LEARNING LANGUAGES IN A UNION

NG YEN KAOW

(B.Sc.(Hons), NUS)

A THESIS SUBMITTED

FOR THE DEGREE OF MASTER OF SCIENCE

SCHOOL OF COMPUTING

NATIONAL UNIVERSITY OF SINGAPORE

2001

Acknowledgements

I need to thank Prof. Jain for not asking of me the same standard which he requires of himself; or otherwise I am sure to not be able to complete this thesis. His contributions to proofs and fixes for proofs (, as well as tolerance in proof-reading this) goes without saying. I owe it to Prof. Tay Tiong Seng for teaching me on how to do geometry of higher dimensions, and for his proof ideas on most (if not all) of the geometrical claims. I am indebted to Prof. Phil Long, Prof. Frank Stephan and Mr. Ono Hirotaka for ideas and suggestions. I also wish to thank Francis Ng, Foo Hee Meng and Prof. Leong Hon Wai for general encouragement. Whatever little contribution to this thesis that is truly mine I dedicate to my family.

Summary

In learning theory for learning languages, a machine is given words in a language and the machine is said to identify the language if it correctly names the language. In this thesis we study classes of languages where the unions of up to a fixed number (say n) of languages from the class are identifiable. We distinguish between two different scenarios: in one scenario, the learner need only to name the language which results from the union; in the other, the learner must individually name the languages which make up the union (we say that the unioned language is discerningly identified). We define three kinds of identification criteria based on this and by the use of some naturally occurring classes of languages, demonstrate that the inferring power of each of these identification criterion decreases as we increase the number of languages allowed in the union, thus resulting in an infinite hierarchy for each identification criterion. A comparison between the different identification criteria also yielded similar hierarchies. We define generalized versions of the identification criteria, and show that the hierarchies continue to hold for these generalized criteria. We show that for each n, there exists a class of disjoint languages where all unions of up to n languages from this class can be discerningly identified, but there is no learner which identifies every union of $n + 1$ languages from this class. We give sufficient conditions for classes of languages where the unions can be discerningly identified. We also present language classes which are complete with respect to weak reduction (in terms of intrinsic complexity) for our identification criteria.

Contents

1 Introduction

What exactly is "*learning*"? When can we say that a person "*has learnt something*"? To illustrate, consider a person who tries to learn the game of poker by watching others play. The person can very well play the game intelligently by simply remembering every hand ever played, with or without understanding the rules of the game. It would be difficult to argue that this person has not learnt poker, especially if this person can consistently beat you. In the spirit of the Turing Test, learning should be considered in the most straight-forwardly measurable manner.

In the most general form, learning is simply a process where a person attempts to provide explanations to an observed phenomenon. This does not say anything about what constitutes successful learning. Much of what we typically mean by "learning" only comes into the picture when we consider the criteria for saying "when" a person has learnt something. In the following text we shall first describe a mathematical framework to simulate the learning process (consisting of a learning environment and a learner), and then give examples of learning criteria.

We assume that the information obtainable by a learner can be encoded into the natural numbers. The justification is that at any one time, we can only record a finite number of datum from nature[1], and each of these only to a finite precision. For this reason, we use formal languages as abstract representation for observable phenomenon. From here onwards, we shall refer to languages and phenomena

[1]Note that the number of data may go to infinity as time approaches infinity.

interchangeably.

Likewise, a learning environment in this regard, can be abstracted to an infinite sequence of (all or part of, and not necessarily limited to) natural numbers for a given language (or phenomena). For realism, the numbers in the sequence are usually allowed to repeat. A common practice is to also allow the sequence to contain a special symbol representing "*null*". This becomes useful for phenomena that are represented by the empty set, where it can then be represented by a sequence which consists only of "null" symbols.

We now turn to the learner. In the fashion of the Church-Turing thesis, we assume that learning is a computable process, and for this reason the learner is a computable device, called an *inductive inference machine* in the literature. Thus we can consider a learner as an algorithmic device which gets as input an infinite sequence of natural numbers. For each finite initial portion of the sequence, the device may output a conjecture, possibly a grammar for the language which the sequence is for. This results in a sequence of conjectures, which we may then use to determine if the learner has successfully learnt the input sequence.

Note that we have not yet defined the following requirement:

1. What is the correspondence between the phenomena and their representation sequences?

   For instance, there may be assumptions in the data that the learner receives. Perhaps the data is presented in an ascending fashion, or perhaps there are inaccuracies in the data.

2. What is the degree of satisfaction required of the machine's answer?

   There may be imposed restrictions on the behaviour of the learner. Perhaps we restrict the learner to only make guesses that are consistent with the input, or perhaps we require the learner to make a correct guess within a fixed number of tries. On the other hand, we may also relax our requirement, for example allowing the learner to make mistakes within a range that we consider tolerable.

Gold [Gol67] introduced the notion of *identification in the limit.* In his original work, a learner is given, one by one, with or without repetition, all the numbers in a language, as well as the special symbol "null". The learner is said to identify the language *in the limit* if and only if regardless of the order and frequency these elements are presented, at some time the learner outputs a grammar for that language and then never changes its mind. The learner is said to *identify a class of languages* if and only if it identifies every language in that class. This criterion is now referred to as **TxtEx**-identification [CL82], and has been the basis for many other learning criteria.

Various variations of Gold's learning criterion have been considered. For example, Blum and Blum [BB75] studied more lenient conditions for a learner to identify a function, and Bārzdiņš [B$\bar{7}$4] imposed restrictions on the conjectures that a learner is allowed to make. In all these studies, a key motivation is to compare the *inferring power* of learners under the different criteria. The inferring power of a learner on a given kind of input and under a given criterion is, roughly speaking, the class of languages on which the learner meets the requirements of the criterion.

The majority of the learning criteria in these studies assume that the input given to the learner consists of elements from a single language. In reality, it is not unusual for learners to be presented with information that is some sort of mixture. For example, children growing up in a multi-lingual environment may be exposed to more than one (natural) language without being explicitly told their origin; or, in a physical experiment, radiations collected by the same detector may originate from many different source processes.

In this dissertation we investigate learning criteria which require a learner to explain the observed phenomena, even when they are presented as a mixture of several phenomena. In our terminology, the learner must be able to identify unions of languages. This adds new complications to the definitions of the learning criteria. For instance, how many languages should be allowed in a union? Does a learner have to name all of the languages included in the union?

In Chapter 2 we give the necessary background in computational learning theory, including definitions for a few well known learning criteria. Readers already acquainted with the theory may skip this chapter.

In Chapter 3 we formalize a few notions of “identifying unions of languages”. This is followed by a survey of related works in computational learning theory. We then provide some languages which fulfills our learning criteria in Chapter 4, and show that under the new learning criteria, inferring power of learners in general lessen when more languages are allowed in the unions.

In Chapter 5 we attempt to find conditions which are sufficient for learning unions of languages. In Chapter 6, we extend well known identification criteria to accommodate learning of unions of languages, and examine how the inferring power of learners is affected under each of these extensions.

Freivalds, Kinber, and Smith [FKS95] proposed a new method for measuring the complexity of learning called *intrinsic complexity*. Since its' conception, the idea has been studied actively in the field of computational learning theory [JS96, KPSW99, JKW00, JK01]. In Chapter 7, we analyse the intrinsic complexity of a few classes of languages with respect to the learning criteria defined in Chapter 4.

2 Notation and Preliminaries

2.1 Notation

Any unexplained recursion-theoretic notation is from [Rog67]. N denotes the set of natural numbers, N_a denotes the set $\{i \in N \mid i \leq a\}$. N^+ denotes the set of positive integers. S, with or without decorations, ranges over subsets of N. *rat* denotes the set of non-negative rational numbers. $\mathbb{R}$ denotes the set of real numbers. For this thesis, functions have their domain and range in N. f and g, with or without decorations, range over total functions. η, with or without decorations, ranges over partial functions.

$\emptyset$, $\in$, $\subset$, $\subseteq$, $\supset$, $\supseteq$ respectively denote empty set, element of, proper subset, subset, proper superset, superset. $\max(.)$, $\min(.)$ denote maximum and minimum of a set, where by convention $\max(\emptyset) = 0$ and $\min(\emptyset) = \infty$. Cardinality of a set S is denoted by $card(S)$. $D_0, D_1, \ldots$ stand for a computable sequence of all finite sets [Rog67]. $A - B$ denotes the set $\{x \mid x \in A \text{ and } x \notin B\}$. $A \triangle B$ denotes the symmetric difference of A and B, that is, $(A - B) \cup (B - A)$. For any two functions η_1 and η_2, $\eta_1 =^n \eta_2$ means that $card(\{x \mid \eta_1(x) \neq \eta_2(x)\}) \leq n$; η_1 and η_2 are called *n-variants*. $\eta_1 =^* \eta_2$ means that $card(\{x \mid \eta_1(x) \neq \eta_2(x)\})$ is finite; η_1 and η_2 are called *finite-variants*. For any two sets S_1 and S_2, $S_1 =^n S_2$ means $card(S_1 \triangle S_2) \leq n$; S_1 and S_2 are called *n-variants*. $S_1 =^* S_2$ denotes $card(S_1 \triangle S_2)$ is finite; S_1 and S_2 are called *finite-variants*.

$\langle \cdot, \cdot \rangle$ stands for an arbitrary, computable bijective mapping from $N \times N$ onto

N. For all x and y, $\pi_1(\langle x, y\rangle) = x$ and $\pi_2(\langle x, y\rangle) = y$. We assume without loss of generality that $\langle\cdot,\cdot\rangle$ is monotonically increasing in both of its arguments. $\langle\cdot,\cdot\rangle$ can be extended to n-tuples in a natural way (including $n = 1$, where $\langle x\rangle$ may be taken to be x). Projection functions $\pi_1, \ldots, \pi_n$ corresponding to n-tuples can be defined similarly (where the tuple size would be clear from context). Due to the above isomorphism between N^n and N, we often identify the tuple $(x_1, \ldots, x_n)$ with $\langle x_1, \ldots, x_n\rangle$.

The quantifiers $\overset{\infty}{\forall}$, $\overset{\infty}{\exists}$ and $\exists!$ denote, for all but finitely many, there exists infinitely many and there exists a unique, respectively.

A computable numbering is a partial computable function from N^2 to N. The symbol ψ ranges over computable numberings. We denote by ψ_i, the partial function, $\lambda x.\psi(i, x)$. Thus ψ_i denotes the partial function computed by the program with index i in the numbering ψ. Ψ denotes an arbitrary Blum complexity measure for ψ. W_i^ψ denotes $domain(\psi_i)$. W_i^ψ is, then, the r.e. set/language ($\subseteq N$) accepted (or equivalently, generated) by the ψ-program i. We also say that i is a ψ-grammar for W_i^ψ. $W_{i,s}^\psi$ denotes the set $\{x \leq s \mid \Psi_i(x) \leq s\}$. We say that numbering ψ is reducible to numbering ψ' (written $\psi \prec \psi'$) if and only if there exists a recursive function h such that for each $i \in N$, $\psi_i = \psi'_{h(i)}$. In this case we say that h witnesses that $\psi \prec \psi'$. An acceptable numbering is a computable numbering to which every computable numbering can be reduced. The symbol φ denotes a standard acceptable numbering [Rog67] and the symbol Φ denotes an arbitrary fixed Blum complexity measure for the φ-system [Blu67]. In this thesis we abbreviate W_i^φ to W_i, and $W_{i,s}^\varphi$ to $W_{i,s}$.

$\mathcal{E}$ denotes the class of all r.e. languages. $\mathcal{R}$ denotes the set of all recursive functions, that is total computable functions. Symbol L, with or without decorations, ranges over $\mathcal{E}$. The symbol $\mathcal{L}$, with or without decorations, ranges over subsets of $\mathcal{E}$.

$\mathcal{K}$ denotes the diagonal halting problem set, that is, $\mathcal{K} = \{x \mid x \in W_x\}$. ($\mathcal{K}$ is a recursively enumerable, non-recursive set.)

We often use the following classes:

> *SINGLE* denotes the class $\{\{x\} \mid x \in N\}$.
>
> *FIN* denotes the class $\{D \subset N \mid D \text{ is finite}\}$.
>
> *INIT* denotes the class $\{\{x \mid x \leq a\} \mid a \in N\}$.

A class $\mathcal{L}$ of r.e. languages is said to be *recursively enumerable* if there is $S \in \mathcal{E}$ such that $\mathcal{L} = \{W_i \mid i \in S\}$; in this case S is said to be an r.e. index set for $\mathcal{L}$ [Rog67]. For each non-empty, recursively enumerable class of languages $\mathcal{L}$, there exists a total recursive function f such that $\mathcal{L} = \{W_{f(i)} \mid i \in N\}$.

$\mathcal{L}$ is said to be 1–1 recursively enumerable if and only if (i) $\mathcal{L}$ is finite or (ii) there exists a recursive function f such that $\mathcal{L} = \{W_{f(i)} \mid i \in N\}$ and $W_{f(i)} \neq W_{f(j)}$, if $i \neq j$. In this latter case we say that $W_{f(0)}, W_{f(1)}, \ldots$ is a 1–1 recursive enumeration of $\mathcal{L}$.[2]

A partial function d from N to N is said to be partial limiting recursive, if and only if there exists a recursive function $\mathbf{F}$ from $N \times N$ to N such that for all

[2] A 1–1 recursively enumerable class of languages is sometimes also called *r.e. indexable without repetition.*

x, $d(x) = \lim_{y\to\infty} \mathbf{F}(x,y)$. Here if $d(x)$ is not defined then $\lim_{y\to\infty} \mathbf{F}(x,y)$ must also be undefined. A partial limiting recursive function d is called (total) limiting recursive, if d is total. $\downarrow$ denotes defined or converges. $\uparrow$ denotes undefined or diverges.

2.2 Preliminaries

2.2.1 Language Identification

We now present concepts from language learning theory. The next definition introduces the concept of a *sequence* of data.

Definition 1 [Gol67]

(a) A *sequence* σ is a mapping from an initial segment of N into $(N \cup \{\#\})$. The empty sequence is denoted by λ.

(b) The *content* of a sequence σ, denoted *content*(σ), is the set of natural numbers in the range of σ.

(c) The *length* of σ, denoted by $|\sigma|$, is the number of elements in σ. So, $|\lambda| = 0$.

(d) For $n \leq |\sigma|$, the initial sequence of σ of length n is denoted by $\sigma[n]$. So, $\sigma[0]$ is λ.

Intuitively, #'s represent pauses in the presentation of data. We let σ, τ, and γ, with or without decorations, range over finite sequences. SEQ denotes the set of all finite sequences.

Definition 2 [Gol67]

(a) A *text* T for a language L is a mapping from N into $(N \cup \{\#\})$ such that L is the set of natural numbers in the range of T.

(b) The *content* of a text T, denoted by *content*(T), is the set of natural numbers in the range of T; that is, the language which T is a text for.

(c) $T[n]$ denotes the finite initial sequence of T with length n.

Unless stated otherwise, we let T, with or without decorations, range over texts. $\mathbf{T}$ denotes the set of all texts.

Definition 3 [Gol67] Let $\sigma, \tau \in$ SEQ be given.

(a) The result of concatenating τ onto the end of σ is denoted by $\sigma \diamond \tau$. Sometimes we abuse notation slightly and write $\sigma \diamond x$ (where $x \in N \cup \{\#\}$) to denote the sequence formed by adding x at the end of σ.

(b) We write "$\sigma \subseteq \tau$" if σ is an initial segment of τ, and "$\sigma \subset \tau$" if σ is a proper initial segment of τ (we also say that τ is an *extension* of σ). Likewise, we write $\sigma \subset T$ if σ is an initial sequence of text T.

(c) Let finite sequences $\sigma^0, \sigma^1, \sigma^2, \ldots$ be given such that $\sigma^0 \subset \sigma^1 \subset \sigma^2 \subset \ldots$ and $\lim_{i \to \infty} |\sigma^i| = \infty$. Then there is a unique text T such that for all $n \in N$, $\sigma^n = T[|\sigma^n|]$. This text is denoted $\bigcup_n \sigma^n$.

Definition 4 [Gol67] An *inductive inference machine* (**IIM**) is an algorithmic device which computes a mapping from SEQ into N.

We let M, with or without decorations, range over the **IIM**s. $M(T[n])$ is interpreted as the grammar (index for an accepting program) conjectured by the machine M on the initial sequence $T[n]$. We say that M converges on T to i (written $M(T)\downarrow = i$) if for all but finitely many n, $M(T[n]) = i$.

Gold [Gol67] introduced the following language learning criterion known as **TxtEx**-identification.

Definition 5 [Gol67]

(a) M **TxtEx**-identifies a text T just in case there exists $i \in N$ such that $W_i = content(T)$, and $M(T)\downarrow = i$.

(b) M **TxtEx**-identifies an r.e. language L (written $L \in \mathbf{TxtEx}(M)$) just in case M **TxtEx**-identifies each text for L.

(c) M **TxtEx**-identifies a class $\mathcal{L}$ of r.e. languages (written $\mathcal{L} \subseteq \mathbf{TxtEx}(M)$) just in case M **TxtEx**-identifies each language from $\mathcal{L}$.

(d) $\mathbf{TxtEx} = \{\mathcal{L} \subseteq \mathcal{E} \mid (\exists M)[\mathcal{L} \subseteq \mathbf{TxtEx}(M)]\}$.

This is easily generalized to the following, called *anomalous* ***TxtEx**-identification.*

Definition 6 [OW82] Let $a \in N \cup \{*\}$.

(a) M $\mathbf{TxtEx}^a$-identifies a text T just in case there exists $i \in N$ such that $W_i =^a content(T)$, and $M(T)\downarrow = i$.

(b) M $\mathbf{TxtEx}^a$-identifies an r.e. language L (written $L \in \mathbf{TxtEx}^a(M)$) just in case M $\mathbf{TxtEx}^a$-identifies each text for L.

(c) M $\mathbf{TxtEx}^a$-identifies a class $\mathcal{L}$ of r.e. languages (written $\mathcal{L} \subseteq \mathbf{TxtEx}^a(M)$)

just in case M $\mathbf{TxtEx}^a$-identifies each language from $\mathcal{L}$.

(d) $\mathbf{TxtEx}^a = \{\mathcal{L} \subseteq \mathcal{E} \mid (\exists M)[\mathcal{L} \subseteq \mathbf{TxtEx}^a(M)]\}$.

It is clear that $\mathbf{TxtEx}^0$ is simply $\mathbf{TxtEx}$. We now introduce some technical notions which are useful in the study of learning capabilities of the **IIM**s.

Definition 7 [BB75] Let $a \in N \cup \{*\}$. σ is a ***TxtEx***^{a}*-locking sequence* for M on L if and only if

(a) $content(\sigma) \subseteq L$,

(b) $W_{M(\sigma)} =^a L$, and

(c) For all extensions τ of σ, if $content(\tau) \subseteq L$, then $M(\tau) = M(\sigma)$.

We often refer to $\mathbf{TxtEx}^a$-locking sequences as simply locking sequences (a will be clear from context). We now present a very important lemma in learning theory due to L. and M. Blum.

Lemma 1 *[BB75] Let $a \in N \cup \{*\}$. If M **TxtEx**^{a}-identifies L, then there is a **TxtEx**^{a}-locking sequence for M on L.*

Osherson and Weinstein [OW82] introduced another infinite hierarchy of identification criterion called **TxtBc**-identification which we describe below. "Bc" stands for *behaviourally correct*. Case and Lynes [CL82] independently introduced a similar notion.

Definition 8 [OW82, CL82] Let $a \in N \cup \{*\}$.

(a) M $\mathbf{TxtBc}^a$-identifies an r.e. language L (written $L \in \mathbf{TxtBc}^a(M)$) just in case for all texts T for L, for all but finitely many n, $W_{M(T[n])} =^a L$.

(b) $\mathbf{TxtBc}^a = \{\mathcal{L} \subseteq \mathcal{E} \mid (\exists M)[\mathcal{L} \subseteq \mathbf{TxtBc}^a(M)]\}$.

(c) $\mathbf{TxtBc}^0$ is abbreviated to **TxtBc**.

Thus a learner M $\mathbf{TxtBc}^a$-identifies language L just in case M, fed any text for L, produces an infinite sequence of hypotheses, all but finitely many of which are for a-variants of L. It is not required of a successful $\mathbf{TxtBc}^a$ learner to converge to a single index. For $a \geq 1$, the successive conjectures of the successful $\mathbf{TxtBc}^a$ learner need not even be for the same language. A stricter criterion would allow the learner to only vacillate around some finite number of conjectures. Case [Cas88] captured this notion through the following language learning criterion.

Definition 9 [Cas88] Let $a \in N \cup \{*\}$ and $b \in N^+ \cup \{*\}$.

(a) We say that M on T *finitely converges to* a finite set D just in case (i) for all but finitely many n, $M(T[n]) \in D$, and (ii) for all $i \in D$, there exists infinitely many n such that $M(T[n]) = i$.

(b) M $\mathbf{TxtFex}_b^a$-identifies L (written $L \in \mathbf{TxtFex}_b^a(M)$) just in case for all texts T for L there is a finite, non-empty set D of cardinality at most b such that M finitely converges to D and, for each $i \in D$, $W_i =^a L$.

(c) $\mathbf{TxtFex}_b^a = \{\mathcal{L} \subseteq \mathcal{E} \mid (\exists M)[\mathcal{L} \subseteq \mathbf{TxtFex}_b^a(M)]\}$.

Thus, M $\mathbf{TxtFex}_b^a$-identifies L just in case for every text T for L there is a non-empty, finite set D_T of no more than b indices for a-variants of L such that all but finitely many of M's conjectures are drawn from D_T.

Lemma 1 has an analogue for **TxtFex** and **TxtBc** learning also.

We now state some of the basic results relating the different types of identification criteria introduced above.

Theorem 2 *[Cas88, CL82] For all $i, n \in N$.*

(a) $\boldsymbol{TxtEx}^{n+1} - \boldsymbol{TxtFex}_*^n \neq \emptyset$.

(b) $\boldsymbol{TxtEx}^{2n+1} - \boldsymbol{TxtBc}^n \neq \emptyset$.

(c) $\boldsymbol{TxtEx}^{2n} \subset \boldsymbol{TxtBc}^n$.

(d) $\boldsymbol{TxtFex}_{i+1}^0 - \boldsymbol{TxtFex}_i^* \neq \emptyset$.

(e) $\bigcup_{n \in N} \boldsymbol{TxtFex}_i^n \subset \boldsymbol{TxtFex}_i^*$.

(f) $\bigcup_{n \in N} \boldsymbol{TxtBc}^n \subset \boldsymbol{TxtBc}^*$.

Parts (a), (d) and (e) are due to Case [Cas88, Cas99]. Parts (b) and (c) are due to Case and Lynes [CL82]. Part (f) is due to Case and Smith [CS83].

Data presented to learners may suffer from inaccuracies. In [FJ89, FJ96] Fulk and Jain considered three kinds of possible corruption to texts. They are, (a) intrusions of erroneous data, (b) omissions of data, and (c) both. Affected texts are called *"noisy"*, *"incomplete"*, and *"imperfect"* respectively.

Definition 10 [FJ96] Let $L \in \mathcal{E}$ and $a \in (N \cup \{*\})$ be given.

(a) A text T is *a-noisy* for L just in case $L \subseteq content(T)$ and $card(content(T) - L) \leq a$.

(b) A text T is *a-incomplete* for L just in case $content(T) \subseteq L$ and $card(L - content(T)) \leq a$.

(c) A text T is *a-imperfect* for L just in case $card(L \triangle content(T)) \leq a$.

Definition 11 [FJ96] Let $a, b \in (N \cup \{*\})$ be given.

(a.1) M $\mathbf{N}^a\mathbf{TxtEx}^b$-identifies $L \in \mathcal{E}$ (written $L \in \mathbf{N}^a\mathbf{TxtEx}^b(M)$) just in case for all a-noisy texts T for L, $M(T)\downarrow$ and $W_{M(T)} =^b L$.

(a.2) $\mathbf{N}^a\mathbf{TxtEx}^b = \{\mathcal{L} \mid (\exists M)[\mathcal{L} \subseteq \mathbf{N}^a\mathbf{TxtEx}^b(M)]\}$

(b.1) M $\mathbf{In}^a\mathbf{TxtEx}^b$-identifies $L \in \mathcal{E}$ (written $L \in \mathbf{In}^a\mathbf{TxtEx}^b(M)$) just in case for all a-incomplete texts T for L, $M(T)\downarrow$ and $W_{M(T)} =^b L$.

