International Journal of Algebra, Vol. 19, 2025, no. 1, 1 - 5
HIKARI Ltd, www.m-hikari.com
https://doi.org/10.12988/ija.2025.91861

A Note on the Disjunctive Decomposition of Dense Languages

Di Zhang

School of Mathematics
Guangdong University of Education
Guangzhou, 510303, Guangdong Province, China

Abstract

The relationship between dense languages and disjunctive languages has been widely studied. Guo et al. in [1] proved that for any finite alphabet X, the dense languages over X can be decomposed into a disjoint union of two disjunctive languages, one of which is discrete. In this note, we improve the algorithm for disjunctive decomposition of dense languages in the case of $|X| \geq 2$, and we present a new algorithm for disjunctive decomposition of dense languages in the case of $|X| = 1$.

Keywords: Disjunctive languages; Dense languages; Discrete; Decomposition

Mathematics Subject Classification: 20M35

1 Introduction and Preliminaries

Let X be a nonempty finite set called an alphabet. The elements in X are called the letters. Let X^* be the free monoid generated by an alphabet X. The elements and subsets of X^* are called the *words* and *languages* over X respectively. The length of a word w over X is the number of letters occurring in w and is denoted by $|w|$. We denote the cardinality of a language L over X by $|L|$. For a given language L over X, the relation P_L on X^* defined by

$$x \equiv y(P_L) \;\Leftrightarrow\; ''(\forall u, v \in X^*)\; uxv \in L \Leftrightarrow uyv \in L''$$

is a congruence on free monoid X^* and is known as the *principal congruence* determined by L. As usual, the set of all positive (nonnegative) integers is denoted by $\mathbb{N}(\mathbb{N}^0)$.

Let L be a language over X. Then L is said to be *dense* if $X^*wX^* \bigcap L \neq \emptyset$ for any $w \in X^*$; L is said to be *disjunctive* if P_L is the equality relation on X^*. A language L over X is called *discrete* if any distinct words $x, y \in L$, $|x| \neq |y|$. A language L over X is called *semi-discrete* if for some positive integer k, L contains at most k words for any given length. Obviously, if $k = 1$, then a semi-discrete language is discrete. The dense languages and disjunctive languages are closely related. From [3], a disjunctive language is dense and the converse is not true. But H. J. Shyr in [4] showed that the dense languages over a finite alphabet X with $|X| \geq 2$ can be decomposed into a disjoint union of infinitely many disjunctive languages. Since then, the disjunctive structure of dense languages have been considered by many authors in the literature, such as, Xu *et al.* in [5] showed that for any finite alphabet X with $|X| \geq 2$, the dense languages over X can be written as a disjoint union of two disjunctive languages, one of which is semi-discrete. And Guo *et al.* in [1] proved the following two facts.

(i) For $|X| = 1$, any dense language over X can be written as a disjoint union of infinitely many disjunctive languages.

(ii) For any finite alphabet X, dense languages over X can be written as a disjoint union of two disjunctive languages with one of which is discrete.

In this note, we improve the proof of the above proposition (ii), specifically, we simplify the algorithm for disjunctive decomposition of dense languages in the case of $|X| \geq 2$, and we give a new algorithm for disjunctive decomposition of dense languages in the case of $|X| = 1$.

In this paper, the free monoid X^* needs to be equipped with a total order, we adopt the *standard total order* $\leq$ which is defined on X^* as follows([3]): For any $u, v \in X^*$, if $|u| < |v|$, then $u < v$; if $|u| = |v|$, then $\leq$ is the lexicographical order on X^n for all positive integer n.

2 Main result

The main result of this paper is the following.

Proposition. *For any finite alphabet X, dense languages D over X can be decomposed into a disjoint union of two disjunctive languages with one of which is discrete.*

Proof of the Proposition. Let D be a dense language. We shall construct a discrete dense subset D_1 of D.

We first consider the case that $|X| \geq 2$. Let $\leq$ be the standard total order.

Suppose $X^* = \{x_1, x_2, \dots, x_n, \dots\}$, where $x_i \leq x_j$ for $i \leq j$. Let

$$A_i = \{(u,v) \in X^* \times X^* | ux_iv \in D\}, \quad i = 1, 2, \dots.$$

Since D is dense, by Proposition 4.19 of [3], we have $|A_i| = \infty$, $i = 1, 2, \dots$.

