On Some Decompositions of r-Disjunctive Languages

1Di Zhang, 2Yuqi Guo and 3K. P. Shum

1, 2School of Mathematics and Statistics, Lanzhou University, Lanzhou, 730000, China
3Institute of Mathematics, Yunnan University, Kunming, 650091, China
1zhdi88@gmail.com, 2yqguo259@swu.edu.cn, 3kpshum@ynu.edu.cn

Abstract. Some kinds of decompositions of r-disjunctive languages on an arbitrary alphabet will be investigated. We will show that an f-disjunctive (t-disjunctive) language can be divided into two parts and either one part of them is an f-disjunctive (t-disjunctive) language or both parts are r-disjunctive but not f-disjunctive (t-disjunctive) languages. Finally, a relevant result of H. J. Shyr and S. S. Yu concerning the disjunctive languages will be improved.

2010 Mathematics Subject Classification: 20M35

Keywords and phrases: Disjunctive languages, f-disjunctive languages, t-disjunctive languages, r-disjunctive languages, r-regular languages.

1. Introduction and preliminaries

Let X be a nonempty finite set called alphabet in which the elements are called letters. Let $X^*$ be the free monoid generated by an alphabet $X$. Then, the elements and subsets of $X^*$ are called the words and languages over $X$. The identity of the free monoid $X^*$ is called an empty word and is denoted by 1. Let $X^* = X^* \setminus \{1\}$ be the free semigroup generated by $X$. The length of a word $w$ over $X$ is the number of letters occurring in $w$ and is denoted by $lg(w)$. We denote the cardinality of a language $L$ over $X$ by $|L|$. For any two languages $A, B$ over $X$, the concatenation $AB$ of $A$ and $B$ is the language $\{xy | x \in A, y \in B\}$ over $X$. For a given language $L$ over $X$, the relation $P_L$ on $X^*$ defined by

$$x \equiv y(P_L) \iff \forall u, v \in X^* \exists u, v \in L \iff uyv \in L$$

is a congruence on free monoid $X^*$ and is known as the principal congruence determined by $L$. The quotient monoid $X^*/P_L$ is called the syntactic monoid of $L$ and is denoted by $\text{Syn}(L)$. For any word $u$ over $X$, we often use $[u]_L$ to denote the $P_L$-class of $X^*$ containing $u$. As usual, the set of all positive (nonnegative) integers is denoted by $\mathbb{N}(\mathbb{N}^0)$.

We call a language $L$ over $X$ disjunctive [9] if $P_L$ is the equality relation on $X^*$. Let $\mathcal{D}$ be the class of all disjunctive languages over $X$. A language $L$ over $X$ is called regular [5, 10] if the index of $P_L$ (i.e., the number of $P_L$-classes of $X^*$) is finite. Let $\mathcal{R}$ be the class of all regular languages over $X$. Then we call a language $L$ over $X$ a midst-language [12] if $L$ is...
D. Zhang, Y. Q. Guo and K. P. Shum

neither regular nor disjunctive. Let $\mathcal{M}$ be the class of all midst-languages over $X$. Then, we have the following proposition.

**Proposition 1.1.** [11] Let $X$ be an alphabet with $|X| = 1$. Then $\mathcal{M} = \emptyset$, that is, over $X$, a language is disjunctive if and only if it is not regular.

But when $|X| \geq 2$, the case is completely different from $|X| = 1$ [6]. In this case, $\mathcal{M} \neq \emptyset$, that is, \{R $\cup$ D\} $\subsetneq X^*$.

We call a language $L$ over $X$ dense if $X^wX^* \cap L \neq \emptyset$ for any $w \in X^*$; otherwise, the language $L$ is said to be thin. According to Reis and Shyr [10], a language $L$ is dense if and only if $L$ contains a disjunctive language. Denote the class of all dense languages over $X$ by $D_d$.

The generalized disjunctive languages have been considered by a number of authors in the literature, such as, Guo, Reis and Thierrin [1] in 1988 called a language $L$ over $X$ relatively f-disjunctive (relatively disjunctive), that is, rf-disjunctive for short (r-disjunctive for short), if there exists a dense language $D$ over $X$ such that for all $u \in X^*$, $|[u]_L \cap D| < \infty$ ($|[u]_L \cap D| \leq 1$). It has been shown in [1] that $L$ is rf-disjunctive if and only if $L$ is r-disjunctive, if and only if either $X^*$ has no dense $P_L$-classes or has infinitely many dense $P_L$-classes. Let $D_r$ be the class of all r-disjunctive languages over $X$. Then, the concept of relatively regular language was first introduced by Liu, Shum and Guo in 2008 (see [6]). They called a language $L$ over $X$ relatively regular, that is, r-regular for short, if Syn($L$) has a finite ideal. Let $R_r$ be the class of all r-regular languages over $X$.

This paper is based on the following background.

(I) Obviously, when $|X| = 1$, $D_r = D$, $R_r = R$. In [6], the authors proved the following fact which forms a generalization of Proposition 1.1 to any alphabet $X$ from $|X| = 1$.

**Proposition 1.2.** [6] Let $X$ be an alphabet. Then a language over $X$ is r-disjunctive if and only if it is not r-regular.