(b.2) $\mathbf{In}^a\mathbf{TxtEx}^b = \{\mathcal{L} \mid (\exists M)[\mathcal{L} \subseteq \mathbf{In}^a\mathbf{TxtEx}^b(M)]\}$

(c.1) M $\mathbf{Im}^a\mathbf{TxtEx}^b$-identifies $L \in \mathcal{E}$ (written $L \in \mathbf{Im}^a\mathbf{TxtEx}^b(M)$) just in case for all a-imperfect texts T for L, $M(T)\downarrow$ and $W_{M(T)} =^b L$.

(c.2) $\mathbf{Im}^a\mathbf{TxtEx}^b = \{\mathcal{L} \mid (\exists M)[\mathcal{L} \subseteq \mathbf{Im}^a\mathbf{TxtEx}^b(M)]\}$

Using the same principle we can define the following identification criteria: $\mathbf{N}^a\mathbf{TxtFex}^b_c$, $\mathbf{In}^a\mathbf{TxtFex}^b_c$, $\mathbf{Im}^a\mathbf{TxtFex}^b_c$, $\mathbf{N}^a\mathbf{TxtBc}^b$, $\mathbf{In}^a\mathbf{TxtBc}^b$, $\mathbf{Im}^a\mathbf{TxtBc}^b$.

Theorem 3 *[FJ96] For all $i \in N$.*

(a) $\boldsymbol{Im}^*\boldsymbol{TxtEx}^{i+1} - \boldsymbol{TxtFex}^i_* \neq \emptyset$.

(b) $\boldsymbol{Im}^*\boldsymbol{TxtEx}^* - \bigcup_{j\in N} \boldsymbol{TxtFex}^j_* \neq \emptyset$.

(c) $\boldsymbol{Im}^*\boldsymbol{TxtFex}_{i+1} - \boldsymbol{TxtFex}^*_i \neq \emptyset$.

(d) $\boldsymbol{Im}^*\boldsymbol{TxtEx}^{2i+1} - \boldsymbol{TxtBc}^i \neq \emptyset$.

(e) $\boldsymbol{Im}^*\boldsymbol{TxtBc} - \boldsymbol{TxtFex}^*_* \neq \emptyset$.

(f) $\boldsymbol{Im}^*\boldsymbol{TxtBc}^{i+1} - \boldsymbol{TxtBc}^i \neq \emptyset$.

(g) $\boldsymbol{Im}^i\boldsymbol{TxtEx} - [\boldsymbol{N}^{i+1}\boldsymbol{TxtBc}^* \cup \boldsymbol{In}^{i+1}\boldsymbol{TxtBc}^*] \neq \emptyset$.

2.2.2 Learning From Informants

What if texts presented to a learner contains not only the information of whether a word is in a language, but also information of whether a word is not in a language? In learning by informant, data presented to the learner is a characteristic function for a language.

Definition 12 [Gol67, BB75, CL82] We say that a text G for a 0–1 valued function is an informant for L just in case $content(G) = \{(x, 1) \mid x \in L\} \cup \{(x, 0) \mid x \notin L\}$.

Definition 13 [Gol67, BB75, CL82] Let $a \in N \cup \{*\}$.

(a) M $\mathbf{InfEx}^a$-identifies L (written $L \in \mathbf{InfEx}^a(M)$) just in case on all informants G for L, $M(G)\downarrow$ and $W_{M(G)} =^a L$.

(b) $\mathbf{InfEx}^a = \{\mathcal{L} \mid (\exists M)[\mathcal{L} \subseteq \mathbf{InfEx}^a(M)]\}$.

Proposition 4 *[Gol67]* $\boldsymbol{InfEx} - \boldsymbol{TxtEx} \neq \emptyset$.

2.2.3 Standard IIM Enumeration

Unless stated otherwise, we let $M_0, M_1, \ldots$ denote a recursive sequence of total **IIM**s, such that every class of languages in every identification criteria introduced in this chapter is identified by at least one of the machines in the sequence [OSW86].

3 Identification of Unions of Languages

3.1 Introduction

We start our analysis of the identification of unions of languages in this chapter.

In the past, the studies of identifying unions of languages were largely motivated by problems in pattern languages [Wri89, SA00, GK99]. These studies are restricted in the sense that the propositions are devised only to answer the specific problems in pattern languages. Since the pattern languages are recursive, their results generally apply only to languages that are recursive.

We claim that in reality, it is not unusual for learners to be presented with a mixture of information consisting of data from many different languages. One argument is that, in a multi-lingual environment, children may very well be exposed to more than one (natural) language at one time; or, in a physical experiment, radiations collected by the same detector may come from many different source processes. We therefore consider it worthwhile to investigate the problem of identifying unions of languages beyond the context of the pattern languages.

To arrive at a general definition of "identification of unions of languages", we need to note first that a few interpretations are possible for what *"identification"* may mean for a language which is the union of other languages. For instance, it could mean *"identifying the unioned language"*, or more demandingly, *"identifying the languages which made up the unioned language"*. In the latter case, suppose there exists more than one possible sets of languages which combines to give the unioned language, should the learner be required to supply all the possibilities, or just one of them, or should such a scenario be considered unidentifiable? In this chapter we will choose a few definitions to work with.

3.2 Defining The Unions of Languages

We now give formal definitions for notions of "unions of languages".

Definition 14 [SA00] Let $k \in N^+$ and $\mathcal{L} \subseteq \mathcal{E}$.

(a) Define *the union language of* $\mathcal{L}$, $L_{\mathcal{L}} = \bigcup_{L \in \mathcal{L}} L$.

(b) Define *the class of at most k unions of* $\mathcal{L}$, $\mathcal{L}^k = \{\, L_{\mathcal{L}'} \mid \mathcal{L}' \subseteq \mathcal{L} \wedge card(\mathcal{L}') \leq k\}$.

3.3 The Paradigm UTxtEx

We define our first notion for "learning the unions of languages" with the following identification criterion.

Definition 15 Let $k \in N^+$ and $\mathcal{L} \subseteq \mathcal{E}$.

(a) M $\mathbf{U}^k\mathbf{TxtEx}$-identifies[3] $\mathcal{L}$ just in case $\mathcal{L}^k \subseteq \mathbf{TxtEx}(M)$.

(b) $\mathbf{U}^k\mathbf{TxtEx} = \{\mathcal{L} \subseteq \mathcal{E} \mid (\exists M)[M\ \mathbf{U}^k\mathbf{TxtEx}\text{-identifies } \mathcal{L}]\}$.

UTxtEx coincides with the definition of "identification of unions of languages" in [Wri89, SA00].

[3]The **U** in **UTxtEx** stands for *Unioned.*

3.3.1 Known Results in UTxtEx Identification

Definition 16 [Ang80b]

A collection of non-empty languages $\mathcal{L} = \{L_i \mid i \in N\}$ is *an indexed family of recursive languages*[4] just in case there exists a computable function f such that for each $i \in N$ and for each $x \in N$,

$$f(i,x) = \begin{cases} 1 & \text{if } x \in L_i \\ 0 & \text{otherwise} \end{cases}$$

The following concept of *finite elasticity*, originally (incorrectly) defined in [Wri89], is from [MSW91].

Definition 17 [Wri89, MSW91] A collection of languages $\mathcal{L}$ has infinite elasticity just in case there exists an infinite sequence of pairwise distinct numbers, $\{w_i \in N \mid i \in N\}$, and an infinite sequence of pairwise distinct languages, $\{A_i \in \mathcal{L} \mid i \in N\}$, such that for each $k \in N$, $\{w_i \mid i < k\} \subseteq A_k$, but $w_k \notin A_k$. $\mathcal{L}$ is said to have finite elasticity just in case $\mathcal{L}$ does not have infinite elasticity.

Theorem 5 *[Wri89] Let $\mathcal{L}$ be an indexed family of recursive languages. If $\mathcal{L}$ has finite elasticity, then for all n, $\mathcal{L} \in$ **U^n TxtEx**.*

Corollary 6 *[Wri89] Let $\mathcal{L}$ be the class of all pattern languages, then for all n, $\mathcal{L} \in$ **U^n TxtEx**.*

---

[4]In [dJK96], de Jongh et. al. generalized the term *indexed family* to refer to any infinite, recursively enumerable class of r.e. languages.

Shinohara and Arimura in [SA00] noted that this result does not apply to the class $\mathcal{L}^*$ of unions of unbounded number of languages, and showed that the class of unbounded unions of pattern languages is not **TxtEx**-identifiable.

Definition 18 [Ang80a] A collection of languages $\mathcal{L}$ has *finite thickness* just in case for each $n \in N$, $card(\{L \in \mathcal{L} \mid n \in L\})$ is finite.

The following is a concept frequently rediscovered in the literature.

Definition 19 [Kru72] Let $\geqq$ be a preorder over a set A. An *infinite anti-chain over A with respect to* $\geqq$ is an infinite sequence $a_0, a_1, \ldots, a_i, \ldots$ such that $a_i \in A$ for any $i \geq 0$, and $i \neq j$ implies a_i and a_j are incomparable; that is, neither $a_i \geqq a_j$ nor $a_j \geqq a_i$.

Theorem 7 *[SA00] Let $\mathcal{L}$ be an indexed family of recursive languages with finite thickness. If $\mathcal{L}$ has no infinite anti-chain with respect to the set inclusion $\subseteq$, then $\mathcal{L} \in$* ***U* TxtEx****.*

3.4 The Paradigm DUTxtEx

We now define an identification criterion, where the learner must furthermore identify each of the languages in the union.

Definition 20 Given $\mathcal{L} \subseteq \mathcal{E}$ where $card(\mathcal{L}) < \infty$.

(a) We say a set of indices $\{x_1, x_2, \ldots, x_{card(\mathcal{L})}\} \subseteq N$ is a *representation index set of* $\mathcal{L}$ just in case $\{W_{x_1}, W_{x_2}, ..., W_{x_{card(\mathcal{L})}}\} = \mathcal{L}$.

(b) Let $\mathcal{I}_{\mathcal{L}} = \{I \mid I$ is a representation index set of $\mathcal{L}\}$.

(c) Let $\mathcal{I} = \{I \mid (\exists \mathcal{L} \subseteq \mathcal{E}, card(\mathcal{L}) < \infty)[I \in \mathcal{I}_{\mathcal{L}}]\}$.

Any representation index set $\{x_1, x_2, \ldots, x_{card(\mathcal{L})}\}$ can be represented by a natural number k where $D_k = \{x_1, x_2, \ldots, x_{card(\mathcal{L})}\}$. This representation is implicit whenever the context requires such an interpretation.

Proposition 8 *Given $\mathcal{L} \subseteq \mathcal{E}$, for every $\mathcal{L}', \mathcal{L}'' \subseteq \mathcal{L}$, $\mathcal{L}' \neq \mathcal{L}''$ implies that $\mathcal{I}_{\mathcal{L}'} \cap \mathcal{I}_{\mathcal{L}''} = \emptyset$.*

Definition 21 Let $k \in N^+$ and $\mathcal{L} \subseteq \mathcal{E}$.

(a) M $\mathbf{DU^kTxtEx}$-identifies[5] $\mathcal{L}$ just in case for each $\mathcal{L}' \subseteq \mathcal{L}$, where $card(\mathcal{L}') \leq k$, for every text T for $L_{\mathcal{L}'}$, $M(T)\downarrow$ and $M(T) \in \mathcal{I}_{\mathcal{L}'}$.

(b) $\mathbf{DU^kTxtEx} = \{\mathcal{L} \subseteq \mathcal{E} \mid (\exists M)[M\ \mathbf{DU^kTxtEx}\text{-identifies } \mathcal{L}]\}$.

Hence a $\mathbf{DU^kTxtEx}$ learner, upon every text for the union of up to k languages, is able to tell exactly what these languages are. As an example, it is easy to verify that the class of all singleton languages, $SINGLE = \{\{x\} \mid x \in N\}$ is in $\mathbf{DU^*TxtEx}$.

We explain a limitation to the kind of classes that can be in **DUTxtEx**, with the following proposition.

Proposition 9 *Given $\mathcal{L} \subseteq \mathcal{E}$. If there exists finite $\mathcal{L}', \mathcal{L}'' \subseteq \mathcal{L}$, $\mathcal{L}' \neq \mathcal{L}''$, but $L_{\mathcal{L}'} = L_{\mathcal{L}''}$, then $\mathcal{L} \notin \mathbf{DU^k TxtEx}$ for $k = \max(card(\mathcal{L}'),\ card(\mathcal{L}''))$.*

---

[5]The **D** in **DUTxtEx** stands for *Discernible*.

PROOF. A text for $L_{\mathcal{L}'}$ is also a text for $L_{\mathcal{L}''}$. But $\mathcal{I}_{\mathcal{L}'} \cap \mathcal{I}_{\mathcal{L}''} = \emptyset$ and an **IIM** cannot converge to indices for both an $I' \in \mathcal{I}_{\mathcal{L}'}$ and an $I'' \in \mathcal{I}_{\mathcal{L}''}$. ∎

For the purpose of avoiding Proposition 9, it is useful to introduce the following terminology.

Definition 22 *Let* $\mathcal{L} \subseteq \mathcal{E}$ *and* $k \in N^+$. The class of languages $\mathcal{L}^k$ is said to be *uniquely definable from* $\mathcal{L}$ just in case for all $L \in \mathcal{L}^k$, there exists a unique $\mathcal{L}' \subseteq \mathcal{L}$, where $card(\mathcal{L}') \leq k$, such that $L_{\mathcal{L}'} = L$.

Hence for a class $\mathcal{L}^k$ that is uniquely definable from $\mathcal{L}$, every collection of up to k languages from $\mathcal{L}$ uniquely generates a language in $\mathcal{L}^k$.

3.5 The Paradigm WDUTxtEx

There are other conceivable ways of defining identification criteria similar to **DUTxtEx** which do not lead to a situation like that in Proposition 9. One possibility is that when there exist distinct classes of languages $\mathcal{L}', \mathcal{L}'' \subseteq \mathcal{L}$, each of cardinality at most k, such that $L_{\mathcal{L}'} = L_{\mathcal{L}''}$, we might allow the learner which learns $\mathcal{L}^k$ the freedom to choose either $\mathcal{L}'$ or $\mathcal{L}''$ to conjecture. The following identification criterion captures precisely this notion.

Definition 23 Let $k \in N^+$ and $\mathcal{L} \subseteq \mathcal{E}$.

(a) M $\mathbf{WDU}^k\mathbf{TxtEx}$-identifies[6] $\mathcal{L}$ just in case for each $L \in \mathcal{L}^k$, for every text

[6]The **W** in **WDUTxtEx** stands for *Weakly*.

T for L, $M(T)\downarrow$, and there exists $\mathcal{L}' \subseteq \mathcal{L}$, where $card(\mathcal{L}') \leq k$, and T is a text for $L_{\mathcal{L}'}$, such that $M(T) \in \mathcal{I}_{\mathcal{L}'}$.

(b) $\mathbf{WDU}^k\mathbf{TxtEx} = \{\mathcal{L} \subseteq \mathcal{E} \mid (\exists M)[M\ \mathbf{WDU}^k\mathbf{TxtEx}\text{-identifies } \mathcal{L}]\}$.

It is clear from the definitions that $\mathbf{WDU}^1\mathbf{TxtEx} \equiv \mathbf{DU}^1\mathbf{TxtEx} \equiv \mathbf{U}^1\mathbf{TxtEx} \equiv \mathbf{TxtEx}$.

3.6 The Paradigm SDUTxtEx

Another way to avoid complications similar to Proposition 9 is to require the learner to suggest all the possible unions for which the text is for. To us, this seems exceedingly demanding, and we include the following identification criterion only for completeness.

Definition 24 Let $k \in N^+$ and $\mathcal{L} \subseteq \mathcal{E}$.

(a) M $\mathbf{SDU}^k\mathbf{TxtEx}$-identifies[7] $\mathcal{L}$ just in case for each $L \in \mathcal{L}^k$, for every text T for L, $M(T)\downarrow$ and for all $\mathcal{L}' \subseteq \mathcal{E}$, if

(1) $\mathcal{L}' \subseteq \mathcal{L}$,

(2) $card(\mathcal{L}') \leq k$, and

(3) T is a text for $L_{\mathcal{L}'}$,

then $card(\mathcal{I}_{\mathcal{L}'} \cap W_{M(T)}) = 1$; otherwise $card(\mathcal{I}_{\mathcal{L}'} \cap W_{M(T)}) = 0$.

(b) $\mathbf{SDU}^k\mathbf{TxtEx} = \{\mathcal{L} \subseteq \mathcal{E} \mid (\exists M)[M\ \mathbf{SDU}^k\mathbf{TxtEx}\text{-identifies } \mathcal{L}]\}$.

[7] The **S** in **SDUTxtEx** stands for *Strongly*.

Hence an $\mathbf{SDU}^k\mathbf{TxtEx}$ learner, upon a given text T, for each set of at most k languages from $\mathcal{L}$ which generates $content(T)$, the learner outputs a representation index set for that set of languages.

4 Basic Hierarchy Results

4.1 DUTxtEx Hierarchy

Proposition 10 ***DU²TxtEx*** $\subset$ ***TxtEx***.

PROOF. That $\mathbf{DU^2TxtEx} \subseteq \mathbf{TxtEx}$ is immediate from definition. Consider the collection of finite languages, *FIN*. It is clear that $FIN \in \mathbf{TxtEx}$. Choose distinct sets $D_1, D_2, D_3 \in FIN$ such that $D_1 \cup D_2 = D_3$. By Proposition 9, $FIN \notin \mathbf{DU^2TxtEx}$. ∎

We now extend this result to $\mathbf{DU}^{n+1}\mathbf{TxtEx} \subset \mathbf{DU}^n\mathbf{TxtEx}$ for any $n \in N$. First some preliminary results need to be obtained.

Definition 25 Let $n \in N$, $n \geq 2$.

(a) A coordinate in an $(n-1)$-space[8] is written $(x_1, ..., x_{n-1})$. For any point A, let $x_i(A)$ (where $1 \leq i \leq n-1$) denote it's position along the x_i axis.

(b) We define a (convex) simplex G in an $(n-1)$-space by the set of all the points $(x_1, ..., x_{n-1})$ in a finite region which satisfy n linear equations [Cox63]

$$\begin{pmatrix} \nu_{k,1} \\ \nu_{k,2} \\ \vdots \\ \nu_{k,n-1} \end{pmatrix} \cdot (x_1, x_2, \ldots, x_{n-1}) \leq b_k \qquad (k = 1, 2, \ldots, n)$$

[8] Unless stated otherwise, we assume Euclidean spaces in this thesis.

where $b_1, \ldots, b_n$ are co-efficients in $\mathbb{R}$. The inequality for each k defines a bounding hyperplane for the simplex, where each vector $(\ \nu_{k,1}\ \ \nu_{k,2}\ \ \cdots\ \ \nu_{k,n-1}\ )$ represents an (assumed unit) outward normal for the bounding hyperplane. The intersections of each hyperplane with G forms a simplex in $(n-2)$-space, and is called a *facet* of G.

(c) Let G_n be an arbitrary, fixed simplex of n vertices in an $(n-1)$-space.

(d) Let $T_n = \{T \mid T \text{ is a (vector) translation in } (n-1)\text{-space}\}$.

(e) For each $T \in T_n$ and each simplex G, $G + T$ denotes the simplex with vertices at $\{A + T \mid A \text{ is a vertex of } G\}$.

(f) Let $\Gamma_n = \{G_n + T \mid T \in T_n\}$. Hence Γ_n is the set of all the simplexes that G_n can transform into under only translations.

(g) For $G \in \Gamma_n$, let $V(G) = \{A \mid A \text{ is a vertex of } G\}$.

The following gives an important result which is used, in part or in full, in several propositions throughout this thesis.

Claim 11 *Let $n \geq 2$. Given $G \in \Gamma_n$ with vertices at $A_1, A_2, \ldots, A_n$. Let C be a point in $G - V(G)$. For each $i \in N$, $1 \leq i \leq n$, let $\vec{\mu_i} = (1/|\overrightarrow{CA_i}|)\,\overrightarrow{CA_i}$ and for $\delta \in \mathbb{R}$, let $G_i(\delta) = G + \delta\vec{\mu_i}$. There exists a collection of*

(1) n simplexes $G_1(\epsilon_1), G_2(\epsilon_2), \ldots, G_n(\epsilon_n)$ where each $\epsilon_i \in \mathbb{R}, \epsilon_i > 0$, and

(2) n numbers $\xi_1, \xi_2, \ldots, \xi_n \in \mathbb{R}$ where each $\xi_i \in \mathbb{R}, 0 < \xi_i \leq \epsilon_i$,

such that for each $k, 1 \leq k \leq n$, for all $\delta_k \in \mathbb{R}$ where $0 \leq \delta_k \leq \xi_k$, $G_k(\delta_k) \subseteq \bigcup_{j=1}^{n} G_j(\epsilon_j)$.

PROOF. For $n = 2$, given G with $V(G) = \{(y_1), (y_2)\}$, let $\epsilon_1 = \epsilon_2 = \xi_1 = \xi_2 = |(y_2 - y_1)/2|$ and we are done. We now show the case for $n \geq 3$. Let $n \geq 3$ be given. We use the notations introduced from the beginning of the claim statement, up to (and including) $G_i(\delta)$. For each i, $1 \leq i \leq n$, let F^i denote the facet with vertices at $V(G) - \{A_i\}$, and let $\vec{\nu_i}$ denote F^i's outward unit normal. We demonstrate these notations with 2-D examples in Figure 1 and Figure 2.

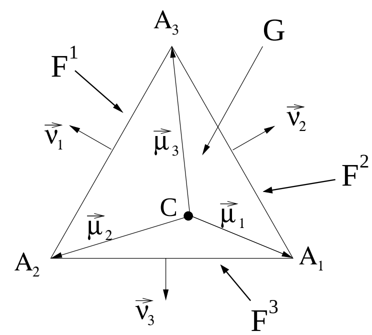

Figure 1: The vectors $\vec{\mu_i}$, $\vec{\nu_i}$ and the facets F^i (for $i = 1, 2, 3$) in 2-D.

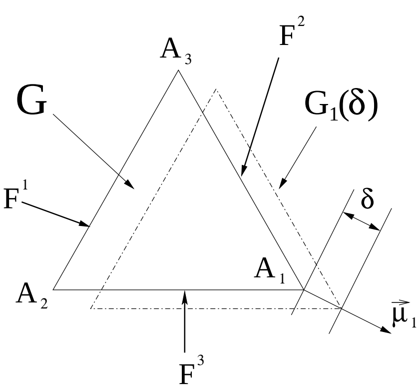

Figure 2: $G_1(\delta)$ in 2-D.

Let $b_1, b_2, \ldots, b_n$ be such that the simplex G are the points X which fulfill the n inequalities $\vec{\nu_i} \cdot X \leq b_i$ for $i = 1, \ldots, n$ as in part (b) of Definition 25.

We now present a few sub-claims for use in the proof.

