Acta Cybernetica **25** (2022) 717–731.

Dense Languages and Non Primitive Words

Toshihiro Koga[ab]

Abstract

In this paper, we are concerned with dense languages and non primitive words. A language L is said to be dense if any string can be found as a substring of element of L. It is known that if a regular language R is dense, then R contains infinitely many non-primitive words. Then it is natural to ask whether this result can be generalized for a wider class of dense languages. In this paper, we actually obtain such generalization.

Keywords: dense languages, primitive words, monoid

1 Introduction

1.1 Density and asymptotic density

Let Σ be a non-empty finite set of distinct symbols with $|\Sigma| \geq 2$. A language $L \subseteq \Sigma^*$ is said to be dense (in Σ^*) iff $\Sigma^* s \Sigma^* \cap L \neq \emptyset$ for any $s \in \Sigma^*$, and L is said to be thin (in Σ^*) iff L is not dense in Σ^*. The concept of dense languages is important in code theory (e.g., [1, 2]), and many classifications and properties of dense languages are already known (e.g., [6, 8, 13]). Next, for $L \subseteq \Sigma^*$ and $n \geq 0$, let $D_n(L) := |L \cap \Sigma^n|/|\Sigma^n|$. Moreover, let $D^*(L) := \lim_n (1/n) \sum_{i=0}^{n-1} D_i(L)$ (asymptotic density of L), provided that the limit exists. Then L is said to have positive asymptotic density iff $D^*(L)$ exists and $D^*(L) > 0$. Although $D^*(L)$ does not necessarily exist in general, we can easily show (by Theorem III.6.1 of [12]) that if R is a regular language, then $D^*(R)$ always exists. Moreover, the same Theorem III.6.1 implies that if R is regular, then $D^*(R) = 0$ iff $\lim_n D_n(R) = 0$. Some other basic properties of asymptotic density can be found in Chapter 13 of [2].

1.2 Dömösi-Horváth-Ito conjecture

Let **REG** be the family of all regular languages over Σ and **CFL** be the family of all context-free languages over Σ. Let $Q_\Sigma \subseteq \Sigma^+$ be the set of all primitive words. In formal language theory, Dömösi-Horváth-Ito conjecture states that $Q_\Sigma \notin$ **CFL**.

[a]#D-804 Purimasitei 4-1-1, Nagatsutaminamidai, Midori-ku, Yokohama-shi, Kanagawa-ken 226-0018, Japan

[b]E-mail: toshihiro1123_f_ma_mgkvv@w7.dion.ne.jp, ORCID: 0000-0001-6016-1306

DOI: 10.14232/actacyb.293457

This conjecture was first suggested in [3], and still remains open. Some strategies for approaching this conjecture can be found in [4]. In 2020, Ryoma Syn'ya [16] suggested a new strategy for approaching $Q_\Sigma \notin \mathbf{CFL}$. He proved that any regular language with positive asymptotic density always contains infinitely many non-primitive words. Precisely, his theorem can be stated as follows:

Theorem 1.1 (Ryoma Sin'ya [16])**.** *Let $R \in \mathbf{REG}$ satisfy $D^*(R) > 0$. Then there exists $z \in \Sigma^+$ and $p \geq 1$ such that $z^{pn+1} \in R$ $(\forall n \geq 0)$. In particular, we cannot have $R \subseteq Q_\Sigma$ for such R.*

This result states that Q_Σ does not have good lower approximations by regular languages. Since $D^*(Q_\Sigma) = 1$, we obtain $Q_\Sigma \notin \mathbf{CFL}$, provide that we have:

Claim 1.1. *Let $L \in \mathbf{CFL}$ satisfy $D^*(L) = 1$. Then there exists $R \in \mathbf{REG}$ such that $R \subseteq L$ and $D^*(R) > 0$.*

If this claim is true, then in view of $D^*(Q_\Sigma) = 1$, we can conclude from Theorem 1.1 and Claim 1.1 that $Q_\Sigma \notin \mathbf{CFL}$. However, in fact, the above Claim 1.1 is actually false. A counter-example is implicitly shown in [16, Theorem 14], and directly shown in [17]. Specifically, let $\Sigma = \{a, b\}$ and $L = \{v \in \Sigma^* \mid |v|_a \leq 2|v|_b\}$. Then we can show that this is a counter-example. Therefore, we need some generalizations of Theorem 1.1 if we continue his strategy. Aside from this, Theorem 1.1 itself is of independent interest, because this result states a non-trivial connection between asymptotic density and primitive words. In this paper, we are concerned with such connections, and we generalize Theorem 1.1 for a wider class of dense languages.

2 Main result

In this section, we state our main result. We first begin with a connection between density and positive asymptotic density for regular languages:

Theorem 2.1 (Ryoma Sin'ya [15])**.** *Let $R \in \mathbf{REG}$. Then $\lim_n D_n(R) = 0$ if and only if R is thin.*

A simple proof of this theorem can also be found in [7]. As we have already mentioned in Section 1, if R is regular, then $\lim_n D_n(R) = 0$ iff $D^*(R) = 0$. Moreover, if R is regular, then $D^*(R)$ always exists. Combining these with Theorem 2.1, it follows that if R is regular, then R is dense iff R has positive asymptotic density. Hence, we can restate Theorem 1.1 as follows:

Theorem 2.2. *Let $R \in \mathbf{REG}$ be dense. Then there exists $z \in \Sigma^+$ and $p \geq 1$ such that $z^{pn+1} \in R$ $(\forall n \geq 0)$.*

As we have just mentioned, Theorem 2.2 is equivalent to Theorem 1.1, Now we generalize Theorem 2.2 for a wider class of dense languages. We first introduce some notations. Let $\mathbf{TL} := \{L \subseteq \Sigma^* \mid L \textit{ is thin}\}$, i.e., $\mathbf{TL}$ is the set of all thin languages over Σ. Next, For any set X, let 2^X denote the power set of X. For any

$\mathcal{N} \subseteq 2^{\Sigma^*}$, we define $\Gamma(\mathcal{N}) \subseteq 2^{\Sigma^*}$ as the regular closure of $\mathcal{N}$. In other words, we define $\Gamma(\mathcal{N})$ as the smallest set such that

$$\mathcal{N} \subseteq \Gamma(\mathcal{N}),\ \forall L_1, L_2 \in \Gamma(\mathcal{N})\ [\ L_1 \cup L_2,\ L_1 L_2,\ L_1^* \in \Gamma(\mathcal{N})\ ].$$

Then, our result can be stated as follows:

Main Theorem 2.3. *Let $L \in \Gamma(\mathbf{TL})$ be dense. Then we have*

$$\forall u, v \in \Sigma^*,\ \exists z \in \Sigma^* v \Sigma^*,\ \exists p \geq 1,\ \forall n \geq 0\ [\ (zu)^{pn} z \in L\ ]. \tag{1}$$

Since $\mathbf{REG} = \Gamma(\{\emptyset\} \cup \{\{a\} \mid a \in \Sigma\})$ and $\{\emptyset\} \cup \{\{a\} \mid a \in \Sigma\} \subseteq \mathbf{TL}$, we have $\mathbf{REG} \subseteq \Gamma(\mathbf{TL})$, and in fact $\mathbf{REG} \subsetneq \Gamma(\mathbf{TL})$. Moreover, if $L \subseteq \Sigma^*$ satisfies the condition (1), then there exists $z \in \Sigma^+$ and $p \geq 1$ such that $z^{pn+1} \in L$ ($\forall n \geq 0$). This implies that Theorem 2.2 is just a special case of Main Theorem 2.3. In other words, Main Theorem 2.3 is a generalization of Theorem 2.2 (and Theorem 1.1).

