Finite identification with positive and with complete data

Dick de Jongh and Ana Lucia Vargas Sandoval
D.H.J.deJongh@uva.nl, ana.varsa@gmail.com

Institute for Logic, Language and Computation
University of Amsterdam, The Netherlands

Abstract. We study the differences between finite identifiability of recursive languages with positive and with complete data. In finite families the difference lies exactly in the fact that for positive identification the families need to be anti-chains, while in in the infinite case it is less simple, being an anti-chain is no longer a sufficent condition. We also show that with complete data there are no maximal learnable families whereas with positive data there usually are, but there do exist positively identifiable familes without a maximal positively identifiable extension. We also investigate a conjecture of ours, namely that each positively identifiable family has either finitely many or continuously many maximal noneffectively positively identifiable extensions. We verify this conjecture for the restricted case of families of equinumerous finite languages.

Keywords: formal learning theory, finite identification, positive data, complete data, indexed family, anti-chains

1 Introduction

The groundbreaking work of Gold [3] in 1967 started a new era for developing mathematical and computational frameworks for studying the formal process of *learning*. Gold's model, *identification in the limit*, has been studied for learning recursive functions, recursively enumerable languages, and recursive languages with *positive data* and with *complete data*. The learning task consists of identifying languages as members of a family of languages, the learning function can output infinitely many conjectures but they need to stabilize in one permanent one. In Gold's model, a huge difference in power between learning with positive data and with complete data is exposed. With positive data a family of languages containing all finite languages and at least one infinite one can't be learnable. With complete data the learning task becomes almost trivial.
Based on Gold's model and results, Angluin's [1] work focuses on indexed families of recursive languages, i.e., families of languages with a uniform decision procedure for membership. Such families naturally occur as the sets of languages generated by types of grammars. In particular, Angluin [1] gave a characterization when Gold's learning task can be executed. Her work shows that many

non-trivial families of recursive languages can be learned by means of positive data only.
A few years later, Mukouchi [7] (and simultaneously Lange and Zeugmann [5]) introduced the framework of *finite identification* "in Angluin's style" for both positive and complete data. The learning task is as in Gold's model with the difference that the learning function can only guess once. Mukouchi presents an Angluin style characterization theorem for positive and complete finite identification. As expected, finite identification with complete data is more powerful than with positive data only. However, the distinction is much less marked than in Gold's framework. His work didn't draw much attention until recently, de Jongh and Gierasimczuk[2] further developed the theory of finite identification.
It is often believed that children do not use negative data when they learn their native language. In opposition to that, a large amount of theoretical and experimental work in computational linguistics (see e.g. [6]) has been conducted to analyze and test the intuition that there is a powerful contribution of "negative" data for improving and speeding up children language acquisition (see e.g. [9, 4]).

In this work, we focus on a more fine-grained theoretical analysis of the distinction between finite identification with positive and with complete data in Angluin-style. Our aim is to formally study the concrete difference: what can we do more with complete information for families of recursive languages than with only positive information.
After a section with preliminaries we start, in Section 3, with finite identification of finite families. Here the distinction between positive and complete data comes out very clearly: the difference is exactly described by the fact that with positive data families can only be identified if they are anti-chains w.r.t. $\subset$.
Then, in Section 4, we question whether any finitely identifiable family is contained in a maximal finitely identifiable one. Maximal learnable families are of special interest because a learner for a maximal learnable family is a learner for all of its subfamilies. First, in Subsection 4.1, we address this in the positive data setting. Simple examples of positively identifiable families are often maximal like the set of all sets of exactly n elements. We provide a mildly positive result w.r.t. the existence of a *non-effectively finitely identifiable* maximal extension for families concerning only finite languages and give some hints about the obstacles to a more general result. Then we present a canonical family which does not have an *effective* maximal *finitely identifiable* extension, i.e. answering the above question negatively. Then, in subsection 4.2, we come to study the complete data setting. Surprisingly, we provide a completely negative result concerning maximal learnable families for effective or non-effective finite identification with complete data: any finitely identifiable family can be extended to a larger one which is also finitely identifiable, ergo maximal identifiable families do not exist in the case of complete data.
After this, in Secion 4, we partially address the question of: how many maximal extensions a positively identifiable family has. We study two particular cases: subfamilies of the family of all pairs and subfamilies of the family of all triples. These two cases allowed us to conclude general results for: subfamilies of the

family of all n-tuples for a fixed n. Our conclusions on these cases cannot be interpreted as answers for the more general -and much more complex- question, but we hope this will bring light towards it. Finally, in Section 6, we fix our attention on families which are anti-chains. We show that infinite anti-chains of infinite languages exist which can be identified with complete information but not with positive information only. For infinite anti-chains of finite languages such an example cannot exist if the indexing of the languages is by canonical indices because such families are always positively identifiable. The case of arbitrary indexing is investigated but not fully solved. We do exhibit an example of an indexable family of finite sets that cannot be given a canonical indexing.

2 Preliminaries

We use standard notions from recursion theory and learning theory (see e.g., [8]), and for "Angluin's style" identification in the limit (see [1], [7]).

Since we can represent strings of symbols by natural numbers, we will always refer to $\mathbb{N}$ as our universal set. Thus *languages* are sets of natural numbers, i.e. $L \subseteq \mathbb{N}$. A *family* $\mathcal{L} = \{L_i | i \in \mathbb{N}\}$ will be *an indexed family of recursive languages*, i.e. the two-place predicate $y \in L_i$ is recursive. In case all languages are finite and there is a recursive function f such that for each i, $f(i)$ is a canonical index[1] for L_i i.e. $L_i = F_{f(i)}$, then we call $\mathcal{L}$ a *canonical family*.

In finite identification a *learner* will be a total recursive function that takes its values in $\mathbb{N} \cup \{\uparrow\}$ where $\uparrow$ stands for *undefined*.

A *positive data presentation* of a language L is an infinite sequence $\sigma^+ := x_1, x_2, \ldots$ of elements of $\mathbb{N}$ such that $\{x_1, x_2, \ldots\} = L$. A *complete data presentation* of a language L is an infinite sequence of pairs $\sigma := (x_1, t_1), (x_2, t_2), \ldots$ of $\mathbb{N} \times \{0, 1\}$ such that $\{x_n \in \mathbb{N} : t_n = 1, n \geq 0\} = L$ and $\{x_m \in \mathbb{N} : t_m = 0, m \geq 0\} = \mathbb{N} \setminus L$. An initial segment of length n of σ and σ^+ is indicated by $\sigma[n]$ and $\sigma^+[n]$.

A family $\mathcal{L}$ of languages is said to be *finitely identifiable from positive data* (*pfi*) (or *finitely identifiable from complete data* (*cfi*)) if there exists a recursive learner φ which satisfies the following: for any language L_i of $\mathcal{L}$ and for any positive data sequence σ^+ (or complete data sequence σ) of L_i as input to φ, φ produces on exactly one initial segment $\sigma^+[n]$ a guess $\varphi(\sigma^+[n]) = j$ such that $L_j = L_i$, and stops. We occasionally relax the condition of the recursivity of the learner φ, in such cases φ is said to be a *non-effective learner* and $\mathcal{L}$ is said to be *non-effectively finitely identifiable from positive data* (*non-effectively pfi*) (or *non-effectively finitely identifiable from complete data* (*non-effectively cfi*)). Clearly a family that is *pfi* it is also *non-effectively pfi*. Similarly for *cfi*. For readability, we will use *nepfi* to refer to the notion of *non-effectively pfi*.

Let $\mathcal{L}$ be a family of languages, and let L be a language in $\mathcal{L}$. A finite set D_L is a *definite tell-tale set* (DFTT) for L if $D_L \subseteq L$ and $\forall L' \in \mathcal{L}, (D_L \subseteq L' \rightarrow L' = L)$.

[1] We write F_n for the finite set with canonical index n

A language L' is said to be *consistent* with a pair of finite sets (B, C) if $B \subseteq L'$ and $C \subseteq \mathbb{N} \setminus L'$. A pair of finite sets D_L, $\overline{D}_L$ is a *definite,co-definite pair of tell-tale sets* (DFTT, co-DFTT) for L if L is consistent with $(D_L, \overline{D}_L)$, and $\forall L' \in \mathcal{L}$, if L' is consistent with $(D_L, \overline{D}_L)$ then, $L' = L$.

Theorem 1. *(Mukouchi's Characterization Theorem [7][5])*
A family $\mathcal{L}$ of languages is finitely identifiable with positive data (pfi) *iff for every $L \in \mathcal{L}$ there is a DFTT set D_L in a uniformly computable way. That is, there exists an effective procedure F that on input i, index of L, produces the canonical index $F(i)$ of some definite finite tell-tale of L.*

A family $\mathcal{L}$ of languages is finitely identifiable with complete data (cfi) *iff for every $L \in \mathcal{L}$ there is a pair of DFTT,co-DFTT sets $(D_L, \overline{D}_L)$ in a uniformly computable way.*

Corollary 1 *If a family $\mathcal{L}$ has two languages such that $L_i \subset L_j$, then $\mathcal{L}$ is not* pfi*.*

Clearly if a family if *pfi* then it is *cfi*. A completely analogous theorem holds for non-effective learners and non-effective procedures for *pfi* and *cfi*.

Theorem 2. *If $\mathcal{L}$ is a canonical family where no $L_i \in \mathcal{L}$ is a proper subset of any other $L_j \in \mathcal{L}$, then $\mathcal{L}$ is pfi.*

Proof. For every $L_i \in \mathcal{L}$. Simply take $D_i = L_i$ as the DFTT.

Similarly, if $\mathcal{L}$ is any family of finite languages that is an anti-chain then $\mathcal{L}$ is non-effectively pfi.

3 Finite families of languages

This section is dedicated to finite families of languages. A pair of simple but striking results already provides a good insight in a feature underlying the difference between finite identification on positive and on complete data.