Sub-claim 1 *For each $i \in N, 1 \leq i \leq n$, and $\delta \in \mathbb{R}$, the simplex $G_i(\delta)$ is defined by the points X which satisfy n inequalities $\vec{\nu_j} \cdot X \leq b_j + \delta(\vec{\mu_i} \cdot \vec{\nu_j})$ for $j = 1, \ldots, n$.*

Proof. Let i, $1 \leq i \leq n$ and $\delta \in \mathbb{R}$ be given. A point X is in $G_i(\delta)$ if and only if it is translated by $-\delta\vec{\mu_i}$ to a point in G. Hence, for each X in $G_i(\delta)$, we have $\vec{\nu_j} \cdot (X - \delta\vec{\mu_i}) \leq b_j$ (or equivalently, $\vec{\nu_j} \cdot X \leq b_j + \delta(\vec{\mu_i} \cdot \vec{\nu_j})$) for $j = 1, \ldots, n$. □

Sub-claim 2 *For any point X in G, the shortest path from X to any hyperplane with normal $\vec{\nu_i}$ where $1 \leq i \leq n$, is (1) along $\vec{\nu_i}$, and (2) this shortest distance is given by $|\vec{\nu_i} \cdot X - b_i|$.*

Proof. Let i, $1 \leq i \leq n$ be given. Part (1) follows from that $\vec{\nu_i}$ is the (outward) normal of F^i. For (2), let T be any translation that translate X, a point in G to some point in F^i, say X'. The component of T along $\vec{\nu_i}$ is $|\vec{\nu_i} \cdot T| = |\vec{\nu_i} \cdot (X' - X)|$ $= |\vec{\nu_i} \cdot (X - X')| = |\vec{\nu_i} \cdot X - \vec{\nu_i} \cdot X'|$. Since $\vec{\nu_i} \cdot X' = b_i$, $|\vec{\nu_i} \cdot T| = |\vec{\nu_i} \cdot X - b_i|$. □

Sub-claim 3 *For each $i, j \in N, 1 \leq i, j \leq n$, $i \neq j$, we have $\vec{\nu_j} \cdot \vec{\mu_i} \geq 0$.*

Proof. Let j, $1 \leq j \leq n$ be given. Any point C in G must satisfy $\vec{\nu_j} \cdot C \leq b_j$. However, for all $i \in N, 1 \leq i \leq n$ where $i \neq j$ we have $\vec{\nu_j} \cdot A_i = b_j$. The two statements combine into $\vec{\nu_j} \cdot C \leq \vec{\nu_j} \cdot A_i$. Hence $\vec{\nu_j} \cdot (A_i - C) \geq 0$. It follows that for all $i, j \in N, 1 \leq i, j \leq n$, $i \neq j$, we have $\vec{\nu_j} \cdot \vec{\mu_i} = \vec{\nu_j} \cdot \ (1/|\overrightarrow{CA_i}|)\overrightarrow{CA_i} \geq 0$. □

Sub-claim 4 *For each $i \in N$, $1 \leq i \leq n$, $\vec{\mu_i} \cdot \vec{\nu_i} < 0$.*

Proof. First observe that for each i, $1 \leq i \leq n$, A_i is the furthest point in G from F^i. Hence by Sub-claim 2, for each point $C \neq A_i$ in G, $|\vec{\nu_i} \cdot A_i - b_i| > |\vec{\nu_i} \cdot C - b_i|$. We note that $\vec{\nu_i} \cdot A_i - b_i$ and $\vec{\nu_i} \cdot C - b_i$ are either zero or negative values (due to the inequality $\vec{\nu_i} \cdot C \leq b_i$ for all points in G), and hence $|\vec{\nu_i} \cdot A_i - b_i| > |\vec{\nu_i} \cdot C - b_i|$ becomes $\vec{\nu_i} \cdot A_i - b_i < \vec{\nu_i} \cdot C - b_i$. Subsequently, this evaluates to $\vec{\nu_i} \cdot (A_i - C) < 0$. Hence for any choice of the point C, $\vec{\nu_i} \cdot \vec{\mu_i} = \vec{\nu_i} \cdot \ (1/|\overrightarrow{CA_i}|)\ \overrightarrow{CA_i} < 0$. □

Sub-claim 5 *Given $i, j \in N$, $1 \leq i, j \leq n$, and non-negative $\delta_i, \delta_j \in \mathbb{R}$. If for all k where $1 \leq k \leq n$, $k \neq i$, $k \neq j$, we have $\delta_i(\vec{\mu_i} \cdot \vec{\nu_k}) \leq \delta_j(\vec{\mu_j} \cdot \vec{\nu_k})$, then the region $G_i(\delta_i) - G_j(\delta_j)$ can be defined as the points X which fulfill the inequalities*

(1) $\vec{\nu_j} \cdot X > b_j + \delta_j(\vec{\mu_j} \cdot \vec{\nu_j})$, and

(2) $\vec{\nu_k} \cdot X \leq b_k + \delta_i(\vec{\mu_i} \cdot \vec{\nu_k})$ for $k = 1, 2, \ldots, n$ (the n inequalities for $G_i(\delta_i)$).

Proof. Let i, j, δ_i, δ_j be as given in sub-claim. By Sub-claim 1, $G_i(\delta_i)$ is the region which fulfills n inequalities $\vec{\nu_k} \cdot X \leq b_k + \delta_i(\vec{\mu_i} \cdot \vec{\nu_k})$ for $k = 1, \ldots, n$, while $G_j(\delta_j)$ is the region which fulfills n inequalities $\vec{\nu_k} \cdot X \leq b_k + \delta_j(\vec{\mu_j} \cdot \vec{\nu_k})$ for $k = 1, \ldots, n$. The region not in $G_j(\delta_j)$, is the union of n regions, each of which satisfy one of the n inequalities

$$\vec{\nu_k} \cdot X > b_k + \delta_j(\vec{\mu_j} \cdot \vec{\nu_k}) \qquad k = 1, 2, \ldots, n. \qquad \text{————— (A)}$$

Since for each k where $k \neq j$, $\delta_j(\vec{\mu_j} \cdot \vec{\nu_k}) \geq \delta_i(\vec{\mu_i} \cdot \vec{\nu_k})$ (note that the case of $k = i, k \neq j$ is given by Sub-claims 3 and 4), each inequality in (A) where $k \neq j$ must have $\vec{\nu_k} \cdot X > b_k + \delta_j(\vec{\mu_j} \cdot \vec{\nu_k}) \geq b_k + \delta_i(\vec{\mu_i} \cdot \vec{\nu_k})$, and hence defines a region not in $G_i(\delta_i)$. Such inequalities are obviously redundant.

It follows that $G_i(\delta_i) - G_j(\delta_j)$ is sufficiently defined by the points X which fulfill $\vec{\nu_j} \cdot X > b_j + \delta_j(\vec{\mu_j} \cdot \vec{\nu_j})$ in addition to the n inequalities for $G_i(\delta_i)$. □

We now proceed with the proof of the claim. We shall carry out this proof in two parts. In the first part, we shall first select a set of n simplexes, $G_1(\epsilon_1), \ldots, G_n(\epsilon_n)$

which fulfill certain properties. These properties will be defined at the end of this part of the proof. We shall make use of these properties in the second part, to demonstrate the existence of a set of n numbers $\xi_1, \xi_2, \ldots, \xi_n$ which fulfills the conditions as in claim.

Part I. Let non-zero, positive $\epsilon'_1, \epsilon'_2, \ldots, \epsilon'_n \in \mathbb{R}$, be given. Let the intersection point of the $n-1$ hyperplanes

$$\vec{\nu_i} \cdot X = b_i + \epsilon'_i(\vec{\mu_i} \cdot \vec{\nu_i}) \qquad i = 1, \ldots, n-1$$

be denoted by Z_n. We illustrate this in Figure 3. When all ϵ'_i for $1 \leq i \leq n-1$ approach 0, Z_n is at A_n. As each of these ϵ'_i increases, Z_n tends away from A_n towards some point in G, until eventually it leaves G entirely.

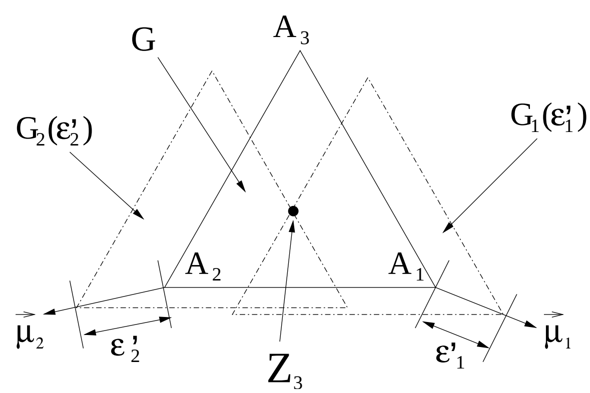

Figure 3: The intersection point Z_n in 2-D.

We denote the distance between Z_n and A_n by η_n. We want to let $\epsilon'_1, \epsilon'_2, \ldots, \epsilon'_n$ be so small such that the following holds

(1) Z_n is within $G_n(\epsilon'_n)$, and

(2) $\epsilon'_n < \frac{|\vec{\nu_n} \cdot A_n - b_n| - \eta_n}{|\vec{\mu_n} \cdot \vec{\nu_n}|}$ (or equivalently, $|\vec{\nu_n} \cdot A_n - b_n| - \epsilon'_n|\vec{\mu_n} \cdot \vec{\nu_n}| > \eta_n$).

Clause (2) is possible since $|\vec{\mu_n} \cdot \vec{\nu_n}| \neq 0$ (by Sub-claim 4), and since $\frac{|\vec{\nu_n} \cdot A_n - b_n| - \eta_n}{|\vec{\mu_n} \cdot \vec{\nu_n}|}$ can only increase as each ϵ_i' decreases. It is clear that both clauses can be fulfilled for sufficiently small ϵ_i's. Note that for any set of $\epsilon_1', \ldots, \epsilon_n'$ where the conditions hold, they continue to hold for any set of positive $\epsilon_1'', \ldots, \epsilon_n''$ where for each i, $\epsilon_i'' \leq \epsilon_i'$. This shows that the following choice of $\epsilon_1, \epsilon_2, \ldots, \epsilon_n \in \mathbb{R}$ is possible.

We first generalize notations. For any $k \in N, 1 \leq k \leq n$ and any set of positive real numbers $\delta = \{\delta_1, \ldots, \delta_n\}$, let the intersection point of the $n-1$ equations

$$\vec{\nu_j} \cdot X = b_j + \delta_j(\vec{\mu_j} \cdot \vec{\nu_j}) \qquad j = 1, \ldots, n, \ j \neq k$$

be denoted $Z_{k,\delta}$. (This generalizes Z_n defined earlier). Figure 4 demonstrates these new notations. For each $Z_{k,\delta}$, we denote the distance between A_k and $Z_{k,\delta}$ by $\eta_{k,\delta}$.

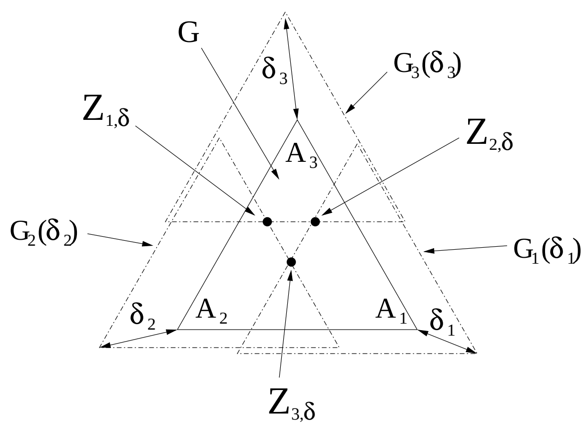

Figure 4: $Z_{i,\delta}$ for $i = 1, 2, 3$ in 2-D.

Finally, let $\epsilon = \{\epsilon_1, \ldots, \epsilon_n\}$ be a set of non-zero positive reals so small such that for each k, $1 \leq k \leq n$,

$Z_{k,\epsilon}$ is within $G_k(\epsilon_k)$, and ———— (Condition I)

$|\vec{\nu_k} \cdot A_k - b_k| - \epsilon_k |\vec{\mu_k} \cdot \vec{\nu_k}| > \eta_{k,\epsilon}$. ———— (Condition II)

These are our selected $G_1(\epsilon_1), \ldots, G_n(\epsilon_n)$.

End of Part I. $\square$

In the next part, we shall first find a set of values $\xi_1, \xi_2, \ldots, \xi_n \in \mathbb{R}$. For arbitrary $k \in N$, $1 \leq k \leq n$, we will consider the two regions: (1) $G_k(\xi_k) - \bigcup_{1 \leq i \leq n, i \neq k} G_i(\epsilon_i)$, and (2) $G_k(\xi_k) - G_k(\epsilon_k)$. We will show, using the values of $\epsilon_1, \ldots, \epsilon_n$ and $\xi_1, \ldots, \xi_n$ chosen, that these two regions do not intersect, resulting in $G_k(\xi_k) - \bigcup_{1 \leq i \leq n} G_i(\epsilon_i) = \emptyset$, and hence $G_k(\xi_k) \subseteq \bigcup_{1 \leq i \leq n} G_i(\epsilon_i)$.

Part II. We will now find the values for $\xi_1, \xi_2, \ldots, \xi_n \in \mathbb{R}$ as in claim. For each i, let ξ_i, $0 < \xi_i \leq \epsilon_i$ be a value so small such that for each j, k, $1 \leq j, k \leq n$, $k \neq j$, $\xi_i(\vec{\mu_i} \cdot \vec{\nu_k}) \leq \epsilon_j(\vec{\mu_j} \cdot \vec{\nu_k})$. (Figure 5 illustrates this condition in 2-D.) Since $\epsilon_1, \ldots, \epsilon_n$ are non-zero, such $\xi_1, \xi_2, \ldots, \xi_n$ exist.

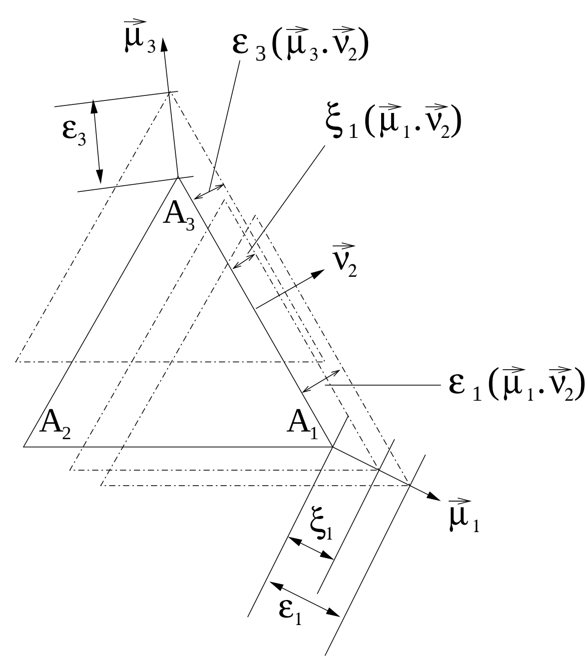

Figure 5: $\xi_1(\vec{\mu_1} \cdot \vec{\nu_2}) \leq \min(\epsilon_1(\vec{\mu_1} \cdot \vec{\nu_2}), \epsilon_3(\vec{\mu_3} \cdot \vec{\nu_2}))$.

Let $k \in N$, $1 \leq k \leq n$ be given.

Consider the region $G_k(\xi_k) - \bigcup_{1\leq i\leq n, i\neq k} G_i(\epsilon_i)$. By Sub-claim 5, this region is defined as the points X which fulfill the set of $(n-1)+n$ inequalities

(A) $\vec{\nu_i} \cdot X > b_i + \epsilon_i(\vec{\mu_i} \cdot \vec{\nu_i})$ where $1 \leq i \leq n$, $i \neq k$, and

(B) $\vec{\nu_i} \cdot X \leq b_i + \xi_k(\vec{\mu_k} \cdot \vec{\nu_i})$ where $i = 1, \ldots, n$.

Intuitively, the set of hyperplanes in (A) bound $G_k(\xi_k) - \bigcup_{1\leq i\leq n, i\neq k} G_i(\epsilon_i)$ to only a region around A_k. This is shown in Figure 6.

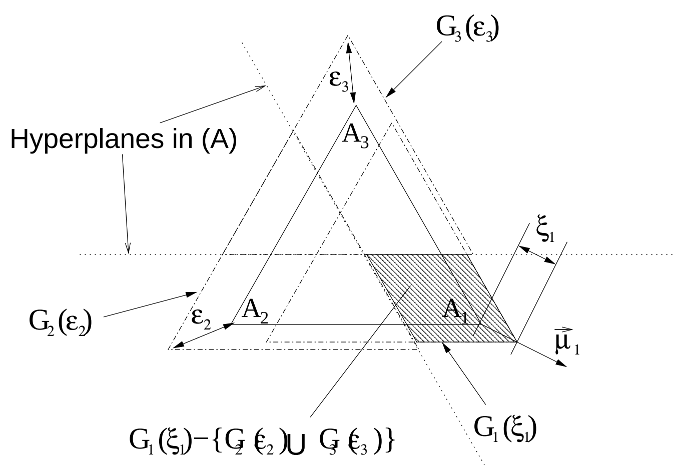

Figure 6: The region $G_1(\xi_1) - \{G_2(\epsilon_2) \cup G_3(\epsilon_3)\}$.

The furthest point in this region from A_k hence, extends only to the intersections of the hyperplanes in (A). Recall that we have, in Part I of the proof, denoted this intersection point by $Z_{k,\epsilon}$, and denoted it's distance to A_k by $\eta_{k,\epsilon}$.

Hence in conclusion, **the region $G_k(\xi_k) - \bigcup_{1\leq i\leq n, i\neq k} G_i(\epsilon_i)$ do not extend beyond a distance of $\eta_{k,\epsilon}$ from A_k.**

Consider now the region $G_k(\xi_k) - G_k(\epsilon_k)$. By Sub-claim 5, this region is defined as the points X which fulfill the set of $n+1$ inequalities

(C) $\vec{\nu_k} \cdot X > b_k + \epsilon_k(\vec{\mu_k} \cdot \vec{\nu_k})$, and

(D) $\vec{\nu_i} \cdot X \leq b_i + \xi_k(\vec{\mu_k} \cdot \vec{\nu_i})$ where $i = 1, \ldots, n$.

Hence intuitively, this region is bounded beyond A_k by the hyperplane in (C). We illustrate this in Figure 7. By Sub-claim 2, **the nearest point in** $G_k(\xi_k) - G_k(\epsilon_k)$ **from** A_k **is at a distance of at least**

$$|\vec{\nu_k} \cdot A_k - b_k| - \epsilon_k |\vec{\mu_k} \cdot \vec{\nu_k}|$$

away, that is, the shortest distance from A_k to the hyperplane in (C).

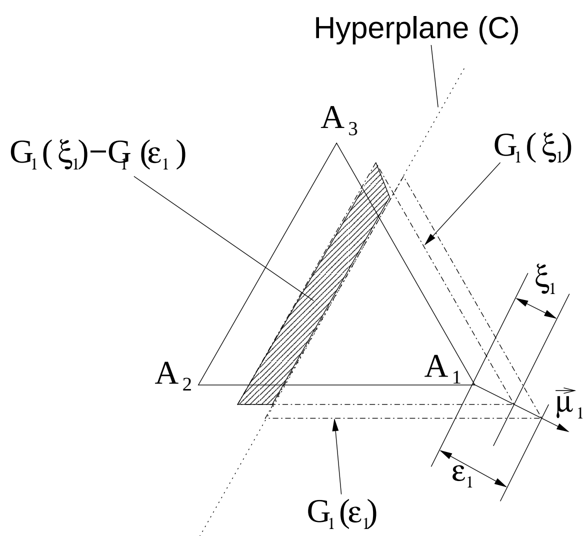

Figure 7: The region $G_1(\xi_1) - G_1(\epsilon_1)$.

However, by our choice of ϵ_k in Part I (Condition II), $|\vec{\nu_k} \cdot A_k - b_k| - \epsilon_k |\vec{\mu_k} \cdot \vec{\nu_k}|$ is strictly larger than $\eta_{k,\epsilon}$!

It follows that the two regions

1. $G_k(\xi_k) - \bigcup_{1 \leq i \leq n, i \neq k} G_i(\epsilon_i)$, and
2. $G_k(\xi_k) - G_k(\epsilon_k)$

do not intersect. Hence, $G_k(\xi_k) - \bigcup_{i=1}^{n} G_i(\epsilon_i) = \emptyset \Rightarrow G_k(\xi_k) \subseteq \bigcup_{i=1}^{n} G_i(\epsilon_i)$.

End of Part II. $\square$

It is easy to verify that the argument in Part II holds for all $k \in N$, $1 \leq k \leq n$, and for all $\delta_k \in \mathbb{R}, \delta_k \leq \xi_k$. Hence, for all k, $1 \leq k \leq n$, and for all $\delta_k \in \mathbb{R}, 0 \leq \delta_k \leq \xi_k$, $G_k(\delta_k) \subseteq \bigcup_{j=1}^{n} G_j(\epsilon_j)$. $\blacksquare$

The following sub-claim gives auxiliary results which we will use in proving the next claim (Claim 12).

Sub-claim 6 *Given $G \in \Gamma_n$ with vertices at $A_1, A_2, \ldots, A_n$. For each $i, 1 \leq i \leq n$, let F^i denote the facet with vertices at $V(G) - \{A_i\}$, and let $\vec{\nu_i}$ denote F^i's outward normal. Let $G' = G + T$ where $T \in T_n$ and $G' \neq G$, then for each $i, 1 \leq i \leq n$, the following holds:*

(a) $T \cdot \vec{\nu_i} < 0 \Rightarrow (V(G) - \{A_i\}) \cap G' = \emptyset$.

(b) $A_i \in G' \Rightarrow T \cdot \vec{\nu_i} < 0$.

(c) $A_i \in G' \Rightarrow$ *for all* $j \neq i$, $1 \leq j \leq n$, $T \cdot \vec{\nu_j} \geq 0$.

(d) $\mathit{card}(V(G) \cap G') \leq 1$.

Proof. Let G, G', T, $A_1, \ldots, A_n$, $F^1, \ldots, F^n$ and $\vec{\nu_1}, \ldots, \vec{\nu_n}$ be as defined in sub-claim. Let $b_1, b_2, \ldots, b_n$ be such that each point X in G fulfills the n inequalities $\vec{\nu_i} \cdot X \leq b_i$ for $i = 1, \ldots, n$ as in part (b) of Definition 25.

Let $i \in N, 1 \leq i \leq n$ be given.

For part (a), note that each point X in $V(G) - \{A_i\}$ lies on F^i, and hence fulfills the hyperplane equation $\vec{\nu_i} \cdot X = b_i$. On the other hand, if $T \cdot \vec{\nu_i} < 0$, then each point X in G' would have to fulfill the inequality $\vec{\nu_i} \cdot X \leq b_i + T \cdot \vec{\nu_i} < b_i$, and hence G' does not include $V(G) - \{A_i\}$.

For part (b), suppose $A_i \in G'$, then there exists a point C in G where $C \neq A_i$ and $C + T = A_i$. Now we know that A_i is the furthest point in G from F^i along $\vec{\nu_i}$, so by Sub-claim 2 in the proof of Claim 11, $|\vec{\nu_i} \cdot A_i - b_i| > |\vec{\nu_i} \cdot C - b_i|$. Since $\vec{\nu_i} \cdot A_i - b_i$ and $\vec{\nu_i} \cdot C - b_i$ are either zero or negative values (due to the inequality $\vec{\nu_i} \cdot X \leq b_i$ for all points $X \in G$), this becomes $\vec{\nu_i} \cdot A_i - b_i < \vec{\nu_i} \cdot C - b_i$. Substituting $C = A_i - T$, we get $\vec{\nu_i} \cdot A_i - b_i < \vec{\nu_i} \cdot (A_i - T) - b_i$, which gives us $\vec{\nu_i} \cdot T < 0$.

For part (c), observe that for each j, $1 \leq j \leq n$, each point X in G' is bounded by the inequality $\vec{\nu_j} \cdot X \leq b_j + T \cdot \vec{\nu_j}$. Since for each $j \neq i$, $\vec{\nu_j} \cdot A_i = b_j$ (that is, since A_i lies on F^j), if $T \cdot \vec{\nu_j} < 0$, then A_i would not be in G'. Hence for each $j \neq i$, $1 \leq j \leq n$, $T \cdot \vec{\nu_j} \geq 0$.

Part (d) follows from part (a) and part (b). ∎

The following claim gives an important result which we shall use, to various extents, throughout this thesis. Basically, we want to show that every set of up to n distinct simplexes in Γ_n uniquely define a region that is definable by at most n simplexes from Γ_n.

Claim 12 *Let* $\Gamma^1, \Gamma^2 \subset \Gamma_n$ *where* $\max(\{card(\Gamma^1), card(\Gamma^2)\}) \leq n$. *Then*

(1) $\bigcup_{G \in \Gamma^1} V(G) \subseteq \bigcup_{G \in \Gamma^2} G$, *and*

(2) $\bigcup_{G \in \Gamma^2} V(G) \subseteq \bigcup_{G \in \Gamma^1} G$

if and only if $\Gamma^1 = \Gamma^2$.

PROOF. The if direction is obvious. We now show the only if part.

By way of contradiction suppose there exist $\Gamma^1, \Gamma^2 \subset \Gamma_n$, each with cardinality at most n, where $\Gamma^1 \neq \Gamma^2$ but the conditions (1) and (2) in claim holds.

Let G be a simplex in $\Gamma^1 - \Gamma^2$ with vertices at $A_1, \ldots, A_n$. For each $i, 1 \leq i \leq n$, let $\vec{\nu_i}$ denote the (unit) outward normal of the facet in G with vertices $V(G) - \{A_i\}$. Let $b_1, \ldots, b_n$ be such that each point X in G fulfills the n inequalities $\vec{\nu_i} \cdot X \leq b_i$ for $i = 1, \ldots, n$ as in part (b) of Definition 25.