Leading up to [1], [6], some special cases of r-disjunctive languages have been defined. In particular, a language $L$ over an alphabet $X$ was first called by Guo, Shyr and Thierrin [2] f-disjunctive if each $P_L$-class of $X^*$ is finite, and later, Mu [8] called a language $L$ over $X$ t-disjunctive if each $P_L$-class of $X^*$ is thin. Denote the class of all f-disjunctive (t-disjunctive) languages over $X$ by $D_f$ ($D_t$).

The following proposition is useful in this paper.

**Proposition 1.3.** [1, 7]

1. If $|X| = 1$, then $D = D_f = D_t = D_r$.

2. If $|X| \geq 2$, then $D \subseteq D_f \subsetneq D_t \subsetneq D_r \subsetneq D^d$.

Some more characterizations of r-disjunctive languages can be found in [1, 2, 6, 8, 9, 11].

(II) The following result is a known result on some decompositions of disjunctive languages.

**Proposition 1.4.** [10] Let $L$ be a disjunctive language over $X$, $L = L_1 \cup L_2$ (i.e., $L = L_1 \cup L_2$ and $L_1 \cap L_2 = \emptyset$). Then, the following statements hold.
On Some Decompositions of r-Disjunctive Languages

729

(1) \( \{L_1, L_2\} \cap D \neq \emptyset \), or
(2) \( \{L_1, L_2\} \subseteq D_r^d \setminus D_r \).

In the case (2) of Proposition 1.4, each one of \( L_1 \) and \( L_2 \) is not r-regular language, for if not, then another one of them must be disjunctive by [6]. This fact clearly contradicts to \( L_1, L_2 \notin D_r \). Hence, Proposition 1.4 can be modified to the following form.

**Proposition 1.4'.** Let \( L \) be a disjunctive language over \( X \), \( L = L_1 \cup L_2 \). Then, the following statements hold.

(1) \( \{L_1, L_2\} \cap D \neq \emptyset \), or
(2) \( \{L_1, L_2\} \subseteq D_r \setminus D_f \).

(III) Proposition 1.4' actually says that the disjoint union decompositions of disjunctive languages have two cases. Any disjunctive language has the decomposition of the case (1) in Proposition 1.4. As to the decomposition in the case (2) of Proposition 1.4, not every disjunctive language has this decomposition, for instance, discrete disjunctive language, that is, the disjunctive language \( L \) with \( |L \cap X^n| \leq 1 \) for any \( n \in \mathbb{N} \) [10]. But Shyr and Yu have shown in [12] that there exists such a disjunctive language \( L \) over \( X \) with \( |X| \geq 3 \) so that \( L = L_1 \cup L_2 \), \( \{L_1, L_2\} \subseteq D_r \setminus D_f \), to be more precise, we have

(1.1) \( \{L_1, L_2\} \subseteq D_f \setminus D_r \).

In this paper, we will further discuss in Section 4 about the existence case mentioned by Shyr and Yu in [12]; we will also discuss the decompositions of languages in \( D_f, D_t \) and \( D_r \) like Proposition 1.4' in Section 2; and in Section 3, we will show that the languages in \( D_f, D_t \) have similar decompositions just as the case (2)' in Proposition 1.4' with (1.1).

Making contact with the above background of this paper, we have started to apply some results of this paper to our following work to describe the disjunctive degree in some sense of languages, this shows one spot of the potential value of this paper.

For terminologies and notations not mentioned in this paper, the reader is referred to [4, 5, 10].

In the remaining part of the paper, we always assume that \( |X| \geq 2 \).

2. Some decompositions of r-disjunctive languages(I)

In the following theorem, we consider the decompositions of languages in \( D_f, D_r \) and \( D_t \) which are similar to Proposition 1.4'.

**Theorem 2.1.** Let \( L \in D_f (D_t, D_r) \), \( L = L_1 \cup L_2 \). Then the following statements hold:

(1) \( \{L_1, L_2\} \cap D_f (D_t, D_r) \neq \emptyset \), or
(2) \( \{L_1, L_2\} \subseteq D_r \setminus D_f (D_r \setminus D_t, D_r \setminus D_r) \).

**Remark 2.1.** If we divide a language \( L \in D_r \) into \( L_1 \) and \( L_2 \), then \( L_1, L_2 \) must satisfy Theorem 2.1(1). Otherwise, both of \( L_1, L_2 \) are not in \( D_r \), by Proposition 1.2, they are in \( D_r \). Hence, \( L \) is in \( D_r \) because by [6], \( D_r \) is closed under the operation of union, this is clearly a contradiction. Here, we write \( \{L_1, L_2\} \subseteq D_r \setminus D_r \) to seek a unity of expression with the \( D_f \) and \( D_t \) languages.

In proving the above theorem, we need the following proposition.
Proposition 2.1.

(1) If \( L \subseteq D_f \), \( R \subseteq L \), and \( R \subseteq \mathcal{R} \), then \( L \setminus R \subseteq D_f \).

(2) If \( L \subseteq D_r \), \( R \subseteq L \), and \( L \cap R = \emptyset \), then \( L \cup R \subseteq D_r \).