The rest of this paper is structured as follows. In Section 3, we provide some lemmas related to monoids. In Section 4, we prove Main Theorem 2.3. In Section 5, we show that Main Theorem 2.3 is a non-trivial generalization of Theorem 2.2. In Section 6, we state some remarks. In Section 7, we state related work.

We assume that the reader is familiar with Regular languages and semigroup theory. For basic information about these topics, see, e.g., [10].

3 Some lemmas related to monoids

In this section, we provide some lemmas related to monoids.

Definition 3.1. *Let X be a monoid. Then $L \subseteq X$ is said to be dense in X iff $XsX \cap L \neq \emptyset$ for any $s \in X$, and L is said to be thin in X iff L is not dense in X.*

Lemma 3.1. *Let X be a monoid. Let $n \geq 1$ and $A_1, \cdots, A_n \subseteq X$. If $\cup_{i=1}^n A_i$ is dense in X, then A_i is dense in X for some $i \in [1, n]$.*

Proof. The proof is essentially the same as [4, Proposition 2.2.1]. □

Lemma 3.2. *Let $A_1, A_2 \subseteq \Sigma^*$. If $A_1 A_2$ is dense in Σ^*, then A_i is dense in Σ^* for some $i \in \{1, 2\}$.*

Proof. Suppose that A_1 and A_2 are thin. Then $\Sigma^* v_i \Sigma^* \cap A_i = \emptyset$ for some $v_i \in \Sigma^*$ ($i = 1, 2$). Since $A_1 A_2$ is dense, we have $\Sigma^* v_1 v_2 \Sigma^* \cap A_1 A_2 \neq \emptyset$, so there exists $x, y \in \Sigma^*$ and $a_i \in A_i$ ($i = 1, 2$) such that $x v_1 v_2 y = a_1 a_2$. Then, v_1 is a substring of a_1, or v_2 is a substring of a_2. This contradicts the definition of v_1 and v_2. □

Lemma 3.3. *Let M be a finite monoid. Then we have the following:*
(i) Let $t, x, y \in M$ satisfy $t = xty$. Then $x^m t = t = t y^m$ for some $m \geq 1$.
(ii) Let $t, u, x, y \in M$ satisfy $t = xtuty$. Then $(tu)^p t = t$ for some $p \geq 1$. In particular, for any $n \geq 1$ we have $(tu)^{pn} t = t$.

Proof. (i): Since M is finite, we have $\exists m \geq 1, \ \forall z \in M$ [z^m *is idempotent*] (see [10, Proposition 6.33]). Now assume that $t = xty$. Then $t = xty = x^2ty^2 = \cdots = x^m ty^m$, so $x^m t = (x^m)^2 ty^m = x^m ty^m = t$. Similarly, $ty^m = t$.
(ii): If $t = xtuty$, then $t = (x)t(uty)$, so $t = t(uty)^m$ for some $m \geq 1$ by (i). Then $t = t(uty)(uty)^{m-1} = (tu)t(y(uty)^{m-1})$, so $(tu)^p t = t$ for some $p \geq 1$ by (i). $\square$

Lemma 3.4. *Let M be a finite monoid. Let X be a monoid. Let $\eta : X \to M$ be a monoid homomorphism. Let $S \subseteq M$. Let $R := \eta^{-1}(S)$ $(\subseteq X)$. If R is dense in X, then we have the following:*

$$\forall u, v \in X, \ \exists z \in XvX, \ \exists p \geq 1, \ \forall n \geq 0 \ [\ (zu)^{pn} z \in R \ ].$$

Proof. Since R is dense in X, we have $R \neq \emptyset$. In view of $R = \eta^{-1}(S)$, we have $S \neq \emptyset$. Next, we have $R = \eta^{-1}(S) = \cup_{t \in S} \eta^{-1}(\{t\})$. Since R is dense, $\cup_{t \in S} \eta^{-1}(\{t\})$ is also dense. Since "$\cup_{t \in S}$" is a non-empty finite union, we can apply Lemma 3.1, so $\eta^{-1}(\{t\})$ is dense for some $t \in S$. Now let $u, v \in X$ be arbitrary. Since $\eta^{-1}(\{t\})$ is dense in X, we have $XvX \cap \eta^{-1}(\{t\}) \neq \emptyset$, so $xvy \in \eta^{-1}(\{t\})$ for some $x, y \in X$. Let $z := xvy$. Then $z \in XvX$ and $\eta(z) = t$. Next, since $\eta^{-1}(\{t\})$ is dense, we have $XzuzX \cap \eta^{-1}(\{t\}) \neq \emptyset$, so $x'zuzy' \in \eta^{-1}(\{t\})$ for some $x', y' \in X$. Then $\eta(x'zuzy') = t$, i.e., $\eta(x')\eta(z)\eta(u)\eta(z)\eta(y') = t$. Keeping in mind $\eta(z) = t$, we have $\eta(x')t\eta(u)t\eta(y') = t$. By assumption on M, we can apply (ii) of Lemma 3.3, so there exists $p \geq 1$ such that $(t\eta(u))^{pn} t = t$ $(\forall n \geq 1)$. Since $\eta(z) = t$, we have $\eta((zu)^{pn} z) = (t\eta(u))^{pn} t = t$ $(\forall n \geq 1)$, so $(zu)^{pn} z \in \eta^{-1}(\{t\}) \subseteq R$ $(\forall n \geq 1)$. In addition, if $n = 0$, then $(zu)^{pn} z = z \in \eta^{-1}(\{t\}) \subseteq R$. In summary,

$$z \in XvX, \ p \geq 1, \ (zu)^{pn} z \in R \ (\forall n \geq 0).$$

Thus we complete the proof. $\square$

Lemma 3.5. *Let X be a monoid. Let $X_0 \subseteq X$ be a submonoid. Let Q be a non-empty finite set. Let $L \subseteq X$ and $s \in X$. Let $R : (Q \times Q) \to 2^X$. Assume that*

(i) $\forall n \geq 1, \ \forall x_1, \cdots, x_n \in X_0, \ \exists p_0, \cdots, p_n \in Q, \ \forall i \in [1, n]$ [$x_i \in R(p_{i-1}, p_i)$],

(ii) $\forall p, q \in Q, \ \exists t_0, t_1 \in X_0$ [$t_0 s R(p, q) s t_1 \subseteq L$],

(iii) $\forall p, q, r \in Q$ [$R(p, q) s R(q, r) \subseteq R(p, r)$].