Now we take $(u_1, v_1) \in A_1$, and let $T_1 = \{(u,v) \in X^* \times X^* || uv| \leq |u_1v_1|\}$. Then we take $(u_2, v_2) \in A_2 \backslash T_1$, this is possible because $|A_2| = \infty$ and $|A_2 \backslash T_1| = \infty$. Assume that for some $k \geq 2$, we have already chosen $(u_k, v_k) \in A_k \setminus T_{k-1}$, where $T_{k-1} = \{(u,v) \in X^* \times X^* || uv| \leq |u_{k-1}v_{k-1}|\}$. Let $T_k = \{(u,v) \in X^* \times X^* || uv| \leq |u_kv_k|\}$. Since $|A_{k+1} \setminus T_k| = \infty$, we can choose $(u_{k+1}, v_{k+1}) \in A_{k+1} \setminus T_k$. Thus, we obtain inductively an infinite subset

$$T = \{(u_i, v_i) \in X^* \times X^* | \ i = 1, 2, \dots\}$$

of $X^* \times X^*$. Denote

$$D_1 = \{u_ix_iv_i | \ i = 1, 2, \dots\}.$$

From the definition of A_i, and the fact $(u_i, v_i) \in A_i$, it follows that D_1 is a dense subset of D. For any $i, j \in \mathbb{N}$ with $j > i$, by the manner of choosing (u_i, v_i), we have $|u_jv_j| > |u_iv_i|$. By the definition of $\leq$, it follows that $|x_j| \geq |x_i|$. Hence

$$|u_jx_jv_j| > |u_ix_iv_i|.$$

Thus, D_1 is discrete. Therefore, D_1 is a disjunctive subset of D by [3].

Now we show that $D \setminus D_1$ is disjunctive. Let $D_2 = D \setminus D_1$. For any $x_i, x_j \in X^*$ with $|x_i| = |x_j|$, $1 \leq i < j$, consider $A_i \setminus T_{j-1}$ and $A_j \setminus T_{j-1}$. If $A_i \setminus T_{j-1} = A_j \setminus T_{j-1}$, then taking (u_j, v_j), we have $u_jx_jv_j \in D_1$, which implies $u_jx_jv_j \notin D_2$. Since $(u_j, v_j) \in A_j \setminus T_{j-1} = A_i \setminus T_{j-1}$, it follows that $u_jx_iv_j \in D$. By $x_i < x_j$ and $|x_i| = |x_j|$, we have $u_jx_iv_j \neq u_jx_jv_j$ and $|u_jx_iv_j| = |u_jx_jv_j|$. Since D_1 is discrete, it follows that $u_jx_iv_j \notin D_1$. Hence $u_jx_iv_j \in D_2$. This shows that $x_i \not\equiv x_j (P_{D_2})$. If $A_i \setminus T_{j-1} \neq A_j \setminus T_{j-1}$, then we have the following two cases:

(1) There exists $(u,v) \in X^* \times X^*$ such that $(u,v) \in A_i \setminus T_{j-1}$, $(u,v) \notin A_j \setminus T_{j-1}$.

(2) There exists $(u,v) \in X^* \times X^*$ such that $(u,v) \in A_j \setminus T_{j-1}$, $(u,v) \notin A_i \setminus T_{j-1}$.

For case (1), since $(u,v) \in A_i \setminus T_{j-1}$, we have $ux_iv \in D$ and $|uv| > |u_{j-1}v_{j-1}|$, where $(u_{j-1}, v_{j-1}) \in T$. Thus, $|ux_iv| > |u_{j-1}x_iv_{j-1}|$. From $|x_i| = |x_j|$, we have $|x_i| = |x_{j-1}|$. Hence $|ux_iv| > |u_{j-1}x_{j-1}v_{j-1}|$. If $|uv| < |u_jv_j|$, where $(u_j, v_j) \in T$, then $|ux_iv| < |u_jx_iv_j| = |u_jx_jv_j|$. By the construction of D_1, ux_iv is not in D_1. Hence $ux_iv \in D \setminus D_1$. If $|uv| \geq |u_jv_j|$, we can choose another $(u'_j, v'_j) \in A_j \setminus T_{j-1}$ such that $|u'_jv'_j| > |uv|$. This can be done because $|A_j \setminus T_{j-1}| = \infty$. Then we replace (u_j, v_j) of T with (u'_j, v'_j), and we reselect (u_k, v_k) of T, $k > j$, according to the selection method stated in the previous paragraph, and replace the corresponding elements of D_1 with the modified $u_kx_kv_k, k = j, j+1, \ldots$. Then D_1 remains a dense and discrete subset of D. Now it is clear that $|ux_iv| < |u_jx_iv_j|$. Therefore, $ux_iv \notin D_1$, which implies $ux_iv \in D \setminus D_1$. On the other hand, since $|uv| > |u_{j-1}v_{j-1}|$ and $(u, v) \notin A_j \setminus T_{j-1}$, we have $(u, v) \notin A_j$. Thus $ux_jv \notin D$. This shows that $x_i \not\equiv x_j(P_{D_2})$.