Proof. By [6], Proposition 2.1 holds for \( L \subseteq D_f \) and \( L \subseteq D_r \). Here we just discuss about \( L \subseteq D_r \).

(1) Suppose that \( L \subseteq D_r \). Then for any dense language \( \{x_1, x_2, \ldots, x_m, \ldots\} \), there exist \( x_i, x_j, i \neq j \) such that \( x_i \equiv x_j(P_L) \).

Since \( R \subseteq \mathcal{R} \), by Lemma 3.3 of [6], there exist an \( w \in X^* \), and an \( n \in \mathbb{N} \) such that \( (wx)^n \) is contained in a \( P_r \)-class, and whence

\[ (wx_1^n) \equiv (wx_2^n) \equiv \cdots \equiv (wx_m^n) \equiv \cdots \equiv (P_R). \]

Notice that \( \{(wx_i^n)^i | i = 1, 2, \ldots\} \) is dense and \( (wx_i^n) \neq (wx_j^n) \), since \( x_i \equiv x_j \) when \( i \neq j \). Then by the definition of \( r \)-disjunctive languages, there exist \( (wx_i^n) \) and \( (wx_j^n) \), \( i \neq j \) such that

\[ (wx_i^n) \neq (wx_j^n)(P_L). \]

That is, for some \( u, v \in X^* \), we have

\[ u(wx_i^n)v \in L, \quad u(wx_j^n)v \notin L, \]

or vice versa. We now suppose that the former case hold, so \( u(wx_i^n)v \notin R \). Then this result leads to \( u(wx_i^n)v \notin R \) since \( (wx_i^n) \equiv (wx_j^n)(P_R) \). Thus \( u(wx_i^n)v \in L \setminus R \), but \( u(wx_j^n)v \notin L \setminus R \). This shows that

\[ (wx_i^n) \neq (wx_j^n)(P_{L \setminus R}), \]

and hence

\[ x_i \equiv x_j(P_{L \setminus R}). \]

Thus, \( L \setminus R \subseteq D_r \).

(2) This part follows directly from (1) because for any language \( L \) over \( X \), \( P_L = P_{\overline{L}} \), where \( \overline{L} \) is the complement of \( L \).

Corollary 2.1. If \( L \subseteq D_r \) and \( R \subseteq L \) and \( R \subseteq \mathcal{R} \), then \( L \setminus R \subseteq D_r \setminus D_f \).

We now return to prove Theorem 2.1.

Proof of Theorem 2.1. We only prove that \( L \subseteq D_f(D_f) \). Assume that neither \( L_1 \) nor \( L_2 \) is \( f \)-disjunctive( \( t \)-disjunctive) language and suppose that \( L_2 \) is not an \( r \)-disjunctive language. Then \( L_2 \) is an \( r \)-regular language. Now, we see that \( L_1 \) is \( f \)-disjunctive( \( t \)-disjunctive)language by Proposition 2.1. This result contradicts to our assumption. Hence, \( L_2 \) is an \( r \)-disjunctive language. Similarly, \( L_1 \) is also an \( r \)-disjunctive language.

Any \( f \)-disjunctive (\( t \)-disjunctive, \( r \)-disjunctive) language has the decomposition of the case (1) in Theorem 2.1 by Proposition 2.1. For the languages in \( D_f \) and \( D_r \), we naturally ask the question: Does the decomposition of the languages that satisfying Theorem 2.1(2) exist? More precisely, we ask whether the decomposition which is similar to (1.1) exists or not? We will give an affirmative answer to the above questions in section 3. In order to simplify our description, we let \( D_1 = D, D_2 = D_f \setminus D, D_3 = D_r \setminus D, D_4 = D_r \setminus D_f \). Hence, \( D_r \) is a disjoint union of \( D_1, D_2, D_3 \) and \( D_4 \).
3. Some decompositions of r-disjunctive languages(II)—Decomposition from $D_i$ to $D_{i+1}$, \( i = 2, 3 \)

First of all, we give the following preparations.

Let $L$ be a nonempty language over an alphabet $X$. We call $L$ contained in $X^+$ a code if for any $x_i, y_j \in L, i = 1, 2, \ldots, m, j = 1, 2, \ldots, n, x_1x_2 \cdots x_m = y_1y_2 \cdots y_n$ implies that $m = n$ and $x_i = y_i, i = 1, 2, \ldots, n$. We call $L$ a prefix language (suffix language) if for any $x \in L, xy \notin L$ for all $y \in X^+(yx \notin L, \text{for all } y \in X^+)$. It is immediate to see that each prefix(suffix) language contained in $X^+$ is a code. Hence, we also call a prefix (suffix) language contained in $X^+$ a prefix (suffix) code. Obviously, a singleton-set of $X^+$ is a prefix code and is a suffix code as well.

For prefix codes and suffix codes, we have the following lemma.

**Lemma 3.1.** [2] Let $L$ be a language over $X$ and $P(S)$ a prefix(suffix) code over $X$. Then for any $u, v \in X^+, u \neq v(P_L)$ implies $u \neq v(P_L(P_S)), i.e. P_L(P_S) \subseteq P_L$.

**Corollary 3.1.** If $L$ is a t-disjunctive language over $X$ and $P(S)$ is a prefix(suffix) code over $X$, then $P_L(L_S)$ is t-disjunctive.