Then we have the following:

$$\forall x, y \in X_0, \ \exists z \in XyX, \ \exists p \geq 1, \ \forall n \geq 0 \ [\ (zx)^{pn} z \in L \ ]. \tag{2}$$

Proof. STEP1: Let c_x $(\forall x \in X_0)$ be new distinct symbols, and let $\Sigma_0 := \{c_x \mid x \in X_0\}$. Note that Σ_0 can be an infinite set (of distinct symbols). We can trivially verify

$$\forall n \geq 1, \ \forall v = v_1 v_2 \cdots v_n \in \Sigma_0^n, \ \exists x_1, \cdots, x_n \in X_0, \ \forall i \in [1, n] \ [\ v_i = c_{x_i} \ ].$$

Next, let M be the set of all maps from 2^Q to 2^Q. For any $f, g \in M$, we define $f \circ g \in M$ as $(f \circ g)(U) := g(f(U))$ $(\forall U \in 2^Q)$. We also define $id_{2^Q} \in M$ as

$id_{2^Q}(U) := U$ ($\forall U \in 2^Q$). Note that $(M, \circ, id_{2^Q})$ is a finite monoid. We define a monoid homomorphism $\eta : \Sigma_0^* \to M$ as follows: Let $\varepsilon \in \Sigma_0^*$ be the empty string. We first define $\eta(\varepsilon) := id_{2^Q}$. Next, for $x \in X_0$, we define $\eta(c_x) \in M$ as

$$\eta(c_x)(U) := \{q \in Q \mid \exists p \in U \; [\; x \in R(p, q) \;] \; \} \; \textit{for} \; U \in 2^Q. \tag{3}$$

Next, for $n \geq 2$ and $v = v_1 v_2 \cdots v_n \in \Sigma_0^n$, we define $\eta(v) := \eta(v_1) \circ \eta(v_2) \circ \cdots \eta(v_n)$. By this definition, we can easily show that $\eta : \Sigma_0^* \to M$ is a monoid homomorphism. Moreover, by induction on $|v| \geq 0$, we can easily verify the following:

$$\forall v \in \Sigma_0^*, \; \forall U_1, U_2 \in 2^Q \; [\; U_1 \subseteq U_2 \Rightarrow \eta(v)(U_1) \subseteq \eta(v)(U_2) \;]. \tag{4}$$

STEP2: For any $p, q \in Q$, we define $A_{p,q} := \{v \in \Sigma_0^* \mid q \in \eta(v)(\{p\})\}$. Note that $\varepsilon \in A_{p,p}$ for any $p \in Q$. Moreover, we can trivially verify that

$$\forall v, w \in \Sigma_0^* \; [\; \eta(v) = \eta(w) \Rightarrow \forall p, q \in Q \; [\; v \in A_{p,q} \Leftrightarrow w \in A_{p,q} \;] \;]. \tag{5}$$

Next, we show

$$\forall p, q, r \in Q \; [\; A_{p,q} A_{q,r} \subseteq A_{p,r} \;]. \tag{6}$$

Let $v \in A_{p,q}$ and $w \in A_{q,r}$ be arbitrary. Then $q \in \eta(v)(\{p\})$ and $r \in \eta(w)(\{q\})$. In particular, $\{q\} \subseteq \eta(v)(\{p\})$. By (4), we have $\eta(w)(\{q\}) \subseteq \eta(w)(\eta(v)(\{p\})) = (\eta(v) \circ \eta(w))(\{p\}) = \eta(vw)(\{p\})$. In view of $r \in \eta(w)(\{q\})$, we have $r \in \eta(vw)(\{p\})$, so $vw \in A_{p,r}$. Thus we obtain (6). Next, we show

$$\forall v \in \Sigma_0^*, \; \exists p, q \in Q \; [\; v \in A_{p,q} \;]. \tag{7}$$

Since $\varepsilon \in A_{p,p}$ for any $p \in Q$, we have only to show

$$\forall v \in \Sigma_0^+, \; \exists p, q \in Q \; [\; v \in A_{p,q} \;].$$

Let $n \geq 1$ and $v = v_1 \cdots v_n \in \Sigma_0^n$ be arbitrary. There exists $x_1, \cdots, x_n \in X_0$ such that $v_i = c_{x_i}$ ($\forall i \in [1, n]$). By (i), there exists $p_0, \cdots, p_n \in Q$ such that $x_i \in R(p_{i-1}, p_i)$ ($\forall i \in [1, n]$). For any $i \in [1, n]$, it follows from (3) that

$$\begin{aligned} &\eta(c_{x_i})(\{p_{i-1}\}) = \{q \in Q \mid \exists p \in \{p_{i-1}\} \; [\; x_i \in R(p, q) \;] \; \} \\ &\quad = \{q \in Q \mid x_i \in R(p_{i-1}, q)\} \ni p_i, \end{aligned}$$

i.e., $p_i \in \eta(c_{x_i})(\{p_{i-1}\})$, so $c_{x_i} \in A_{p_{i-1}, p_i}$. In view of (6), we have

$$v = c_{x_1} c_{x_2} \cdots c_{x_n} \in A_{p_0, p_1} A_{p_1, p_2} \cdots A_{p_{n-1}, p_n} \subseteq A_{p_0, p_n}.$$

Thus we obtain (7).

STEP3: We define $g : \Sigma_0^+ \to X$ as follows: For any $x \in X_0$, we define $g(c_x) := x$. For any $n \geq 2$ and $v = v_1 v_2 \cdots v_n \in \Sigma_0^n$, we define $g(v) := g(v_1) s g(v_2) s \cdots s g(v_n)$. Note that we have $g(vw) = g(v) s g(w)$ for any $v, w \in \Sigma_0^+$. Then we can easily verify

$$\forall v, w \in \Sigma_0^+, \; \forall n \geq 1 \; [\; g(v^n w) = (g(v) s)^n g(w) \;]. \tag{8}$$

Next, we show

$$\forall v \in \Sigma_0^+,\ \forall p,q \in Q\ [\ v \in A_{p,q} \Rightarrow g(v) \in R(p,q)\ ]. \tag{9}$$

The proof is by induction on $|v| \geq 1$. We first show the case $|v| = 1$. Let $v \in \Sigma_0^1$ and $p, q \in Q$ satisfy $v \in A_{p,q}$. We have $v = c_x$ for some $x \in X_0$. In view of $v \in A_{p,q}$, we have $q \in \eta(v)(\{p\})$. In addition,

$$\begin{aligned}&\eta(v)(\{p\}) = \eta(c_x)(\{p\}) = \{q' \in Q \mid \exists p' \in \{p\}\ [\ x \in R(p',q')\ ]\ \}\\ &= \{q' \in Q \mid x \in R(p,q')\}\,,\end{aligned}$$

so $q \in \{q' \in Q \mid x \in R(p,q')\}$, i.e., $x \in R(p,q)$. Since $g(v) = g(c_x) = x$, we obtain $g(v) \in R(p,q)$. Thus we obtain (9) for $|v| = 1$. Next, let $n \geq 1$ be arbitrary. Assume that (9) holds for $|v| = n$. Let $v \in \Sigma_0^{n+1}$ and $p, q \in Q$ satisfy $v \in A_{p,q}$. We can write $v = wc_x$ for some $w \in \Sigma_0^n$ and $x \in X_0$. In view of $v \in A_{p,q}$, we have

$$\begin{aligned}&q \in \eta(v)(\{p\}) = \eta(wc_x)(\{p\}) = (\eta(w) \circ \eta(c_x))(\{p\}) = \eta(c_x)(\eta(w)(\{p\}))\\ &= \{q' \in Q \mid \exists p' \in \eta(w)(\{p\})\ [\ x \in R(p',q')\ ]\ \}\,,\end{aligned}$$

so there exists $p' \in \eta(w)(\{p\})$ such that $x \in R(p',q)$. In view of $p' \in \eta(w)(\{p\})$, we have $w \in A_{p,p'}$. By inductive hypothesis, we have $g(w) \in R(p,p')$. Then $g(v) = g(wc_x) = g(w)sg(c_x) = g(w)sx \in R(p,p')sR(p',q)$. By (iii), we have $R(p,p')sR(p',q) \subseteq R(p,q)$, so $g(v) \in R(p,q)$. Thus we obtain (9) for $|v| = n+1$. By induction, we obtain (9).