Theorem 3. *A finite family of languages $\mathcal{L}$ is finitely identifiable from positive data iff no language $L \in \mathcal{L}$ is a proper subset of another $L' \in \mathcal{L}$.*

Proof. From left to right follows straightforwardly by contraposition of corollary 1. From right to left take L_i any language in $\mathcal{L}$. Since $L_i \nsubseteq L_j$ for any $j \neq i$, choose $n_{ij} = \mu n\{n \in L_i \setminus L_j\}$ and let $D_i = \{n_{ij} | j \neq i\}$. Let us verify that D_i is a DFTT for L_i: Clearly it is finite because the family is finite, so $\{n_{ij} | j \neq i\}$ is finite and $D_i \subseteq L_i$. By construction, if $D_i \subseteq L_k \in \mathcal{L}$ then $i = k$. □

Theorem 4. *Any finite collection of languages $\mathcal{L} = \{L_1, \ldots, L_n\}$ is finitely identifiable with complete data.*

Proof. Let $\mathcal{L}$ be any family of languages satisfying this condition and L_i any language in $\mathcal{L}$. Take any $j \neq i$, then $L_i \setminus L_j \neq \emptyset$ or $L_j \setminus L_i \neq \emptyset$. If $L_i \setminus L_j \neq \emptyset$, take the first $n_{ij} \in L_i \setminus L_j$ to be in D_i. If $L_j \setminus L_i \neq \emptyset$, take $m_{ij} \in L_j \setminus L_i$ to be in $\overline{D}_i$. This pair of sets is consistent with L_i by construction, in fact they are DFTT, co-DFTT sets for L_i. Note that this pair cannot be consistent with any other language $L_k \in \mathcal{L}$ such that $L_k \neq L_i$ simply by construction. Since i was arbitrary, by Mukouchi's characterization theorem for complete data we have that $\mathcal{L}$ is cfi. □

4 Looking for maximal learnable families

In this section we question whether any finitely identifiable family is contained in a maximal finitely identifiable one. A positive answer to such question allows an alternative classification of finitely learnable families. We first address the question for maximal *nepfi* families and later for maximal *pfi*. We provide positive result for maximal *nepfi* extensions of families with finite languages. For families containing infinite languages this question remains open. For the usual *pfi* families, maximal *pfi* extensions exist. But we do give an example of a canonical family which does not have a maximal *pfi* extension. The case of *cfi* is rather different, as we will show in Subsection 4.2 that maximal extensions for *cfi* families never exist.

Theorem 5. *Every indexed family of finite languages which is* pfi *is contained in a maximal family of languages which is* nepfi*.*

Proving theorem 5 is by a classical Zorn lemma construction. If infinite languages are present in the family, such a Zorn lemma construction cannot be applied since not every family which is an anti-chain is *nepfi*.

4.1 The existence of maximal *pfi* families

In this subsection we address the above questions for *pfi* families.

The following example shows that not every *pfi* family can be extended into a maximal *pfi* family. First we present the following known definition of *recursively inseparable sets.*

Definition 1 *We say that $A, B \subseteq \mathbb{N}$, $A \cap B = \emptyset$, are recursively inseparable iff there is no recursive set $C \subseteq \mathbb{N}$ such that ($C \supseteq A$ and $C \cap B = \emptyset$).*

It is well-known that r.e. sets A, B exist which are recursively inseparable [10].

Theorem 6. *Let $A \subseteq \mathbb{N}$ and $B \subset \mathbb{N} \setminus A$ be two recursively inseparable r.e sets. Let $\mathcal{L} := \{\{a\} : a \in A\} \cup \{\{b, c\} : b, c \in B\}$. The family $\mathcal{L}$ is* pfi *and there is no canonical maximal* pfi *extension of $\mathcal{L}$.*

Proof. First note that since both A and B are r.e., $\mathcal{L}$ is a canonical indexed family. It is easy to see that it is *pfi* since any language serves as its own DFTT. Now by contradiction, suppose there is a maximal canonical *pfi* family extension of $\mathcal{L}$, say $\mathcal{L}'$. Note that because of maximality and canonicity of $\mathcal{L}'$, for each finite set $Y \subseteq \mathbb{N}$, we can decide whether $Y \in \mathcal{L}'$ or $Y \notin \mathcal{L}'$. This can be done by checking for Y with each $L_i \in \mathcal{L}$ whether $Y = L_i$, $Y \subset L_i$ or $L_i \subset Y$. One of the three has to happen for some L_i, otherwise Y can be added to $\mathcal{L}$ as a new element without impairing positive identifiability. Thus $\mathcal{L}'$ is decidable. Therefore we can construct a set $A' \supseteq A$ of singletons in $\mathcal{L}'$ that is recursive and $A' \cap B = \emptyset$, so A' separates A from B. This contradicts the inseparability of A and B. □

We can strengthen Theorem 6 to conclude that $\mathcal{L}$ has no maximal *pfi* extension at all.

Theorem 7. *The family $\mathcal{L}$ of Theorem 6 has no maximal pfi extension at all.*

Proof. Let $\mathcal{L}' \supseteq \mathcal{L}$ be a maximal pfi family. We define recursive A', B' such that $A \subseteq A'$ and $B \subseteq B'$, $A' \cap B' = \emptyset$, $A' \cup B' = \mathbb{N}$. Let $\mathcal{L}' = \{L_n : n \in \mathbb{N}\}$. For each i determine whether $i \in A'$ or $i \in B'$ as follows: Find the first n such that $i \in L_n$. By maximality and indexicality such n exists. Now consider D_n the DFTT of L_n. We distinguish two possibilities: (1) $D_n = \{i\}$, and (2) $D_n \neq \{i\}$. In case (1) put $i \in A'$. Note that $A \subseteq A'$, because if $i \in A$ then $\{i\} \in \mathcal{L}$ and thus $\{i\} \in \mathcal{L}'$. Note that $\{i\} \subset L_n$ is impossible since $\mathcal{L}'$ is pfi. Thus $\{i\} = L_n$ and so $D_n = \{i\}$. In case (2) put $i \in B'$. Note that $B \subseteq B'$, because if $i \in B$ then $\{i, j\} \in \mathcal{L} \subseteq \mathcal{L}'$ for some $j \neq i$. So $D_n \neq \{i\}$ because $\mathcal{L}'$ is pfi. Therefore A', B' have been constructed as required, contradiction. □

This theorem applies not only to families with infinite members but also to non-canonical families of only finite languages. In theorem 14 from section 6 we show that such families exist.

4.2 Do maximal *cfi* families exist?

In this section we address the question whether every *cfi* family is contained in a maximal one. Or in other words, if we can always find *cfi* extensions for *cfi* families. Surprisingly, we show that the latter is indeed always possible, the question whether maximal *cfi* families exist is answered negatively.

First observe the following. Let $\overline{\mathcal{L}}$ be the complement family of any cfi family $\mathcal{L}$, i.e. $\overline{\mathcal{L}} = \{\overline{L} : L \in \mathcal{L}\}$ where $\overline{L} = \mathbb{N} \setminus L$. Note that for every sequence σ of complete data for a family $\overline{\mathcal{L}}$ there is mirror image of σ, say sequence $\overline{\sigma}$ (presented in exactly the same order), for the cfi family $\mathcal{L}$ with inverted values of 0's and 1's. So $(k, 1)_j \in \sigma$ iff $(k, 0)_j \in \overline{\sigma}$ for any $j \in \mathbb{N}$. We obtain the following result.

Proposition 1 *If a family $\mathcal{L}$ is cfi then $\overline{\mathcal{L}}$ is cfi as well.*

Proof. Let φ be a learner for $\mathcal{L}$ we can define a learner $\overline{\varphi}$ for $\overline{\mathcal{L}}$ as follows:

$$\overline{\varphi}(\sigma[n]) = \overline{L} \text{ iff } \varphi(\overline{\sigma}[n]) = L$$

Clearly $\overline{\varphi}$ is a recursive learner for $\overline{\mathcal{L}}$. □

Corollary 2 *If either $\mathcal{L}$ or $\overline{\mathcal{L}}$ is cfi then $\mathcal{L}$ and $\overline{\mathcal{L}}$ are cfi.*

This is not the case for *pfi* families, since for instance the family of all singletons $\mathcal{L}^s$ is *pfi* but its complement family, namely the family of all co-singletons, is clearly not *pfi*.
Consider any language L, then a *direct successor* of L is $L \cup \{n\}$ with n not in L. For every non-cofinite language $L_i \subseteq \mathbb{N}$ let $Suc(L_i)$ be the set of all direct successors of L_i and $Suc_{\mathcal{L}}(L_i) = Suc(L_i) \cap \mathcal{L}$.

Proposition 2 *If $\mathcal{L}$ is cfi then $Suc_{\mathcal{L}}(L_i)$ is finite for every L_i in $\mathcal{L}$.*

Proof. Let $\mathcal{L}$ be a cfi family and $L \in \mathcal{L}$. By contradiction suppose there are infinitely many direct successors of L in $\mathcal{L}$. Thus we assume that $Suc_{\mathcal{L}}(L)$ has infinitely many elements.
Since $\mathcal{L}$ is cfi we have a DFTT and co-DFTT sets for L, namely D_L and $\overline{D}_L$. First note that $D_L \subseteq L_i$ for any $L_i \in Suc(L)$. Since $\overline{D}_L$ is finite, the contradiction will follow by showing that $\overline{D}_L$ only serves to disambiguate between a finite number of direct successors of L in $\mathcal{L}$. We prove the following: $\overline{D}_L \cap L_j \neq \emptyset$ only for finitely many $L_j \in Suc_{\mathcal{L}}(L)$.
First note that $\overline{D}_L$ is finite and for all disjoint $L_j, L_i \in Suc_{\mathcal{L}}(L)$, $L_j = L \cup \{k_j\} \neq L \cup \{k_i\} = L_i$ for some $k_j, k_i \in \mathbb{N}$. Since $\overline{D}_L$ is co-DFTT of L, $\overline{D}_L \cap L = \emptyset$. Thus, if $L_j \in Suc_{\mathcal{L}}(L)$ and $\overline{D}_L \cap L_j \neq \emptyset$ then $\overline{D}_L \cap L_j = \{k_j\}$. Since for each $L_i, L_j \in Suc_{\mathcal{L}}(L)$ we have that $k_j \neq k_i$ and $\overline{D}_L$ is finite, $\overline{D}_L$ can only intersect finitely many $L_j \in Suc_{\mathcal{L}}(L)$.