Similar consequence for disjunctive languages and f-disjunctive languages can be found in [10] and [2] respectively.

**Proposition 3.1.** Let $L$ be a language over $X$. If $L$ is not f-disjunctive, then for any finite language $F$ of $X^*$, the language $FL$ and $LF$ are not f-disjunctive.

**Proof.** We just consider the language $FL$, the conclusion for $LF$ can be dually obtained. Suppose that $F \neq \emptyset$ and $F \neq \{1\}$ (the conclusion is trivial when $F = \emptyset$ and $F = \{1\}$).

Let $m = \max\{l(x)\mid x \in F\} \in \mathbb{N}$. Suppose that $L$ is not an f-disjunctive language. Then there exist an infinite language $\{x_1, x_2, \ldots, x_n, \ldots\}$ such that

$$x_1 \equiv x_2 \equiv \cdots \equiv x_n \equiv \cdots (P_L).$$

Since $P_L$ is a congruence,

$$w^m x_1 \equiv w^m x_2 \equiv \cdots \equiv w^m x_n \equiv \cdots (P_L)$$

for any $w \in X^+$. We now show that

$$w^m x_1 \equiv w^m x_2 \equiv \cdots \equiv w^m x_n \equiv \cdots (P_{FL}).$$

Suppose that there exist $w^m x_i, w^m x_j \in \{w^m x_1, w^m x_2, \ldots, w^m x_n, \ldots\}$ such that $w^m x_i = w^m x_j(P_{FL})$. Then, there exist $u, v \in X^*$ such that $u w^m x_i v \in FL$ and $u w^m x_j v \notin FL$ or vice versa. Without loss of generality, we may let $u w^m x_i v \in FL$ and $u w^m x_j v \notin FL$.

Consider the following two cases:

1. $u = u_1u_2$, for some $u_1 \in X^+, u_2 \in X^*$ such that $u_1 \in F, u_2 w^m x_i v \in L$. Clearly, $u_2 w^m x_i v \notin L$. Hence, $x_i \notin x_j(P_L)$, which is a contradiction.

2. $uw^k1w_1 \in F, w_2w^k2x_jv \in L$, where $w_1 \in X^+, w_2 \in X^*, w = w_1w_2, k_1, k_2 \in \mathbb{N}^0$ and $k_1 + k_2 + 1 = m$. Again $w_2w^k2x_jv \notin L$ and $x_i \notin x_j(P_L)$, and so a contradiction.

This shows that the conclusion

$$w^m x_1 \equiv w^m x_2 \equiv \cdots \equiv w^m x_n \equiv \cdots (P_{FL})$$

holds and hence $FL$ is not an f-disjunctive language.

The following proposition is a similar proposition for t-disjunctive languages.
Proposition 3.2. Let $L$ be a language over $X$. If $L$ is not a $t$-disjunctive language, then for any finite language $F$ of $X^*$, the language $LF$ and $FL$ are not $t$-disjunctive.

Proof. The proof of this proposition is similar to the proof of Proposition 3.1.

Similar consequence of non-disjunctive languages can be found in [12].

In the following lemma, we will show that there are languages in $\mathcal{D}_2$ which are unions of two disjoint languages in $\mathcal{D}_3$, see the following Proposition 3.3 and Proposition 3.4.

We now use $w_x$ to denote the number of letters $x$ occurring in the word $w$ over $X$. Then, we establish the following lemma.

Lemma 3.2. Let $X = \{x_1, x_2, \ldots, x_r\}$, $r \geq 2$,

$$L_{x_i} = \{ w \in X^* | w_{x_1}, w_{x_2}, \ldots, w_{x_{r-1}}, w_{x_r} \in I \},$$

where $I = \{ 2^0, 2^1, 2^2, \ldots, 2^m, \ldots \}$. Then $L_{x_i} \in \mathcal{D}_3$, $i = 1, 2, \ldots, r$.

Proof. For any $i \in \{1, 2, \ldots, r\}$, we assert that the $P_{L_{x_i}}$-classes are

$$C_{j_1j_2j_3j_4j_5j_6j_7j_8j_9} = \{ w \in X^* | w_{j_1} = j_2, \ldots, w_{j_r} = j_1 \},$$

for some letter $x_i$ with $x_i \in X \setminus \{x_i\}$, where $p, q \in \mathbb{N}^0$, $p \neq q$. Since $p \neq q$, we may let $p - q = k, k \in \mathbb{N}$. Then for a sufficient large $m, m \in \mathbb{N}$, we can find $x_i^m \in X^*, n \in \mathbb{N}$ such that

$$(x_i^n u)_x = n + p = 2m, \quad 2m - 2m - 1 > k.$$  

Hence, we have $(x_i^n v)_x = n + q = n + p - k = 2m - k$, and $2m - 1 < 2m - k < 2m$.