By Part (d) of Sub-claim 6, unless $card(\Gamma^2) = n$ (so that each simplex in Γ^2 includes a point in $V(G)$), $V(G) \not\subseteq \bigcup_{G' \in \Gamma^2} G'$. For this reason, let Γ^2 be a collection of n simplexes, $\{G + T^1, G + T^2, \ldots, G + T^n\}$, where for each $i, 1 \leq i \leq n$, T^i is a translation such that $A_i \in G + T^i$.

By Part (b) of Sub-claim 6, for each i, $T^i \cdot \vec{\nu_i} < 0$. Note first that each translation $-T^i$ translates the simplex $G + T^i$ to G. Since $-T^i \cdot \vec{\nu_i} > 0$, by Part (b) of Sub-claim 6, the vertex in $G + T^i$ at the position $A_i + T^i$, is not in G. Figure 8 illustrates these vertices in 2-D.

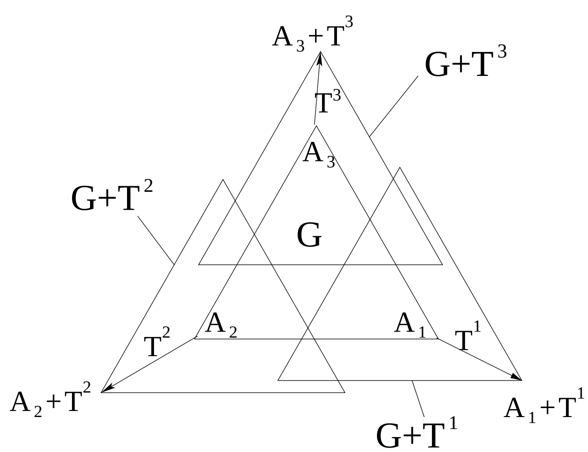

Figure 8: The vertices $A_i + T^i$ for $i = 1, 2, 3$.

Each $G + T^i$ hence contributes a vertex positioned at $A_i + T^i$, which is not in G. We claim that there is no simplex in Γ_n that can include two vertices in $\{A_i + T^i \mid 1 \leq i \leq n\}$.

To see this, towards a contradiction, let i, j be where $1 \leq i, j \leq n$ and $i \neq j$, and let simplex G', translation T' be where $G' = G + T'$, such that the vertices at $A_i + T^i$ and $A_j + T^j$ are in G'. Since $A_i + T^i$ is in G', by Part (b) of Sub-claim 6,

$$
\begin{aligned}
& (T' - T^i) \cdot \vec{\nu_i} < 0 && (\because T' - T^i \text{ translates } G + T^i \text{ to } G') \\
\Rightarrow\ & T' \cdot \vec{\nu_i} < T^i \cdot \vec{\nu_i} && \\
\Rightarrow\ & T' \cdot \vec{\nu_i} < 0 && (\because T^i \cdot \vec{\nu_i} < 0)
\end{aligned}
$$

Now each point X in G' must fulfill n inequalities, one of which being

$$
\begin{aligned}
& \vec{\nu_i} \cdot X \leq b_i + T' \cdot \vec{\nu_i} && \\
\Rightarrow\ & \vec{\nu_i} \cdot X < b_i && (\because T' \cdot \vec{\nu_i} < 0 \text{ from the previous result})
\end{aligned}
$$

However, the point $A_j + T^j$ lies on the hyperplane defined by the equality

$$
\begin{aligned}
& \vec{\nu_i} \cdot X = b_i + T^j \cdot \vec{\nu_i} && \\
\Rightarrow\ & \vec{\nu_i} \cdot X \geq b_i && (\because T^j \cdot \vec{\nu_i} \geq 0, \text{ by using } A_j \in G + T^j \\
& && \qquad\qquad \text{on Part (c) of Sub-claim 6})
\end{aligned}
$$

This shows that $A_j + T^j$ is not in G', a contradiction.

Hence, any simplex in Γ_n can contain at most one vertex in $\{A_i + T^i \mid 1 \leq i \leq n\}$. However, there are less than n simplexes in the set $\Gamma^1 - \{G\}$, hence not all of the n vertices in $\{A_i + T^i \mid 1 \leq i \leq n\}$ can be in $\bigcup_{G' \in \Gamma^1} G'$. Claim follows. ∎

It is clear that Claim 12 implies that every set of up to n distinct simplexes in Γ_n uniquely define a region that is definable by at most n simplexes from Γ_n.

We now restrict our geometrical definitions to allow only rational values.

Definition 26 Let $n \in N$, $n \geq 2$.

(a) Let $v_1, v_2, \ldots, v_{n-1}$ be unit vectors along each axis of an $(n-1)$-space.

(b) Let $\mathcal{G}_n$ be a simplex with vertices at $v_1, v_2, ..., v_{n-1}$ and $-v_1$ from origin O.

(c) Let $\mathcal{T}_n = \{\sum_{1 \leq i < n} \alpha_i v_i \mid \alpha_i \in rat\}$. Hence $\mathcal{T}_n$ is the set of all (vector) translations with only rational valued components. $\mathcal{T}_n$ is a proper subset of T_n in Definition 25.

(d) Let $\Lambda_n = \{\mathcal{G}_n + T \mid T \in \mathcal{T}_n\}$.

(e) Let RAT_n denote the set of all the points in $(n-1)$-space with only rational valued coordinates. Let $\text{coderat}_n(.)$ be an effective bijective mapping from RAT_n to N. Let $\text{decoderat}_n(.)$ be the inverse function of $\text{coderat}_n(.)$.

(f) For $G \in \Lambda_n$, define the language of G, $L(G) = \{\text{coderat}_n(C) \mid C \in G \cap RAT_n\}$.

(g) For $\Lambda \subseteq \Lambda_n$, $\mathcal{L}(\Lambda) = \{L(G) \mid G \in \Lambda\}$.

(h) Let $TRANSIM_n = \mathcal{L}(\Lambda_n)$.

We now extend Proposition 10 to all $n \in N$.

Proposition 13 *For all $n \in N^+$, $\boldsymbol{DU^{n+1}TxtEx} \subset \boldsymbol{DU^{n}TxtEx}$.*

PROOF. The case of $n = 1$ is shown in Proposition 10. We now show the case for $n \geq 2$. Fix an $n \geq 2$, that $TRANSIM_n \in \mathbf{DU^nTxtEx}$ is seen by M below, where for each text T and each $m \in N$,

$M(T[m])$:

Let $S^m \leftarrow \{A \mid \text{coderat}_n(A) \in \mathit{content}(T[m])\}$.

(This step converts input into coordinates.)

If there exists a collection $\Lambda^m \subset \Lambda_n$ of at most n simplexes such that

(A) $\bigcup_{G\in\Lambda^m} V(G) \subseteq S^m$, and

(B) $S^m \subseteq \bigcup_{G\in\Lambda^m} G$.

(Note that since S^m is finite, this check is recursive.)

Then output $\{L(G) \mid G \in \Lambda^m\}$.

Otherwise, output $\emptyset$.

To see that M $\mathbf{DU}^n\mathbf{TxtEx}$-identifies $\mathit{TRANSIM}_n$, let $\Lambda' \subset \Lambda_n$ be any collection of n simplexes. By Claim 12, is is easy to verify that at some stage t when all of $\bigcup_{G\in\Lambda'} V(G)$ has appeared in S^t, the only set of (at most n) simplexes (from Λ_n) that can fulfill conditions (A) and (B) in the definition of M is Λ', hence $\Lambda^t = \Lambda'$. Thus M, given a text for $L_{\mathcal{L}(\Lambda')}$, outputs $\mathcal{L}(\Lambda')$ in the limit.

To show that $\mathit{TRANSIM}_n \notin \mathbf{DU}^{n+1}\mathbf{TxtEx}$, note that by Claim 11, there exists $\Lambda, \Lambda' \subset \Lambda_n$ where $\mathit{card}(\Lambda) = \mathit{card}(\Lambda') \leq n+1$ such that $\Lambda \neq \Lambda'$ but $L_{\mathcal{L}(\Lambda)} = L_{\mathcal{L}(\Lambda')}$. Hence by Proposition 9, $\mathit{TRANSIM}_n \notin \mathbf{DU}^{n+1}\mathbf{TxtEx}$. ∎

4.2 UTxtEx Hierarchy

Proposition 14 *For all $n \in N^+$, $\boldsymbol{U}^{n+1}\boldsymbol{TxtEx} \subset \boldsymbol{U}^n\boldsymbol{TxtEx}$.*

PROOF. That for all $n \in N^+$, $\mathbf{U}^{n+1}\mathbf{TxtEx} \subseteq \mathbf{U}^n\mathbf{TxtEx}$ is immediate from definition. To show $\mathbf{TxtEx} - \mathbf{U}^2\mathbf{TxtEx} \neq \emptyset$, let $\mathcal{L} = \{\mathcal{K}\} \cup \mathit{SINGLE}$. It is easy to see that $\mathcal{L} \in \mathbf{TxtEx}$. However, it is known that $\{\mathcal{K} \cup \{x\} \mid x \in N\} \notin \mathbf{TxtEx}$ (Proposition 4.7 in [JORS99]). Since $\{\mathcal{K} \cup \{x\} \mid x \in N\} \subseteq \mathcal{L}^2$, $\mathcal{L} \notin \mathbf{U}^2\mathbf{TxtEx}$.

We now show the case for $n \geq 2$. Let PRIMES be the set of all the prime numbers and $p_1, p_2, \ldots$ be an enumeration of PRIMES in ascending order. Let ψ be a computable numbering for which for all $i \in N$, $W^{\psi}_{p_i} = W_i$.

For each $G = \mathcal{G}_n + T$ where $T \in \mathcal{T}_n$, let $X_1(G) = T \cdot v_1$, and let $\mathit{Lang}(G) = \{\langle 0, x\rangle \mid x \in L(G)\} \cup \{\langle 1, y\rangle \mid y \in W^{\psi}_{h(X_1(G))}\}$, where $h(a)$ is the denominator of rational a in reduced form. Clearly, h is a recursive function. Let $\mathit{ExtTRANSIM}_n = \{\mathit{Lang}(G) \mid G \in \Lambda_n\}$.

To see that $\mathit{ExtTRANSIM}_n \in \mathbf{U}^n\mathbf{TxtEx}$, consider each $L \in \mathit{ExtTRANSIM}_n$ as consisting of two parts, $A = \{x \in L \mid \pi_1(x) = 0\}$ and $B = \{x \in L \mid \pi_1(x) = 1\}$. It is easy to verify that an index for B can be obtained given an index for A. Now using the proof for Proposition 13, it is easy to see that we can obtain n indices for all the A parts for any given n languages from $\mathit{ExtTRANSIM}_n$. The respective B parts can then be obtained from these indices. It follows that all n languages can be identified, and hence $\mathit{ExtTRANSIM}_n \in \mathbf{U}^n\mathbf{TxtEx}$.

Let the simplex G in Claim 11 be $\mathcal{G}_n$ and let C be the origin O. Let $\Lambda \subset \Lambda_n$ be a collection of n simplexes as in that claim. Without loss of generality, we require that the numbering ψ has it that for all $G \in \Lambda$, $W^{\psi}_{h(X_1(G))} = \emptyset$. Let $\xi \in \mathit{rat}$, $\xi > 0$ be such that for all $\delta \in \mathbb{R}, 0 \leq \delta \leq \xi$, $\mathcal{G}_n + \delta v_1 \subseteq \bigcup_{G \in \Lambda} G$. Such ξ exists by Claim 11. Let $\Lambda' = \{\mathcal{G}_n + \alpha v_1 \mid \alpha \in rat \wedge 0 \leq \alpha \leq \xi\}$. (It is clear that $\Lambda' \subset \Lambda_n$, and that for each simplex $\mathcal{G}_n + \alpha v_1$ in Λ', $X_1(\mathcal{G}_n + \alpha v_1) = \alpha$.) Let $\mathcal{L}' = \{Lang(G') \cup \bigcup_{G \in \Lambda} Lang(G) \mid G' \in \Lambda'\}$.

Note that for each $G' \in \Lambda'$, $Lang(G') \cup \bigcup_{G \in \Lambda} Lang(G) = \{\langle 0, y\rangle \mid y \in \bigcup_{G \in \Lambda} L(G)\} \cup \{\langle 1, y\rangle \mid y \in W^{\psi}_{h(X_1(G'))}\}$. Hence for each $z \in rat$ where $X_1(\mathcal{G}_n) \leq z \leq X_1(\mathcal{G}_n + \xi v_1)$, there exists a language in $\mathcal{L}'$ which differ from $\bigcup_{G \in \Lambda} Lang(G)$ only by the set $\{\langle 1, y\rangle \mid y \in W^{\psi}_{h(z)}\}$.

Now there exists an $m \in N$ such that $(\forall p \in PRIMES\ where\ p > m)(\exists\, l \in N \mid l$ is co-prime with $p)[X_1(\mathcal{G}_n) \leq \frac{l}{p} \leq X_1(\mathcal{G}_n + \xi v_1)]$. Thus the set $\{W^{\psi}_{h(z)} \mid z \in rat,$ $X_1(\mathcal{G}_n) \leq z \leq X_1(\mathcal{G}_n + \xi v_1)\}$ includes all the r.e. languages. Thus if $\mathcal{L}'$ is in **TxtEx**, then the set of all the r.e. languages would be in **TxtEx**. It is known that no **IIM** can **TxtEx**-identify the set of all the r.e. languages [Gol67]. Hence $\mathcal{L}'$ cannot be in **TxtEx**. Since $\mathcal{L}' \subseteq ExtTRANSIM_n{}^{n+1}$, $ExtTRANSIM_n \notin \mathbf{U}^{n+1}\mathbf{TxtEx}$. ∎

Proposition 15 *For all $n \in N^+$, $\boldsymbol{DU^n TxtEx} - \boldsymbol{U^{n+1} TxtEx} \neq \emptyset$.*

PROOF. The case of $n = 1$ is shown by that $\{\mathcal{K}\} \cup SINGLE \in \mathbf{TxtEx} - \mathbf{U}^2\mathbf{TxtEx}$.

Let $ExtTRANSIM_n$ be as defined in the proof for Proposition 14. Using the proof method in Proposition 14, it is easy to verify that for all $n \geq 2$, $ExtTRANSIM_n$ is in $\mathbf{DU}^n\mathbf{TxtEx}$ but not in $\mathbf{U}^{n+1}\mathbf{TxtEx}$. ∎

4.3 WDUTxtEx Hierarchy

Proposition 16 *For all $n \in N^+$, $\boldsymbol{WDU^{n+1}TxtEx} \subset \boldsymbol{WDU^{n}TxtEx}$.*

Proof. That for all $n \in N^+$, $\mathbf{WDU}^{n+1}\mathbf{TxtEx} \subseteq \mathbf{WDU}^{n}\mathbf{TxtEx}$ is immediate from definition. Proposition 15 then completes the proof. ∎

4.4 Inter-paradigm Comparisons

Proposition 17 *Let $n \in N^+$.*

(a.1) $(\boldsymbol{WDU^{*}TxtEx} \cap \boldsymbol{DU^{n}TxtEx}) - \boldsymbol{DU^{n+1}TxtEx} \neq \emptyset$.

(a.2) $(\boldsymbol{U^{*}TxtEx} \cap \boldsymbol{DU^{n}TxtEx}) - \boldsymbol{DU^{n+1}TxtEx} \neq \emptyset$.

(a.3) $(\boldsymbol{U^{*}TxtEx} \cap \boldsymbol{WDU^{n}TxtEx}) - \boldsymbol{WDU^{n+1}TxtEx} \neq \emptyset$.

(b.1) $\boldsymbol{DU^{n}TxtEx} - \boldsymbol{WDU^{n+1}TxtEx} \neq \emptyset$.

(b.2) $\boldsymbol{WDU^{n}TxtEx} - \boldsymbol{U^{n+1}TxtEx} \neq \emptyset$.

Proof. Part (b.1) and (b.2) are corollaries of Proposition 15. Part (a.2) is a corollary of part (a.1).

We now show part (a.1). For $n = 1$, it is easy to see that FIN witnesses the separation. Let $n \in N$ and $n \geq 2$. Let $\{G_1, G_2, \ldots, G_n\} \subset \Lambda_n$ be a collection of n simplexes and let $G_0 \notin \{G_1, G_2, \ldots, G_n\}$ be such that $G_0 \subseteq \bigcup_{i=1}^{n} G_i$. (Such $G_0, G_1, \ldots, G_n$ exist by Claim 11.) Let $\mathcal{L} = \{L(G_i) \mid 0 \leq i \leq n\}$. It is easy to verify that $\mathcal{L} \in \mathbf{DU}^{n}\mathbf{TxtEx} \cap \mathbf{WDU}^{*}\mathbf{TxtEx}$. However, since a text for $\bigcup_{i=0}^{n} L(G_i)$ is also a text for $\bigcup_{i=1}^{n} L(G_i)$, by Proposition 9, $\mathcal{L} \notin \mathbf{DU}^{n+1}\mathbf{TxtEx}$.

We now show part (a.3). We first observe that while a **UTxtEx** learner, in learning a class of languages $\mathcal{L}$, is allowed to conjecture languages outside of $\mathcal{L}$, a **WDUTxtEx** learner is allowed to conjecture only languages in $\mathcal{L}$. The following proof exploits this weakness in **WDUTxtEx** identification.

Let $n \in N^+$. For $i, k \in N$, let $A_{i,k} = \{\langle \lfloor i/(n+1) \rfloor \cdot (n+1) + j, \langle i, k \rangle \rangle \mid j \in N,\ 0 \leq j \leq n\} \cup \{\langle i, x \rangle \mid x \in N\}$. Intuitively, each $A_{i,k}$ has the information $\langle i, k \rangle$ deposited on $n+1$ tracks, that is, from track $\lfloor i/(n+1) \rfloor \cdot (n+1)$ to track $\lfloor i/(n+1) \rfloor \cdot (n+1) + n$; and has the track i completely filled.

Given total $g : N \to N$ and $i \in N$, let $L_{i,g} = A_{i,g(i)}$. Let $\mathcal{L}_g = \{L_{i,g} \mid i \in N\}$. Since for any g, for any collection of n languages $\mathcal{L} \subseteq \mathcal{L}_g$, each $\langle i, g(i) \rangle$ can be easily obtained from a text for $L_{\mathcal{L}}$, it is easy to verify that for all g, $\mathcal{L}_g \in$ $\mathbf{U^*TxtEx} \cap \mathbf{WDU^nTxtEx}$.

Note that for all $e \in N$, and total g, g', $\bigcup_{0 \leq j \leq n} L_{(n+1)*e+j,g} = \bigcup_{0 \leq j \leq n} L_{(n+1)*e+j,g'}$ $= \{\langle i, x \rangle \mid (n+1)*e \leq i < (n+1)*(e+1) \text{ and } x \in N\}$. That is, intuitively, track $(n+1)*e$ to track $(n+1)*e+n$ are completely filled.

We now define f such that for all e, $\{L_{(n+1)*e+j,f} \mid 0 \leq j \leq n\}$ is not the set of languages to which M_e converges on a text T_e for $\{\langle i, x \rangle \mid (n+1)*e \leq i < (n+1)*(e+1) \text{ and } x \in N\}$. For this purpose, let total g, g' be such that for all $x \in N$, $\{g(x) \mid (n+1)*e \leq x < (n+1)*(e+1)\} \neq \{g'(x) \mid (n+1)*e \leq x < (n+1)*(e+1)\}$. For each $e \in N$, for all $x \in N$ where $(n+1)*e \leq x < (n+1)*(e+1)$, let

$$f(x) = \begin{cases} g(x) & \text{if } M_e(T_e) \text{ converges to a representation} \\ & \qquad\qquad\qquad \text{index set for } \{L_{(n+1)*e+j,g'} \mid 0 \leq j \leq n\} \\ & \\ g'(x) & \text{otherwise} \end{cases}$$

Since for all e, M_e on the text T_e does not converge to a representation index set for $\{L_{(n+1)*e+j,f} \mid 0 \leq j \leq n\}$, $\mathcal{L}_f$ is not in $\mathbf{WDU}^{n+1}\mathbf{TxtEx}$. ∎

Corollary 18 *For all $n \in N^+, n \geq 2$, $\boldsymbol{DU^n TxtEx} \subset \boldsymbol{WDU^n TxtEx} \subset \boldsymbol{U^n TxtEx}$.*

PROOF. That $\mathbf{DU}^n\mathbf{TxtEx} \subseteq \mathbf{WDU}^n\mathbf{TxtEx} \subseteq \mathbf{U}^n\mathbf{TxtEx}$ is trivial. Proposition then follows from part (a.1) and (a.3) of Proposition 17. ∎

5 Sufficient Conditions For DUTxtEx Identification

5.1 Disjointness

It may be argued that languages in a union only fail to be discerningly identifiable as a result of crucial information regarding one language being lost within the other languages; that is, when all the "important" members (such as $ExtTRANSIM_n$ in the proof of Proposition 14) of one language are also members of some other languages in the union. It is natural to ask if disjointness would be a sufficient condition for unions of languages to be discerningly identifiable. The following result answers this in the negative.

Lemma 19 *There exists* $\mathcal{L} \in \boldsymbol{TxtEx}$ *where*

(a) $\emptyset \notin \mathcal{L}$, *and*

(b) for all $L, L' \in \mathcal{L}$, $L \cap L' = \emptyset$

such that $\mathcal{L} \notin \boldsymbol{U}^2\boldsymbol{TxtEx}$. [9]

Proof. Unless stated otherwise, let e, i, with or without decorations, range over N. Let σ and τ, with or without decorations, range over SEQ. For each **IIM** M_e, we construct finite set S_e and languages $L_e, L'_e \subseteq \{\langle e, x\rangle \mid x \in N\}$ such that

$$L_e = \{\langle e, 0\rangle\} \cup \{\langle e, x\rangle \mid x \in S_e\},$$

and L'_e satisfies the following two properties:

[9] *Note that* $\mathcal{L}^2$ *is uniquely definable from* $\mathcal{L}$.

1. $L'_e = \{\langle e, x\rangle \mid x \in W_{\min(\{\pi_2(y) | y \in L'_e\})}\}$,

2. $\min(\{\pi_2(x) \mid x \in L'_e\}) > \max(S_e)$.

Let $\mathcal{L} = \{L_e, L'_e \mid \ e \in N\}$. It is clear that for any pair of languages $L, L' \in \mathcal{L}$, $L \cap L' = \emptyset$. Since given any e, (1) L_e is finite and contains $\langle e, 0\rangle$, while (2) L'_e is (in a sense) *self-describing*[10] and does not contain $\langle e, 0\rangle$, it is easy to verify that $\mathcal{L} \in \mathbf{TxtEx}$.

We now show that $\mathcal{L} \notin \mathbf{U^2TxtEx}$. For each M_e here is the construction to show that M_e does not $\mathbf{U^2TxtEx}$-identify $\mathcal{L}$. By Kleene's Recursion Theorem [Rog67] there exists an index $e' > 0$ such that $W_{e'}$ may be defined in stages $s = 0, 1, 2 \ldots$, as below. For each s, $W^s_{e'}$ denotes the finite portion of $W_{e'}$ enumerated just before stage s.

Stage 0: Let $W^1_{e'} = \{e'\}$ and $\sigma^1 = \langle e, 0\rangle \diamond \langle e, e'\rangle$. Go to stage 1.

Stage s: Search for τ where $content(\tau) \subseteq \{\langle e, i\rangle \mid i > \max(W^s_{e'})\}$ such that $M_e(\sigma^s) \neq M_e(\sigma^s \diamond \tau)$. If and when τ is found, enumerate $\{i \mid \langle e, i\rangle \in content(\tau)\}$ into $W^{s+1}_{e'}$, and let $\sigma^{s+1} = \sigma^s \diamond \tau$. Go to stage $s + 1$.

If the search for τ failed at any stage s, let $L_e = content(\sigma^s)$, let an index $e'' > \max(W^s_{e'})$ be such that $\min(W_{e''}) = e''$ and let $L'_e = \{\langle e, i\rangle \mid i \in W_{e''}\}$. Since stage s does not succeed, M_e does not identify at least one of L_e and $L_e \cup L'_e$.

---

[10] In inductive inference, classes such as $\{L \subseteq N \mid W_{\min(L)} = L\}$ are called *self-describing* since grammar for languages in these classes are coded inside the language in a simple way. L'_e here satisfies similar properties.

If the search is successful in all stages, then let $L_e = \{\langle e, 0\rangle\}$ and $L'_e = \{\langle e, x\rangle \mid x \in W_{e'}\}$; now M_e fails to converge on the input $\bigcup_s \sigma^s$, a text for $L_e \cup L'_e$. ∎

The following extends this result to $\mathbf{U}^n\mathbf{TxtEx}$ for any $n \in N$. The proof uses similar ideas as those in the proof we just described.