Continuing with the general proof. Take $L_i \in Suc_{\mathcal{L}}(L)$ such that $L_i \cap \overline{D}_L = \emptyset$. We can take such a language $L_i \in Suc_{\mathcal{L}}(L)$ because of the previous claim and our initial assumption that the set $Suc_{\mathcal{L}}(L)$ is infinite. Then L_i is a witness for showing that $\overline{D}_L, D_L$ are not co-DFTT and DFTT sets for L which is a contradiction. This is because $D_L \subseteq L_i$ and $\overline{D}_L \cap L_i = \emptyset$, so disambiguation between L and L_i is not possible.
Since the choice of $\overline{D}_L$ was arbitrary as a co-DFTT for L it follows that $Suc_{\mathcal{L}}(L)$ must be finite. □

Next comes the crucial result, the non-existence of maximal cfi families.

Theorem 8. *Let $\mathcal{L}$ be an indexed* cfi *family and $L \in \mathcal{L}$. For any pair DFTT, co-DFTT $D_L, \overline{D}_L$ of L, if $D_L \cup \{n\}$ is such that $n \notin \overline{D}_L \cup L$ then $\mathcal{L} \cup \{D_L \cup \{n\}\}$ is* cfi.

Proof. The strategy is to extend $\mathcal{L}$ with $D_L \cup \{n\}$ where D_L is the DFTT of L selected by the DFTT-function for the family $\mathcal{L}$. W.l.o.g. we can assume that $\mathcal{L}$ has a non cofinite language, and such language we fix as L. This is simply because of the following: if all languages in $\mathcal{L}$ are cofinite, then all languages in the complement family $\overline{\mathcal{L}}$ are finite. Thus we can find a cfi extension $\mathcal{L}'$ for $\overline{\mathcal{L}}$ by applying the theorem to $\overline{\mathcal{L}}$. By proposition 1 we know that $\overline{\mathcal{L}'}$ is cfi. Note that $\overline{\mathcal{L}'}$ is an extension of our original family $\mathcal{L}$. Therefore, we can assume that $L \in \mathcal{L}$ is not cofinite.

Let $\mathcal{L}$ be a countable family with a non cofinite language $L \in \mathcal{L}$ and $\overline{D}_L$ the co-DFTT for L obtained by the dftt/co-dftt function on $\mathcal{L}$. Since L is not cofinite we know that L has infinitely many direct successors. By Proposition 2 we have that $\mathcal{L}$ contains only finitely many direct successors of L. Note that since $\overline{D}_L$ is finite, we can choose infinitely many $m \in \mathbb{N}$ such that $m \notin \overline{D}_L$ and $m \notin L' \in Succ_{\mathcal{L}}(L)$. Take $n \in \mathbb{N}$ satisfying these characteristics and $D_L \cup \{n\}$ the respective direct successor for D_L.

Claim: The family $\mathcal{L}' = \mathcal{L} \cup \{D_L \cup \{n\}\}$ is cfi.

In fact we claim that the finite sets $D'_{D_L \cup \{n\}} = D_L \cup \{n\}$, $\overline{D}'_{D_L \cup \{n\}} = \overline{D}_L$, and $D'_L = D_L$, $\overline{D}'_L = \overline{D}_L \cup \{n\}$ are DFTT, co-DFTT sets for $D_L \cup \{n\}$ and L respectively in $\mathcal{L}'$. If $D_L \cup \{n\} = D_j \cup \{n'\}$ happens to be the case for some $L_j \in \mathcal{L}$, then fix $D'_j = D_j$ and $\overline{D}'_j = \overline{D}_j \cup \{m \in D_L \cup \{n\} \;:\; m \notin L_j\}$. For the rest of the languages in $\mathcal{L}'$ the DFTT's and the co-DFTT's will be exactly the ones chosen by the function f initially for $\mathcal{L}$, i.e, for all the rest of $L_i \in \mathcal{L}'$, $D'_i = D_i$ and $\overline{D}'_i = \overline{D}_i$.

Proof of claim: By construction of D'_L, $\overline{D}'_L$, and $D_L \cup \{n\}$, the pair D'_L, $\overline{D}'_L$ cannot be consistent with $D_L \cup \{n\}$. It cannot be consistent with any other $L_j \in \mathcal{L}'$ either because that will contradict that $\mathcal{L}$ is cfi. Now we need to show that $D'_{D_L \cup \{n\}}$ and $\overline{D}'_{D_L \cup \{n\}}$ are not consistent with any other $L_j \in \mathcal{L}'$. By contradiction, suppose $D'_{D_L \cup \{n\}}$ and $\overline{D}'_{D_L \cup \{n\}}$ are consistent with a language $L_j \in \mathcal{L}'$ and $L_j \neq D_L \cup \{n\}$. Since $L_j \neq D_L \cup \{n\}$ and by definition of $\mathcal{L}'$ we obtain $L_j \in \mathcal{L}$. By definition of $\mathcal{L}$ and since $L, L_j \in \mathcal{L}$ we have $L_j \neq L$. Thus,

$$D_L \subseteq D'_{D_L \cup \{n\}} \subseteq L_j$$

and

$$\overline{D}'_{D_L \cup \{n\}} = \overline{D}_L \subseteq \mathbb{N} \setminus L_j.$$

This implies that D_L and $\overline{D}_L$ were not DFTT, co-DFTT sets for L w.r.t. $\mathcal{L}$, contradicting that $\mathcal{L}$ is cfi. Similarly we cannot have any pair $D_i, \overline{D}_i$ chosen by the function f consistent with any other $L_j \neq L_i \in \mathcal{L}'$. Therefore $\mathcal{L} \cup \{D_L \cup \{n\}\}$ is cfi. □

Corollary 3 *Maximal* cfi *extensions do not exist for any* cfi *family* $\mathcal{L}$*.*

There are other ways of extending a *cfi* family than the one described in Theorem 8 as the following example shows.

Example 1 *Take the family* $\mathcal{L} = \{\{0\}, \{0,1\}, \{0,1,2\}, \ldots, \{0,1,2,3,\ldots,n\}, \ldots\}$. *This family is* cfi. *Note that for* $L = \{0\}$ *we can extend* $\mathcal{L}$ *with* $L \cup \{2\}$ *and preserve* cfi *even though a co-DFTT is* $\{1,2\}$. *Moreover we can extend it with* $L \cup \{3\}$, $L \cup \{4\}$ *and so on, and preserve* cfi.

5 Counting maximal extensions

We are also interested in the follow-up question: How many maximal *nepfi* extensions can a *pfi* family have? It is our guess that every pfi family has finitely many maximal *nepfi* extensions or continuously many. We ignore indexability of the family in investigating this question, and, since we do not care in this section whether we have a *pfi* or *nepfi* maximal extension because we are hert after structural properties only we may be less careful about the distinction.

First consider the following example: Let $\mathcal{L}^s$ be the family of all singletons. Clearly it is maximal with respect to *pfi*. However if we take out one of the singletons, say $\{0\}$, we obtain a *pfi* subfamily $\mathcal{L}^s_0$ which is no longer maximal and its only *pfi* extension is $\mathcal{L}^s$. If we remove $\{1\}$ from $\mathcal{L}^s_0$, we can maximally extend this family in two different ways, either adding $\{0, 1\}$ or adding $\{0\}$ and $\{1\}$. Thus we have two independent maximal *pfi* extensions for $\mathcal{L}^s_1$. We can repeat this effective deletion-procedure finitely many times and still obtain finitely many extensions. For regaining maximality, we are indeed "restricted" in the structural sense. The following lemma illustrates this.

Lemma 1. *Let $\mathcal{L}$ be a maximal* pfi *family and $\mathcal{L}'$ is a maximal* pfi *extension of $\mathcal{L} \setminus \{x\}$ where $\{x\} \in \mathcal{L}$. Then for all $L \in \mathcal{L}'$ which are not in $\mathcal{L} \setminus \{x\}$, L is of the form $\{x\} \cup A$ for some $A \subseteq L_i \in \mathcal{L} \setminus \{x\}$.*

Proof. Note first that in order to achieve pfi maximality in an extension $\mathcal{L}'$ of $\mathcal{L} \setminus \{\{x\}\}$, any new language $L \in \mathcal{L}'$ in the extension needs to have x as an element. Thus any $L \in \mathcal{L}' \setminus (\mathcal{L} \setminus \{\{x\}\})$ is such that $x \in L$. Let $A = L \setminus \{\{x\}\}$. We will prove that $A \subset L_i$ for some $i \in \mathbb{N}$. By maximality of $\mathcal{L}$, A itself could not be added to $\mathcal{L}$ and preserve pfi. Thus, either $A \subset L_i$ or $L_i \subseteq A$ for some $L_i \in \mathcal{L}$. The latter cannot be since if $L_i \subseteq A$ then $L_i \subseteq A \cup \{x\}$ and $\mathcal{L}'$ should be an anti-chain. Therefore $A \subset L_i$. □

In the following example we see that even when the languages are all finite, we may still regain uncountably many maximal *pfi* extensions.

Example 2 *Let $\mathcal{L} = \{\{0\} \cup \mathcal{L}'\}$ where $\mathcal{L}' = \{\{i, j, k\} : i, j, k \in \mathbb{N} \setminus \{0\}\}$. Clearly $\mathcal{L}$ is maximal* pfi *family. Consider $\mathcal{L}'$, by lemma 1 in order to regain maximality, the languages to add must be of the form $\{0\} \cup A$ for some $A \subseteq L_i$ and some $L_i \in \mathcal{L}'$. Therefore we have the following procedure for constructing continuously many maximal* pfi *extensions of $\mathcal{L}'$: For each $B \subseteq \mathbb{N} \setminus \{0\}$ add the triples of the form $\{0, n, m\}$ with $n \neq m$ and $n, m \in B$ and all the pairs of the form $\{0, c\}$ with $c \notin B$. This construction is for all $B \subseteq \mathbb{N} \setminus \{0\}$, thus $\mathcal{L}'$ has continuously many maximal* pfi *extensions.*

5.1 Continuously many maximal pfi extensions

We dedicate this subsection to study cases in which we can recover uncountably many maximal extensions of a given family. We first address some cases of families with finite languages similar to example 2. After studying these cases, we exhibit some sufficient conditions for a family in order to have continuously many maximal extensions.

Consider the following example.