Consider

$$z = x_i^{n_1} x_i^{n_2} \cdots x_i^{n_p} u$$

and

$$z' = x_i^{n_1} x_i^{n_2} \cdots x_i^{n_p} y' v,$$

where $t_1, t_2, \ldots, t_{r-2}$ is an arrangement of $\{1, 2, \ldots, r\} \setminus \{s, i\}$, and $n_1, n_2, \ldots, n_{r-2} \in \mathbb{N}^0$. Choose $n_1, n_2, \ldots, n_{r-2}$ such that $z_{x_1}, \ldots, z_{x_{r-2}} \in I$, and by the above discussion, we have $z_{x_1} = 2m$ and $z_{x_2} \in I$, but $z_{x_{r-1}} = 2m - k, z_{x_{r-2}} \notin I$. By the construction of $L_{x_i}$, we have $z \in L_{x_i}$ and $z' \notin L_{x_i}$, and hence we conclude that $u \neq v(P_{L_{x_i}})$. This result shows that each $P_{L_{x_i}}$-class is contained in some $C_{j_1j_2j_3j_4j_5j_6j_7j_8j_9}, j_1, \ldots, j_6, j_7, \ldots, j_r \in \mathbb{N}^0$. On the other hand, for any $w, w' \in C_{j_1j_2j_3j_4j_5j_6j_7j_8j_9}, C_{j_1j_2j_3j_4j_5j_6j_7j_8j_9} \in \{ C_{j_1j_2j_3j_4j_5j_6j_7j_8j_9} \mid j_1, \ldots, j_6, j_7, \ldots, j_r \in \mathbb{N}^0 \},$ if $u, v \in X^*$ and $uvw \in L_{x_i}$, then by the construction of $L_{x_i}$, we have $(uvw)_x = u_{x_i} + w_{x_i} + v_{x_i} \in I, s = 1, 2, \ldots, i - 1, i + 1, \ldots, r$. Since $w_{x_i} = w_{x_i}$ for $s = 1, 2, \ldots, i - 1, i + 1, \ldots, r$, we have $(uvw')_x = u_{x_i} + w_{x_i} + v_{x_i} = u_{x_i} + w_{x_i} + v_{x_i} \in I, s = 1, 2, \ldots, i - 1, i + 1, \ldots, r$. This implies that $uvw' \in L_{x_i}$. Dually, for all $u, v \in X^*$, we can deduce that $uvw \in L_{x_i}$ from the fact $uvw' \in L_{x_i}$. Hence, $w \equiv w'(P_{L_{x_i}}).$ Thus the assertion holds.

For each $C_{j_1j_2j_3j_4j_5j_6j_7j_8j_9} \in \{ C_{j_1j_2j_3j_4j_5j_6j_7j_8j_9} \mid j_1, \ldots, j_6, j_7, \ldots, j_r \in \mathbb{N}^0 \}$, we observe that for every word $w$ in $C_{j_1j_2j_3j_4j_5j_6j_7j_8j_9}$ the $w_{x_i}$ is not restricted, and so $w_{x_i}$ can be any number in $\mathbb{N}^0$. Since $\mathbb{N}^0$ is infinite, $C_{j_1j_2j_3j_4j_5j_6j_7j_8j_9}$ is infinite. Moreover, by the definition of $C_{j_1j_2j_3j_4j_5j_6j_7j_8j_9}$, for all $w \in C_{j_1j_2j_3j_4j_5j_6j_7j_8j_9}, w_{x_1} = j_1$, where $j_1$ is a given number in $\mathbb{N}^0$, we have for any word $u \in X^*$ with $u_{x_1}$ greater than $j_1$, $u$ is not a subword of any word.
in \(C_{j_1j_2\cdots j_{i-1}j_{i+1}\cdots j_r}\). This shows that \(C_{j_1j_2\cdots j_{i-1}j_{i+1}\cdots j_r}\) is not dense. Thus \(C_{j_1j_2\cdots j_{i-1}j_{i+1}\cdots j_r}\) is an infinite but not dense language. This shows that \(L_{x_i} \in \mathcal{D}_3\).

Consider the following languages.

\[
L_{x_i} = \{w \in X^* | w_{x_1}, w_{x_2}, \ldots, w_{x_{r-1}} \in I\}, \text{ where } I = \{2^0, 2^1, 2^2, \ldots, 2^n, \ldots\}, \\
L_{x_1} = \{w \in X^* | w_{x_2}, \ldots, w_{x_{r-1}}, w_{x_1} \in I\}, \text{ where } I = \{2^0, 2^1, 2^2, \ldots, 2^n, \ldots\}.
\]

By Lemma 3.2, \(L_{x_i}, L_{x_1} \in \mathcal{D}_3\). Let

\[
L_1 = L_{x_i}\{x_r\}, \quad L_2 = L_{x_1}\{x_1\}.
\]

Since \(\{x_r\}, \{x_1\}\) are suffix codes, by Corollary 3.1, \(L_1, L_2 \in \mathcal{D}_1\) and \(L_1, L_2 \not\in \mathcal{D}_f\) by Proposition 3.1. Thus, we arrive at the following proposition.

**Proposition 3.3.** \(L_1, L_2 \in \mathcal{D}_3\) and \(L_1 \cap L_2 = \emptyset\).

We also have the following proposition.