STEP4: Let $F := \eta(\Sigma_0^*) \subseteq M$. Then F is a non-empty finite set. Moreover, we trivially have $\eta : \Sigma_0^* \to F$, so $\Sigma_0^* = \cup_{f\in F}\eta^{-1}(\{f\})$. In particular, $\cup_{f\in F}\eta^{-1}(\{f\})$ is dense in Σ_0^*. By Lemma 3.1, $\eta^{-1}(\{f\})$ is dense in Σ_0^* for some $f \in F$. At this point, we obtain the following:

- M is a finite monoid, Σ_0^* is a monoid, $\eta : \Sigma_0^* \to M$ is a monoid homomorphism, $\{f\} \subseteq M$, and $\eta^{-1}(\{f\}) \subseteq \Sigma_0^*$ is dense in Σ_0^*.

Then we can apply Lemma 3.4, and we obtain

$$\forall u,v \in \Sigma_0^*,\ \exists z \in \Sigma_0^* v \Sigma_0^*,\ \exists p \geq 1,\ \forall n \geq 0\ [\ (zu)^{pn}z \in \eta^{-1}(\{f\})\ ].$$

Since $f \in F = \eta(\Sigma_0^*)$, we have $f = \eta(w)$ for some $w \in \Sigma_0^*$. Then, for any $w' \in \eta^{-1}(\{f\})$, we trivially have $\eta(w') = \eta(w)$. Thus we obtain

$$\forall u,v \in \Sigma_0^*,\ \exists z \in \Sigma_0^* v \Sigma_0^*,\ \exists p \geq 1,\ \forall n \geq 0\ [\ \eta((zu)^{pn}z) = \eta(w)\ ]. \tag{10}$$

Next, by (7), we have $w \in A_{p,q}$ for some $p, q \in Q$. By (ii), we have $t_0 sR(p,q)st_1 \subseteq L$ for some $t_0, t_1 \in X_0$. Now let $x', y' \in X_0$ be arbitrary. Let $x := t_1 x' t_0$. Since $x', t_0, t_1 \in X_0$ and X_0 is a monoid, we have $x \in X_0$. Then $c_x, c_{y'} \in \Sigma_0 \subseteq \Sigma_0^*$, so we can apply (10), i.e., there exists $z \in \Sigma_0^* c_{y'} \Sigma_0^*$ and $p \geq 1$ such that $\eta((zc_x)^{pn}z) = \eta(w)$ $(\forall n \geq 0)$. Since $w \in A_{p,q}$, we obtain $(zc_x)^{pn}z \in A_{p,q}$ $(\forall n \geq 0)$ by (5). Since $z \in \Sigma_0^* c_{y'} \Sigma_0^* \subseteq \Sigma_0^+$, we have $(zc_x)^{pn}z \in \Sigma_0^+$ $(\forall n \geq 0)$, so $g((zc_x)^{pn}z) \in$

$R(p,q)$ ($\forall n \geq 0$) by (9). Then $t_0sg((zc_x)^{pn}z)st_1 \in t_0sR(p,q)st_1 \subseteq L$ ($\forall n \geq 0$). In short,

$$\forall n \geq 0 \; [\; t_0sg((zc_x)^{pn}z)st_1 \in L \;]. \tag{11}$$

Let $z' := t_0sg(z)st_1$. We show $t_0sg((zc_x)^{pn}z)st_1 = (z'x')^{pn}z'$ ($\forall n \geq 0$). If $n = 0$, then $t_0sg((zc_x)^{pn}z)st_1 = t_0sg(z)st_1 = z' = (z'x')^{pn}z'$. If $n \geq 1$, then keeping in mind $x = t_1x't_0$ and (8), we have

$$\begin{aligned} &g((zc_x)^{pn}z) = (g(zc_x)s)^{pn}g(z) = (g(z)sg(c_x)s)^{pn}g(z) \\ &= (g(z)sxs)^{pn}g(z) = (g(z)st_1x't_0s)^{pn}g(z), \end{aligned}$$

so

$$\begin{aligned} &t_0sg((zc_x)^{pn}z)st_1 = t_0s(g(z)st_1x't_0s)^{pn}g(z)st_1 \\ &= (t_0sg(z)st_1x')^{pn}t_0sg(z)st_1 = (z'x')^{pn}z'. \end{aligned}$$

Thus we obtain $t_0sg((zc_x)^{pn}z)st_1 = (z'x')^{pn}z'$ ($\forall n \geq 0$). Combining this with (11), we obtain $(z'x')^{pn}z' \in L$ ($\forall n \geq 0$). Moreover, since $z \in \Sigma_0^*c_{y'}\Sigma_0^*$, we can write $z = \alpha c_{y'}\beta$ for some $\alpha, \beta \in \Sigma_0^*$. If $\alpha, \beta \in \Sigma_0^+$, then $g(z) = g(\alpha)sy'sg(\beta) \in Xy'X$. Similarly, we obtain $g(z) \in Xy'X$ in the case of $\alpha = \varepsilon$ or $\beta = \varepsilon$. Then $z' = t_0sg(z)st_1 \in Xy'X$. In summary, we obtain

$$\forall x', y' \in X_0, \; \exists z' \in Xy'X, \; \exists p \geq 1, \; \forall n \geq 0 \; [\; (z'x')^{pn}z' \in L \;].$$

Thus we complete the proof. □

4 Proof of Main Theorem 2.3

In this section, we prove Main Theorem 2.3. For $L \subseteq \Sigma^*$, we define $H_1(L)$ as

$$H_1(L): \; \forall u, v \in \Sigma^*, \; \exists z \in \Sigma^*v\Sigma^*, \; \exists p \geq 1, \; \forall n \geq 0 \; [\; (zu)^{pn}z \in L \;].$$

Note that $H_1(L)$ is exactly the same statement as (1). Next, we define

$$\mathcal{M}_1 := \mathbf{TL} \cup \{L \subseteq \Sigma^* \mid H_1(L)\}.$$

As for this $\mathcal{M}_1$, we can show the following Lemma:

Lemma 4.1. *$\mathcal{M}_1$ is closed under regular operations, i.e., we have $L_1 \cup L_2$, L_1L_2, $L_1^* \in \mathcal{M}_1$ for any $L_1, L_2 \in \mathcal{M}_1$.*

Once we have obtained this lemma, we can show Main Theorem 2.3 as follows:

Proof. We prove Main Theorem 2.3, provided that Lemma 4.1 is already proved. First, we trivially obtain $\mathbf{TL} \subseteq \mathcal{M}_1$. Moreover, $\mathcal{M}_1$ is closed under regular operations by Lemma 4.1. By the minimality of $\Gamma(\mathbf{TL})$, we obtain $\Gamma(\mathbf{TL}) \subseteq \mathcal{M}_1$. Now let $L \in \Gamma(\mathbf{TL})$ be dense. Since $\Gamma(\mathbf{TL}) \subseteq \mathcal{M}_1$, we have $L \in \mathcal{M}_1$. Then $L \in \mathbf{TL}$ or $H_1(L)$. Since L is dense, we must have $H_1(L)$, i.e., L satisfies (1). □

At this point, we have only to show Lemma 4.1. Therefore, the rest of this section is devoted to showing Lemma 4.1. First, one can trivially verify the closure of $\mathcal{M}_1$ under union by applying Lemma 3.1 and the following basic fact:

$$\forall A, B \subseteq \Sigma^* \; [\; [\; A \subseteq B,\; H_1(A)\;] \Rightarrow H_1(B)\;]. \tag{12}$$

Next, we show the closure under concatenation:

Proof. We first show the following:

$$\forall L_1, L_2 \subseteq \Sigma^* \; [\; [\; H_1(L_1) \vee H_1(L_2)\;] \Rightarrow [\; L_1L_2 = \emptyset \vee H_1(L_1L_2)\;]\;]. \tag{13}$$

Let $L_1, L_2 \subseteq \Sigma^*$ satisfy $H_1(L_1) \vee H_1(L_2)$. If $L_1L_2 = \emptyset$, then we obtain (13). Now we may assume $L_1L_2 \neq \emptyset$. Then $L_1 \neq \emptyset$ and $L_2 \neq \emptyset$, so we can take $l_1 \in L_1$ and $l_2 \in L_2$. If $H_1(L_1)$ holds, then let $u, v \in \Sigma^*$ be arbitrary. We apply $H_1(L_1)$ with l_2u and v. Then there exists $z \in \Sigma^*v\Sigma^*$ and $p \geq 1$ such that $(z(l_2u))^{pn}z \in L_1$ $(\forall n \geq 0)$. Let $z' := zl_2$. Then $z' \in \Sigma^*v\Sigma^*$. Moreover, $(z'u)^{pn}z' = (zl_2u)^{pn}zl_2 = ((zl_2u)^{pn}z)l_2 \in L_1l_2 \subseteq L_1L_2$ $(\forall n \geq 0)$. Thus we obtain $H_1(L_1L_2)$. Next, if $H_1(L_2)$ holds, then let $u, v \in \Sigma^*$ be arbitrary. We apply $H_1(L_2)$ with ul_1 and v. Then there exists $z \in \Sigma^*v\Sigma^*$ and $p \geq 1$ such that $(z(ul_1))^{pn}z \in L_2$ $(\forall n \geq 0)$. Let $z' := l_1z$. Then $z' \in \Sigma^*v\Sigma^*$. In general, we have $(xy)^nx = x(yx)^n$ for any $x, y \in \Sigma^*$ and $n \geq 0$, so $(z'u)^{pn}z' = (l_1zu)^{pn}l_1z = l_1(zul_1)^{pn}z \in l_1L_2 \subseteq L_1L_2$ $(\forall n \geq 0)$. Thus we obtain $H_1(L_1L_2)$, and we complete the proof of (13). Now the closure of $\mathcal{M}_1$ under concatenation trivially follows from (13), Lemma 3.2, and $\emptyset \in \mathbf{TL}$. □

Finally, we show the closure under Kleene star. For that, we need the following:

Lemma 4.2. *Let $A \subseteq \Sigma^*$. If A is thin and A^* is dense, then we have $H_1(A^*)$.*

Proof. STEP1: Let $\varepsilon \in \Sigma^*$ be the empty string. Let $A \subseteq \Sigma^*$. Assume that A is thin, A^* is dense, and $\varepsilon \notin A$. We show $H_1(A^*)$ in this case.[1] If $A = \emptyset$, then $A^* = \{\varepsilon\}$. However, since A^* is dense, this is a contradiction. Thus we obtain $A \neq \emptyset$. Next, since A is thin, we have $\Sigma^*t\Sigma^* \cap A = \emptyset$ for some $t \in \Sigma^*$. If $t = \varepsilon$, then $\Sigma^*\Sigma^* \cap A = \emptyset$, so we must have $A = \emptyset$, which is a contradiction. Thus we obtain $t \neq \varepsilon$. Since A^* is dense, we have $\Sigma^*t\Sigma^* \cap A^* \neq \emptyset$, so $t'tt'' \in A^*$ for some $t', t'' \in \Sigma^*$. Let $s := t'tt''$. Then $s \neq \varepsilon$ and $s \in A^*$. If $\Sigma^*s\Sigma^* \cap A \neq \emptyset$, then in view of $\Sigma^*s\Sigma^* \subseteq \Sigma^*t\Sigma^*$, we have $\Sigma^*t\Sigma^* \cap A \neq \emptyset$, which is a contradiction. Thus we obtain $\Sigma^*s\Sigma^* \cap A = \emptyset$. Next, let S_{pre} be the set of all prefixes of s and S_{suf} be the set of all suffixes of s. Note that we have $\varepsilon \in S_{pre}$ and $\varepsilon \in S_{suf}$. Let

$$S'_{pre} := \{\beta \in S_{pre} \mid \exists w \in \Sigma^* \; [\; w\beta \in A\;]\;\},$$
$$S'_{suf} := \{\gamma \in S_{suf} \mid \exists w \in \Sigma^* \; [\; \gamma w \in A\;]\;\}.$$

Since $A \neq \emptyset$, it is easy to verify that $\varepsilon \in S'_{pre}$ and $\varepsilon \in S'_{suf}$. Next, let

$$Q := \left\{(\beta, \alpha, \gamma) \mid \beta \in S'_{pre},\; \gamma \in S'_{suf},\; \alpha \in A^*,\; \beta\alpha\gamma = s\right\}.$$

[1] In fact, the additional assumption $\varepsilon \notin A$ is not essential, but we adopt this assumption for simplicity.

Note that Q is a finite set. In addition, since $s \in A^*$, we have $(\varepsilon, s, \varepsilon) \in Q$, so $Q \neq \emptyset$. Next, for $p = (\beta, \alpha, \gamma) \in Q$ and $q = (\beta', \alpha', \gamma') \in Q$, we define $R(p, q) := \{x \in \Sigma^* \mid \gamma x \beta' \in A^*\}$. Let $X := \Sigma^*$, $X_0 := \Sigma^*$, and $L := A^*$. We show (i), (ii), and (iii) of Lemma 3.5.

(i): Let $n \geq 1$ and $x_1, \cdots, x_n \in X_0$. We have to show there exists $p_0, \cdots, p_n \in Q$ such that $x_i \in R(p_{i-1}, p_i)$ $(\forall i \in [1, n])$. We first deal with the case $n = 1$. Then $x_1 \in X_0$ is given, and we have to show there exists $p_0, p_1 \in Q$ such that $x_1 \in R(p_0, p_1)$. First, since A^* is dense, we have $\Sigma^* s x_1 s \Sigma^* \cap A^* \neq \emptyset$, so $usx_1sv \in A^*$ for some $u, v \in \Sigma^*$. This implies that we can decompose the whole string usx_1sv into concatenations of strings in A. Keeping in mind $\Sigma^* s \Sigma^* \cap A = \emptyset$, the decomposition for each s (in usx_1sv) is like Fig. 1. Therefore, the decomposition for usx_1sv is like Fig. 2, so $p_0 := (\beta, \alpha, \gamma) \in Q$ and $p_1 := (\beta', \alpha', \gamma') \in Q$ in Fig. 2 satisfies $x_1 \in R(p_0, p_1)$. See also Fig. 3.