Example 3 *Let $\mathcal{L}$ be the family $\{0\}, \{1\} \cup \{\{i,n\} : i,n \in \mathbb{N} \setminus \{0,1\}\}$. This is clearly a pfi family because every language is mutually incomparable with any other language in the family. Moreover it is maximal (w.r.t. pfi) precisely because any other subset of $\mathbb{N}$ is either a subset or a superset of $\{\{i,n\}|i,n \in \mathbb{N} \setminus \{0,1\}\}$, or a superset of $\{0\}, \{1\}$. Now consider the subfamily $\mathcal{L}' = \mathcal{L} \setminus \{\{0\}, \{1\}\}$, $\mathcal{L}'$ has continuously many maximal pfi extensions. Clearly $\mathcal{L}$ is one, and for every $B \subseteq \mathbb{N}$, the family $\mathcal{L}' \cup \{\{0,1,b\}|b \in B\} \cup \{\{0,c\},\{1,c\}|c \notin B\}$ is a maximal pfi extension of $\mathcal{L}'$. Since we have continuously many $B \subseteq \mathbb{N}$, we have continuously many maximal pfi extensions of $\mathcal{L}'$.*
However, if we take the similar maximal pfi *family $\{0\} \cup \{\{i,n\} : i,n \in \mathbb{N} \setminus \{0\}\}$ and consider the* pfi *subfamily $\{\{i,n\} : i,n \in \mathbb{N} \setminus \{0\}\}$ it turns out that it has only two maximal* pfi *extensions, namely $\mathcal{L}^2$ and $\{0\} \cup \{\{i,n\} : i,n \in \mathbb{N} \setminus \{0\}\}$ itself.*

By a similar combinatorial argument as in example 3, we straightforwardly obtain the following result .

Proposition 3 *For every finite set $\{0,1,\ldots,m\}$ with $m > 0$, the subfamily $\mathcal{L}^2_{\setminus\{0,1,\ldots,m\}} = \{\{i,n\}|i,n \in \mathbb{N} \setminus \{0,1,\ldots,m\}\}$, of the family of all pairs $\mathcal{L}^2$, has continuously many maximal pfi extensions.*

Proof. Simply because $\mathcal{L}^2_{\setminus\{0,1,\ldots,m\}} \subset \mathcal{L}^2 \setminus \{\{0,a,\},\{1,b\} : a,b \in \mathbb{N}\}$ and, by example 3, the latter has continuously many maximal pfi extensions. □

By example 2 we know that the subfamily $\mathcal{L}^3 \setminus \{\{0,a,b\} : a,b \in \mathbb{N}\}$ already has continuously many maximal pfi extensions. Therefore any subfamily $\mathcal{L}^3_{\setminus\{0,1,\ldots,m\}}$ obtained by removing all triples of the form $\{i,a,b\}$ with $a,b \in \mathbb{N}$ and $i \in \{0,1,\ldots,m\}$ has continuously many maximal pfi extensions for any $m \in \mathbb{N}$. Since a similar combinatorial argument works for any subfamily of quadruples, quintuples etc, we can generalize this result to all $n \geq 3 \in \mathbb{N}$.

Proposition 4 *Let $n \geq 3 \in \mathbb{N}$ and $\mathcal{L}^n$ be the class of all n-tuples. Any subfamily $\mathcal{L}^n_{\setminus\{0,1,\ldots,m\}}$ obtained by removing all n-tuples of the form $\{i,x_1,\ldots,x_{n-1}\}$ with $x_j \in \mathbb{N}$ and $i \in \{0,1,\ldots,m\}$ has continuously many maximal pfi extensions for any $m \in \mathbb{N}$.*

5.2 The class of all pairs $\mathcal{L}^2$

In this subsection we study subfamilies of the family of all pairs. This will also bring some general insights for equinumerous families with more than two elements. First we provide the following definition. In it and further on we will write *n-tuple* for and unordered n-tuple, i.e. just a set of n elements.

Definition 2

- *Let $\mathcal{Y} = \{Y_1, \dots Y_n\}$ be any set of pairs in $\mathcal{L}^2$, $NUM(\mathcal{Y})$ to be the set of all numbers which appear in the pairs $Y_1, \dots Y_n$, and $PAIRS(\mathcal{Y})$ the set of all pairs formed by elements in $NUM(\mathcal{Y})$. Let $\mathcal{L}^{\mathcal{Y}}$ to be the subfamily of all pairs which are not in $PAIRS(\mathcal{Y})$, i.e. $\mathcal{L}^{\mathcal{Y}} = \mathcal{L}^2 \setminus PAIRS(\mathcal{Y})$.*
- *We can easily generalize the definition above to the family $\mathcal{L}^n$ of* all n-tuples *for $n \in \mathbb{N}$. We denote as $nTUP(\mathcal{Y})$ the set of all n-tuples formed by elements in $NUM(\mathcal{Y})$ and $\mathcal{L}^{\mathcal{Y}} = \mathcal{L}^n \setminus nTUP(\mathcal{Y})$.*

The combinatorial notion of Sperner familiy explains why for every finite set of pairs $\mathcal{Y} = Y_1, \dots Y_n$, the subfamily $\mathcal{L}^{\mathcal{Y}}$ has finitely many maximal *pfi* extensions.

Definition 3 *A* Sperner family *(or Sperner system) is a family of subsets of A in which none of the sets is contained in any other. Equivalently, a Sperner family is an* anti-chain *in the inclusion lattice over the power set of A.*

From here on we will refer to Sperner families as anti-chains. The number of different anti-chains on a set of n elements is counted by the so-called Dedekind numbers. Determining these numbers is known as the Dedekind problem. The number amount of anti-chains on $\{0, 1, 2, \dots, n\}$ for $n \in \mathbb{N}$ are $2, 3, 6, 20, 168, 7581, \dots$ respectively.

Lemma 2. *Let $\mathcal{L}^{\mathcal{Y}} \subseteq \mathcal{L}^2$ be the family corresponding to some finite set of pairs $\mathcal{Y} = \{Y_1, \dots Y_n\}$. For every maximal pfi extension $\mathcal{L}$ of $\mathcal{L}^{\mathcal{Y}}$ and every $L \in (\mathcal{L} \setminus \mathcal{L}^{\mathcal{Y}})$, $L \subseteq NUM(\mathcal{Y})$.*

Proof. To obtain a contradiction suppose there is a maximal pfi extension $\mathcal{L} \neq \mathcal{L}^2$ of $\mathcal{L}^{\mathcal{Y}}$ such that for some $L \in (\mathcal{L} \setminus \mathcal{L}^{\mathcal{Y}})$, $L \not\subseteq NUM(\mathcal{Y})$. Thus, there is $z \in L$ such that $z \notin NUM(\mathcal{Y})$. Clearly L cannot be a singleton, thus a $w \neq z$ exists in L such that $\{w, z\} \notin PAIRS(\mathcal{Y})$. Therefore $\{w, z\} \in \mathcal{L}^{\mathcal{Y}}$ simply by definition of $\mathcal{L}^{\mathcal{Y}}$. But since $\{w, z\} \subseteq L \in \mathcal{L}$, $\mathcal{L}$ cannot be a maximal pfi extension of $\mathcal{L}^{\mathcal{Y}}$ contradicting our initial assumption. □

Proposition 5 *For every finite set of pairs $\mathcal{Y} = \{Y_1, \dots Y_n\}$, the number of maximal pfi extensions of the subfamily $\mathcal{L}^{\mathcal{Y}}$ is bounded by the Dedekind number of the set $NUM(\mathcal{Y}) = \{y_1, \dots, y_m\}$ or in other words, by the number of anti-chains in $NUM(\mathcal{Y}) = \{y_1, \dots, y_m\}$. Moreover, the maximal pfi extensions of $\mathcal{L}^{\mathcal{Y}}$ correspond to the maximal singleton-free anti-chains on $NUM(\mathcal{Y})$.*

Proof. Let $\mathcal{Y} = \{Y_1, \dots Y_n\}$ be any finite set of pairs and $\mathcal{L}^{\mathcal{Y}} \subseteq \mathcal{L}^2$ the corresponding family. By lemma 2 we know that for every $\mathcal{L}$ maximal pfi extension of $\mathcal{L}^{\mathcal{Y}}$, if $L \in (\mathcal{L} \setminus \mathcal{L}^{\mathcal{Y}})$ then $L \subseteq NUM(\mathcal{Y}) = \{y_1, \dots, y_m\}$. Therefore, for every $\mathcal{L}$ maximal pfi extension of $\mathcal{L}^{\mathcal{Y}}$ we have that $(\mathcal{L} \setminus \mathcal{L}^{\mathcal{Y}}) \subseteq \mathcal{P}(NUM(\mathcal{Y}))$ which is finite. By Mukouchi's corollary (corollary 1) we know that $(\mathcal{L} \setminus \mathcal{L}^{\mathcal{Y}})$ must be an anti-chain in $\mathcal{P}(NUM(\mathcal{Y}))$. Therefore every $\mathcal{L}$ maximal pfi extension of $\mathcal{L}^{\mathcal{Y}}$ corresponds to some anti-chain in $\mathcal{P}(NUM(\mathcal{Y}))$ without singletons. Moreover since $\mathcal{L} \supseteq \mathcal{L}^{\mathcal{Y}}$ is maximal pfi then $(\mathcal{L} \setminus \mathcal{L}^{\mathcal{Y}})$ is precisely a maximal singleton-free anti-chain in $\mathcal{P}(NUM(\mathcal{Y}))$. For the other direction, if we extend $\mathcal{L}^{\mathcal{Y}}$ with any maximal singleton-free anti-chain in $\mathcal{P}(NUM(\mathcal{Y}))$ then clearly the resulting family $\mathcal{L}$ is a maximal pfi extension. Simply because any $L \in (\mathcal{L} \setminus \mathcal{L}^{\mathcal{Y}})$ has L itself as a DFTT set. □

We can straightforwardly generalize Proposition 5 for a subfamily of the family of all n-tuples, $\mathcal{L}^n$, for every $n \in \mathbb{N}$.