Theorem 20 *For all $n \in N^+$, there exists $\mathcal{L} \in$ **DU^{n}TxtEx** where*

(a) $\emptyset \notin \mathcal{L}$, and

(b) for all $L, L' \in \mathcal{L}$, $L \cap L' = \emptyset$

*such that $\mathcal{L} \notin$ **U^{n+1}TxtEx**.*

Proof. Let $n \in N^+$ be given. Unless stated otherwise, let e, i, j, k, with or without decorations, range over N, and S, with or without decorations, range over finite sets. σ and τ, with or without decorations, range over SEQ. For each **IIM** M_e, we construct S_e, L_e^0, $L_e^1, \ldots, L_e^n$ where

$$L_e^0 = \{\langle e, 0, 0\rangle\} \cup \{\langle e, i, j\rangle \mid 1 \leq i \leq n,\ j \in S_e\}$$

and for $1 \leq i \leq n$, L_e^i satisfies the following two properties:

(1) $L_e^i = \{\langle e, i, j\rangle \mid j \in W_{\min(\{\pi_3(x) \mid x \in L_e^i\})}\}$

(2) $\min(\{\pi_3(x) \mid x \in L_e^i\}) > \max(S_e)$.

Let $\mathcal{L} = \{L_e^0, L_e^1, ..., L_e^n \mid e \in N\}$. It is clear that for all $L, L' \in \mathcal{L}$, $L \cap L' = \emptyset$. We shall first show that $\mathcal{L} \in \mathbf{DU}^n\mathbf{TxtEx}$. We define an auxiliary recursive function

$g : N^3 \mapsto N$ as follow. For each e and j,

$$W_{g(e,0,j)} = \{\langle e, 0, 0\rangle\} \cup \{\langle e, i, k\rangle \mid 1 \le i \le n \wedge k \in D_j\}$$

and for each $i \ge 1$ and e, j,

$$W_{g(e,i,j)} = \{\langle e, i, k\rangle \mid k \in W_j\}.$$

Now $\mathcal{L} \in \mathbf{DU}^n\mathbf{TxtEx}$ is witnessed by following M. For each text T and each $m \in N$,

$M(T[m])$:

$S \leftarrow \emptyset$.

$A \leftarrow \{j \mid (\exists w \in content(T[m]))[\pi_1(w) = j]\}$.

For each $e \in A$ do

  $B \leftarrow content(T[m])$.

  If $\langle e, 0, 0\rangle \in content(T[m])$ then

    $C \leftarrow \{j \mid (\forall i, 1 \le i \le n)[\langle e, i, j\rangle \in content(T[m])]\}$.

    Let j be such that $D_j = C$.

    $S \leftarrow S \cup \{g(e, 0, j)\}$.

    $B \leftarrow B - W_{g(e,0,j)}$.

  For $i \leftarrow 1$ to n do

    If exists j_0 such that $\langle e, i, j_0\rangle \in B$, then

      For minimum such j_0, let $S \leftarrow S \cup \{g(e, i, j_0)\}$.

Output S.

It is easy to verify that M $\mathbf{DU}^n\mathbf{TxtEx}$-identifies $\mathcal{L}$.

We now show that $\mathcal{L} \notin \mathbf{U}^{n+1}\mathbf{TxtEx}$. For each M_e here is the construction to show that M_e does not $\mathbf{U}^{n+1}\mathbf{TxtEx}$-identify $\mathcal{L}$. By Kleene's Recursion Theorem [Rog67] there exists an index e' such that $W_{e'}$ may be defined in stages $s = 0, 1, 2\ldots$, as below. For each s, $W_{e'}^s$ denotes the finite portion of $W_{e'}$ enumerated just before stage s.

Stage 0: Let $\sigma^1 = \langle e, 0, 0\rangle \diamond \langle e, 1, e'\rangle \diamond \langle e, 2, e'\rangle \diamond \ldots \diamond \langle e, n, e'\rangle$. Let $W_{e'}^1 = \{e'\}$. Go to stage 1.

Stage s: Search for τ where $\mathit{content}(\tau) \subseteq \{\langle e, i, j\rangle \mid 1 \leq i \leq n \;\wedge\; j > \max(W_{e'}^s)\}$ such that $M_e(\sigma^s) \neq M_e(\sigma^s \diamond \tau)$. If and when τ is found, enumerate $\{j \mid (\exists i', 1 \leq i' \leq n)[\langle e, i', j\rangle \in \mathit{content}(\tau)]\}$ into $W_{e'}$, and let σ^{s+1} be an extension of σ^s such that $\mathit{content}(\sigma^{s+1}) = \{\langle e, 0, 0\rangle\} \cup \{\langle e, i, j\rangle \mid 1 \leq i \leq n \wedge j \in W_{e'}$ enumerated up to now$\}$. Go to stage $s+1$.

If the search for τ failed at any stage s, then let $L_e^0 = \mathit{content}(\sigma^s)$, let an index $e'' > \max(W_{e'}^s)$ be such that $\min(W_{e''}) = e''$. For each $i \in N$, $1 \leq i \leq n$, let $L_e^i = \{\langle e, i, j\rangle \mid j \in W_{e''}\}$. Since stage s does not succeed, M_e does not identify at least one of L_e^0 and $(L_e^0 \cup \bigcup_{i=1}^n L_e^i)$.

If the search is successful at all stages, let $L_e^0 = \{\langle e, 0, 0\rangle\}$ and for each $i \in N$, $1 \leq i \leq n$ let $L_e^i = \{\langle e, i, x\rangle \mid x \in W_{e'}\}$, then M_e fails to converge on the input

$\bigcup_s \sigma^s$, a text for $L_e^0 \cup \bigcup_{i=1}^n L_e^i$. ∎

In obtaining the above result we have used a class of languages that is not recursively enumerable. It remains to be seen if for recursively enumerable classes of languages, disjointness can be sufficient as a condition for $\mathbf{U}^n\mathbf{TxtEx}$ identification for all $n > 0$. The following example, however, shows the contrary.

Example 21 For $x, y \in N$, let

$$L_{x,0} = \{\langle x,0\rangle\} \cup \{\langle x,y\rangle \mid (\forall z \in N, z \leq y)[\varphi_x(z)\downarrow]\}$$

$$L_{x,y+1} = \begin{cases} \{\langle x,y+1\rangle\} & \text{if } \langle x,y+1\rangle \notin L_{x,0} \\ L_{x,0} & \text{otherwise} \end{cases}$$

Let $\mathcal{L} = \{L_{x,y} \mid x, y \in N\}$. It is clear that $\mathcal{L}$ is recursively enumerable and pairwise disjoint. That $\mathcal{L} \in \mathbf{TxtEx}$ is observed by M below. For each sequence σ,

$M(\sigma)$:

Search for $x \in N$ such that $\langle x,0\rangle \in content(\sigma)$.

If and when found, return the standard index for the language $L_{x,0}$.

If no such x exists, return the standard index for the language $content(\sigma)$.

Towards a contradiction suppose there exists M' such that $\mathcal{L}^2 \subseteq \mathbf{TxtEx}(M')$. Now for each $x \in N$, if

$\varphi_x \in \mathcal{R}$, then $L_{x,0} = \{\langle x,y\rangle \mid y \in N\}$. (More precisely, for all $z \in N$, $L_{x,z} = \{\langle x,y\rangle \mid y \in N\}$.) Since $L_{x,0} \in \mathbf{TxtEx}(M')$, by Lemma 1, there exists a sequence σ where $content(\sigma) \subseteq L_{x,0}$, such that for all extensions τ of σ where $content(\tau) \subseteq \{\langle x,y\rangle \mid y \in N\}$, $M'(\tau) = M'(\sigma)$.

$\varphi_x \notin \mathcal{R}$, then there exists $z \in N$ such that $L_{x,z} \not\subseteq L_{x,0}$. Since M' **TxtEx**-identifies both $L_{x,0}$ and $(L_{x,0} \cup L_{x,z})$, for all sequences σ where $content(\sigma) \subseteq L_{x,0}$,

1. if $W_{M'(\sigma)} \neq L_{x,0}$, then there exists an extension τ of σ, where $content(\tau) \subseteq L_{x,0}$, such that $W_{M'(\tau)} = L_{x,0}$, and
2. if $W_{M'(\sigma)} = L_{x,0}$, then there exists an extension τ of σ, where $content(\tau) \subseteq (L_{x,0} \cup L_{x,z})$, such that $W_{M'(\tau)} = (L_{x,0} \cup L_{x,z})$.

In summary, if $\varphi_x \notin \mathcal{R}$, then for all sequences σ where $content(\sigma) \subseteq L_{x,0}$, there exists an extension τ of σ, where $content(\tau) \subseteq \{\langle x, y\rangle \mid y \in N\}$, such that $M'(\tau) \neq M'(\sigma)$.

Hence we have

$$\begin{aligned}\varphi_x \in \mathcal{R} \Leftrightarrow\ & (\exists \sigma \in \mathrm{SEQ} \mid content(\sigma) \subseteq L_{x,0}) \\ & \qquad [(\forall \tau \supset \sigma \mid content(\tau) \subseteq \{\langle x, y\rangle \mid y \in N\})\ [M'(\tau) = M'(\sigma)]]\end{aligned}$$

Now the condition on the right hand side is Σ_2 to check. However, the set $\{x \mid \varphi_x \text{ is recursive}\}$ is not Σ_2, a contradiction. □

5.2 Functions That Enumerate Distinguishing Elements

Let recursively enumerable $\mathcal{L} \subseteq \mathcal{E}$ be given. Suppose for all $L \in \mathcal{L}$, there is an effective procedure to enumerate an element which is uniquely in L, that is, no other language in $\mathcal{L}$ contains this element. Can we then identify every collection of languages drawn from $\mathcal{L}$? An answer is attempted in the following proposition.

Proposition 22 *Let $\mathcal{L}$ be a 1–1 recursively enumerable class of languages as wit-*

nessed by the numbering ψ. If there exist a limiting recursive function $d : N \to N$ and total recursive $F : N^2 \to N$ for which $d(i) = \lim_{t\to\infty} F(i,t)$ such that

(a) for all $i \in N$, $d(i) \in W_i^{\psi}$,

(b) for all $i, j \in N$, $d(i) \in W_j^{\psi} \Rightarrow i = j$, and

(c) for all $j \in N$, $card(range(F) \cap W_j^{\psi}) < \infty$.

*Then $\mathcal{L} \in$ **DU* TxtEx**.*

PROOF. Let $\mathcal{L}$, ψ, F, d be as in proposition. Let recursive function h witnesses that $\psi \prec \varphi$. Unless stated otherwise, we let i, j, with or without decorations, range over N. Define M as follows, such that for each text T and for each $m \in N$,

$M(T[m])$:

$S \leftarrow \emptyset$.

For $i = 0$ to m do

If $[F(i,m) \in content(T[m]) \cap W_{i,m}^{\psi}]$

and $\neg[(\exists i', j')[i' < m \ \wedge \ j' < m \wedge i' \neq j'$

$\wedge \ F(i,m) \in W_{i',m}^{\psi} \cap W_{j',m}^{\psi}]]$

Then insert $h(i)$ into S.

Output S.

We claim that M **DU*TxtEx**-identifies $\mathcal{L}$. Let $\mathcal{L}' \subseteq \mathcal{L}$ be a finite collection of languages. Let D be such that $\{W_i^{\psi} \mid i \in D\} = \mathcal{L}'$. Let T be a text for $L = \bigcup_{i\in D} W_i^{\psi}$, and let $A = \text{range}(F) \cap L$. By clause (c) in the proposition, $card(A) < \infty$. Intuitively, A contains all the potential "distinguishing element"s M will encounter during the identification process. Since D and A are finite, there exists $n \in N$ so large that

(1) For all $i \in D$, $(\forall t \in N, t > n)$ $[F(i,t) = d(i) \wedge d(i) \in \mathit{content}(T[t]) \cap W^{\psi}_{i,t}]$.

(2) For all $x \in A - \{d(k) \mid k \in D\}$, $(\forall n' \in N, n' > n)$ $[(\exists j \in N - D)[x \in W^{\psi}_{j,n'}]$

$\Rightarrow (\exists i', j' < n)\ [i' \neq j' \wedge x \in W^{\psi}_{i',n'} \cap W^{\psi}_{j',n'}]]$

Clause (1) ensures that all $i \in D$ will eventually be output by M. Clause (2) ensures that all programs $j \notin D$, which enumerate some element in A are excluded from consideration (note that every element in A is enumerated by some program in D).

Hence for all $n' \in N$ where $n' > n$, $i \in D$ if and only if $i \in M(T[n'])$. It follows that M **DU*****TxtEx**-identifies $\mathcal{L}$. ∎

Corollary 23 *Let $\mathcal{L}$ be a class of languages for which there exists a 1–1 numbering and that*

(a) $\emptyset \notin \mathcal{L}$.

(b) for all $L, L' \in \mathcal{L}$, $L \neq L' \Rightarrow L \cap L' = \emptyset$

*then $\mathcal{L} \in$ **DU*** **TxtEx**.*

Proof. Let $\mathcal{L} = \{W^{\psi}_i \mid i \in N\}$ where ψ is a 1–1 numbering for $\mathcal{L}$. For $i, t \in N$, let $F(i,t) = \min(W^{\psi}_{i,t})$ and let $d(i) = \lim_{t \to \infty} F(i,t)$. Clearly,

(a) for all $i \in N$, $d(i) \in W^{\psi}_i$,

(b) for all $i, j \in N$, $d(i) \in W^{\psi}_j \Rightarrow i = j$, and

(c) for all $j \in N$, $\mathit{card}(\mathit{range}(F) \cap W^{\psi}_j) < \infty$.

d thus fulfills all the conditions for Proposition 22. Hence $\mathcal{L} \in$ **DU*****TxtEx**. ∎

In Proposition 22, some weaker conditions for (a) and (b) may not be sufficient, even if d is a recursive function. For instance, if we have only the following conditions:

(a) for all $i \in N$, $d(i) \in W_i^\psi$,

(b) for all $i \in N$, $card(\{j \in N \mid d(i) \in W_j^\psi\}) < \infty$,

(c) d is recursive.

then identifiability for $\mathcal{L}^2$ cannot be guaranteed, as the following example shows.

Example 24 For $i \in N$, let

$$
\begin{aligned}
L_0 &= \{\langle 0, x\rangle \mid x \in N\} \cup \{\langle 1, x\rangle \mid x \in \mathcal{K}\} \\
L_{i+1} &= \begin{cases} \{\langle 0, i+1\rangle\} \cup \{\langle 1, i\rangle\} \cup \{\langle 2, i\rangle\} & \text{if } i \in \mathcal{K} \\ \{\langle 0, i+1\rangle\} \cup \{\langle 1, i\rangle\} & \text{otherwise} \end{cases}
\end{aligned}
$$

Let $\mathcal{L} = \{L_i \mid i \in N\}$ and define d such that for all $x \in N$, $d(x) = \langle 0, x\rangle$. It is easy to verify that (a) $\mathcal{L}$ is 1–1 recursively enumerable, (b) $\mathcal{L} \in \mathbf{TxtEx}$, (c) $\mathcal{L}^2$ is uniquely definable from $\mathcal{L}$, and (d) d satisfies all the conditions given above for $\mathcal{L}$. However, for all $k \in N$, the language $\{\langle 0, i\rangle \mid i \in N\} \cup \{\langle 1, x\rangle \mid x \in \mathcal{K} \cup \{k\}\}$ is in $\mathcal{L}^2$, hence (by Proposition 4.7 in [JORS99]) $\mathcal{L}^2$ is not in **TxtEx**. □

A similar set of weaker conditions for (a) and (b), where instead of a single unique element d is required to name only a set of elements which is unique to each language in the class, as in the following:

(a) for all $i \in N$, $D_{d(i)} \subseteq W_i^\psi$,

(b) for all $i, j \in N$, $D_{d(i)} \subseteq W_j^\psi \Rightarrow i = j$.

(c) d is recursive.

then such a function will also fail to guarantee that $\mathcal{L}^2 \in \mathbf{TxtEx}$, as demonstrated by the following example.

Example 25 For $i \in N$, let

$$
\begin{aligned}
L_0 &= \{\langle 0,0\rangle\} \cup \{\langle 1,x\rangle \mid x \in N\} \\
L_1 &= \{\langle 1,1\rangle\} \cup \{\langle 0,x\rangle \mid x \in N\} \cup \{\langle 2,x\rangle \mid x \in \mathcal{K}\} \\
L_{i+2} &= \begin{cases} \{\langle 0,i+2\rangle\} \cup \{\langle 1,x\rangle \mid x \in N\} \cup \{\langle 2,i\rangle\} \cup \{\langle 3,i\rangle\} & \text{if } i \in \mathcal{K} \\ \{\langle 0,i+2\rangle\} \cup \{\langle 1,x\rangle \mid x \in N\} \cup \{\langle 2,i\rangle\} & \text{otherwise} \end{cases}
\end{aligned}
$$

Let $\mathcal{L} = \{L_i \mid i \in N\}$ and define d such that for all $x \in N$, $d(x) = \{\langle 0,x\rangle, \langle 1,x\rangle\}$. It is easy to verify that $\mathcal{L}$ is a 1–1 recursively enumerable class of languages in $\mathbf{TxtEx}$ where all the languages in $\mathcal{L}^2$ are uniquely definable from $\mathcal{L}$, and that d satisfies all the conditions above for $\mathcal{L}$. However, for all $k \in N$, the language $\{\langle 0,i\rangle, \langle 1,i\rangle \mid i \in N\} \cup \{\langle 2,x\rangle \mid x \in \mathcal{K} \cup \{k\}\}$ is in $\mathcal{L}^2$, thus (by Proposition 4.7 in [JORS99]) $\mathcal{L}^2$ is not in $\mathbf{TxtEx}$. □

5.3 Restrictions On Structures Of Languages

Proposition 26 *Given $n \in N^+$. Let $\mathcal{L}$ be a class of languages such that*

(a) every language in $\mathcal{L}^n$ is uniquely definable from $\mathcal{L}$,

(b) for all $L \in \mathcal{L}$, card$(\{L' \in \mathcal{L} \mid L' \cap L \neq \emptyset\}) < \infty$,

(c) there exists a computable numbering ψ for $\mathcal{L}$ such that:

(1) for all $L \in \mathcal{L}$, card$(L) = \infty \Rightarrow$ card$(\{i \mid W_i^\psi = L\}) = 1$.

(2) for all $L \in \mathcal{L}$, card$(L) < \infty \Rightarrow$ card$(\{i \mid W_i^\psi = L\}) < \infty$.

*then $\mathcal{L} \in$ **DU**^{n}**TxtEx**.*

PROOF. Let $n \in N^+$ be given. Let $\mathcal{L}$ be as in proposition. Unless stated otherwise, we let i, j, k, m, n, with or without decorations, range over N. We let A and B, with or without decorations, range over *FIN*. Let h witnesses that $\psi \prec \varphi$. Define **IIM** M as follows such that for each text T,

$M(T[m])$:

Let $\mathcal{C}^m = \{i \mid i \leq m \ \wedge \ W^{\psi}_{i,m} \cap content(T[m]) \neq \emptyset\}$.

Let $Candidates^m = \{S \subseteq \mathcal{C}^m \mid card(S) \leq n\}$.

Let $s_0 = \max(\{s \in N \mid (\exists S \in Candidates^m)[\ \bigcup_{i \in S} W^{\psi}_{i,s} \subseteq content(T[m])$

$\wedge\ \bigcup_{i \in S} W^{\psi}_{i,m} \supseteq content(T[s])]\})$.

Output $\{h(i) \mid i \in D_{k_0}\}$, where $k_0 = \min(\{k \mid D_k \in Candidates^m$

$\wedge\ \bigcup_{i \in D_k} W^{\psi}_{i,s_0} \subseteq content(T[m])$

$\wedge\ \bigcup_{i \in D_k} W^{\psi}_{i,m} \supseteq content(T[s_0])\})$.

Intuitively, M outputs the seemingly best grammar set in $Candidates^m$ which describes the input text. We claim that M $\mathbf{DU}^n\mathbf{TxtEx}$-identifies $\mathcal{L}$. Let $\mathcal{L}' \subseteq \mathcal{L}$ be where $card(\mathcal{L}') \leq n$. Let B be such that $\{W^{\psi}_i \mid i \in B\} = \mathcal{L}'$. Let T be a text for $L = \bigcup_{i \in B} W^{\psi}_i$. We divide B into two groups, $B_1 = \{i \in B \mid W^{\psi}_i \text{ is finite}\}$ and $B_2 = \{i \in B \mid W^{\psi}_i \text{ is infinite}\}$. By the requirement of ψ, for each $i \in B_1$, there exist only finitely many j such that $W^{\psi}_i = W^{\psi}_j$, and for each $i \in B_2$, for all $j \neq i$, $W^{\psi}_i \neq W^{\psi}_j$. Let $\mathcal{A} = \{A \mid \bigcup_{i \in A} W^{\psi}_i = \bigcup_{i \in B_1} W^{\psi}_i\}$. Since $\mathcal{L}^n$ is uniquely definable from $\mathcal{L}$, the only sets of languages which are capable of generating L are $\{B_2 \cup A \mid A \in \mathcal{A}\}$. Let $CorrectInd = \{B_2 \cup A \mid A \in \mathcal{A}\}$.

Let $\mathcal{C}' = \{i \mid W_i^\psi \cap content(T) \neq \emptyset\}$. Since each language in $\{W_i^\psi \mid i \in B\}$ intersects with only finitely many other languages in $\mathcal{L}$, $\mathcal{C}'$ is finite. It is easy to verify that there exists n_0 such that for all $n' > n_0$, $\mathcal{C}^{n'} = \mathcal{C}^{n'+1} = \mathcal{C}'$. Let $n_1 > n_0$ be so large that

$$(\forall i \in \mathcal{C}')[W_i^\psi \text{ is finite} \Rightarrow W_i^\psi = W_{i,n_1}^\psi$$
$$\wedge\ W_i^\psi \cap content(T[n_1]) = W_i^\psi \cap content(T)]$$

Let $Candidates' = Candidates^{n_1}$. Clearly, for all $n' > n_0$, $Candidates^{n'} = Candidates^{n'+1} = Candidates'$. Let $n_2 > n_1$ be so large that

$$\neg[(\exists A \in Candidates' - CorrectInd)[\ (\bigcup_{i \in A} W_{i,n_2}^\psi \subseteq L)$$
$$\wedge\ (\bigcup_{i \in A} W_i^\psi \supseteq content(T[n_2]))]]$$

Let $n_3 > n_2$ be so large that

$$[\bigcup_{i \in B} W_{i,n_2+1}^\psi \subseteq content(T[n_3]) \wedge \bigcup_{i \in B} W_{i,n_3}^\psi \supseteq content(T[n_2+1])]$$

Clearly, for all $n' > n_3$, $\{D \in Candidates' \mid \bigcup_{i \in D} W_{i,n_2+1}^\psi \subseteq content(T[n']) \wedge \bigcup_{i \in D} W_{i,n'}^\psi \supseteq content(T[n_2+1])\} = CorrectInd$. Hence for all $n' > n_3$, M outputs $\min(\{k \mid D_k \in CorrectInd\})$. It follows that M $\mathbf{DU}^n\mathbf{TxtEx}$-identifies $\mathcal{L}$. ∎

Corollary 27 *Fix $n \in N^+$. Let $\mathcal{L} = \{L_i \mid i \in N\}$ be a 1–1 recursively enumerable class of languages where*

(a) every language in $\mathcal{L}^n$ is uniquely definable from $\mathcal{L}$.

(b) for all $i \in N$, $card(\{j \mid L_i \cap L_j \neq \emptyset\}) < \infty$.

Then $\mathcal{L} \in \boldsymbol{DU^n TxtEx}$.

The conditions in Proposition 26 are not necessary — as is evident by $TRANSIM_n$, which is 1–1 recursively enumerable but every language in the class intersects with infinitely many other languages within the class.

6 Hierarchies For Other Criteria of Identification

In this chapter we extend our study of the learning of unions of languages, beyond the context of **TxtEx** identification, to other identification criteria. We are mainly interested in identification criteria similar to **TxtBc**, **TxtFex**, **NTxtEx**, **InTxtEx**, **ImTxtEx** and **InfEx**.

6.1 Generalizing The U, DU and WDU Paradigms

For the purpose of this chapter, we assume an identification criterion to include (1) a procedure of relating a language with a text which is used as an (infinite) input stream to an **IIM**, and (2) a definition of what constitute an acceptable sequence of output by the **IIM**, upon given the input text.

