting $\mathcal{L}_0 \cup \mathcal{L}_1$ can be translated back into a classifier splitting $\mathcal{L}$ into two parts, one of which is complete. This complete part is reduced into either $\mathcal{L}_0$ or $\mathcal{L}_1$ and thus one of these two classes is complete for behaviourally correct learning with respect to $\leq_{strong}$. $\square$

As just seen, any splitting $\mathcal{L}_0, \mathcal{L}_1$ of a class which is complete for behaviourally correct learning satisfies either $\mathcal{L}_0 \equiv_{strong} \mathcal{L}_1$ or $\mathcal{L}_0 <_{strong} \mathcal{L}_1$ or $\mathcal{L}_1 <_{strong} \mathcal{L}_0$. As the class is strong mitotic, it can happen that the two halves of a split are equivalent although this is not always the case. The next result gives a class where the two halves of a splitting are always comparable but never equal.

Theorem 17. *There is a recursively enumerable and behaviourally correctly learnable class which is not weak mitotic such that every splitting $\mathcal{L}_0, \mathcal{L}_1$ of the class satisfies either $\mathcal{L}_0 \leq_{strong} \mathcal{L}_1$ or $\mathcal{L}_1 \leq_{strong} \mathcal{L}_0$.*

Proof. A minor modification of the construction of Post's simple set [17], gives the following: There is a recursive partition of the odd natural numbers into sets $I_0, I_1, I_2, \ldots$ and a recursively enumerable set L_1 of odd natural numbers such that $1 \in L_1$, $I_n \not\subseteq L_1$ for all n and L_1 intersects every recursively enumerable set which contains infinitely many odd natural numbers. Now complete the definition of $\{L_0, L_1, L_2, \ldots\}$ for the $n \neq 1$ as follows: $L_n = \{n\}$ if n is even and $L_n = L_1 \cup \{n\}$ if n is odd.

Clearly $\{L_0, L_1, L_2, \ldots\}$ is behaviourally correct learnable. The learner conjectures content(σ) if σ does not contain an odd number and content(σ) $\cup\, L_1$ if σ contains an odd number.

Let $\mathcal{L}_0, \mathcal{L}_1$ be a splitting of $\{L_0, L_1, L_2, \ldots\}$. L_1 is in one of these classes, say in $\mathcal{L}_1$. Let C be the classifier witnessing the split. Then there is a locking sequence σ for C on L_1 such that content(σ) $\subseteq L_1$ and $C(\tau) = 1$ for all extensions τ of σ with content(τ) $\subseteq L_1$. Let T be a text of L_1. Now for every $a \in L_1$ and every n, $C(\sigma a T[n]) = 1$. Since L_1 is simple, it follows that the set

$$D = \{a \mid a \text{ is odd and } \exists n\, (M(\sigma a T[n]) = 0)\}$$

is finite and thus $L_a \in \mathcal{L}_1$ for all odd $a \notin D$. Let n be an index such that $D \cap I_m = \emptyset$ for all $m \geq n$ and let $a_0, a_1, a_2, \ldots$ be an ascending enumeration of all even numbers plus members of D. Note that $\mathcal{L}_0 \subseteq \{L_{a_0}, L_{a_1}, L_{a_2}, \ldots\}$.

Let $b_m = \min(I_{n+m} - L_1)$, note that b_m always exists since $I_{n+m} \not\subseteq L_1$ for all m. Note that $L_1 \cup \{b_m\}$ is in $\mathcal{L}_1$ for all m.

There is an operator Θ which translates a text T of any set in $\mathcal{L}_0$ into a text for the set L_1 in the case that T does not contain any a_m and into a text for $L_1 \cup \{b_m\}$ in the case that a_m is the first member of the sequence $a_0, a_1, a_2, \ldots$ appearing in T. The operator Θ exists since one just copies an enumeration of L_1 until some a_m shows up in T. After one has found m, one keeps inserting into the enumeration of L_1 the least element of I_{n+m} which has not yet appeared in the enumeration of L_1 and continues to output this modified enumeration. For Ψ, given a sequence converging to an index e of some $L_1 \cup \{b_m\}$, one can find the m of b_m in the limit from a simulated enumeration of W_e and thus translate this sequence into one converging to an index of L_{a_m}. So the reduction (Θ, Ψ) witnesses that $\mathcal{L}_0 \leq_{strong} \mathcal{L}_1$.