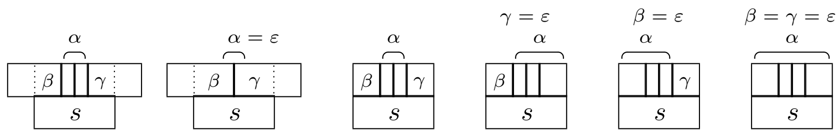

Figure 1: Six examples of decomposition for each s in usx_1sv.

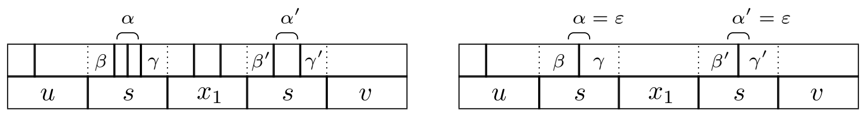

Figure 2: Two examples of decomposition for usx_1sv.

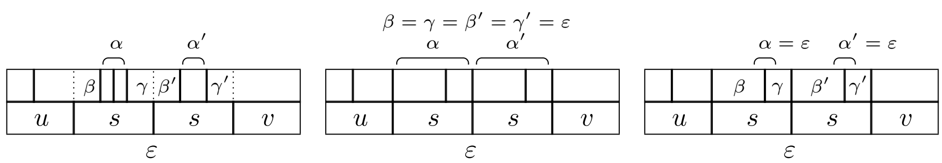

Figure 3: Three examples of decomposition for usx_1sv with $x_1 = \varepsilon$.

In general case $n \geq 1$, we have $\Sigma^* s x_1 s \cdots s x_n s \in \Sigma^* \cap A^* \neq \emptyset$, so $usx_1s \cdots sx_nsv \in A^*$ for some $u, v \in \Sigma^*$. This implies that we can decompose the whole string $usx_1s \cdots sx_nsv$ into concatenations of strings in A. Keeping in mind $\Sigma^* s \Sigma^* \cap A = \emptyset$, we can easily show that there exists $p_0, \cdots, p_n \in Q$ such that $x_i \in R(p_{i-1}, p_i)$ $(\forall i \in [1, n])$. See also Fig. 4 and Fig. 5.

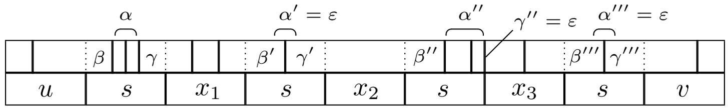

Figure 4: An example of decomposition.

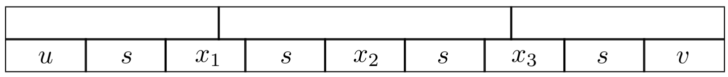

Figure 5: Decomposition like above is impossible due to $\Sigma^* s \Sigma^* \cap A = \emptyset$.

(ii): Let $p = (\beta, \alpha, \gamma) \in Q$ and $q = (\beta', \alpha', \gamma') \in Q$ be arbitrary. Since $\beta \in S'_{pre}$ and $\gamma' \in S'_{suf}$, we have $t_0\beta, \gamma' t_1 \in A$ for some $t_0, t_1 \in \Sigma^*$ $(= X_0)$. Let $x \in R(p, q)$ be arbitrary. Then $\gamma x \beta' \in A^*$, so $(t_0\beta)\alpha(\gamma x \beta')\alpha'(\gamma' t_1) \in A^*$. Since $\beta\alpha\gamma = s$ and $\beta'\alpha'\gamma' = s$, we obtain $t_0 s x s t_1 \in A^*$. Since $x \in R(p, q)$ is arbitrary, we have $t_0 s R(p, q) s t_1 \subseteq A^*$ $(= L)$, so we obtain (ii).
(iii): Let $p = (\beta, \alpha, \gamma) \in Q$, $q = (\beta', \alpha', \gamma') \in Q$, and $r = (\beta'', \alpha'', \gamma'') \in Q$ be arbitrary. Let $x \in R(p, q)$ and $y \in R(q, r)$. Then $\gamma x \beta' \in A^*$ and $\gamma' y \beta'' \in A^*$. In particular, $(\gamma x \beta')\alpha'(\gamma' y \beta'') \in A^*$. Since $\beta'\alpha'\gamma' = s$, we have $\gamma x s y \beta'' \in A^*$, so $xsy \in R(p, r)$. This implies $R(p, q) s R(q, r) \subseteq R(p, r)$. Thus we obtain (iii).
Consequently, we can apply Lemma 3.5, and we obtain (2). In other words,

$$\forall x, y \in \Sigma^*,\ \exists z \in \Sigma^* y \Sigma^*,\ \exists p \geq 1,\ \forall n \geq 0\ [\ (zx)^{pn} z \in A^*\ ].$$

This implies $H_1(A^*)$.
STEP2: Let $A \subseteq \Sigma^*$. Assume that A is thin and A^* is dense. Let $B := A - \{\varepsilon\}$. In general, we have $(A - \{\varepsilon\})^* = A^*$, so $B^* = A^*$. Since A^* is dense, it follows that B^* is dense. If B is dense, then in view of $B \subseteq A$, it follows that A is dense, which is a contradiction. Therefore, B is thin. Moreover, we have $\varepsilon \notin B$. Hence, by STEP1, we have $H_1(B^*)$. Since $B^* = A^*$, we obtain $H_1(A^*)$. $\square$

The closure of $\mathcal{M}_1$ under Kleene star trivially follows from Lemma 4.2 and (12). Hence, we complete the proof of Lemma 4.1.

5 On Theorem 2.2 and Main Theorem 2.3

In this section, we prove the following theorem:

Theorem 5.1. *Let $\Sigma = \{a, b\}$. Then there exists a dense $L \in \Gamma(\mathbf{TL})$ such that there is no dense $R \in \mathbf{REG}$ with $R \subseteq L$.*

In view of this theorem, we can say that Main Theorem 2.3 is a non-trivial generalization of Theorem 2.2.

Proof. Let $\mathbb{N}$ be the set of all positive integers. Let $I = \{(pqn)^4 + q \mid p, q, n \geq 1\} \subseteq \mathbb{N}$. We show the following:

(i) $\forall p, q \geq 1$ [$pm + q \in I$ *for infinitely many* $m \geq 1$].

(ii) $\forall p, q \geq 1$ [$pm + q \in \mathbb{N} - I$ *for infinitely many* $m \geq 1$].