Proposition 6 *For every finite set of n-tuples $\mathcal{Y} = \{Y_1, \dots Y_n\}$, the number of maximal pfi extensions of the subfamily $\mathcal{L}^{\mathcal{Y}}$ is bounded by the number of anti-chains in the finite set $NUM(\mathcal{Y})$. Moreover, the maximal pfi extensions of $\mathcal{L}^{\mathcal{Y}}$ correspond to the maximal anti-chains in $NUM(\mathcal{Y})$ and such anti-chains contain no k-cardinality sets for any $k \leq n-1$.*

So far we know the following about subfamilies of $\mathcal{L}^2$: (1) By proposition 5, any subfamily of $\mathcal{L}^2$ obtained by removing finitely many pairs from $\mathcal{L}^2$ has only finitely many maximal pfi extensions; (2) by Example 3 and proposition 3 we know that any subfamily of $\mathcal{L}^2$ obtained by removing all pairs of the form $\{i, n\}$ with $n \in \mathbb{N}$ and $i \in \{0, 1, \dots, m\}$ (of which there are infinitely many) has either 2 maximal pfi extensions (when $0 = m$) or continuously many (when $0 < m$). But what happens when we consider subfamilies obtained by removing finitely or infinitely many arbitrary pairs? The answer to this question will also clarify what happens to subfamilies of all n-tuples $\mathcal{L}^n$ for any $n \in \mathbb{N}$. In this section we will first study what happens when we remove from $\mathcal{L}^2$ any finite group of pairs. Then we will study the case of removing infinitely many pairs. We provide a complete overview of our investigation on the number of maximal pfi extensions of subfamilies of $\mathcal{L}^2$ and will be able to conclude that every subfamily of $\mathcal{L}^2$ has either finitely or continuously many maximal pfi extensions.

Definition 4 *We say that $\mathcal{G} \subseteq \mathcal{L}^2$ is a* cluster *in $\mathcal{L}^2$ if $PAIRS(\mathcal{G}) = \mathcal{G}$ (see Definition 2) and $\|\mathcal{G}\| > 1$.*

Clearly for every $\mathcal{Y} \subseteq \mathcal{L}^2$, $PAIRS(\mathcal{Y})$ is a cluster in $\mathcal{L}^2$. The minimal-in-size clusters of $\mathcal{L}^2$ are the ones that contain three pairs.

Lemma 3. *For any finite set $\mathcal{Y} \subseteq \mathcal{L}^2$, $\mathcal{Y} \subseteq PAIRS(\mathcal{Y})$ and this is the minimal cluster that contains $\mathcal{Y}$.*

To illustrate the lemma above, let $\mathcal{Y} = \{\{1, 2\}, \{2, 3\}\}$. Then the minimal cluster that contains $\mathcal{Y}$ is $PAIRS(\mathcal{Y}) = \{\{1, 2\}, \{1, 3\}, \{2, 3\}\}$. We have many finite clusters that contain $PAIRS(\mathcal{Y})$ and therefore $\mathcal{Y}$. For instance the cluster $\mathcal{G} = \{\{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 4\}, \{2, 4\}, \{3, 4\}\}$.

Proposition 7 *Let $\mathcal{G}_1, \dots, \mathcal{G}_n$ be a finite set of clusters. Then the family*

$$\mathcal{L}^2 \setminus (\mathcal{G}_1 \cup \dots \cup \mathcal{G}_n)$$

has finitely many maximal pfi extensions.

Proof. Take $\mathcal{G} \supseteq \mathcal{G}_1 \cup \ldots \cup \mathcal{G}_n$ the minimal cluster that contains all $\mathcal{G}_1, \ldots, \mathcal{G}_n$, which exists by lemma 3. By Proposition 5, $\mathcal{L}^{\mathcal{G}}$ has finitely many maximal pfi extensions, and $\mathcal{L}^2 \setminus (\mathcal{G}_1 \cup \ldots \cup \mathcal{G}_n)$ as well, since $\mathcal{L}^2 \setminus (\mathcal{G}_1 \cup \ldots \cup \mathcal{G}_n) \supseteq \mathcal{L}^{\mathcal{G}}$. □

Definition 5 *We say that $\{\mathcal{G}_1, \ldots, \mathcal{G}_n\}$ is a maximal set of clusters in $\mathcal{L}^2$ outside $\mathcal{L}$, if they are pairwise disjoint, $\mathcal{G}_1 \cup \ldots \cup \mathcal{G}_n \subseteq \mathcal{L}^2 \setminus \mathcal{L}$ and for any cluster $\mathcal{G} \subseteq \mathcal{L}^2 \setminus \mathcal{L}$, $\mathcal{G} \subseteq \mathcal{G}_i$ for some $i \in \{1, \ldots, n\}$.*

Theorem 9. *Let $\mathcal{L} \subseteq \mathcal{L}^2$.*

1. *If there are only finitely many clusters $\mathcal{G}_1, \ldots, \mathcal{G}_n$ such that $(\bigcup_{i=1}^n \mathcal{G}_i) \subseteq (\mathcal{L}^2 \setminus \mathcal{L})$ and for at most one $k \in \mathbb{N}$ $\{\{k, m\} : m \in \mathbb{N} \setminus \{k\}\} \cap \mathcal{L} = \emptyset$, then $\mathcal{L}$ has finitely many maximal pfi extensions.*
2. *If $\{\mathcal{G}_i\}_{i \in \mathbb{N}}$ is a countable sequence of disjoint clusters such that $\bigcup_{i=1}^\infty \mathcal{G}_i \subseteq (\mathcal{L}^2 \setminus \mathcal{L})$, or if for more than one $k \in \mathbb{N}$ we have that $\{\{k, m\} : m \in \mathbb{N} \setminus \{k\}\} \cap \mathcal{L} = \emptyset$, then $\mathcal{L}$ has continuously many maximal pfi extensions.*

Proof. 1. Let $\mathcal{G}_1, \ldots, \mathcal{G}_n$ be the maximal finite sequence of disjoint clusters such that $(\bigcup_{i=1}^n \mathcal{G}_i) \subseteq (\mathcal{L}^2 \setminus \mathcal{L})$ and for at most one $k \in \mathbb{N}$, $\{\{k, m\} : m \in \mathbb{N} \setminus \{k\}\} \cap \mathcal{L} = \emptyset$. By treating the following cases we exhaust all the possibilities.

(a) There are no clusters contained in $(\mathcal{L}^2 \setminus \mathcal{L})$ and for exactly one $k \in \mathbb{N}$, $\{\{k, m\} : m \in \mathbb{N} \setminus \{k\}\} \cap \mathcal{L} = \emptyset$.

(b) There are no clusters contained in $(\mathcal{L}^2 \setminus \mathcal{L})$ and for no $k \in \mathbb{N}$, $\{\{k, m\} : m \in \mathbb{N} \setminus \{k\}\} \cap \mathcal{L} = \emptyset$, i.e. there are no clusters contained in $(\mathcal{L}^2 \setminus \mathcal{L})$ and for all $k \in \mathbb{N}$, $\{\{k, m\} : m \in \mathbb{N} \setminus \{k\}\} \cap \mathcal{L} \neq \emptyset$.

(c) There is a non-empty finite sequence of clusters $\mathcal{G}_1, \ldots, \mathcal{G}_n$ such that $(\bigcup_{i=1}^n \mathcal{G}_i) \subseteq (\mathcal{L}^2 \setminus \mathcal{L})$ and for at most one $k \in \mathbb{N}$, $\{\{k, m\} : m \in \mathbb{N} \setminus \{k\}\} \cap \mathcal{L} = \emptyset$.

Note that the case (a) when there are no clusters contained in $(\mathcal{L}^2 \setminus \mathcal{L})$ and for exactly one $k \in \mathbb{N}$, $\{\{k, m\} : m \in \mathbb{N} \setminus \{k\}\} \cap \mathcal{L} = \emptyset$ then each pair $\{a, b\} \in (\mathcal{L}^2 \setminus \mathcal{L})$ is of the form $\{k, m\}$ for some $m \in \mathbb{N}$. This is because otherwise, the pairs $\{a, b\}, \{k, a\}, \{k, b\}$ would form a cluster contained in $(\mathcal{L}^2 \setminus \mathcal{L})$. Therefore the only possibility is that $(\mathcal{L}^2 \setminus \mathcal{L}) = \{\{k, m\} : m \in \mathbb{N} \setminus \{k\}\}$ and we already discussed this case in example 3. So in this case $\mathcal{L}$ has only two possible maximal extensions.

Now we will prove case (b). Note that since there are no clusters contained in the complement of $\mathcal{L}$, we cannot add any language larger than a pair. To see this, take any triple of elements in $\mathbb{N}$, $\{a, b, c\}$. By our assumption, the cluster $\mathcal{G} = \{\{a, b\}, \{b, c\}, \{a, c\}\}$ is not contained in $\mathcal{L}^2 \setminus \mathcal{L}$. Therefore one of these pairs is already in $\mathcal{L}$. Thus we cannot add the language $\{a, b, c\}$ to extend $\mathcal{L}$. The same reasoning applies to any set larger than a triple. Since for all $k \in \mathbb{N}$, $\{\{k, m\} : m \in \mathbb{N} \setminus \{k\}\} \cap \mathcal{L} \neq \emptyset$, we cannot add singletons either: for every $i \in \mathbb{N}$ there is a language $L_i := \{i, m\} \in \mathcal{L}$ such that $\{i\} \subseteq \{i, m\}$ which prevents pfi. Therefore the only maximal pfi extension is $\mathcal{L}^2$.

Now we will prove case (c). Let $\mathcal{G} = \mathcal{G}_1, \ldots, \mathcal{G}_m$ be the maximal set of clusters outside $\mathcal{L}$, i.e. $(\bigcup_{i=1}^n \mathcal{G}_i) \subseteq (\mathcal{L}^2 \setminus \mathcal{L})$, and let $k \in \mathbb{N}$ be such that $\{\{k, m\} : m \in \mathbb{N} \setminus \{k\}\} \cap \mathcal{L} = \emptyset$. The strategy of the proof is to show that