6.1.1 Input Text

Given an identification criterion $\mathcal{I}$, *a text* T *is* $\mathcal{I}$*-admissible for language* L just in case in $\mathcal{I}$ identification, T is used as input (to **IIM**s) to represent the language L.

6.1.2 Sequence of Grammars

We say that an infinite sequence of grammars, G, is $\mathcal{I}$*-admissible for language* L *on text* T just in case G is an infinite sequence of grammars witnessing $\mathcal{I}$-identification of L on input T. We denote the j-th grammar in the sequence G by $G(j)$.

In this chapter, we further make the following assumption.

Assumption 1 An identification criterion $\mathcal{I}$ in this chapter is assumed to satisfy the following conditions.

(a) An **IIM** M is said to $\mathcal{I}$-identify a text T for language L just in case T is an $\mathcal{I}$-admissible text for L, and $M(T[0]), M(T[1]), M(T[2]), \ldots$ is an $\mathcal{I}$-admissible sequence of grammars for L on T.

(b) An **IIM** M is said to $\mathcal{I}$-identify a language L just in case M $\mathcal{I}$-identify each $\mathcal{I}$-admissible text T for L.

(c) An **IIM** M is said to $\mathcal{I}$-identify a class of languages $\mathcal{L}$ just in case M $\mathcal{I}$-identify each $L \in \mathcal{L}$.

This assumption immediately precludes identification criteria where the learner is expected to identify languages outside of the identifiable class. For $n \in N$, the identification criteria $\mathbf{U}^{n+2}\mathbf{TxtEx}$, $\mathbf{DU}^{n+2}\mathbf{TxtEx}$ and $\mathbf{WDU}^{n+2}\mathbf{TxtEx}$ fails the assumption.

This assumption also precludes some identification criteria where a learner is expected to respond in certain ways even on texts for languages outside of the identifiable class. For example, reliable learning [Min76] or refutable learning [MA93, LW94, MA95, Jai98].

We also note that for all $i, m, n \in N$, this assumption is valid for the identification criteria $\mathbf{TxtEx}^n$, $\mathbf{TxtBc}^n$, $\mathbf{TxtFex}^n_m$, $\mathbf{N}^i\mathbf{TxtEx}^n$, $\mathbf{In}^i\mathbf{TxtEx}^n$ and $\mathbf{Im}^i\mathbf{TxtEx}^n$.

6.1.3 Text Translation and Text Decomposable Identification Criteria

Given a set of languages $\mathcal{L}$ and a sequence σ for $L_{\mathcal{L}}$, we want to be able to express a sequence which is relevant to only one language in $\mathcal{L}$, which preserves the order of data presentation in σ. Towards this end we define the function tr below.

Given a sequence σ, $f, g, h \in \mathcal{R}$ and $L, L' \in \mathcal{E}$, let $\mathrm{tr}(\sigma, f, g, h, L, L')$ denote a sequence where for all $i \in N$,

$$\mathrm{tr}(\sigma, f, g, h, L, L')(i) = \begin{cases} f(\sigma(i)) & \text{if } \sigma(i) \in L \\ g(\sigma(i)) & \text{if } \sigma(i) \in L' - L \\ h(\sigma(i)) & \text{otherwise} \end{cases}$$

For any given text T, we also let $tr(T, f, g, h, L, L')$ be a text where for all $i \in N$,

$$\mathrm{tr}(T, f, g, h, L, L')(i) = \begin{cases} f(T(i)) & \text{if } T(i) \in L \\ g(T(i)) & \text{if } T(i) \in L' - L \\ h(T(i)) & \text{otherwise} \end{cases}$$

Let $ERASERS = \{f \in \mathcal{R} \mid (\forall x \in N)[f(x) \text{ is either } x \text{ or } \#]\}$. We say that an identification criterion $\mathcal{I}$ is *text decomposable* just in case for all $\mathcal{L} \subseteq \mathcal{E}$, if T is an $\mathcal{I}$-admissible text for $L_{\mathcal{L}}$, then for all $L \in \mathcal{L}$ and for all $f \in ERASERS$, $\mathrm{tr}(T, \lambda x[x], \lambda x[\#], f, L, L_{\mathcal{L}})$ is an $\mathcal{I}$-admissible text for L. Intuitively, $ERASERS$ is used to allow adversary flexibility outside the relevant $(L_{\mathcal{L}})$ region.

It is easy to verify that for all $i, m, n \in N$, the identification criteria $\mathbf{TxtEx}^n$, $\mathbf{TxtBc}^n$, $\mathbf{TxtFex}^n_m$, $\mathbf{N}^i\mathbf{TxtEx}^n$, $\mathbf{In}^i\mathbf{TxtEx}^n$, $\mathbf{Im}^i\mathbf{TxtEx}^n$ are text decomposable.

6.1.4 Sequence of Conjectures (Sequence of Sets of Grammars)

In attempting to discern the languages in a union, each conjecture of the learner may consist of several grammars. Given $i, j \in N$, we denote the j-th conjecture in a sequence $\mathcal{G}$ by $\mathcal{G}(j)$, and denote the i-th grammar in the j-th conjecture by $\mathcal{G}(j)[i]$. A *sequence of conjectures* is then an infinite sequence of sets of grammars $\{\mathcal{G}(0)[0], \mathcal{G}(0)[1], \mathcal{G}(0)[2], \ldots\}, \{\mathcal{G}(1)[0], \mathcal{G}(1)[1], \mathcal{G}(1)[2], \ldots\}, \{\mathcal{G}(2)[0], \mathcal{G}(2)[1], \mathcal{G}(2)[2], \ldots\}, \ldots$. Unless stated otherwise, we let $\mathcal{G}$, with or without decorations, range over sequences of conjectures. For any conjecture $\mathcal{G}(j)$, $card(\mathcal{G}(j))$ is the number of grammars conjectured in $\mathcal{G}(j)$.

Given text decomposable identification criterion $\mathcal{I}$ and a finite collection of languages $\mathcal{L} = \{L_1, L_2, \ldots, L_{card(\mathcal{L})}\}$. We say that a sequence of conjectures $\mathcal{G}$ is *$\mathcal{I}$-admissible for $\mathcal{L}$ on input T* just in case there exists $n \in N$ such that

(1) for all $m \in N$ where $m > n$, $card(\mathcal{G}(m)) = card(\mathcal{L})$,

(2) for each $m \in N$, $m > n$, there exists a permutation $i_1^m, i_2^m, \ldots, i_{card(\mathcal{L})}^m$ of $0, 1, 2, \ldots, card(\mathcal{L}) - 1$, such that for each $k \in N$, $1 \leq k \leq card(\mathcal{L})$, $\mathcal{G}(n+1)[i_k^{n+1}], \mathcal{G}(n+2)[i_k^{n+2}], \mathcal{G}(n+3)[i_k^{n+3}], \ldots$ is $\mathcal{I}$-admissible for L_k on the texts $\mathrm{tr}(T, \lambda x[x], \lambda x[\#], f, L_k, L_{\mathcal{L}})$ for all $f \in ERASERS$.[11]

Intuitively, this says that any sequence of conjectures can be re-arranged into n sequences of grammars, such that each sequence of grammars is $\mathcal{I}$-admissible for a language in $\mathcal{L}$.

[11]Note that since $\mathcal{I}$ is text decomposable, each text $\mathrm{tr}(T, \lambda x[x], \lambda x[\#], f, L_k, L_{\mathcal{L}})$ where $f \in ERASERS$ is $\mathcal{I}$-admissible for L_k.

6.1.5 The Generalized U Paradigm

Definition 27 Given $n \in N^+$, text decomposable $\mathcal{I}$, and $\mathcal{L} \subseteq \mathcal{E}$.

(a) M $\mathbf{U}^n\mathcal{I}$-identifies $\mathcal{L}$ just in case for all $L \in \mathcal{L}^n$, for all $\mathcal{I}$-admissible T for L, $M(T[0]), M(T[1]), M(T[2]), \ldots$ is an $\mathcal{I}$-admissible sequence of grammars for L on T.

(b) $\mathbf{U}^n\mathcal{I} = \{\mathcal{L} \mid (\exists M)[M\ \mathbf{U}^n\mathcal{I}\text{-identifies } \mathcal{L}]\}$.

6.1.6 The Generalized DU Paradigm

Definition 28 Given $n \in N^+$, text decomposable $\mathcal{I}$, and $\mathcal{L} \subseteq \mathcal{E}$.

(a) M $\mathbf{DU}^n\mathcal{I}$-identifies $\mathcal{L}$ just in case for all $\mathcal{L}' \subseteq \mathcal{L}$ where $card(\mathcal{L}') \leq n$, for all $\mathcal{I}$-admissible text T for $L_{\mathcal{L}'}$, $M(T[0]), M(T[1]), M(T[2]), \ldots$ is an $\mathcal{I}$-admissible sequence of conjectures for $\mathcal{L}'$ on T.

(b) $\mathbf{DU}^n\mathcal{I} = \{\mathcal{L} \mid (\exists M)[M\ \mathbf{DU}^n\mathcal{I}\text{-identifies } \mathcal{L}]\}$.

6.1.7 The Generalized WDU Paradigm

Definition 29 Given $n \in N^+$, text decomposable $\mathcal{I}$, and $\mathcal{L} \subseteq \mathcal{E}$.

(a) M $\mathbf{WDU}^n\mathcal{I}$-identifies $\mathcal{L}$ just in case for all $\mathcal{L}' \subseteq \mathcal{L}$ where $card(\mathcal{L}') \leq n$, for all $\mathcal{I}$-admissible text T for $L_{\mathcal{L}'}$, there exists $\mathcal{L}'' \subseteq \mathcal{L}$ where $card(\mathcal{L}'') \leq n$ such that T is an $\mathcal{I}$-admissible text for $L_{\mathcal{L}''}$, and $M(T[0]), M(T[1]), M(T[2]), \ldots$ is an $\mathcal{I}$-admissible sequence of conjectures for $\mathcal{L}''$ on T.

(b) $\mathbf{WDU}^n\mathcal{I} = \{\mathcal{L} \mid (\exists M)[M\ \mathbf{WDU}^n\mathcal{I}\text{-identifies } \mathcal{L}]\}$.

It is easy to verify that these definitions are consistent with our earlier definitions for **UTxtEx**, **DUTxtEx** and **WDUTxtEx**.

In this chapter we consider identifications dependent only on total **IIM**s. We let $M_0, M_1, \ldots$ denote a recursive sequence of total **IIM**s, such that every class of languages in every identification criteria in this chapter is identified by at least one of the machines in the sequence [OSW86].

6.2 General Results

It is difficult to derive results for all identification criteria in general since their definitions may be conceivably erratic. Much of this section is hence, devoted to abstracting attributes of identification criteria required for our results.

6.2.1 Text Insensitivity

Definition 30 An identification criterion $\mathcal{I}$ is said to be *text insensitive* just in case for all sequences of grammar G and all $L \in \mathcal{E}$, if there exists $\mathcal{I}$-admissible text T for L such that G is $\mathcal{I}$-admissible for L on T, then for all $\mathcal{I}$-admissible text T' for L, G is $\mathcal{I}$-admissible for L on T'.

Hence for a text insensitive identification criterion $\mathcal{I}$, saying that *a sequence of grammars is $\mathcal{I}$-admissible for L* means the same thing as saying that the sequence of grammars is $\mathcal{I}$-admissible for L on input T.

It can be verified that for a text decomposable, text insensitive identification criterion $\mathcal{I}$, for all $n \in N^+$, $\mathcal{K} \in \{\mathbf{U}, \mathbf{DU}, \mathbf{WDU}\}$, for all sequences of conjectures $\mathcal{G}$ and all $\mathcal{L} \subseteq \mathcal{E}$, if there exists $\mathcal{K}^n\mathcal{I}$-admissible text T for $L_{\mathcal{L}}$ such that $\mathcal{G}$ is

$\mathcal{K}^n\mathcal{I}$-admissible for $\mathcal{L}$ on T, then for all $\mathcal{K}^n\mathcal{I}$-admissible text T' for $L_{\mathcal{L}}$, $\mathcal{G}$ is $\mathcal{K}^n\mathcal{I}$-admissible for $\mathcal{L}$ on T'.

We note that for all $i, m, n \in N$, $\mathbf{TxtEx}^n$, $\mathbf{TxtBc}^n$, $\mathbf{TxtFex}^n_m$, $\mathbf{N}^i\mathbf{TxtEx}^n$, $\mathbf{In}^i\mathbf{TxtEx}^n$, $\mathbf{Im}^i\mathbf{TxtEx}^n$ are all text insensitive, while identification criteria which places restrictions on the conjectures learners are allowed to make based on the input (e.g., consistent [Ang80b] and set-driven learning [WC80]) are typically not.

6.2.2 extCyl-cylindrification

Below we define a cylindrification method required for our result.

Definition 31 Given $k \in N^+$ and a class of languages $\mathcal{L} = \{L_i \mid i \in N\}$.

(a) Let $\mathrm{extCyl}(L, k) = \{\langle k, x\rangle \mid x \in L\} \cup \{\langle 0, y\rangle \mid y \in N\}$.

(b) Let $\mathrm{extCyl}(\mathcal{L}, k) = \{\mathrm{extCyl}(L_i, k) \mid i \in N\}$.

(c) An identification criteria $\mathcal{I}$ is *unaidable by extCyl-cylindrification* just in case $\mathcal{L} \notin \mathcal{I}$ implies that for all $k > 0$, $\mathrm{extCyl}(\mathcal{L}, k) \notin \mathcal{I}$.

(d) A text insensitive identification criteria $\mathcal{I}$ is *unhindered by extCyl-cylindrification* just in case there exists recursive function $f : N^2 \mapsto N$ (we say f witnesses that $\mathcal{I}$ is unhindered by extCyl-cylindrification) such that if $j_1, j_2, j_3 \ldots$ is an $\mathcal{I}$-admissible sequence of grammars for L, then for all $k > 0$, $f(j_1, k), f(j_2, k),\ f(j_3, k), \ldots$ is an $\mathcal{I}$-admissible sequence of grammars for $\mathrm{extCyl}(L, k)$.

(e) An identification criteria $\mathcal{I}$ is said to *use extCyl-decylindrifiable text* just in case for all $k > 0$, if T is an $\mathcal{I}$-admissible text for $\mathrm{extCyl}(L, k)$, then $\mathrm{tr}(T, \lambda x[\pi_2(x)], \lambda x[\#], \lambda x[\#], \{x \in N \mid \pi_1(x) = k\}, \emptyset)$ is an $\mathcal{I}$-admissible text for L.

It is easy to verify that for all $i, m, n \in N$, $\mathbf{TxtEx}^n$, $\mathbf{TxtBc}^n$, $\mathbf{TxtFex}^n_m$, $\mathbf{N}^i\mathbf{TxtEx}^n$, $\mathbf{In}^i\mathbf{TxtEx}^n$, $\mathbf{Im}^i\mathbf{TxtEx}^n$ are unaidable by extCyl-cylindrification. Suppose otherwise. Let $\mathcal{I} \in \{$ $\mathbf{TxtEx}^n$, $\mathbf{TxtBc}^n$, $\mathbf{TxtFex}^n_m$, $\mathbf{N}^i\mathbf{TxtEx}^n$, $\mathbf{In}^i\mathbf{TxtEx}^n$, $\mathbf{Im}^i\mathbf{TxtEx}^n\}$, $k \in N^+$ and $\mathcal{L} \notin \mathcal{I}$ be given, where $\mathrm{extCyl}(\mathcal{L}, k) \in \mathcal{I}$ as witnessed by M. For all $\sigma \in \mathrm{SEQ}$, let τ_σ be a sequence where $content(\tau_\sigma)$ $= \{\langle k, x\rangle \mid x \in content(\sigma)\} \cup \{\langle 0, x\rangle \mid x \leq |\sigma|\}$, and let $M'(\sigma) = e$, where $W_e = \{x \mid \langle k, x\rangle \in W_{M(\tau_\sigma)}\}$.[12] Clearly, if M $\mathcal{I}$-identifies $\mathrm{extCyl}(\mathcal{L}, k)$, then M' $\mathcal{I}$-identifies $\mathcal{L}$, a contradiction.

To see that for all $i, m, n \in N$, $\mathbf{TxtEx}^n$, $\mathbf{TxtBc}^n$, $\mathbf{TxtFex}^n_m$, $\mathbf{N}^i\mathbf{TxtEx}^n$, $\mathbf{In}^i\mathbf{TxtEx}^n$, $\mathbf{Im}^i\mathbf{TxtEx}^n$ are unhindered by extCyl-cylindrification, let total function f be such that for each $j, k \in N$, $W_{f(j,k)} = \{\langle 0, x\rangle \mid x \in N\} \cup \{\langle k, x\rangle \mid x \in W_j\}$. It is easy to verify that, for $\mathcal{I} \in \{\mathbf{TxtEx}^n$, $\mathbf{TxtBc}^n$, $\mathbf{TxtFex}^n_m$, $\mathbf{N}^i\mathbf{TxtEx}^n$, $\mathbf{In}^i\mathbf{TxtEx}^n$, $\mathbf{Im}^i\mathbf{TxtEx}^n\}$ and $k \in N^+$, if the sequence of grammars $j_1, j_2, j_3, \ldots$ is $\mathcal{I}$-admissible for L, then $f(j_1, k), f(j_2, k), f(j_3, k), \ldots$ is a sequence of grammars $\mathcal{I}$-admissible for $\mathrm{extCyl}(L, k)$.

It is easy to verify that for all $i, m, n \in N$, $\mathbf{TxtEx}^n$, $\mathbf{TxtBc}^n$, $\mathbf{TxtFex}^n_m$, $\mathbf{N}^i\mathbf{TxtEx}^n$, $\mathbf{In}^i\mathbf{TxtEx}^n$, $\mathbf{Im}^i\mathbf{TxtEx}^n$ uses extCyl-decylindrifiable text.

6.2.3 Clean Text

Definition 32 An identification criterion $\mathcal{I}$ is said to *use clean text* just in case for all texts T, for all $L \in \mathcal{E}$, T is $\mathcal{I}$-admissible for $L \Rightarrow content(T) \subseteq L$.

[12] To ensure syntactic convergence, here we assume e is dependent only on $M(\tau_\sigma)$.

As example, for $i, n \in N$, $\mathbf{In}^i\mathbf{TxtEx}^n$ uses clean text, while $\mathbf{N}^i\mathbf{TxtEx}^n$ and $\mathbf{Im}^i\mathbf{TxtEx}^n$ do not.

Proposition 28 *Let $\mathcal{I}$ and $\mathcal{J}$ be two identification criteria where $\mathcal{I}$*

(a) is text decomposable,

(b) unhinderd by extCyl-cylindrification (and thus text insensitive), and

(c) uses clean and extCyl-decylindrifiable text,

while $\mathcal{J}$ is unaidable by extCyl-cylindrification. Then $\mathcal{I} - \mathcal{J} \neq \emptyset \Rightarrow (\exists \mathcal{L} \subseteq \mathcal{E})[\mathcal{L} \in \boldsymbol{DU}^*\mathcal{I} - \mathcal{J}]$.

Proof. Given $\mathcal{L} \in \mathcal{I} - \mathcal{J}$, we show how to construct $\mathcal{L}'$ such that $\mathcal{L}' \in \mathbf{DU}^*\mathcal{I} - \mathcal{J}$.

Suppose $\mathcal{L} = \{L_i \mid i \in N\}$ (if $\mathcal{L}$ is finite, let $\mathcal{L} = \{L_i \mid 0 \leq i < card(\mathcal{L})\}$ instead). Since $\mathcal{L} \notin \mathcal{J}$, for each $m \in N$, M_m does not $\mathcal{J}$-identify $\mathcal{L}$. Since $\mathcal{J}$ is unaidable by extCyl-cylindrification, M_m does not $\mathcal{J}$-identify $\text{extCyl}(\mathcal{L}, m+1)$. Since M_m does not identify $\text{extCyl}(\mathcal{L}, m+1)$, Assumption 1 implies that there exists at least one language in $\text{extCyl}(\mathcal{L}, m+1)$ which M_m does not $\mathcal{J}$-identify. Let total f be such that for all $m \in N$, M_m does not $\mathcal{J}$-identify $\text{extCyl}(L_{f(m)}, m+1)$. The collection $\mathcal{L}' = \{\text{extCyl}(L_{f(m)}, m+1) \mid m \in N\}$ is then $\mathcal{J}$-identified by no **IIM**.

To show that $\mathcal{L}' \in \mathbf{DU}^*\mathcal{I}$, let M be an **IIM** which $\mathcal{I}$-identifies $\mathcal{L}$, let g witness that $\mathcal{I}$ is unhindered by extCyl-cylindrification, and define M' as below, such that for each $\sigma \in \text{SEQ}$,

$M'(\sigma)$:

$S \leftarrow \emptyset$.

Let $A = \{i \in N \mid i \neq 0 \wedge (\exists x \in content(\sigma))[\pi_1(x) = i]\}$.

For each $i \in A$ do

$$\tau^i(\sigma) = \mathrm{tr}(\sigma, \lambda x[\pi_2(x)], \lambda x[\#], \lambda x[\#], \{x \in content(\sigma) \mid \pi_1(x) = i\}, \{x \in content(\sigma) \mid \pi_1(x) \neq i\}).$$

Insert $g(M(\tau^i(\sigma)), i)$ into S.

Return S.

Let finite $\mathcal{L}'' \subseteq \mathcal{L}'$ be given, and let T be an $\mathcal{I}$-admissible text for $L_{\mathcal{L}''}$. Since $\mathcal{I}$

1. uses clean text,
2. is text decomposable, and
3. uses extCyl-decylindrifiable text,

it can be verified that for each $m \in N$, $\bigcup_s \tau^{m+1}(T[s])$ is a text $\mathcal{I}$-admissible for $L_{f(m)}$. Subsequently, since M $\mathcal{I}$-identifies each $L_{f(m)} \in \mathcal{L}$, it outputs an $\mathcal{I}$-admissible sequence of grammars for $L_{f(m)}$. From here it is easy to see that M' will output an $\mathcal{I}$-admissible sequence of conjectures for $\mathcal{L}''$. ∎

Corollary 29 *For all $m \in N^+$, $\boldsymbol{DU^*TxtFex}^0_{m+1} - \boldsymbol{TxtFex}^*_m \neq \emptyset$.*

PROOF. Follows from Proposition 28 and that for all $m \in N^+$, $\mathbf{TxtFex}^0_{m+1} - \mathbf{TxtFex}^*_m \neq \emptyset$.[Cas88] ∎

6.2.4 Identification By Conjecturing Grammar For Text

We now introduce a concept required for our next result.

Definition 33 An identification criterion $\mathcal{I}$ is said to be *based upon conjecturing a grammar for the input text* just in case:

(a) for all texts T, for all $L \in \mathcal{E}$, T is $\mathcal{I}$-admissible for $L \Rightarrow content(T) = L$.

(b) for all $L \in \mathcal{E}$, for all $j \in N$, $W_j = L \Rightarrow$ (for each sequence of grammars G) $[(\stackrel{\infty}{\forall} i \in N)[G(i) = j] \Rightarrow G$ is $\mathcal{I}$-admissible for $L]$.

It is easy to verify that $\mathbf{TxtEx}^n$, $\mathbf{TxtBc}^n$, $\mathbf{TxtFex}^n_m$ are based upon conjecturing a grammar for the input text, but $\mathbf{N}^i\mathbf{TxtEx}^n$, $\mathbf{In}^i\mathbf{TxtEx}^n$, $\mathbf{Im}^i\mathbf{TxtEx}^n$ are not.

Proposition 30 *Let $n \in N$, $n \geq 2$. Given $\mathcal{I}$, $\mathcal{J}$, two identification criteria where $\mathcal{I}$ is based upon conjecturing a grammar for the input text and $\mathcal{J}$ is unaidable by extCyl-cylindrification. Then $(\exists \mathcal{L} \subseteq \mathcal{E} \mid \mathcal{L} \notin \mathcal{J}) \Rightarrow (\exists \mathcal{L}' \subseteq \mathcal{E} \mid \mathcal{L}' \in \boldsymbol{DU}^{n-1}\mathcal{I} - \boldsymbol{U}^n\mathcal{J})$.*

Proof. Suppose $\mathcal{L} = \{L_i \mid i \in N\}$ (if $\mathcal{L}$ is finite, let $\mathcal{L} = \{L_i \mid 0 \leq i < card(\mathcal{L})\}$ instead). Since $\mathcal{L} \notin \mathcal{J}$, for each $m \in N$, M_m does not $\mathcal{J}$-identify $\mathcal{L}$. Since $\mathcal{J}$ is unaidable by extCyl-cylindrification, M_m does not $\mathcal{J}$-identify extCyl$(\mathcal{L}, m+1)$. Since M_m does not identify extCyl$(\mathcal{L}, m+1)$, Assumption 1 implies that there exists at least one language in extCyl$(\mathcal{L}, m+1)$ which M_m does not $\mathcal{J}$-identify. Let total f be such that for all $m \in N$, M_m does not $\mathcal{J}$-identify extCyl$(L_{f(m)}, m+1)$.