(i): This is obvious.
(ii): For any $t \geq 1$, we can easily show $I \cap [1, t] \subseteq \{(pqn)^4 + q \mid 1 \leq p, q, n \leq t^{1/4}\}$. In particular, $|I \cap [1, t]| \leq t^{3/4}$, so $\lim_{t\to+\infty} |I \cap [1, t]|/t = 0$. Now let $p, q \geq 1$. We show $pm+q \in \mathbb{N}-I$ for infinitely many $m \geq 1$. Supposing the contrary, there exists $m_0 \geq 1$ such that $pm + q \in I$ $(\forall m \geq m_0)$. Let $J := \{pm + q \mid m \geq m_0\}$, for short. Then we have $J \subseteq I$. Combining this inclusion with $\lim_{t\to+\infty} |I \cap [1, t]|/t = 0$, we have $\lim_{t\to+\infty} |J \cap [1, t]|/t = 0$. However, since $J = \{pm + q \mid m \geq m_0\}$, we have $\lim_{t\to+\infty} |J \cap [1, t]|/t = 1/p$. This is a contradiction. Thus we obtain (ii).
Next, we define $f : a\Sigma^*b \cup \{\varepsilon\} \to \mathbb{N} \cup \{0\}$ as follows: For any $v \in a\Sigma^*b$, there exists unique $k \geq 1$ and unique $n_1, m_1, \cdots, n_k, m_k \geq 1$ such that $v = a^{n_1}b^{m_1} \cdots a^{n_k}b^{n_k}$. Then we define $f(v) := |\{i \in [1, k] \mid n_i \in I\}|$. We also define $f(\varepsilon) := 0$. As for this f, we can easily verify the following:

$$\forall v, w \in a\Sigma^*b \cup \{\varepsilon\} \ [\ vw \in a\Sigma^*b \cup \{\varepsilon\}\,,\ f(vw) = f(v) + f(w)\ ]. \tag{14}$$

Next, let $r \in \{0, 1\}$ be arbitrary. We define

$$L_r := \{v \in a\Sigma^*b \cup \{\varepsilon\} \mid f(v) \equiv r\ (mod\ 2)\} \subseteq a\Sigma^*b \cup \{\varepsilon\}\,.$$

We show L_r satisfies the desired property. We first show that L_r is dense in Σ^*. Take an $n_1 \in I$, and let $\alpha := a^{n_1}b$. Let $v \in \Sigma^*$ be arbitrary. Then $\alpha, avb \in a\Sigma^*b$. By (14), we have $f(avb\alpha^k) = f(avb) + kf(\alpha) = f(avb) + k$ $(\forall k \geq 1)$. In particular, we have $f(avb\alpha^{k_1}) \equiv r\ (mod\ 2)$ for some $k_1 \in \{1, 2\}$. Then $avb\alpha^{k_1} \in L_r$, i.e., we have $\Sigma^*v\Sigma^* \cap L_r \neq \emptyset$. Hence, L_r is dense. Next, suppose that there exists a dense $R \in \mathbf{REG}$ such that $R \subseteq L_r$. Let G be a deterministic finite automaton which represents R. Let $t \geq 1$ be the number of all states of G. Since R is dense, we have $\Sigma^*b^2a^{t+9}ba^2\Sigma^* \cap R \neq \emptyset$. Then we have $u'b^2a^{t+9}ba^2v' \in R$ for some $u', v' \in \Sigma^*$. Let $u := u'b$ and $v := av'$. Then $u, v \in \Sigma^+$ and $uba^{t+9}bav \in R$. By the definition of $t \geq 1$, we can apply a standard pumping argument, and we can show that there exists $p, q \geq 1$ such that $uba^{pn+q}bav \in R$ $(\forall n \geq 1)$. Since $R \subseteq L_r$, we have $uba^{pn+q}bav \in L_r$ $(\forall n \geq 1)$. Since $L_r \subseteq a\Sigma^*b \cup \{\varepsilon\}$, we have $uba^{pn+q}bav \in a\Sigma^*b$ $(\forall n \geq 1)$. Then, the first character of u must be a, and the last character of v must be b. In particular, we have $ub, av, a^{pn+q}b \in a\Sigma^*b$. By (14), we have $f(uba^{pn+q}bav) = f(ub) + f(a^{pn+q}b) + f(av)$. Since $uba^{pn+q}bav \in L_r$, we have $f(uba^{pn+q}bav) \equiv r\ (mod\ 2)$, so $f(ub) + f(a^{pn+q}b) + f(av) \equiv r\ (mod\ 2)$. By (i) and (ii), there exists $m, m' \geq 1$ such that $pm + q \in I$ and $pm' + q \in \mathbb{N} - I$. Then $f(ub) + 1 + f(av) \equiv r\ (mod\ 2)$ and $f(ub) + 0 + f(av) \equiv r\ (mod\ 2)$, which is a contradiction. Hence, there is no dense $R \in \mathbf{REG}$ with $R \subseteq L_r$. Finally, we show $L_r \in \Gamma(\mathbf{TL})$. Let

$$A := \{a^nb^m \mid n \in I,\ m \geq 1\}\,,\ B := \{a^nb^m \mid n \in \mathbb{N} - I,\ m \geq 1\}\,.$$

Consider the following language equations:

$$X_0 = AX_1 \cup BX_0 \cup \{\varepsilon\}\,,\ X_1 = AX_0 \cup BX_1. \tag{15}$$

Let $Y_0, Y_1 \subseteq \Sigma^*$ be the least solution of (15). In fact, we can explicitly write $Y_0 = (AB^*A \cup B)^*$ and $Y_1 = B^*A(AB^*A \cup B)^*$. Since $A, B \in \mathbf{TL} \subseteq \Gamma(\mathbf{TL})$, we have $Y_0, Y_1 \in \Gamma(\mathbf{TL})$. Moreover, we can easily show that L_0, L_1 is also the least solution of (15). Hence, we must have $L_0 = Y_0$ and $L_1 = Y_1$, so $L_r \in \Gamma(\mathbf{TL})$. $\square$

6 Some remarks

In this section, we give some remarks.

6.1 On Lemma 4.1

Let $H_2(L)$ be a statement defined as

$$H_2(L):\ \exists z \in \Sigma^+,\ \exists p \geq 1,\ \forall n \geq 0\ [\ z^{pn+1} \in L\ ].$$

Let $\mathcal{M}_2 := \mathbf{TL} \cup \{L \subseteq \Sigma^* \mid H_2(L)\}$. It is natural to consider $\mathcal{M}_2$ instead of $\mathcal{M}_1$ in Lemma 4.1. We would like to show that $\mathcal{M}_2$ is closed under regular operations. However, $\mathcal{M}_2$ is not closed under concatenation. For example, let $\Sigma = \{a, b\}$, $L_1 := \left\{va^{10|v|} \mid v \in \Sigma^+\right\}$, and $L_2 := \{b\}$. We can show that $L_1, L_2 \in \mathcal{M}_2$ and $L_1L_2 \notin \mathcal{M}_2$. This is why we have considered $H_1(L)$ instead of $H_2(L)$.