each maximal extension of $\mathcal{L}$ is uniquely characterized by some maximal anti-chain in $NUM(\bigcup_{i=1}^{m} \mathcal{G}_i) \cup \{k\}$, and the number of maximal anti-chains is bounded by the Dedekind number of $NUM(\bigcup_{i=1}^{m} \mathcal{G}_i) \cup \{k\}$, which is finite. In order to achieve this we need to prove the following: For any maximal pfi extension $\mathcal{L}_m$ of $\mathcal{L}$ and any $A \in \mathcal{L}_m \setminus \mathcal{L}'$ we have that either $A \subseteq NUM(\bigcup \mathcal{G}) \cup \{k\}$ or $A = \{m, n\}$ for some $\{m, n\} \notin \mathcal{L}$.
We prove this by contradiction. Suppose there is $\mathcal{L}_m \supseteq \mathcal{L}$ such that $\mathcal{L}_m \neq \mathcal{L}^2$ and $A \in \mathcal{L}_m$ is such that $A \not\subseteq NUM(\bigcup \mathcal{G}) \cup \{k\}$ and $A \neq \{m, n\}$ for any $m, n \in \mathbb{N}$. Since $A \not\subseteq NUM(\bigcup \mathcal{G}) \cup \{k\}$, there is $y \in A$ such that $y \notin NUM(\bigcup \mathcal{G}) \cup \{k\}$. Therefore $y \neq k$. Note that A cannot be a singleton, say $\{y\}$, simply because $\{y\} \subseteq \{y, m+1\} \in \mathcal{L}$ where $m = max\{NUM(\mathcal{G}) \cup \{k\}\}$. Thus, there is $z \neq y$ such that $\{z, y\} \subseteq A$. Note that $A \neq \{z, y\}$ since we are supposing $A \neq \{m, n\}$ for any $\{m, n\} \notin \mathcal{L}$. Thus we have that $x \in A$ exists with $x \neq z, y$. Note that if $PAIRS(\{x, z, y\}) \subseteq (\mathcal{L}^2 \setminus \mathcal{L})$ then $PAIRS(\{x, z, y\}) \subseteq \mathcal{G}_i$ for some $i \in \{1, \dots, m\}$, but this cannot be since $y \notin NUM(\bigcup \mathcal{G}) \cup \{k\}$. Therefore $PAIRS(\{x, z, y\}) \cap \mathcal{L} \neq \emptyset$, i.e. there is a pair $\{a, b\} \subseteq \{x, z, y\} \subseteq A$ such that $\{a, b\} \in \mathcal{L}$, but this contradicts that $A \in \mathcal{L}_m \setminus \mathcal{L}$ where $\mathcal{L}_m$ is a maximal pfi extension of $\mathcal{L}$.

Continuing with the proof of this case we obtain that each maximal pfi extension of $\mathcal{L}$ which is not $\mathcal{L}^2$ is characterised by some anti-chain of $NUM(\bigcup \mathcal{G}) \cup \{k\}$ of which are just finitely many. This means that $\mathcal{L}_m = \mathcal{L} \cup P \cup Q$ where Q is an anti-chain of $NUM(\bigcup \mathcal{G}) \cup \{k\}$.

2. First we will study the case when for more than one $k \in \mathbb{N}$ we have that $\{\{k, m\} : m \in \mathbb{N}\} \cap \mathcal{L} = \emptyset$. Note that we already proved in example 3 that the family $\mathcal{L}' = \{\{i, n\} : i, n \in \mathbb{N} \setminus \{0, 1\}\}$ subfamily of $\mathcal{L}^2$ has continuously many maximal pfi extensions. Therefore any subfamily of $\mathcal{L}'$ has continuously many as well. Clearly every $\mathcal{L}$ satisfying the condition just mentioned will be a subfamily of $\mathcal{L}'$. Therefore $\mathcal{L}$ has continuously many maximal pfi extensions. Finally, we prove the remaining case. Let $\mathcal{G}_1, \dots, \mathcal{G}_n, \dots$ be a countable sequence of clusters such that $\bigcup_{i=1}^{\infty} \mathcal{G}_i \subseteq (\mathcal{L}^2 \setminus \mathcal{L})$. w.l.o.g. suppose each cluster is finite (this case is enough since every family $\mathcal{L}$ in which an infinite cluster H_i was "left out" of $\mathcal{L}$, is contained in a family $\mathcal{L}''$ in which every "left-out" cluster is finite), and suppose the clusters are pair-wise disjoint. For each $\mathcal{G}_i$ consider $NUM(G_i)$, the set of all numbers that appear in each pair of the cluster $\mathcal{G}_i$, then we can extend $\mathcal{L}$ in two different ways: (1) by adding the cluster $PAIRS(G_i) = G_i$, and (2) by adding the set $NUM(G_i)$. Note that these two ways of extending $\mathcal{L}$ are mutually exclusive. Therefore and since we have countably many clusters G_i, by a well-known combinatorial argument, we have continuously many maximal pfi extensions for $\mathcal{L}$. □

By Proposition 7 and Theorem 9 we have the following result.

Theorem 10. *Any subfamily $\mathcal{L}$ of $\mathcal{L}^2$ has either finitely many maximal pfi extensions or continuously many.*

Trivially, every $\mathcal{L} \subseteq \mathcal{L}^2$ has an indexable maximal effective pfi extension, namely $\mathcal{L}^2$ itself, but one can also see that, if there are finitely many maximal extensions all of them are indexable maximal effective pfi extensions and if there are continuously many, countably many of those are.

Theorem 9 for $\mathcal{L}^2$ allows us to obtain rather straightforwardly a similar general result for subclasses of the family of all n-tuples $\mathcal{L}^n$ for any $n \in \mathbb{N}$. But there are some subtle details so that we need to tread carefully. Therefore we dedicate the following section to the class of all subfamilies of $\mathcal{L}^n$.

5.3 The class of all n-tuples $\mathcal{L}^n$

Here we generalize all the notions and results we obtained for $\mathcal{L}^2$.

Definition 6 *We say that $\mathcal{G} \subseteq \mathcal{L}^n$ is an n-cluster in $\mathcal{L}^n$ if the set $nTUP(\mathcal{G})$ of all n-tuples formed by numbers in $NUM(\mathcal{G})$ (see Definition 2) is exactly $\mathcal{G}$, i.e., if $nTUP(\mathcal{G}) = \mathcal{G}$.*

Clearly for every $\mathcal{Y} \subseteq \mathcal{L}^n$, $nTUP(\mathcal{Y})$ is an n-cluster in $\mathcal{L}^n$.

Lemma 4. *For any finite set $\mathcal{Y} \subseteq \mathcal{L}^n$, $\mathcal{Y} \subseteq nTUP(\mathcal{Y})$, and this is the minimal n-cluster that contains $\mathcal{Y}$.*

Proposition 8 *Let $\mathcal{G}_1, \ldots, \mathcal{G}_m$ be a finite set of n-clusters. Then the family $\mathcal{L}^n \setminus (\mathcal{G}_1 \cup \ldots \cup \mathcal{G}_m)$ has finitely many maximal pfi extensions.*

Proof. The proof goes as in the case for $\mathcal{L}^2$, taking the minimal n-cluster that contains all $\mathcal{G}_1, \ldots, \mathcal{G}_n$, which by lemma 4 we know exists. By Proposition 6 we know that $\mathcal{L}^{\mathcal{G}}$ has finitely many maximal pfi extensions, and $\mathcal{L}^n \setminus (\mathcal{G}_1 \cup \ldots \cup \mathcal{G}_m) \supseteq \mathcal{L}^{\mathcal{G}}$, so $\mathcal{L}^n \setminus (\mathcal{G}_1 \cup \ldots \cup \mathcal{G}_m)$ has finitely many as well. $\square$

In what follows we will consider the subclasses of $\mathcal{L}^3$. This clarifies straightforwardly what happens in the general case for every $n > 3$.

What is different in $\mathcal{L}^3 \setminus \{\{0, a, b\} : a, b \in \mathbb{N}\}$ that it allows for continuously many maximal pfi extensions, whereas $\mathcal{L}^3 \setminus \{\{0, 1, b\} : b \in \mathbb{N}\}$ does not? The difference lies in the elements that are fixed and the ones that remain "free" in the triples that are discarded from the families. Whenever we fix two elements in the triples we are preventing the combinatorics to act, since with only one "free" element in the triple, there is not much that combinatorics can do. However, with two non-fixed entries we can build continuously many pfi extensions as in case (4) of example 2.

The following result generalizes what is expressed in this idea.

Proposition 9 *The family $\mathcal{L}^3 \setminus \bigcup_1^n \{\{k_i, m_i, a\} : a \in \mathbb{N} \setminus \{k_i, m_i\}\}$ for a finite number of pairs $\{k_i, m_i\} \in \mathcal{L}^2$ has finitely many maximal extensions.*

Proof. Let $\mathcal{L} := \mathcal{L}^3 \setminus \bigcup_1^n \{\{k_i, m_i, a\} : a \in \mathbb{N} \setminus \{k_i, m_i\}\}$. Consider n families of the form $\{\{k_i, m_i, a\} : a \in \mathbb{N} \setminus \{k_i, m_i\}\}$ such that for any $i \in \{1, 2, \ldots, n\}$, $\mathcal{L} \cap \{\{k_i, m_i, a\} : a \in \mathbb{N} \setminus \{k_i, m_i\}\} = \emptyset$. First we will prove that we cannot add any other N-tuple with $N \geq 5$ to $\mathcal{L}$, for this it suffices to see that we can only add finitely many quadruples, namely $\{k_i, m_i, k_j, m_j\}$ for any $i, j \in \{1, 2, \ldots, n\}$. This is because if we could add an N-tuple for $N \geq 5$, then we could add any quadruple of elements in the tuple. Let us prove it then by contradiction, suppose there is a quadruple $\{a, b, c, d\} \neq \{k_i, m_i, k_j, m_j\}$ for any $i, j \in \{1, 2, \ldots, n\}$ and such quadruple can extend $\mathcal{L}$ and preserve pfi. It is sufficient to verify the worst case scenario in which it differs in one element from all the admissible quadruples $\{k_i, m_i, k_j, m_j\}$. Suppose $a \neq k_1, k_2, \ldots, k_n$. It suffices to verify the case for $a \neq k_1$ because the others follows similarly. Note that $\{a, b, c, d\} = \{a, m_1, k_2, m_2\}$. Thus the triple $\{a, m_1, k_2\} \in \mathcal{L}$ because otherwise the 3-cluster $TRI(\{\{a, m_1, k_2\}, \{k_1, m_1, k_2, \}\}) \subseteq \mathcal{L}^3 \setminus \mathcal{L}$ which contradicts our assumption on $\mathcal{L}$. Thus since $\{a, m_1, k_2\} \in \mathcal{L}$, we cannot use $\{a, m_1, k_2, m_2\}$ to extend $\mathcal{L}$ and remain pfi. ◻

The result above does not apply when we consider infinitely many distinctive pairs $\{k_i, m_i\} \in \mathcal{L}^2$. We prove this in the following theorem where we cover all possible cases of subfamilies of $\mathcal{L}^3$. As we mentioned before, the proof for the generalization of Theorem 9 needs to be treated carefully since there are cases that do not correspond exactly to the ones for $n = 2$. For instance in the proof of the following theorem for $n = 3$, there are more cases of a similar kind in which the subfamily has continuously many maximal extensions. The general proof for $n \geq 3$ is basically the same as for $n = 3$.