For $n, i, m, k \in N$, $n \geq 2$, define $\mathcal{A}^n_{i,m,k} = \{\langle i+j, k\rangle \mid 0 \leq j < n\} \cup \{\langle l*n+m, x\rangle \mid l, x \in N\}$. Intuitively, $\mathcal{A}^n_{i,m,k}$ has the number k on n tracks (that is, from track i to track $i+n-1$), and has tracks $m, n+m, 2n+m, \ldots$ completely filled.

Fix an $n \geq 2$. For $i, m \in N$, $0 \leq m < n$, let $k_{i,m}$ be such that $W_{k_{i,m}} = \{\langle i{+}1, x\rangle \mid x \in L_{f(i)}\} \cup \{\langle 0, y\rangle \mid y \in \mathcal{A}^n_{i+1,m,k_{i,m}}\}$ and let $W_{k_{i,m}}$ be denoted $L'_{i,m}$. (That for each $i, m \in N$, such $k_{i,m}$ exists is given by Kleene's Recursion Theorem. [Rog67])

Let $\mathcal{L}' = \{L'_{i,m} \mid 0 \leq m < n \wedge i \in N\}$. That $\mathcal{L}'$ is in $\mathbf{DU}^{n-1}\mathcal{I}$ (irrespective of f chosen) is observed by M below:

$M(\sigma)$:

$S \leftarrow \emptyset$.

Let $A = \{i \in N \mid i \neq 0 \wedge (\exists x \in \mathit{content}(\sigma))[\pi_1(x) = i]\}$.

For each $x \in A$ do

If exists $y \in N$ such that $(\forall z \in N, x \leq z < x + n)[\langle 0, \langle z, y\rangle\rangle \in \mathit{content}(\sigma)]$.

Then insert y into S.

Return S.

Since $\mathcal{I}$ is based upon conjecturing a grammar for the input text, it is easy to verify that for all $\mathcal{L}'' \subseteq \mathcal{L}'$ where $\mathit{card}(\mathcal{L}'') < n$, for each $L \in \mathcal{L}''$, M will output an $\mathcal{I}$-admissible sequence of conjectures for L.

However, for all $i \in N$, $\bigcup_{0 \leq m < n} L'_{i,m} = \mathrm{extCyl}(L_{f(i)}, i + 1)$ is not $\mathcal{J}$-identified by M_i. Hence $\mathcal{L}' \notin \mathbf{U}^n\mathcal{J}$. ∎

Corollary 31 *For all $n \in N^+$, $\boldsymbol{DU^n TxtEx} - \boldsymbol{U^{n+1} TxtBc^*} \neq \emptyset$.*

6.3 Learning Unions of Languages In The Presence of Inaccuracies

For $i \in N$, $n \in N^+$, let $\mathbf{DU}^n\mathbf{N}^i\mathbf{TxtEx}$, $\mathbf{DU}^n\mathbf{In}^i\mathbf{TxtEx}$ and $\mathbf{DU}^n\mathbf{Im}^i\mathbf{TxtEx}$ be as given by Definition 28 for $\mathcal{I} \in \{\mathbf{N}^i\mathbf{TxtEx}, \mathbf{In}^i\mathbf{TxtEx}, \mathbf{Im}^i\mathbf{TxtEx}\}$.

Definition 34 Given $i \in N^+$, $L \in \mathcal{E}$, $\mathcal{L} \subseteq \mathcal{E}$.

(a) Let $\mathtt{COPY}(L, i) = \{\langle k, x\rangle \mid x \in L \wedge k \in N, 1 \leq k \leq i\}$.

(b) Let $\mathtt{COPY}(\mathcal{L}, i) = \{\mathtt{COPY}(L, i) \mid L \in \mathcal{L}\}$.

Intuitively, for any language L and $i \in N^+$, $\mathtt{COPY}(L, i)$ makes i copies of L, each on a different track.

We first look at how having more noise will affect **DUTxtEx** learning.

Proposition 32 *For all $i \in N$, $\boldsymbol{DU^*N^iTxtEx} - \boldsymbol{N^{i+1}TxtEx} \neq \emptyset$.*

PROOF. Fix $i \in N$. Let $\mathcal{L} = \mathtt{COPY}(SINGLE, i{+}1)$. To see that $\mathcal{L} \in \mathbf{DU}^*\mathbf{N}^i\mathbf{TxtEx}$, define M, where for each $\sigma \in$ SEQ,

$M(\sigma)$:

$S \leftarrow \emptyset$.

Let $A = \{x \in N \mid (\forall k \in N, 1 \leq k \leq i+1)[\langle k, x\rangle \in \mathit{content}(\sigma)]\}$.

For each $x \in A$ do

Insert the standard index for the language $\mathtt{COPY}(\{x\}, i+1)$ into S.

Return S.

It is easy to verify that M $\mathbf{DU^*N^iTxtEx}$-identifies $\mathcal{L}$. We now show that $\mathcal{L} \notin \mathbf{N^{i+1}TxtEx}$. Let $L_1 = \mathtt{COPY}(\{0\}, i+1)$ and $L_2 = \mathtt{COPY}(\{1\}, i+1)$. It is easy to see that there exists at least one text T, where $content(T) = L_1 \cup L_2$, such that T is $\mathbf{N^{i+1}TxtEx}$-admissible for both L_1 and L_2. However, no **IIM** can converge on T to an index for both L_1 and L_2. It follows that $\mathcal{L} \notin \mathbf{N^{i+1}TxtEx}$. ∎

Proposition 33 *Given $n \in N^+$ and $\mathcal{L} \in \boldsymbol{DU^n TxtEx}$. For all $i \in N$,*

(a) $\mathtt{COPY}(\mathcal{L}, i+1) \in \boldsymbol{DU^n N^i TxtEx}$.

(b) $\mathtt{COPY}(\mathcal{L}, i+1) \in \boldsymbol{DU^n In^i TxtEx}$.

(c) $\mathtt{COPY}(\mathcal{L}, i+1) \in \boldsymbol{DU^n Im^{\lfloor i/2 \rfloor} TxtEx}$.

PROOF. Let M which $\mathbf{DU^nTxtEx}$-identifies $\mathcal{L}$ be given. We construct M^1, M^2, M^3, such that M^1 $\mathbf{DU^nN^iTxtEx}$-identifies $\mathtt{COPY}(\mathcal{L}, i+1)$, M^2 $\mathbf{DU^nIn^iTxtEx}$-identifies $\mathtt{COPY}(\mathcal{L}, i+1)$, and M^3 $\mathbf{DU^nIm^{\lfloor i/2 \rfloor}TxtEx}$-identifies $\mathtt{COPY}(\mathcal{L}, i+1)$. On input σ, for all j, $1 \le j \le i+1$, let $W^j = \{\pi_2(x) \mid x \in content(\sigma) \wedge \pi_1(x) = j\}$. Let τ^1, τ^2, τ^3 be sequences where

$$content(\tau^1) = \{x \in N \mid card(\{j \mid 1 \le j \le i+1 \wedge x \in W^j\}) > i\},$$

$$content(\tau^2) = \{x \in N \mid (\exists j \in N \mid 1 \le j \le i+1)[x \in W^j]\}, \text{ and}$$

$$content(\tau^3) = \{x \in N \mid card(\{j \mid 1 \le j \le i+1 \wedge x \in W^j\}) > \lfloor i/2 \rfloor\}.$$

(These are, intuitively, attempts at restoring "accurate" sequences.)

M^1, then, outputs a standard representation index set of $\mathtt{COPY}(\{W_i \mid i \in M(\tau^1)\}, i+1)$; M^2, a standard representation index set of $\mathtt{COPY}(\{W_i \mid i \in M(\tau^2)\}, i+1)$; and M^3, a standard representation index set of $\mathtt{COPY}(\{W_i \mid i \in M(\tau^3)\}, i+1)$. ∎

Corollary 34 *For all $i \in N$, $n \in N^+$.*

(a) $\boldsymbol{DU^n N^i TxtEx} - \boldsymbol{U^{n+1} TxtEx} \neq \emptyset$.

(b) $\boldsymbol{DU^n In^i TxtEx} - \boldsymbol{U^{n+1} TxtEx} \neq \emptyset$.

(c) $\boldsymbol{DU^n Im^i TxtEx} - \boldsymbol{U^{n+1} TxtEx} \neq \emptyset$.

PROOF. Let $\mathcal{L} \in \mathbf{DU}^n\mathbf{TxtEx} - \mathbf{U}^{n+1}\mathbf{TxtEx}$. Then $\mathtt{COPY}(\mathcal{L}, i+1) \in \mathbf{DU}^n\mathbf{N}^i\mathbf{TxtEx} \cap \mathbf{DU}^n\mathbf{In}^i\mathbf{TxtEx}$ and $\mathtt{COPY}(\mathcal{L}, 2i+1) \in \mathbf{DU}^n\mathbf{Im}^i\mathbf{TxtEx}$. However, it is easy to show (by contradiction) that neither $\mathtt{COPY}(\mathcal{L}, i+1)$ nor $\mathtt{COPY}(\mathcal{L}, 2i+1)$ is in $\mathbf{U}^{n+1}\mathbf{TxtEx}$. ∎

6.4 Learning Unions of Languages From Informants

Since the text presentation of an informant is not decomposable in the exact sense defined earlier, we are unable to use the generalized definitions to obtain **UInfEx** and **DUInfEx**. Hence we make the following definitions.

Definition 35 Let $k \in N^+$ and $\mathcal{L} \subseteq \mathcal{E}$.

(a.1) M $\mathbf{U}^k\mathbf{InfEx}$-identifies $\mathcal{L}$ just in case for all $L \in \mathcal{L}^k$, for all informants G for L, $M(G)\downarrow$ and $W_{M(G)} = L$.

(a.2) $\mathbf{U}^k\mathbf{InfEx} = \{\mathcal{L} \subseteq \mathcal{E} \mid (\exists M)[M\ \mathbf{U}^k\mathbf{InfEx}\text{-identifies } \mathcal{L}]\}$.

(b.1) M $\mathbf{DU}^k\mathbf{InfEx}$-identifies $\mathcal{L}$ just in case for all $\mathcal{L}' \subseteq \mathcal{L}$, $card(\mathcal{L}') \leq k$, for all informants G for $L_{\mathcal{L}'}$, $M(G)\downarrow$ and $M(G) \in \mathcal{I}_{\mathcal{L}'}$.

(b.2) $\mathbf{DU}^k\mathbf{InfEx} = \{\mathcal{L} \subseteq \mathcal{E} \mid (\exists M)[M\ \mathbf{DU}^k\mathbf{InfEx}\text{-identifies } \mathcal{L}]\}$.

By extending Assumption 1, as well as the notions of text decomposability, text insensitivity, text extCyl-decylindrifiability, use of clean text, and identification criterion based upon conjecturing a grammar for the input text, to informants, one can extend Propositions 28 and 30, to get the following corollary.

Corollary 35 *For all $n \in N^+$.*

(a) $\boldsymbol{DU^n InfEx} - \boldsymbol{U^{n+1} InfEx} \neq \emptyset$.

(b) $\boldsymbol{DU^* InfEx} - \boldsymbol{TxtEx} \neq \emptyset$.

6.4.1 (In)Sufficient Condition for DUInfEx Identification

The following result is analogous to Theorem 20.

Proposition 36 *For all $n \in N^+$, there exists $\mathcal{L} \in \boldsymbol{DU^n InfEx}$ where*

(a) $\emptyset \notin \mathcal{L}$.

(b) for all $L, L' \in \mathcal{L}$, $L \cap L' = \emptyset$

such that $\mathcal{L} \notin \boldsymbol{U^{n+1} InfEx}$.

PROOF. The following proof, while tedious, follows almost exactly the proof of Theorem 20, and is included here in entirety only for completeness.

Let $n \in N^+$ be given. Unless stated otherwise, let e, i, j, k, with or without decorations, range over N, and S, with or without decorations, range over finite sets. σ and τ, with or without decorations, range over informants. For each **IIM** M_e, we construct S_e, L_e^0, $L_e^1, \ldots, L_e^n$ where

$$L_e^0 = \{\langle e,0,0\rangle\} \cup \{\langle e,i,j\rangle \mid 1 \le i \le n,\ j \in S_e\}$$

and for $1 \le i \le n$, L_e^i satisfies the following two properties:

(1) $L_e^i = \{\langle e,i,j\rangle \mid j \in W_{\min(\{\pi_3(x)\mid x \in L_e^i\})}\}$

(2) $\min(\{\pi_3(x) \mid x \in L_e^i\}) > \max(S_e)$.

Let $\mathcal{L} = \{L_e^0, L_e^1, ..., L_e^n \mid e \in N\}$. It is clear that for all $L, L' \in \mathcal{L}$, $L \cap L' = \emptyset$.

Since $\mathcal{L} \in \mathbf{DU}^n\mathbf{TxtEx}$ (recall the proof of Theorem 20), it is easy to see that $\mathcal{L} \in \mathbf{DU}^n\mathbf{InfEx}$.

We now show that $\mathcal{L} \notin \mathbf{U}^{n+1}\mathbf{InfEx}$. The following construction follows exactly the same method as in the proof for Theorem 20; except that where previously we constructed texts, here we construct informants.

For each M_e here is the construction to show that M_e does not $\mathbf{U}^{n+1}\mathbf{InfEx}$-identify $\mathcal{L}$. By Kleene's Recursion Theorem [Rog67] there exists an index e' such that $W_{e'}$ may be defined in stages $s = 0, 1, 2\ldots,$ as below. (We highlight the positive data to enhance readability.)

Stage 0: Let $\sigma^1 = (\langle \mathbf{e},\mathbf{0},\mathbf{0}\rangle, \mathbf{1}) \diamond (\langle e,1,0\rangle,0) \diamond (\langle e,1,1\rangle,0) \diamond \ldots \diamond (\langle e,1,e'-1\rangle,0) \diamond (\langle e,2,0\rangle,0) \diamond (\langle e,2,1\rangle,0)\ldots \diamond (\langle e,2,e'-1\rangle,0)\ldots \diamond (\langle e,n,0\rangle,0) \diamond (\langle e,n,1\rangle,0)\ldots \diamond (\langle e,n,e'-1\rangle,0) \diamond (\langle \mathbf{e},\mathbf{1},\mathbf{e'}\rangle,\mathbf{1}) \diamond (\langle \mathbf{e},\mathbf{2},\mathbf{e'}\rangle,\mathbf{1}) \diamond \ldots \diamond (\langle \mathbf{e},\mathbf{n},\mathbf{e'}\rangle,\mathbf{1})$. Let $W_{e'}^1 = \{e'\}$. Go to stage 1.

Stage s: Search for τ where

(a) $content(\tau) \subseteq \{(\langle \mathbf{e}, \mathbf{i}, \mathbf{j}\rangle, x_j) \mid 1 \leq i \leq n \wedge j \in N \wedge x_j \in \{0, \mathbf{1}\}\}$ $\cup \{(\langle e, i, j\rangle, 0) \mid i, j \in N \wedge (i > n \vee (i = 0 \wedge j \neq 0))\} \cup$ $\{(\langle e'', i, j\rangle, 0) \mid e'' \neq e \wedge i, j \in N\}$, and

(b) $(\forall (\langle e, i, j\rangle, x), (\langle e, i', j\rangle, x') \in content(\sigma^s \diamond \tau) \mid 1 \leq i, i' \leq n)$ $[x = x']$. Note that this is sufficient to avoid all possible inconsistencies in $content(\sigma^s \diamond \tau)$.

such that $M_e(\sigma^s) \neq M_e(\sigma^s \diamond \tau)$. If and when τ is found,

1. enumerate $\{j \mid (\exists i', 1 \leq i' \leq n)\ [(\langle \mathbf{e}, \mathbf{i'}, \mathbf{j}\rangle, \mathbf{1}) \in content(\tau)]\}$ into $W_{e'}$.

2. Let $a^{s+1} = \max(\{j \mid (\exists i', 1 \leq i' \leq n)\ (\exists k)[(\langle e, i', j\rangle, k) \in content(\sigma^s \diamond \tau)]\})$.

   Let X be $W_{e'}$ enumerated up to now.

3. Let σ^{s+1} be an extension of σ^s such that $content(\sigma^{s+1}) =$ $\{(\langle \mathbf{e}, \mathbf{0}, \mathbf{0}\rangle, \mathbf{1})\} \cup \{(\langle \mathbf{e}, \mathbf{i}, \mathbf{j}\rangle, \mathbf{1}) \mid 1 \leq i \leq n \wedge j \in X\} \cup$ $\{(\langle e, i, j\rangle, 0) \mid 1 \leq i \leq n \wedge 0 \leq j \leq a^{s+1} \wedge j \notin X\} \cup$ $\{(\langle e, i, j\rangle, 0) \mid (n < i \leq s \wedge 0 \leq j \leq s) \vee (i = 0 \wedge 0 < j \leq s)\}$ $\cup \{(\langle e'', i, j\rangle, 0) \mid e'' \neq e \wedge 0 \leq i, j \leq s\}$.

Go to stage $s + 1$.

If the search for τ failed at any stage s, then let $L_e^0 = \{x \mid (x, 1) \in content(\sigma^s)\}$, let $e'' > a^s$ be such that $\min(W_{e''}) = e''$. For each $i \in N$, $1 \leq i \leq n$, let $L_e^i = \{\langle e, i, j\rangle \mid j \in W_{e''}\}$. Since stage s does not succeed, M_e does not identify at least one of L_e^0 and $(L_e^0 \cup \bigcup_{i=1}^n L_e^i)$.

If the search is successful at all stages, then let $L_e^0 = \{\langle e, 0, 0\rangle\}$ and for each $i \in N$, $1 \leq i \leq n$, let $L_e^i = \{\langle e, i, x\rangle \mid x \in W_{e'}\}$; now M_e fails to converge on the input $\bigcup_s \sigma^s$, an informant for $L_e^0 \cup \bigcup_{i=1}^n L_e^i$. ∎

7 Intrinsic Complexity

7.1 Introduction

The concept of *intrinsic complexity* is an attempt to describe the relative hardness of identifying a class of languages under the requirement given by an identification criterion. The idea is to reduce the task of $\mathcal{I}$-identifying a class of languages to the task of $\mathcal{J}$-identifying another class. To be able to reduce the $\mathcal{I}$-identification of $\mathcal{L}$ to that of $\mathcal{J}$-identifying $\mathcal{L}'$, we should be able to transform $\mathcal{I}$-admissible texts T for languages in $\mathcal{L}$ to $\mathcal{J}$-admissible texts T' for languages in $\mathcal{L}'$ and further transform $\mathcal{J}$-admissible sequences for T' into $\mathcal{I}$-admissible sequences for T.

7.2 Preliminaries

The following definitions are from [FKS95, JS96]. We define an *enumeration operator* (or just operator), Θ, to be an algorithmic mapping from SEQ into SEQ such that for all σ, $\tau \in$ SEQ, if $\sigma \subseteq \tau$, then $\Theta(\sigma) \subseteq \Theta(\tau)$. We further assume that for all texts T, $\lim_{n\to\infty} |\Theta(T[n])| = \infty$. By extension, we think of Θ as also defining a mapping from $\mathbf{T}$ to $\mathbf{T}$ such that $\Theta(T) = \bigcup_n \Theta(T[n])$.

7.3 Weak Reductions

Jain and Sharma [JS96] distinguished between two kinds of reductions, called *weak* and *strong reductions*. We consider only the former here.

7.3.1 Definitions

Definition 36 [JS96] Let $\mathcal{L}_1, \mathcal{L}_2 \subseteq \mathcal{E}$ be given. Let $\mathcal{I}_1$ and $\mathcal{I}_2$ be two identification criteria. Let $\mathcal{T}_1 = \{T \mid T$ is a text for $L \in \mathcal{L}_1\}$. Let $\mathcal{T}_2 = \{T \mid T$ is a text for $L \in \mathcal{L}_2\}$. We say that $\mathcal{L}_1 \leq_{weak}^{\mathcal{I}_1,\mathcal{I}_2} \mathcal{L}_2$ just in case there exist operators Θ and Ω such that for all $T \in \mathcal{T}_1$ and for all infinite sequences of grammars G the following hold:

(a) $\Theta(T) \in \mathcal{T}_2$, and

(b) if G is an $\mathcal{I}_2$-admissible sequence for $\Theta(T)$, then $\Omega(G)$ is an $\mathcal{I}_1$-admissible sequence for T.

We say that $\mathcal{L}_1 \leq_{weak}^{\mathcal{I}} \mathcal{L}_2$ if and only if $\mathcal{L}_1 \leq_{weak}^{\mathcal{I},\mathcal{I}} \mathcal{L}_2$.

We extend the above definition for our generalized **U**, **DU** and **WDU** paradigm as follows, so that instead of just reducing the task of $\mathcal{I}$-identifying every language in a class $\mathcal{L}_1$, to tasks of $\mathcal{J}$-identifying languages in another class $\mathcal{L}_2$, we want to reduce the task for $\mathcal{I}$-identifying every language in $\mathcal{L}_1^n$, to tasks of $\mathcal{J}$-identifying languages in $\mathcal{L}_2^m$, for some $m, n \in N$.

Definition 37 Let $\mathcal{L}_1, \mathcal{L}_2 \subseteq \mathcal{E}$ be given. Let $\mathcal{I}_1$ and $\mathcal{I}_2$ be two text decomposable identification criteria where Assumption 1 (in Chapter 6) applies. Let $\mathcal{K}_1, \mathcal{K}_2 \in \{\mathbf{U}, \mathbf{DU}, \mathbf{WDU}\}$ and $n, m \in N^+$ be given. Let $\mathcal{T}_1 = \{T \mid T$ is a text for $L \in \mathcal{L}_1^n\}$. Let $\mathcal{T}_2 = \{T \mid T$ is a text for $L \in \mathcal{L}_2^m\}$. We say that $\mathcal{L}_1 \leq_{weak}^{\mathcal{K}_1^n\mathcal{I}_1,\mathcal{K}_2^m\mathcal{I}_2} \mathcal{L}_2$ just in case there exist operators Θ and Ω such that for all $T \in \mathcal{T}_1$ and for all infinite sequences of conjectures $\mathcal{G}$ the following hold:

(a) $\Theta(T) \in \mathcal{T}_2$, and

(b) if $\mathcal{G}$ is an $\mathcal{K}_2^m\mathcal{I}_2$-admissible sequence for $\Theta(T)$, then $\Omega(\mathcal{G})$ is an $\mathcal{K}_1^n\mathcal{I}_1$-admissible sequence for T.

We say that $\mathcal{L}_1 \leq_{weak}^{\mathcal{K}^n\mathcal{I}} \mathcal{L}_2$ if and only if $\mathcal{L}_1 \leq_{weak}^{\mathcal{K}^n\mathcal{I},\mathcal{K}^n\mathcal{I}} \mathcal{L}_2$.

Definition 38 [JKW00] Let $\mathcal{I}$ be an identification criterion. Let $\mathcal{L} \subseteq \mathcal{E}$ be given.

(a) If for all $\mathcal{L}' \in \mathcal{I}$, $\mathcal{L}' \leq_{weak}^{\mathcal{I}} \mathcal{L}$, then $\mathcal{L}$ is $\leq_{weak}^{\mathcal{I}}$-hard.

(b) If $\mathcal{L}$ is $\leq_{weak}^{\mathcal{I}}$-hard and $\mathcal{L} \in \mathcal{I}$, then $\mathcal{L}$ is $\leq_{weak}^{\mathcal{I}}$-complete.

7.3.2 Hard/Complete Languages

Proposition 37 *For all $n \in N^+$,*

(a) INIT is $\leq_{\mathrm{weak}}^{\mathbf{U}^n\mathbf{TxtEx}}$-complete.

(b) INIT is $\leq_{\mathrm{weak}}^{\mathbf{WDU}^n\mathbf{TxtEx}}$-complete.

(c) INIT is $\leq_{\mathrm{weak}}^{\mathbf{DU}^n\mathbf{TxtEx}}$-hard.