6.2 On Main Theorem 2.3

For any $\mathcal{L} \subseteq 2^{\Sigma^*}$, consider the following claim:

Claim 6.1. *Let $R \in \mathcal{L}$ be dense. Then there exists $z \in \Sigma^+$ and $p \geq 1$ such that $z^{pn+1} \in R$ $(\forall n \geq 0)$.*

Note that Claim 6.1 with $\mathcal{L} = \mathbf{REG}$ is exactly Theorem 2.2. Moreover, Claim 6.1 with $\mathcal{L} = \Gamma(\mathbf{TL})$ is also true, as we have already shown. Keeping in mind Dömösi-Horváth-Ito conjecture, it is desirable to prove Claim 6.1 for $\mathcal{L} = \mathbf{CFL}$, because in this case we trivially obtain Dömösi-Horváth-Ito conjecture (by considering $R = Q_\Sigma$). However, in fact, Claim 6.1 does not hold even if $\mathcal{L} = \mathbf{DCFL}$ (deterministic context-free languages). Here we provide a counter-example. Let $\Sigma = \{\ \langle\ ,\ \rangle\ \}$. Let $L \subseteq \Sigma^*$ be the Dyck language over Σ. Let $R := \{\ v\rangle\ \mid v \in L\}$. It is easy to show that R is dense, $R \subseteq Q_\Sigma$, and $R \in \mathbf{DCFL}$. Therefore, this R is a counter-example of Claim 6.1 for $\mathcal{L} = \mathbf{DCFL}$. This fact implies that extending Theorem 2.2 is a hard problem in general. This situation is already indicated in our proofs: we have proved non-trivial lemmas to obtain Main Theorem 2.3.

7 Related work

In this section, we briefly state some related work. For $L \subseteq \Sigma^*$, let $\sim_L$ be the syntactic equivalence of L, and let $\Sigma^*/\sim_L$ be the syntactic monoid of L. By Myhill-Nerode Theorem, we can easily show that L is a regular language iff $\Sigma^*/\sim_L$ is a finite monoid (see also [10, Proposition 3.18]).

7.1 Related work for dense and disjunctive language

A language $L \subseteq \Sigma^*$ is said to be disjunctive iff $\forall u, v \in \Sigma^*[\ u \sim_L v \Leftrightarrow u = v\ ]$. In particular, if L is a disjunctive language, then $\Sigma^*/\sim_L$ is an infinite set.

The concept of disjunctive languages is closely related to dense languages. For example, it is shown in [11, PROPOSITION 2.5] that a language L is dense iff there exists a disjunctive language L' such that $L' \subseteq L$. Many properties of disjunctive languages are already known (e.g., [11, 14]). Moreover, many connections between dense and disjunctive languages are known (e.g., [5, 6]).

7.2 Related work for Theorem 1.1, 2.2, and Main Theorem 2.3

As for Theorem 1.1, 2.2, and Main Theorem 2.3, we refer to [6, 9] as direct related works. Theorem 2.2 is exactly the same as [9, Corollary 4.6]. Next, it is shown in [6] that if $R \subseteq \Sigma^*$ is a dense regular language, then $R \cap Q_\Sigma$ and $R - (R \cap Q_\Sigma)$ are disjunctive. Note that this result implies the following:

Proposition 7.1. *If $R \in \mathbf{REG}$ is dense, then R contains infinitely many non primitive words.*

This is because of the following reasons: let $R \in \mathbf{REG}$ be dense. By [6], $L := R - (R \cap Q_\Sigma)$ is disjunctive. In particular, $\Sigma^*/\sim_L$ is an infinite set. If L is a finite set, then L is regular, so $\Sigma^*/\sim_L$ is a finite monoid. Then $\Sigma^*/\sim_L$ is a finite set, which is a contradiction. Thus, L is an infinite set, i.e., R contains infinitely many non primitive words, so we obtain Proposition 7.1.

Note that Proposition 7.1 is almost the same as Theorem 2.2 (and Theorem 1.1). The only difference is that Theorem 2.2 (and Theorem 1.1) tells us specific examples of non primitive words, i.e., R contains infinitely many non primitive words of the form z^{pn+1}, while Proposition 7.1 does not tell us such examples. As for Main Theorem 2.3, we have proved that if $L \in \Gamma(\mathbf{TL})$ is dense, then we have the condition (1), so there exists $z \in \Sigma^+$ and $p \geq 1$ such that $z^{pn+1} \in L$ ($\forall n \geq 0$). In addition, Main Theorem 2.3 is a non-trivial generalization of Theorem 2.2, as we have already proved in Section 5.

References

[1] Berstel, Jean and Perrin, Dominique. *Theory of Codes.* Academic Press, 1985.

[2] Berstel, Jean, Perrin, Dominique, and Reutenauer, Christophe. *Codes and Automata*, volume 129. Cambridge University Press, 2010.

[3] Dömösi, Pál, Horváth, S, and Ito, M. Formal languages and primitive words. *Publ. Math. Debrecen*, 42(3-4):315–321, 1993.

[4] Dömösi, Pál and Ito, Masami. *Context-free Languages and Primitive Words.* World Scientific, 2015.

[5] Guo, YQ, Xu, GW, and Thierrin, Gabriel. Disjunctive decomposition of languages. *Theoretical Computer Science*, 46:47–51, 1986. DOI: `10.1016/0304-3975(86)90020-4`.

[6] Ito, Masami. Dense and disjunctive properties of languages. In *International Symposium on Fundamentals of Computation Theory*, pages 31–49. Springer, 1993. DOI: `10.1007/3-540-57163-9_3`.

[7] Koga, Toshihiro. On the density of regular languages. *Fundamenta Informaticae*, 168(1):45–49, 2019. DOI: `10.3233/FI-2019-1823`.

[8] Li, Zheng-Zhu, Shyr, Huei-Jan, and Tsai, YS. Classifications of dense languages. *Acta Informatica*, 43(3):173–194, 2006. DOI: `10.1007/s00236-006-0015-y`.

[9] Liu, YJ and Xu, ZB. Monoid algorithms and semigroup properties related to dense regular languages. In *Proceedings of the International Conference on Algebra and Its Applications, Bangkok*, pages 203–214. Citeseer, 2002.

[10] Pin, Jean-Éric. *Mathematical Foundations of Automata Theory*, volume 7. 2010.

[11] Reis, CM and Shyr, Huei-Jan. Some properties of disjunctive languages on a free monoid. *Information and Control*, 37(3):334–344, 1978. DOI: `10.1016/S0019-9958(78)90578-8`.

[12] Salomaa, Arto and Soittola, Matti. *Automata-Theoretic Aspects of Formal Power Series.* Springer Science & Business Media, 2012.

[13] Shyr, HJ and Tseng, Din-Chang. Some properties of dense languages. *Soochow J. Math*, 10:127–131, 1984.

[14] Shyr, Huei-Jan. Disjunctive languages on a free monoid. *Information and Control*, 34(2):123–129, 1977. DOI: `10.1016/S0019-9958(77)80008-9`.

[15] Sin'ya, Ryoma. An automata theoretic approach to the zero-one law for regular languages: Algorithmic and logical aspects. *Electronic Proceedings in Theoretical Computer Science*, 193:172–185, Sep 2015. DOI: `10.4204/eptcs.193.13`.

[16] Sin'ya, Ryoma. Asymptotic approximation by regular languages. In *SOFSEM 2021: Theory and Practice of Computer Science*, pages 74–88, Cham, 2021. Springer International Publishing. DOI: `10.1007/978-3-030-67731-2_6`.

[17] Sin'ya, Ryoma. Asymptotic approximation by regular languages, 2021. Online Worldwide Seminar on Logic and Semantics, `https://www.cs.bham.ac.uk/~vicaryjo/owls/slides/sinya.pdf`.

Received 14th June 2021