**Proposition 3.4.** Let \(L = L_1 \cup L_2\). Then \(L \in \mathcal{D}_2\).

**Proof.** To proceed with the proof, we assert that if \(u \equiv v(P_L)\) for \(u, v \in X^*, u \neq v\), then \(u_{x_i} = v_{x_i}\) for all \(x_i \in X, i = 1, 2, \ldots, r\). In fact, if \(u, v \in X^*, u_{x_i} \neq v_{x_i}, x_i \in X \setminus \{x_r\}\), then by the proof of Lemma 3.2, there exist \(w_1, w_2 \in X^*\) such that \(w_1u_{x_1} \in L_{x_1}, w_1v_{x_1} \not\in L_{x_1}\), and so \(w_1uv_{x_2} \in L_1, w_1v_{x_2} \not\in L_1\), also by the construction of \(L_2, w_1v_{x_2} \not\in L_2\), we have \(u \neq v(P_L)\). If \(u, v \in X^*, u_{x_i} \neq v_{x_i}\), then by the proof of Lemma 3.2 again, there exist \(w_1, w_2 \in X^*\) such that \(w_1u_{x_2} \in L_{x_1}, w_1v_{x_2} \not\in L_{x_1}\), and so we have \(w_1uv_{x_2} \in L_1, w_1v_{x_2} \not\in L_1\). Now, by the construction of \(L_1, w_1v_{x_2} \not\in L_1\), we see that \(u \neq v(P_L)\). This shows that each \(P_L\)-class is contained in some \(C_{j_1j_2\cdots j_r}\) with \(j_1, j_2, \ldots, j_r \in \mathbb{N}^0\), where \(C_{j_1j_2\cdots j_r} = \{w \in X^* | w_{x_1} = j_1, \ldots, x_{i-1} = x_{i+1}, \ldots, w_{x_r} = j_r\}\).

Moreover, for \(C_{j_1j_2\cdots j_r} \in \{C_{j_1j_2\cdots j_r} | j_1, \ldots, j_r \in \mathbb{N}^0\}\) with \(|C_{j_1j_2\cdots j_r}| = 1\), by the above discussions, we have \(C_{j_1j_2\cdots j_r}\) is a \(P_L\)-class. For \(C_{j_1j_2\cdots j_r} \in \{C_{j_1j_2\cdots j_r} | j_1, \ldots, j_r \in \mathbb{N}^0\}\) with \(|C_{j_1j_2\cdots j_r}| \geq 2\). Suppose that \(u, v \in C_{j_1j_2\cdots j_r}, u\) ends at letter \(x_1\), and \(v\) does not end at letter \(x_1\). Then we have the following two cases:

1. \(v\) ends at letter \(x_r\),
2. \(v\) does not end at letter \(x_r\).

For case (1), in view of the proof in Lemma 3.2, we are able to find some word \(w_1 \in X^*\) such that

\[
(w_1u)_{x_i} \in I, \quad i = 2, \ldots, r, \quad (w_1u)_{x_1} \not\in I.
\]

This shows that \(w_1u \in L_{x_1}\), by the above assumption and the construction of \(L_2, w_1u \in L_2\). On the other hand, by \(u, v \in C_{j_1j_2\cdots j_r}\), we have

\[
(w_1v)_{x_i} \in I, \quad i = 2, \ldots, r, \quad (w_1v)_{x_1} \not\in I.
\]

This shows that \(w_1v \in L_{x_1}\) and \(w_1v \not\in L_{x_1}\). Since \(v\) does not end at \(x_1\), we have \(w_1v \not\in L_2\), and by the construction of \(L_1, w_1v \not\in L_1\). Hence, we have \(u \neq v(P_L)\).

For case (2), by using similar arguments as those in case (1), we can obtain \(u \neq v(P_L)\). Similarly, if \(u, v \in C_{j_1j_2\cdots j_r}, u\) ends at \(x_r\), and \(v\) does not end at \(x_r\), then we also have \(u \neq v(P_L)\). Thus, for \(C_{j_1j_2\cdots j_r} \in \{C_{j_1j_2\cdots j_r} | j_1, \ldots, j_r \in \mathbb{N}^0\}\) with \(|C_{j_1j_2\cdots j_r}| \geq 2\), \(C_{j_1j_2\cdots j_r}\) can be divided into three parts:

\[
C_{j_1j_2\cdots j_r} = \{w \in X^* | w_{x_1} = j_1, \ldots, w_{x_r} = j_r, \quad \text{and } w \text{ ends at } x_i, i = 2, \ldots, r - 1\},
\]
and if $P_L$-class is contained in $C_{j_1,j_2,...,j_r}$ with $|C_{j_1,j_2,...,j_r}| ≥ 2$, $j_1,j_2,...,j_r ∈ \mathbb{N}^0$, then it must be in some $C_{j_1,j_2,...,j_r}^I$, $J ∈ \{I,II,III\}$. On the other hand, if $u,v$ are any two words in $C_{j_1,j_2,...,j_r}^I$, $|C_{j_1,j_2,...,j_r}^I| ≥ 2$, $j_1,j_2,...,j_r ∈ \mathbb{N}^0$, that is, $u,v ∈ C_{j_1,j_2,...,j_r}$ and both of $u$ and $v$ end neither at $x_1$ nor at $x_r$, then for any $w_1,w_2 ∈ X^*$, $(w_1uw_2)_x = (w_1vw_2)_x$, $i = 1,2,...,r$. So $w_1uw_2 ∈ L_{x_1}$ if and only if $w_1vw_2 ∈ L_{x_1}$, and $w_1uw_2 ∈ L_{x_r}$ if and only if $w_1vw_2 ∈ L_{x_r}$. If $w_2 = 1$, then by previous assumption, both of $w_1uw_2$ and $w_1vw_2$ belong neither to $L_1$, nor to $L_2$. If $w_2 = 1$, then $w_1uw_2 ∈ L_1$ if and only if $w_1vw_2 ∈ L_1$, and $w_1uw_2 ∈ L_2$ if and only if $w_1vw_2 ∈ L_2$. This implies that $w_1uw_2 ∈ L$ if and only if $w_1vw_2 ∈ L$. Both of the two cases show that $u ≡ v(P_L)$. So $C_{j_1,j_2,...,j_r}, |C_{j_1,j_2,...,j_r}| ≥ 2$, $j_1,j_2,...,j_r ∈ \mathbb{N}^0$, is a $P_L$-class. Similarly, we can show that $C_{j_1,j_2,...,j_r}^I \cap C_{j_1,j_2,...,j_r}^J \cap C_{j_1,j_2,...,j_r}^K = \emptyset$, and $|L| = |C_{j_1,j_2,...,j_r}| = |C_{j_1,j_2,...,j_r}^I| + |C_{j_1,j_2,...,j_r}^J| + |C_{j_1,j_2,...,j_r}^K|$. In view of the above facts, we deduce that for $C_{j_1,j_2,...,j_r} ∈ \{C_{j_1,j_2,...,j_r}|j_1,j_2,...,j_r \in \mathbb{N}^0\}$ with $|C_{j_1,j_2,...,j_r}| ≥ 2$, the $P_L$-classes are

$$C_{j_1,j_2,...,j_r}^I = \{w ∈ X^* | w_{x_i} = j_1,...,w_{x_i} = j_i,...,w_{x_r} = j_r,$$

and $w$ ends at $x_1, i = 2,...,r - 1\},$

$$C_{j_1,j_2,...,j_r}^I = \{w ∈ X^* | w_{x_1} = j_1,...,w_{x_1} = j_i,...,w_{x_r} = j_r,$$

and $w$ ends at $x_r\},$

$$C_{j_1,j_2,...,j_r}^I = \{w ∈ X^* | w_{x_1} = j_1,...,w_{x_1} = j_i,...,w_{x_r} = j_r,$$

and $w$ ends at $x_r\}.$

Clearly,

$$|C_{j_1,j_2,...,j_r}^I| + |C_{j_1,j_2,...,j_r}^I| + |C_{j_1,j_2,...,j_r}^I| = |C_{j_1,j_2,...,j_r}|$$

and

$$|C_{j_1,j_2,...,j_r}| ≤ r_{j_1+j_2+...+j_r}.$$ 

This shows that $L ∈ \mathcal{D}_2.$

We will construct a language of the form $L = L_1 \cup L_2$, where $L$ is in $\mathcal{D}_3$ and $L_1,L_2$ are two disjoint languages in $\mathcal{D}_4$. To this aim, we need some preparations. Recall that a nonempty language $L$ over $X$ is an infix language if for all $x,y,u ∈ X^*$, $u ∈ L$ and $xyvx ∈ L$ together imply $x = y = 1$. Clearly, each infix language contained in $X^+$ is a code, we usually call this code an infix code.

For any $x ∈ X^+$ with $|X| ≥ 2$, we let

$$P_n(x) = \{w ∈ X^+ | x = wu for some u ∈ X^+\},$$

$$S_n(x) = \{w ∈ X^+ | x = uw for some u ∈ X^+\},$$

and

$$I(x) = \{w ∈ X^+ | x = uwv for some u,v ∈ X^+\}.$$
Remark 3.1. We denote the set of all non-trivial prefixes(suffixes) of word $x$ by $P_n(x)$ ($S_m(x)$) and the set of all infixes of word $x$ by $I(x)$.

Definition 3.1. [7, 13] Let $L \subseteq X^+$, $L \neq \emptyset$. Then, we call $L$ a solid code if $L$ is an infix code and $P_n(u) \cap S_m(v) = \emptyset$ for every $u, v \in L$.

Corollary 3.2. [13] Any nonempty subset of a solid code is also a solid code.

Definition 3.2. [13] Let $L \subseteq X^+$, $L \neq \emptyset$ and $w \in X^*$. Then, we call the factorization

\[ w = x_1y_1x_2y_2 \cdots x_ny_nx_{n+1} \]

an $L$-representation of $w$ if $y_i \in L$, $I(x_j) \cap L = \emptyset$, for $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, n + 1$, $n \in \mathbb{N}^0$. And we call the case $n = 0$ in (3.1), that is, $w = x_1$, the trivial $L$-representation of $w$. Obviously, $w$ has the trivial $L$-representation if and only if $I(w) \cap L = \emptyset$, and at this time, $w$ has only the trivial $L$-representation.