Theorem 11. *Let $\mathcal{L} \subseteq \mathcal{L}^3$.*

1. *If there are only finitely many maximal 3-clusters $\mathcal{G}_1, \ldots, \mathcal{G}_n$ such that $(\bigcup_{i=1}^n \mathcal{G}_i) \subseteq (\mathcal{L}^3 \setminus \mathcal{L})$, for all $k \in \mathbb{N}$ $\{(k, a, b) : a, b \in \mathbb{N} \setminus \{k\}\} \cap \mathcal{L} \neq \emptyset$ and there are at most finitely many families of the form $\{\{k_i, m_i, a\} : a \in \mathbb{N} \setminus \{k_i, m_i\}\}$ such that $\bigcup_1^n \{\{k_i, m_i, a\} : a \in \mathbb{N} \setminus \{k_i, m_i\}\} \subseteq \mathcal{L}^3 \setminus \mathcal{L}$, then $\mathcal{L}$ has finitely many maximal pfi extensions.*
2. *If $\{\mathcal{G}_i\}_{i \in \mathbb{N}}$ is a countable sequence of maximal 3-clusters such that $\bigcup_{i=1}^\infty \mathcal{G}_i \subseteq (\mathcal{L}^2 \setminus \mathcal{L})$, or if for at least one $k \in \mathbb{N}$ we have that $\{(k, a, b) : a, b \in \mathbb{N} \setminus \{k\}\} \cap \mathcal{L} = \emptyset$, or if there are infinitely many families of the form $\{\{k_i, m_i, a\} : a \in \mathbb{N} \setminus \{k_i, m_i\}\}$ such that $\bigcup_1^\omega \{\{k_i, m_i, a\} : a \in \mathbb{N} \setminus \{k_i, m_i\}\} \subseteq \mathcal{L}^3 \setminus \mathcal{L}$, then $\mathcal{L}$ has continuously many maximal pfi extensions.*

Proof. 1. Let $\mathcal{G}_1, \ldots, \mathcal{G}_N$ be the finite sequence of all the maximal disjoint 3-clusters contained in $\mathcal{L}^3 \setminus \mathcal{L}$ and let $\mathcal{L} \subseteq \mathcal{L}^3$ as described for this case. We want to show that there are only finitely many maximal pfi extensions of $\mathcal{L}$. W.l.o.g. it is sufficient to show it for when there are two families $\{\{k_1, m_1, a\} : a \in \mathbb{N} \setminus \{k_1, m_1\}\}$, $\{\{k_2, m_2, a\} : a \in \mathbb{N} \setminus \{k_2, m_2\}\}$ such that $\{\{k_i, m_i, a\} : a \in \mathbb{N} \setminus \{k, m\}\} \subseteq \mathcal{L}^3 \setminus \mathcal{L}$ for $i \in \{1, 2\}$. The general case follows the same reasoning. First we will prove that we cannot add any n-tuple with $n \geq 5$ to $\mathcal{L}$, for this it suffices to see that we can only add

finitely many quadruples different from $\{k_1, m_1, k_2, m_2\}$. This is because if we could add an n-tuple for $n \geq 5$, then we could add any quadruple of elements in the tuple. Let us prove it then by contradiction, suppose there are infinitely many quadruples $\{a_i, b_i, c_i, d_i\} \neq \{k_1, m_1, k_2, m_2\}$ for $i \in \mathbb{N}$ which can extend $\mathcal{L}$ and preserve pfi. It is sufficient to check the case in which they differ from $\{k_1, m_1, k_2, m_2\}$ on one element only. W.l.o.g. suppose $a_i \neq k_1$ for every $i \in \mathbb{N}$ and so $\{a_i, b_i, c_i, d_i\} = \{a_i, m_1, k_2, m_2\}$. Then the triple $\{a_i, m_1, k_2\} \in \mathcal{L}$ or $\{a_i, m_1, k_2\} \notin \mathcal{L}$. The former cannot be the case since then $\{a_i, m_1, k_2, m_2\}$ cannot extend $\mathcal{L}$. So the latter is the case, then there is a 3-cluster $TRI(\{\{a_i, m_1, k_2\}, \{k_1, m_1, k_2,\}\}) \subseteq \mathcal{L}^3 \setminus \mathcal{L}$ so for every $i \in \mathbb{N}$, $TRI(\{\{a, m_1, k_2\}, \{k_1, m_1, k_2,\}\}) \subseteq \mathcal{G}_j$ for some $j \in \{1, \ldots, N\}$ which is a contradiction since from $NUM(\bigcup(\mathcal{G}_1, \ldots, \mathcal{G}_N))$ we can obtain only finitely many triples. Thus we can only add finitely many triples, i.e. finitely many quadruples, i.e finitely many n-tuples. We can only add finitely many pairs, namely formed by some elements in $NUM(\bigcup(\mathcal{G}_1, \ldots, \mathcal{G}_N)) \cup \{k_1, m_2, k_2, m_2\}$. We cannot add any singleton because of our initial assumption. As in the case of $\mathcal{L}^2$, for any maximal pfi extension $\mathcal{L}_m$ of $\mathcal{L}$ and any $A \in \mathcal{L}_m \setminus \mathcal{L}'$ we have that either $A \subseteq NUM(\bigcup(\mathcal{G}_1, \ldots, \mathcal{G}_N)) \cup \{k_1, m_2, k_2, m_2\}$ or $A = \{a, b, c\}$ for some $\{a, b, c\} \notin \mathcal{L}$.
The proof is by contradiction as for $\mathcal{L}^2$.

2. If $\{\mathcal{G}_i\}_{i \in \mathbb{N}}$ is a countable sequence of maximal 3-clusters such that $\bigcup_{i=1}^{\infty} \mathcal{G}_i \subseteq (\mathcal{L}^2 \setminus \mathcal{L})$, or if for at least one $k \in \mathbb{N}$ we have that $\{(k, a, b) : a, b \in \mathbb{N} \setminus \{k\}\} \cap \mathcal{L} = \emptyset$ the proof goes exactly as for $\mathcal{L}^2$. We will just prove the remaining case. There are infinitely many families of the form $\{\{k_i, m_i, a\} : a \in \mathbb{N} \setminus \{k_i, m_i\}\}$ such that $\bigcup_1^{\omega}\{\{k_i, m_i, a\} : a \in \mathbb{N} \setminus \{k_i, m_i\}\} \subseteq \mathcal{L}^3 \setminus \mathcal{L}$. Note that we can then have an infinite set of the form $\{k_1, m_1, k_2, m_2, \ldots, k_n, m_n, \ldots\}$. Note that there is no difference between k_i and m_i, but we distinguish them since they are paired together and this will be relevant for our proof.
For each quadruple $\{k_i, m_i, k_j, m_j\} \subseteq \{k_1, m_1, k_2, m_2, \ldots, k_n, m_n, \ldots\}$, we can maximally extend $\mathcal{L}$ with either the pairs $\{k_i, m_i\}$, $\{k_j, m_j\}$ or with $\{k_i, m_i, k_j, m_j\}$ itself.The rest of the pairs can be extended as
$\{k_l, m_l\} \subseteq \{k_1, m_1, k_2, m_2, \ldots, k_n, m_n, \ldots\}$. Since there are countably many quadruples of this form, by a straightforward combinatorial argument we obtain continuously many maximal extensions. Note that if the families $\{\{k_i, m_i, a\} : a \in \mathbb{N} \setminus \{k_i, m_i\}\}$ are disjoint or not does not matter for the argument of the proof. For the worst case scenario, suppose they all share the element $\{0\}$ i.e. $0 = k_i$ for every $i \in \mathbb{N}$. Then there will still be infinitely many m's which are different from each other. Therefore we will have that for every triple $\{0, m_1, m_2\}$ we can add either $\{0, m_1, m_2\}$ or $\{0, m_1\}$, $\{0, m_2\}$. Since we have countably many of these triples, by combinatorics we obtain continuously many maximal pfi extensions.

The following proposition generalizes what happens in example 2 for $n \geq 3$.

Proposition 10 *Let $n \geq 3$ and $\{a_0, a_1, \ldots, a_k\}$ a fixed k-tuple of elements in $\mathbb{N}$ for some $k \leq n - 2$.*

1. *If $k \leq n-3$, the family $\mathcal{L}^n \setminus \{\{a_0, a_1, \ldots, a_k, x_{k+1}, \ldots, x_{n-1}\} \in \mathcal{L}^n : x_i \in \mathbb{N} \setminus \{a_1, \ldots, a_k\}\}$ has continuously many maximal pfi extensions.*
2. *The family $\mathcal{L}^n \setminus \bigcup_{i=1}^{m} \{\{a_{i,0}, a_{i,1}, \ldots, a_{i,n-2}, b\} \in \mathcal{L}^m : b \in \mathbb{N} \setminus \{a_{i,0}, a_{i,1}, \ldots, a_{i,n-2}\}\}$ for finitely many $(n-2)$-tuples $\{a_{i,0}, a_{i,1}, \ldots, a_{i,n-2}\}$ in $\mathcal{L}^{n-2}$ (note that $k = n-2$) has finitely many maximal pfi extensions.*

The following is a straightforward generalization of Theorem 11.