On the other hand, if an operator translates texts of sets in $\mathcal{L}_1$ into texts of sets in $\mathcal{L}_0$ then it has to map some text of L_1 to some text of some L_{a_m} and there is an initial segment σ of this text such that a_m appears on the output when σ is fed into the operator. There is only the set L_{a_m} in $\mathcal{L}_0$ containing a_m but infinitely many sets in $\mathcal{L}_1$ have a text starting with σ. Thus the translation maps some texts of different sets to texts of L_{a_m}. So the translation cannot be used for a weak reduction from $\mathcal{L}_1$ to $\mathcal{L}_0$. Hence $\mathcal{L}_1 \not\leq_{weak} \mathcal{L}_0$. □

6 Autoreducibility

Trakhtenbrot [20] defined that a set A is autoreducible iff one can reduce A to itself such that $A(x)$ is obtained by accessing A only at places different to x. Ladner [14] showed that a recursively enumerable set is mitotic iff it is autoreducible. Ambos-Spies pointed this result out to the authors and asked whether the same holds in the setting of inductive inference. Unfortunately, this characterisation fails for both of the major variants of autoreducibility. These variants are the ones corresponding to strong and weak reducibility.

Definition 18. A class $\mathcal{L}$ is strong (weak) autoreducible iff there is a strong (weak) reduction (Θ, Ψ) from $\mathcal{L}$ to itself such that for all sets $L \in \mathcal{L}$ and all texts T for L, $\Theta(T)$ is a text for a language in $\mathcal{L} - \{L\}$.

Example 19. *Let A be a maximal set and $\mathcal{L}$ contain the following sets:*

- $\{3x\}, \{3x+1\}, \{3x+2\}$ *for all* $x \notin A$;
- $\{3y : y \in A\}$, $\{3y+1 : y \in A\}$, $\{3y+2 : y \in A\}$.

Then the class $\mathcal{L}$ is neither strong mitotic nor weak mitotic. But $\mathcal{L}$ is autoreducible via some (Θ, Ψ) where Θ maps any text T to a text T' such that all elements of the form $3y$ in T have the form $3y+1$ in T', all elements of the form $3y+1$ in T have the form $3y+2$ in T' and all elements of the form $3y+2$ have the form $3y$ in T'.

So even the implication "strong autoreducible $\Rightarrow$ weak mitotic" fails. The remaining question is whether at least the converse direction is true in inductive inference. This is still unknown, but there is some preliminary result on sets which are complete for $\leq_{weak}$.

Theorem 20. *If a class $\mathcal{L}$ is weak complete than it is weak autoreducible.*

Proof. Let $\mathcal{L}$ be weak complete and M be a learner for $\mathcal{L}$ which satisfies the conditions from Remark 5. As $\mathcal{L}$ is weak complete, by Proposition 6, there is a reduction (Θ', Ψ') from the class $\mathcal{I}$ to $\mathcal{L}$ such that for any set $I_x = \{0, 1, \ldots, x\} \in \mathcal{I}$ and any text T for I_x, $\Theta'(T)$ is a text for a set on which M does not converge to an index in I_x. Now, an autoreduction (Θ, Ψ) is constructed.

For this, one first defines Θ'' as follows and then concatenates it with Θ'. The operator Θ'' translates every text T for a set L into a text for $I_{2^n(1+2m)}$ where m, n are chosen such that n is the value to which M converges on T and m is so large that all the elements put into $\Theta''(T)$ when following intermediate hypotheses of M on T are contained in the set $I_{2^n(1+2m)}$. It is easy to verify that this can be done. Then Θ is given as $\Theta(T) = \Theta'(\Theta''(T))$. The sequence $\Theta(T)$ is a text for a set in $\mathcal{L}$ with the additional property that M converges on it to an index larger than $2^n(1+2m)$; this index is therefore different from n and $\text{content}(\Theta(T)) \neq \text{content}(T)$.

The reverse operator Ψ can easily be generated from Ψ'. If E converges to an index for $\text{content}(\Theta(T))$ then $\Psi'(E)$ converges to some index for $I_{2^n(1+2m)}$. The number $2^n(1+2m)$ can be determined in the limit from this index by enumerating the corresponding finite set; thus Ψ can translate E via $\Psi'(E)$ to a sequence which converges to n. $\square$

Example 21. *The class $\mathcal{L}$ from Theorem 9 is weak complete and weak autoreducible but not strong autoreducible.*