For case (2), since $|uv| > |u_{j-1}v_{j-1}|$ and $(u, v) \notin A_i \setminus T_{j-1}$, we have $ux_iv \notin D$. Now we turn to ux_jv. From $(u, v) \in A_j \setminus T_{j-1}$, we have $ux_jv \in D$ and $|ux_jv| > |u_{j-1}x_jv_{j-1}| = |u_{j-1}x_{j-1}v_{j-1}|$. If $|uv| < |u_jv_j|$, then $|ux_jv| < |u_jx_jv_j|$. This shows $ux_jv \notin D_1$, whence $ux_jv \in D \setminus D_1$. If $|uv| \geq |u_jv_j|$, then similar to case(1), we also choose another $(u'_j, v'_j) \in A_j \setminus T_{j-1}$ to replace (u_j, v_j) of T such that $|u'_jv'_j| > |uv|$. And the (u_k, v_k) in T, $k > j$, will be reselected, the corresponding elements of D_1 will be replaced to ensure that D_1 remains a dense and discrete subset of D. Then we have $|ux_jv| < |u'_jx_jv'_j|$, whence $ux_jv \notin D_1$. Therefore, $ux_jv \in D \setminus D_1$. This shows that $x_i \not\equiv x_j(P_{D_2})$.

Thus D_2 is disjunctive.

For $|X| = 1$, suppose $X = \{a\}$, we present a construction that differs from the one in [1]. Since D is dense, D is infinite. Denote $D = \{a^{i_1}, a^{i_2}, \ldots\}$, where $i_1, i_2, \ldots \in \mathbb{N}^0$, $i_1 < i_2 < \ldots$. Construct

$$
\begin{aligned}
A_0 &= \{a^{i_{2^0}}\},\\
A_1 &= \{a^{i_{2^1}}, a^{i_{2^1+1}}\},\\
A_2 &= \{a^{i_{2^2}}, a^{i_{2^2+1}}, \ldots, a^{i_{2^3-1}}\},\\
&\vdots\\
A_k &= \{a^{i_{2^k}}, a^{i_{2^k+1}}, \ldots, a^{i_{2^{k+1}-1}}\},\\
&\ddots
\end{aligned}
$$

Obviously, $D = \cup_{k=0}^{\infty} A_k$ and $A_h \cap A_l = \emptyset$ for all $h, l \in \mathbb{N}^0$, $h \neq l$. Let $D_1 = \cup_{n=0}^{\infty} A_{2^n}$, $D_2 = D \setminus D_1$. For any $m \in \mathbb{N}$, there exists a $t \in \mathbb{N}$ such that $2^t > m$. Consider $a^{i_{2^{2^t+1}-1}}$ and $a^{i_{2^{2^t+1}}}$. It is clearly that $a^{i_{2^{2^t+1}-1}}$ is the longest element in A_{2^t}, $a^{i_{2^{2^t+1}}}$ is the shortest element in $A_{2^{t+1}}$. So $a^{i_{2^{2^t+1}-1}}$ and $a^{i_{2^{2^t+1}}}$ are in D_1, and by the construction of D_1, D_1 does not contain any word with a length between $i_{2^{2^t+1}-1}$ and $i_{2^{2^{t+1}}}$. We also have

$$
i_{2^{2^{t+1}}} - i_{2^{2^t+1}-1} \geq 2^{2^{t+1}} - 2^{2^t+1} + 1 = 2^{2^t}(2^{2^t} - 2) + 1.
$$

Sinse $2^t > m$, it follows that $2^{2^t} > 2^m$. And by $t \in \mathbb{N}$, we have $2^{2^t} - 2$ as a positive integer. Thus $2^{2^t}(2^{2^t} - 2) + 1 > 2^m > m$. Therefore, by Lemma 1 of [1], D_1 is disjunctive. Similar to the proof of D_1, we can show that D_2 is disjunctive.

References

[1] Y. Q. Guo, L. Li and G. W. Xu, On the disjunctive Structure of dense languages, *Science in China, Series A*, **28** (1985), 1233-1238.

[2] Y. Q. Guo, G. Thierrin and G. W. Xu, Disjunctive Decomposition of Languages, *Theoretical Computer Science*, **46** (1986), 47-51. https://doi.org/10.1016/0304-3975(86)90020-4

[3] H. J. Shyr, *Free Monoids and Languages*, Third Edition, Hon Min Book Company, Taichung, Taiwan, 2001.

[4] H. J. Shyr, A Characterization of Dense Languages, *Semigroup Forum*, **30** (1984), 237-240. https://doi.org/10.1007/bf02573454

[5] G. W. Xu, Y. Q. Guo and L. Li, A union decomposition of dense languages with respect to disjunctive languages, *Acta Math. Sinica*, **29** (1986), 184-188.

Received: August 27, 2024; Published: January 3, 2025