Theorem 12. *Let $\mathcal{L} \subseteq \mathcal{L}^n$. In what follows we exhaust all the possible cases for subfamilies of $\mathcal{L}^n$:*

1. *If $\mathcal{L}$ satisfies the following,*
   - *there are at most finitely many n-clusters $\mathcal{G}_1, \ldots, \mathcal{G}_N$ such that $(\bigcup_{i=1}^{N} \mathcal{G}_i) \subseteq \mathcal{L}^n \setminus \mathcal{L}$, and*
   - *for at most finitely many tuples $\{a_0, a_1, \ldots, a_{n-2}\} \in \mathcal{L}^{n-2}$ is that $\{\{a_0, a_1, \ldots, a_{n-2}, x\} \in \mathcal{L}^n : x \in \mathbb{N} \setminus \{a_0, \ldots, a_{n-2}\}\} \cap \mathcal{L} = \emptyset$, and*
   - *for all $m \leq n-3$ and tuples $\{a_0, a_1, \ldots, a_m\} \in \mathcal{L}^m$ is that $\{\{a_0, a_1, \ldots, a_k, x_{m+1}, \ldots, x_{n-1}\} \in \mathcal{L}^n : x \in \mathbb{N} \setminus \{a_0, \ldots, a_{n-2}\}\} \cap \mathcal{L} \neq \emptyset$,*

   then $\mathcal{L}$ has finitely many maximal pfi extensions.
2. *If $\mathcal{L}$ satisfies one of the following cases:*
   - *there is an infinite sequence of n-clusters $\{\mathcal{G}_i\}_{i \in \mathbb{N}}$ such that $(\bigcup_{i=1}^{\infty} \mathcal{G}_i) \subseteq \mathcal{L}^n \setminus \mathcal{L}$, or*
   - *for infinitely many tuples $\{a_0, a_1, \ldots, a_{n-2}\} \in \mathcal{L}^{n-2}$ is that $\{\{a_0, a_1, \ldots, a_{n-2}, x\} \in \mathcal{L}^n : x \in \mathbb{N} \setminus \{a_0, \ldots, a_{n-2}\}\} \cap \mathcal{L} = \emptyset$, or*
   - *for some $m \leq n-3$ and some tuple $\{a_0, a_1, \ldots, a_m\} \in \mathcal{L}^m$ is that $\{\{a_0, a_1, \ldots, a_k, x_{m+1}, \ldots, x_{n-1}\} \in \mathcal{L}^n : x \in \mathbb{N} \setminus \{a_0, \ldots, a_{n-2}\}\} \cap \mathcal{L} = \emptyset$,*

   then $\mathcal{L}$ has continuously many maximal pfi extensions.

Therefore we obtain the following result.

Theorem 13. *Let $n \in \mathbb{N}$. Any subfamily $\mathcal{L}$ of the family of all n-tuples $\mathcal{L}^n$ has either finitely many maximal pfi extensions or continuously many.*

6 Infinite anti-chains

Contrary to the results of Section 3, *cfi* identification is more powerful on infinite anti-chains of infinite languages than *pfi* identification. This is exposed in the following result.

Proposition 1. *The family of all co-singletons, $\{\mathbb{N} \setminus \{i\} | i \in \mathbb{N}\}$, is an anti-chain which is* cfi *but not* pfi.

The family of co-singletons is not even *nepfi*, because no DFTT's exist.
The case of infinite anti-chains of finite languages is less clear. It is a trivial fact that canonical families which are anti-chains are always *pfi*. Families of finite sets which are not canonical already occur in [2] (a simple example is the family $\mathcal{L} = \{L_i : L_i = \{i\} \cup \{y : Tiiy\}\}$ where T is Kleene's T-predicate). By following a diagonalization strategy, we can construct a non-canonical family (but still indexable) of finite languages which is an anti-chain but is not *pfi*.

Theorem 14. *There is an indexed anti-chain $\mathcal{L}$ of finite languages for which there is no canonically indexed family $\{D_{f(n)} : n \in \omega\}$ such that $D_{f(i)} \subseteq L_i$ for all $i \in \mathbb{N}$ and $D_{f(i)} \nsubseteq L_j$ for all $j \neq i$, i.e. this anti-chain is not* pfi.

Proof. The strategy of the construction is by diagonalization. We diagonalize against all r.e. families of canonical finite sets, that is, families of the form $\{F_e : e \in B\}$ with B r.e. Note that this includes all the families of the form $\{F_{f(n)} : n \in \omega\}$ with f computable simply by definition of computable languages and r.e. languages. We will abuse our notation a bit and refer to the canonical families as $\{D_e : e \in B\}$ instead. Thus, we construct a uniformly computable family of finite languages $\{L_i : i \in \omega\}$ such that the following requirement is satisfied for each e:

$$\begin{aligned}(\mathrm{R}_e)\text{: If } &\{D_n : n \in W_e\} \text{ is an anti-chain}\\ &\text{then the following does not hold}\\ &D_i \subseteq L_i \text{ for all } i \in \mathbb{N}, \text{ and } D_i \nsubseteq L_j \text{ for all } j \neq i.\end{aligned}$$

where, as usual, W_e denotes the e-th r.e. set. The strategy for meeting the requirement (R_e) is as follows: First let $\mathcal{A}$ be a strictly increasing countable sequence of indices of the empty set. The procedure is different in case $e \in \mathcal{A}$ or not. If $e \notin \mathcal{A}$ then let $L_e = \{e\}$ until, if ever, we see a canonical code n with $D_n = \{e\}$ appear in W_e. In that stage s we change L_e to be $\{e, a_{2s}\}$ where a_{2s} is the $2s$-th element of $\mathcal{A}$. Moreover, we force $L_{a_{2s}}$ to be $\{a_{2s}, a_{2s+1}\}$ and $L_{a_{2s+1}}$ to be $\{a_{2s+1}, e\}$ in order to satisfy the anti-chain condition. We also ensure that $L_i = \{e\}$ only if $i = e$ and that for every stage $s' < s$, all $L_{a_{2s'}}, L_{a_{2s'+1}}$ have been already established for $i \leq a_{2s}$. We say that (R_e) *requires attention at stage s* if the requirement (R_e) threatens to be violated because there exists $n \in W_{e,s}$ such that $D_n = \{e\}$. In this case, we will modify L_e in the way described above such that $D_n \subset L_e$ and $D_n \subset L_{a_{2s+1}}$, for $a_{2s+1} \in \mathcal{A}$. Thus (R_e) is guaranteed since in later stages we will not change L_e again.

Now the explicit construction. Let $\mathcal{A} = \{a_k\}_{k\in\omega}$ be a strictly increasing sequence of indices of the empty set. At stage s, we have already determined whether $x \in L_i$ for every $i < a_{2s}$ and $x < a_{2s}$. Let $e < a_{2s}$ be the least number such that (R_e) requires attention. We put $e, a_{2s} \in L_e$, $a_{2s}, a_{2s+1} \in L_{a_{2s}}$ and $a_{2s+1}, e \in L_{a_{2s+1}}$ and nothing else. In this case we can be sure (R_e) will be satisfied. For all other $j < a_{2s}$ we put only $j \in L_j$. For all $s' < a_{2s}$, $L_{a_{2s'}}, L_{a_{2s'+1}}$ have already been established and for all $i < a_{2s}$ with $i \neq e$ we have $a_{2s}, a_{2s+1} \notin L_i$. If there is no $e < a_{2s}$ such that (R_e) requires attention then we do the latter for all $i < a_{2s+1}$. Indeed, we have now determined whether $x \in L_i$ for all $x, i \leq a_{2(s+1)}$. This construction defines a uniformly computable anti-chain of finite languages, since at every stage s, whether $x \in L_i$ is effectively determined for all $x, i \leq a_{2s+1}$ and every L_e is either $\{e\}$ or $\{e, a_{2s}\}$. Clearly it is an anti-chain because whenever L_e is a singleton, it has an empty intersection with the rest of the languages. And, whenever L_e is a pair, say $\{e, a_{2s}\}$, by construction there are only two other languages $L_{a_{2s}}, L_{a_{2s+1}}$ such that $L_e \cap L_{a_{2s}} \neq \emptyset \neq L_e \cap L_{a_{2s+1}}$, but these intersections are singletons.

To verify that our construction satisfies all requirements we follow a case-by-case procedure:
Case 1: If there is no $n \in \mathbb{N}$ such that $D_n = \{e\}$ and $n \in W_e$, then (R_e) will never require attention. Then $L_e = \{e\}$, but there exists no $n \in W_e$ such that $D_n = \{e\}$. Thus (1) is satisfied.
Case 2: If there is $n \in W_e$ for which (R_e) requires attention, (i.e, exists $n \in W_e$ such that $D_n = \{e\}$), then there are $n, s' \in \mathbb{N}$ such that $D_n = \{e\}$ and $n \in W_{e,s'}$. Note that there are only finitely many (R_i) with $i < e$ and each time (R_i) receives attention, it receives attention at most once (precisely for i). It follows that there is a stage $s > s'$ at which (R_e) receives attention. We then have $L_e = \{e, a_{2s}\}$, $D_n = \{e\}$ for some $n \in W_e$ but also $e \in L_{a_{2s+1}}$, $a_{2s} \in L_{a_{2s}}$. This ensures that (R_e) gets satisfied for L_e because $D_n = \{e\} \subseteq L_e$, $D_n = \{e\} \subseteq L_{a_{2s+1}}$ and $L_e \neq L_{a_{2s+1}}$. Thus we have that (R_e) is satisfied. Note that the only possible DFTTs for L_e are $D_n = \{e\}$ or L_e itself. But by construction L_e cannot be canonically represented, i.e. it cannot be a DFTT (by definition of DFTT). Therefore $\mathcal{L}$ is an anti-chain of finite languages which is not pfi. $\square$

This particular example happens to be not *cfi* either. The question remains open whether there exists such a family which is *cfi* and not *pfi*.

Bibliography

[1] Dana Angluin. Inductive inference of formal languages from positive data. *Information and control*, 45(2):117–135, 1980.

[2] Nina Gierasimczuk and Dick de Jongh. On the complexity of conclusive update. *The Computer Journal*, pages 56(3):365–377, 2012.

[3] Mark E Gold. Language identification in the limit. *Information and control*, 10(5):447–474, 1967.

[4] Sarah Hiller and Raquel Fernández. A data-driven investigation of corrective feedback on subject omission errors in first language acquisition. *Proceedings of the 20th SIGNLL, Conference on Computational Natural Language Learning (CoNLL)*, pages 105–114, 2016.

[5] Steffen Lange and Thomas Zeugmann. Set-driven and rearrangement-independent learning of recursive languages. *Theory of Computing Systems*, 29(6):599–634, 1996.

[6] Ruslan Mitkov. *The Oxford handbook of computational linguistics*. Oxford University Press, 2005.

[7] Yasuhito Mukouchi. Characterization of finite identification. In *Analogical and Inductive Inference*, pages 260–267. Springer, 1992.

[8] Daniel N Osherson, Michael Stob, and Scott Weinstein. *Systems that learn: An introduction to learning theory for cognitive and computer scientists*. The MIT Press, 1986.

[9] Matthew Saxton, Phillip Backley, and Clare Gallaway. Negative input for grammatical errors: Effects after a lag of 12 weeks. *Journal of child language*, 32(3):643–672, 2005.

[10] Robert I Soare. *Recursively enumerable sets and degrees: A study of computable functions and computably generated sets*. Springer Science & Business Media, 1999.