Proof. Fix $n \in N^+$. We first note that $INIT^n = INIT$.

Given $\mathcal{L} \in \mathbf{U}^n\mathbf{TxtEx}$. Since $INIT$ is $\leq_{weak}^{\boldsymbol{TxtEx}}$-complete [JS96], there exist Θ and Ω which observe that $\mathcal{L}^n \leq_{weak}^{\boldsymbol{TxtEx}} INIT$. It follows that $\mathcal{L} \leq_{weak}^{\boldsymbol{U^n\,TxtEx}} INIT$. Clearly, $INIT \in \mathbf{U}^n\mathbf{TxtEx}$. Thus part (a) is proven.

Parts (b) and (c) can be proved using a technique similar to the one in the proof for Theorem 9 in [JS96]. We only sketch the idea here. Note first that whereas in [JS96], only a single index needs to be dealt with, here there are up to n indices. This is, however, easily solved as follows. Given M which $\mathbf{DU}^n\mathbf{TxtEx}$-identifies $\mathcal{L}$, we combine the n indices conjectured by M using an n-ary pairing function. The problem hence becomes identical to that in [JS96], and the same technique can be applied. Restoration of the indices can be easily done via projection functions $\pi_1, \ldots, \pi_n$. Every $\mathcal{L}' \subseteq \mathcal{L}$ where $card(\mathcal{L}') \leq n$ can hence be shown to reduce to a language in $INIT$. Thus part (c) follows.

Similarly it can be shown that $INIT$ is $\leq_{weak}^{\boldsymbol{WDU^n TxtEx}}$-hard. Part (b) then follows from that $INIT \in \mathbf{WDU^n TxtEx}$. ∎

The following geometrical property is needed in our next theorem. Terms and notations not defined here are from Definitions 25 and 26.

Definition 39 Let $n \in N$, $n \geq 2$.

(a) Let $v_1, v_2, \ldots, v_{n-1}$ be as in part (a), O and $\mathcal{G}_n$ be as in part (b) of Definition 26. Let $A_1, A_2, \ldots, A_n$ be the vertices of $\mathcal{G}_n$ at $v_1, v_2, \ldots, v_{n-1}, -v_1$ respectively.

(b) For each i, $1 \leq i \leq n$, let $\vec{\nu_i}$ denote the (unit) outward normal of the facet in $\mathcal{G}_n$ with vertices $V(\mathcal{G}_n) - \{A_i\}$. Let $b_1, \ldots, b_n$ be such that each point X in $\mathcal{G}_n$ fulfills the n inequalities $\vec{\nu_i} \cdot X \leq b_i$ for $i = 1, \ldots, n$, as in part (b) of Definition 25.

(c) For $1 \leq i \leq n$, let $\vec{\mu_i} = (1/|\overrightarrow{OA_i}|)\,\overrightarrow{OA_i}$. Since for each i, $|\overrightarrow{OA_i}| = 1$, it is clear that $\vec{\mu_i} = \overrightarrow{OA_i}$. For $1 \leq i \leq n$, $\epsilon \in \mathbb{R}$, let $\mathcal{G}_i(\epsilon) = \mathcal{G}_n + \epsilon\vec{\mu_i}$.

(d) For each $\delta \in \mathbb{R}$, $\delta > 0$, let $\mathcal{G}_n^\delta$ be a simplex with vertices at $(1+\delta)\,\overrightarrow{OA_i}$ from O for $i = 1, \ldots, n$. Intuitively, $\mathcal{G}_n^\delta$ is an enlargement of $\mathcal{G}_n$ by a factor of $1 + \delta$.

Note that by letting G and C in Claim 11 be $\mathcal{G}_n$ and O respectively, we can let our definitions of $\vec{\mu_i}$, $\vec{\nu_i}$, b_i and $\mathcal{G}_i(\epsilon)$ here coincide with the definitions of $\vec{\mu_i}$, $\vec{\nu_i}$, b_i and $G_i(\epsilon)$ in the proof of Claim 11 (for each i, $1 \leq i \leq n$). This allows us to use the sub-claims therein in a straight-forward manner.

Claim 38 *Let $\delta, \delta' \in \mathbb{R}$ where $0 \leq \delta' \leq \delta$ be given. For each i, j, k where $1 \leq i, j, k \leq n$, $i \neq k, j \neq k$, we have $\delta'(\vec{\mu_i} \cdot \vec{\nu_k}) \leq \delta(\vec{\mu_j} \cdot \vec{\nu_k})$.*

Proof. Let $k \in N$ where $1 \leq k \leq n$, and $\delta, \delta' \in \mathbb{R}$ where $0 \leq \delta' \leq \delta$, be given. Let F_0^k be the hyperplane given by the equality $\vec{\nu_k} \cdot X = b_k$, and F_1^k be the hyperplane in $\mathcal{G}_n^\delta$ with normal $\vec{\nu_k}$. Since for each i where $1 \leq i \leq n$ and $i \neq k$, $\delta\vec{\mu_i}$ translates a point in F_0^k (namely, A_i) to a point in F_1^k, the distance between F_0^k and F_1^k along $\vec{\nu_k}$ can be written $\delta(\vec{\mu_i} \cdot \vec{\nu_k})$ for any i where $1 \leq i \leq n$, $i \neq k$ (see Figure 9). Hence, each of the $n-1$ inequalities

$$\vec{\nu_k} \cdot X \leq b_k + \delta(\vec{\mu_j} \cdot \vec{\nu_k}) \qquad j = 1, 2, \ldots, n,\ j \neq k$$

defines the same bounding hyperplane in $\mathcal{G}_n^\delta$. It follows that for each i, j, $i \neq k, j \neq k$, $\delta(\vec{\mu_i} \cdot \vec{\nu_k}) = \delta(\vec{\mu_j} \cdot \vec{\nu_k})$.

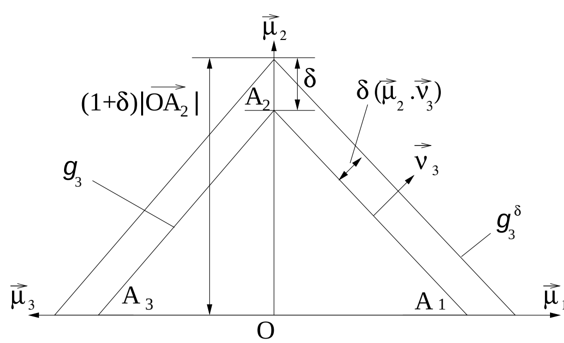

Figure 9: The hyperplane with normal $\vec{\nu_3}$ in $\mathcal{G}_3^\delta$.

Similarly, the hyperplane in $\mathcal{G}_n^{\delta'}$ with normal ν_k is equivalently defined by the $n-1$ inequalities

$$\vec{\nu_k} \cdot X \leq b_k + \delta'(\vec{\mu_j} \cdot \vec{\nu_k}) \qquad j = 1, 2, \ldots, n,\ j \neq k.$$

Thus for each i, j, $i \neq k, j \neq k$, $\delta'(\vec{\mu_i} \cdot \vec{\nu_k}) = \delta'(\vec{\mu_j} \cdot \vec{\nu_k})$.

Since $\delta' \leq \delta$, and for all $j \neq k$, $\vec{\mu_j} \cdot \vec{\nu_k} \geq 0$ (by Sub-claim 3 in the proof of Claim 11), it is trivially true that for all $j \neq k$, $\delta'(\vec{\mu_j} \cdot \vec{\nu_k}) \leq \delta(\vec{\mu_j} \cdot \vec{\nu_k})$. It follows that for all i, j, $1 \leq i, j \leq n$, $i \neq k, j \neq k$, $\delta'(\vec{\mu_i} \cdot \vec{\nu_k}) \leq \delta(\vec{\mu_j} \cdot \vec{\nu_k})$.

Since this holds for each k, $1 \leq k \leq n$, claim follows. □

Claim 39 *There exists $\omega^* \in \mathbb{R}$, $\omega^* > 0$ such that for all $\omega, \omega' \in \mathbb{R}$ where $0 \leq \omega' \leq \omega \leq \omega^*$, $\bigcup_{i=1}^{n} \mathcal{G}_i(\omega') \subseteq \bigcup_{i=1}^{n} \mathcal{G}_i(\omega)$.*

PROOF. Let $\mathcal{G}_1(\epsilon_1), \ldots, \mathcal{G}_n(\epsilon_n)$ where each $\epsilon_i \in \mathbb{R}, \epsilon_i > 0$, and $\xi_1, \ldots, \xi_n \in \mathbb{R}$ where each $\xi_i > 0$, be such that for all $i, 1 \leq i \leq n$, for all δ_i, $0 \leq \delta_i \leq \xi_i$, $\mathcal{G}_i(\delta_i) \subseteq \bigcup_{j=1}^{n} \mathcal{G}_j(\epsilon_j)$ (such values exist by Claim 11). Let $\omega^* \in \mathbb{R}$ be a non-zero positive value smaller than $\min(\{\xi_1, \ldots, \xi_n\})$.

Let $\omega, \omega' \in \mathbb{R}$ be where $0 \leq \omega' \leq \omega \leq \omega^*$. For each i, since $\omega' \leq \xi_i$, by Claim 11, $\mathcal{G}_i(\omega') \subseteq \bigcup_{j=1}^{n} \mathcal{G}_j(\epsilon_j)$. We show that $\mathcal{G}_i(\omega') \subseteq \bigcup_{j=1}^{n} \mathcal{G}_j(\omega)$.

Let $i \in N$, $1 \leq i \leq n$ be given.

By Claim 38, for all j, k where $1 \leq j, k \leq n$ and $i \neq k, j \neq k$, $\omega'(\vec{\mu_i} \cdot \vec{\nu_k}) \leq \omega(\vec{\mu_j} \cdot \vec{\nu_k})$. Hence, for any j, $1 \leq j \leq n$, Sub-claim 5 (in the proof of Claim 11) applies, and the points in $\mathcal{G}_i(\omega') - \mathcal{G}_j(\omega)$ is bounded by the inequality

$$\vec{\nu_j} \cdot X > b_j + \omega(\vec{\mu_j} \cdot \vec{\nu_j}), \qquad \text{———————} \ (1)$$

plus the n inequalities for $\mathcal{G}_i(\omega')$. Similarly, for any j, $1 \leq j \leq n$ the region $\mathcal{G}_i(\omega')$ $-$ $\mathcal{G}_j(\epsilon_j)$ is given by the inequality

$$\vec{\nu_j} \cdot X > b_j + \epsilon_j(\vec{\mu_j} \cdot \vec{\nu_j}), \qquad \text{———————} \ (2)$$

plus the n inequalities for $\mathcal{G}_i(\omega')$.

Given j, $1 \leq j \leq n$, since $\omega \leq \epsilon_j$ and $\vec{\mu_j} \cdot \vec{\nu_j} < 0$ (Sub-claim 4 in the proof of Claim 11), $\epsilon_j(\vec{\mu_j} \cdot \vec{\nu_j}) \leq \omega(\vec{\mu_j} \cdot \vec{\nu_j})$. The region in (2) hence includes the region in (1), and thus $\mathcal{G}_i(\omega') - \mathcal{G}_j(\omega)$ is a subset of $\mathcal{G}_i(\omega') - \mathcal{G}_j(\epsilon_j)$. We give an illustration of a 2-D example in Figure 10.

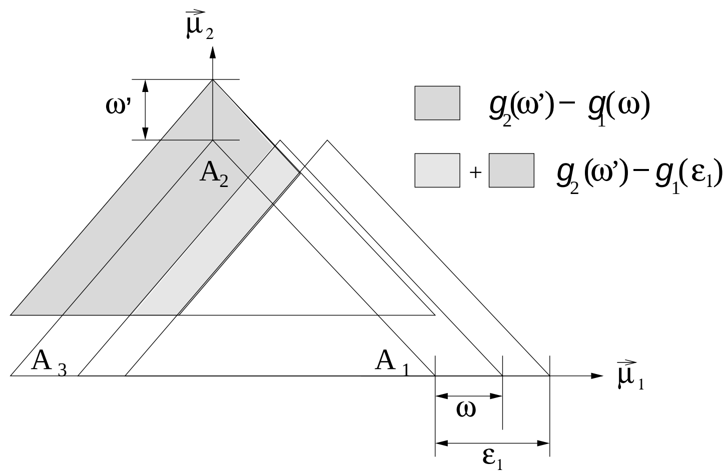

Figure 10: $\mathcal{G}_i(\omega') - \mathcal{G}_j(\omega) \subseteq \mathcal{G}_i(\omega') - \mathcal{G}_j(\epsilon_j)$.

Repeating this on each $\mathcal{G}_j(\omega)$ and $\mathcal{G}_j(\epsilon_j)$ for $j = 1, \ldots, n$, we have $\mathcal{G}_i(\omega') - \bigcup_{j=1}^n \mathcal{G}_j(\omega) \subseteq \mathcal{G}_i(\omega') - \bigcup_{j=1}^n \mathcal{G}_j(\epsilon_j)$. However, by Claim 11, $\mathcal{G}_i(\omega') - \bigcup_{j=1}^n \mathcal{G}_j(\epsilon_j) = \emptyset$. Hence $\mathcal{G}_i(\omega') \subseteq \bigcup_{j=1}^n \mathcal{G}_j(\omega)$. Since this holds for each $i = 1, \ldots, n$, it follows that $\bigcup_{i=1}^n \mathcal{G}_i(\omega') \subseteq \bigcup_{i=1}^n \mathcal{G}_i(\omega)$. ▮

We now give a $\leq_{weak}^{\boldsymbol{DU^n TxtEx}}$-complete language. Terms and notations not defined here are from Definition 26.

Theorem 40 *For all $n \in N, n \geq 2$, $TRANSIM_n$ is $\leq_{\mathrm{weak}}^{\mathbf{DU}^n\mathbf{TxtEx}}$-complete.*

Proof. Let $n \in N$, $n \geq 2$. Let $\langle .,. \rangle_p$ be a 1–1 pairing function with range in the prime numbers. For any $\mathcal{L} \subseteq \mathbf{DU}^n\mathbf{TxtEx}$, we construct Θ and Ω which witness that $\mathcal{L} \leq_{weak}^{\boldsymbol{DU^n TxtEx}} TRANSIM_n$. Let M $\mathbf{DU}^n\mathbf{TxtEx}$-identifies $\mathcal{L}$. Let $\mathcal{L}' \subseteq \mathcal{L}$, $card(\mathcal{L}') \leq n$ and let T be a text for $L_{\mathcal{L}'}$. Without loss of generality, assume that $M(T[0]) =?$.

For each $\epsilon \in rat$, let $XL(\epsilon)$ denote the language $\bigcup_{i=1}^{n} L(\mathcal{G}_i(\epsilon))$. Let $XTRANSIM$ $= \{XL(\epsilon) \mid \epsilon \in rat\}$. It is clear that $XTRANSIM \subset TRANSIM_n{}^n$.

Let ω^* be as in Claim 39, and $\omega \in rat$ be such that $0 < \omega \leq \omega^*$.

Define Θ as follow, such that for each text T and each $s \in N$,

$\Theta(T[0])$:

Let $\omega^0 = 0$.

Return $\#$.

$\Theta(T[s+1])$:

If $M(T[s+1]) = M(T[s])$ return $\Theta(T[s])$.

Else

Find least $m \in N$, and corresponding

$l \in N$ where l is co-prime with $\langle M(T[s+1]), m\rangle_p$

such that $\omega^s \leq \frac{l}{\langle M(T[s+1]),m\rangle_p} < \omega$.

Let $\omega^{s+1} = \frac{l}{\langle M(T[s+1]),m\rangle_p}$.

Return $\Theta(T[s]) \diamond \sigma$, where $content(\sigma) = \{x \mid x \leq s+1 \wedge x \in XL(\omega^{s+1})\}$.

For any $i, j \in N$ where $i < j$, since $\omega^i \leq \omega^j$, by Claim 39, $XL(\omega^i) \subseteq XL(\omega^j)$. Hence at each stage s, $content(\Theta(T[s])) \subseteq XL(\omega^s)$. If M $\mathbf{DU}^n\mathbf{TxtEx}$-identifies $L_{\mathcal{L}'}$, then at some stage t, M stops changing it's mind (that is, $M(T[t]) = M(T)$), and $\Theta(T)$ is a text for the language $XL(\omega^t)$.

To obtain operator Ω transforming a sequence of conjectures for $XL(\omega^t)$ into a sequence of conjectures for $L_{\mathcal{C}'}$, observe that it is possible to restore the value $M(T)$ from a sequence of conjectures for $XL(\omega^t)$. Let $\mathcal{C} = \mathcal{C}(0)\mathcal{C}(1)\mathcal{C}(2)\ldots$ be an infinite sequence of conjectures, define $\Omega(\mathcal{C}) = \mathcal{C}'$ where for each s, $\mathcal{C}'(s)$ is defined as follows. Let $z_s = \max(\{x_1(\text{decoderat}_n(w)) - 1 \mid w \in \bigcup_{i=0}^{n-1} W_{\mathcal{C}(s)[i],s}\})$ (the function x_1 is defined in Definition 25 and decoderat_n is defined in Definition 26). Intuitively, here z_s attempts to restore the value ω^s from $\mathcal{C}(s)$, a conjecture for $XL(\omega^s)$.

Finally, let $\mathcal{C}'(s) = \pi_1(h(z_s))$ where $h(a)$ is the denominator of rational a in reduced form. It is easy to verify that if M $\mathbf{DU}^n\mathbf{TxtEx}$-identifies T, and $\mathcal{C}$ converges to a conjecture for $\mathit{content}(\Theta(T))$, then $\Omega(\mathcal{C})$ converges to $M(T)$.

That $\mathit{TRANSIM}_n \in \mathbf{DU}^n\mathbf{TxtEx}$ then completes the proof. ∎

References

[Ang80a] D. Angluin. Finding patterns common to a set of strings. *Journal of Computer and System Sciences*, 21:46–62, 1980.

[Ang80b] D. Angluin. Inductive inference of formal languages from positive data. *Information and Control*, 45:117–135, 1980.

[B$\bar{7}$4] J. Bārzdiņš. Inductive inference of automata, functions and programs. In *Int. Math. Congress, Vancouver*, pages 771–776, 1974.

[BB75] L. Blum and M. Blum. Toward a mathematical theory of inductive inference. *Information and Control*, 28:125–155, 1975.

[Blu67] M. Blum. A machine-independent theory of the complexity of recursive functions. *Journal of the ACM*, 14:322–336, 1967.

[Cas88] J. Case. The power of vacillation. In D. Haussler and L. Pitt, editors, *Proceedings of the First Annual Workshop on Computational Learning Theory*, pages 133–142. Morgan Kaufmann, 1988.

[Cas99] J. Case. The power of vacillation in language learning. *SIAM Journal on Computing*, 28(6):1941–1969, 1999.

[CL82] J. Case and C. Lynes. Machine inductive inference and language identification. In M. Nielsen and E. M. Schmidt, editors, *Proceedings of the 9th International Colloquium on Automata, Languages and Pro-*

gramming, volume 140 of *Lecture Notes in Computer Science*, pages 107–115. Springer-Verlag, 1982.

[Cox63] H.S.M. Coxeter. *Regular Polytopes (2nd ed.)*. Methuen & Co. Ltd., London, 1963.

[CS83] J. Case and C. Smith. Comparison of identification criteria for machine inductive inference. *Theoretical Computer Science*, 25:193–220, 1983.

[dJK96] D. de Jongh and M. Kanazawa. Angluin's thoerem for indexed families of r.e. sets and applications. In *Proceedings of the Ninth Annual Conference on Computational Learning Theory*, pages 193–204. ACM Press, July 1996.

[FJ89] M. Fulk and S. Jain. Learning in the presence of inaccurate information. In R. Rivest, D. Haussler, and M. Warmuth, editors, *Proceedings of the Second Annual Workshop on Computational Learning Theory*, pages 175–188. Morgan Kaufmann, 1989.

[FJ96] M. Fulk and S. Jain. Learning in the presence of inaccurate information. *Theoretical Computer Science A*, 161:235–261, 1996.

[FKS95] R. Freivalds, E. Kinber, and C. Smith. On the intrinsic complexity of learning. In Paul Vitányi, editor, *Second European Conference on Computational Learning Theory*, volume 904 of *Lecture Notes in Artificial Intelligence*, pages 154–169. Springer-Verlag, 1995.

[GK99] S. Goldman and S. Kwek. On learning unions of pattern languages and tree patterns. *Algorithmic Learning Theory: Tenth International Conference (ALT' 99)*, 1720:347–363, 1999.

[Gol67] E. M. Gold. Language identification in the limit. *Information and Control*, 10:447–474, 1967.

[Jai98] S. Jain. Learning with refutation. *Journal of Computer and System Sciences*, 57(3):356–365, 1998.

[JK01] S. Jain and E. Kinber. On intrinsic complexity of learning geometrical concepts from texts. In *Proceedings of the Fourteenth Annual Conference on Computational Learning Theory*, pages 177–193. Springer-Verlag, 2001.

[JKW00] S. Jain, E. Kinber, and R. Wiehagen. Language learning from texts: Degrees of intrinsic complexity and their characterizations. In Nicolo Cesa-Bianchi and Sally Goldman, editors, *Proceedings of the Thirteenth Annual Conference on Computational Learning Theory*, pages 47–58. Morgan Kaufmann, 2000.

[JORS99] S. Jain, D. Osherson, J. Royer, and A. Sharma. *Systems that Learn: An Introduction to Learning Theory.* MIT Press, Cambridge, Mass., second edition, 1999.

[JS96] S. Jain and A. Sharma. The intrinsic complexity of language identification. *Journal of Computer and System Sciences*, 52:393–402, 1996.

[KPSW99] E. Kinber, C. Papazian, C. Smith, and R. Wiehagen. On the intrinsic complexity of learning infinite objects from finite samples. Technical Report LSA-99-01E, Centre for Learning Systems and Applications, Department of Computer Science, University of Kaiserslautern, Germany, 1999.

[Kru72] J. B. Kruskal. The theory of well-quasi-ordering: A frequently discovered concept. *Journal of Combinatorial Theory (A)*, 13:297–305, 1972.

[LW94] S. Lange and P. Watson. Machine discovery in the presence of incomplete or ambiguous data. In S. Arikawa and K. Jantke, editors, *Algorithmic Learning Theory: Fourth International Workshop on Analogical and Inductive Inference (AII '94) and Fifth International Workshop on Algorithmic Learning Theory (ALT '94)*, volume 872 of *Lecture Notes in Artificial Intelligence*, pages 438–452. Springer-Verlag, 1994.

[MA93] Y. Mukouchi and S. Arikawa. Inductive inference machines that can refute hypothesis spaces. In K.P. Jantke, S. Kobayashi, E. Tomita, and T. Yokomori, editors, *Algorithmic Learning Theory: Fourth International Workshop (ALT '93)*, volume 744 of *Lecture Notes in Artificial Intelligence*, pages 123–136. Springer-Verlag, 1993.

[MA95] Y. Mukouchi and S. Arikawa. Towards a mathematical theory of machine discovery from facts. *Theoretical Computer Science*, 137:53–84, 1995.

[Min76] E. Minicozzi. Some natural properties of strong identification in inductive inference. *Theoretical Computer Science*, 2:345–360, 1976.

[MSW91] T. Motoki, T. Shinohara, and K. Wright. The correct definition of finite elasticity: Corrigendum to identification of unions. In L. Valiant and M. Warmuth, editors, *Proceedings of the Fourth Annual Workshop on Computational Learning Theory*, page 375. Morgan Kaufmann, 1991.

[OSW86] D. Osherson, M. Stob, and S. Weinstein. *Systems that Learn: An Introduction to Learning Theory for Cognitive and Computer Scientists.* MIT Press, 1986.

[OW82] D. Osherson and S. Weinstein. Criteria of language learning. *Information and Control*, 52:123–138, 1982.

[Rog67] H. Rogers. *Theory of Recursive Functions and Effective Computability.* McGraw-Hill, 1967. Reprinted by MIT Press in 1987.

[SA00] T. Shinohara and H. Arimura. Inductive inference of unbounded unions of pattern languages from positive data. *Theoretical Computer Science A*, 241:191–209, 2000.

[WC80] K. Wexler and P. Culicover. *Formal Principles of Language Acquisition.* MIT Press, 1980.

[Wri89] K. Wright. Identification of unions of languages drawn from an identifiable class. In R. Rivest, D. Haussler, and M. Warmuth, editors, *Proceedings of the Second Annual Workshop on Computational Learning Theory*, pages 328–333. Morgan Kaufmann, 1989.