Proof. Let $\mathcal{L}$ and $a_0, a_1, a_2, \ldots$ as in Theorem 9. Assume that (Θ, Ψ) is a strong autoreduction. Then Θ has to preserve inclusions and therefore map infinite sets in $\mathcal{L}$ to infinite sets. So, $\text{content}(\Theta(a_0\,(a_0+1)\,(a_0+2)\,\ldots))$ is an infinite set in $\mathcal{L}$ different from $\{a_0, a_0+1, a_0+2, \ldots\}$. By induction, one can show that

$$\text{content}(\Theta(a_n\,(a_n+1)\,(a_n+2)\,\ldots)) \subseteq \{a_{n+1}, a_{n+1}+1, a_{n+1}+2, \ldots\} \text{ and}$$
$$\text{content}(\Theta(a_n\,(a_n+1)\,(a_n+2)\,\ldots)) \subset \{a_n, a_n+1, a_n+2, \ldots\}.$$

But in Theorem 9 it was shown that no recursive operator has the second of these properties. That $\mathcal{L}$ is weak complete was shown in Theorem 9 and that $\mathcal{L}$ is weak autoreducible follows from Theorem 20. $\square$

References

1. Klaus Ambos-Spies. P-mitotic sets. *Logic and Machines*. Springer LNCS 177:1–23, 1984.
2. Dana Angluin. Inductive inference of formal languages from positive data. *Information and Control*, 45:117–135, 1980.
3. Janis Bārzdiņš. Two theorems on the limiting synthesis of functions. In *Theory of Algorithms and Programs*, Volume 1, Latvian State University, Riga, 210:82–88, 1974.
4. Lenore Blum and Manuel Blum. Towards a mathematical theory of inductive inference. *Information and Control*, 28:125–155, 1975.
5. John Case and Christopher Lynes. Inductive inference and language identification. *Ninth International Colloquium on Automata, Languages and Programming* (ICALP), Aarhus, Denmark, July 12-16, 1982, Proceedings. Springer, *Lecture Notes in Computer Science*, 140:107–115, 1982.
6. Rusins Freivalds, Efim Kinber and Carl Smith. On the Intrinsic Complexity of Learning. *Computational Learning Theory*, Second European Conference, EuroCOLT'95, Barcelona, Spain, March 1995, Proceedings. Springer, *Lecture Notes in Artificial Intelligence*, 904:154–169, 1995.
7. Richard Friedberg. Three theorems on recursive enumeration. *The Journal of Symbolic Logic*, 23:309–316, 1958.
8. Mark Fulk. Prudence and Other Conditions on Formal Language Learning. *Information and Computation*, 85:1–11, 1990.
9. Christian Glaßer, Aduri Pavan, Alan Selman and Liyu Zhang. Mitosis in Computational Complexity. *Theory and Applications of Models of Computation*: Third International Conference, TAMC 2006, Beijing, China, May 15-20, 2006. Proceedings. Springer LNCS, 3959:61–67, 2006.
10. Christian Glaßer, Mitsunori Ogihara, Aduri Pavan, Alan Selman and Liyu Zhang. Autoreducibility, Mitoticity and Immunity. *Mathematical Foundations of Computer Science*: Thirtieth International Symposium, MFCS 2005, Gdansk, Poland, August 29 - September 2, 2005, Proceedings. Springer LNCS, 3618:387–398, 2005.
11. E. Mark Gold. Language identification in the limit. *Information and Control*, 10:447–474, 1967.
12. Sanjay Jain, Efim Kinber and Rolf Wiehagen. Language Learning from Texts: Degrees of Intrinsic Complexity and Their Characterizations. *Journal of Computer and System Sciences*, 63:305–354, 2001.
13. Sanjay Jain and Arun Sharma. The Intrinsic Complexity of Language Identification. *Journal of Computer and System Sciences*, 52:393–402, 1996.
14. Richard Ladner. Mitotic recursively enumerable sets. *The Journal of Symbolic Logic*, 38:199–211, 1973.
15. Piergiorgio Odifreddi. *Classical Recursion Theory.* North-Holland, Amsterdam, 1989.
16. Daniel N. Osherson and Scott Weinstein. Criteria of Language Learning. *Information and Control*, 52:123–138, 1982.

17. Emil Post. Recursively enumerable sets of positive integers and their decision problems, *Bulletin of the American Mathematical Society*, 50:284–316, 1944.
18. Gerald E. Sacks. On the degrees less than $\mathbf{0}'$. *Annals of Mathematics*, 77:211–231, 1977.
19. Robert Soare. *Recursively Enumerable Sets and Degrees. A Study of Computable Functions and Computably Generated Sets*. Springer-Verlag, Heidelberg, 1987.
20. Boris A. Trakhtenbrot. On autoreducibility. *Soviet Mathematics, Doklady* (*Doklady Akademii Nauk SSSR*), 11:814–817, 1970.