We call $n$ in (3.1) an $L$-length of $w$, denote the set of $L$-lengths of $w$ by $L(w)$ and call $(x_1, x_2, \ldots, x_n, x_{n+1})$ in (3.1) an $L$-coefficient of $w$, and the set of $L$-coefficients of $w$ is denoted by $C_L(w)$.

Proposition 3.5. [12, 13] Let $L$ be a nonempty language in $X^+$. Then the following statements hold.

1. Any word $w$ over $X$ has $L$-representation.
2. Any word $w$ over $X$ has unique $L$-representation if and only if $L$ is a solid code.

Obviously, if $L$ is a solid code, then for any $w \in X^*$, the $L$-length of $w$ and $L$-coefficient of $w$ are unique, at this time, we denote them by $l_L(w)$ and $c_L(w)$ respectively.

In the following, we let $|X| \geq 2$, and $\{a, b\} \subseteq X$. Consider $\{ba\}$, by the definition of solid code, $\{ba\}$ is a solid code. Now, by Proposition 3.5, for any word $w \in X^*$, $w$ has unique $\{ba\}$-representation

\[ w = x_1y_1 \cdots x_ny_nx_{n+1}, \]

where $y_i = ba$, $i = 1, 2, \ldots, n$, and $ba \notin I(x_j)$, $j = 1, 2, \ldots, n + 1$. Let $\overline{w} = x_1x_2 \cdots x_nx_{n+1}$, where $(x_1, x_2, \ldots, x_{n+1})$ is the $\{ba\}$-coefficient of $w$. Notice that $ba$ may be in $I(\overline{w})$, for example, if $w = bbaa$, then $\overline{w} = ba$. We use $\overline{w}_a$ and $\overline{w}_b$ to denote the numbers of the letter $a$ and $b$ occurring in word $\overline{w}$ respectively. Clearly, for any word $w \in X^*$, $\overline{w}$ is unique, and hence $\overline{w}_a$ and $\overline{w}_b$ are unique.

Proposition 3.6. Let $L_1 = \{w \in X^* \mid \overline{w}_a = \overline{w}_b\}$. Then $L_1 \in \mathcal{D}_4$.

Proof. We assert that the $P_{L_1}$-classes are

\[ C_i = \{w \in X^* \mid \overline{w}_a = \overline{w}_b + i\}, \quad i = 0, \pm 1, \pm 2, \ldots. \]

In fact, if $u, v$ are any two words over $X$ with $\overline{u}_a - \overline{u}_b = i$, $\overline{v}_a - \overline{v}_b = j$, $i \neq j$, without loss of generality, we may let $i \geq 0$, then, obviously, for $x = 1, y = b^j$, we have

\[ xuy = ub^j, \quad xvy = vb^j, \]

by the definition of $xu$, $xv$, we have

\[ xuy = ub^j, \quad xvy = vb^j. \]

And so

\[ xuy_a = \overline{u}_a, \quad xuy_b = \overline{u}_b + i, \quad \text{and} \quad xvy_a = xvy_b \quad \text{by assumption}, \]

On Some Decompositions of r-Disjunctive Languages
\[ \bar{xy} = \bar{y}_a, \quad \bar{xy}_b = \bar{v}_b + i, \quad \text{and} \quad \bar{xy}_a \neq \bar{xy}_b \text{ by assumption.} \]

By the construction of \( L_1 \), we have \( xy \in L_1, xy \notin L_1 \). Hence, \( u \neq v(P_{L_1}) \).

On the other hand, for every two words \( u, v \in C_i, i = 0, \pm 1, \pm 2, \ldots \), by the definition of \( C_i \), we have

\[
\overline{u}_a - \overline{u}_b = \overline{v}_a - \overline{v}_b = i.
\]

Consider \( wu \) and \( wv, w \in X^* \). Then we divide our discussion into the following cases:

1. Both of \( u, v \) begin with letter \( a \). In this case, let \( u = au', v = av' \), where \( u', v' \in X^* \). If \( w = w'b, w' \in X^* \), then \( \overline{wu} = \overline{w'}bau' \) and \( \overline{wv} = \overline{w'}bav' \), hence

\[
\overline{wu}_a = \overline{w'}a + \overline{u}_a = \overline{w'}a + \overline{u}_a - 1, \\
\overline{wu}_b = \overline{w'}b + \overline{u}_b = \overline{w'}b + \overline{u}_b,
\]

and

\[
\overline{wv}_a = \overline{w'}a + \overline{v}_a = \overline{w'}a + \overline{v}_a - 1, \\
\overline{wv}_b = \overline{w'}b + \overline{v}_b = \overline{w'}b + \overline{v}_b,
\]

so

\[
\overline{wu}_a - \overline{wu}_b = \overline{w'}a - \overline{w'}b + i - 1, \\
\overline{wv}_a - \overline{wv}_b = \overline{w'}a - \overline{w'}b + i - 1.
\]

If \( w = w'a, w' \in X^* \), and \( \overline{w}_a = \overline{w'}a + 1 \), then \( \overline{wu} = \overline{w'}aau' \) and \( \overline{wv} = \overline{w'}aav' \), hence

\[
\overline{wu}_a = \overline{w'}a + \overline{u}_a + 1, \\
\overline{wu}_b = \overline{w'}b + \overline{u}_b,
\]

and

\[
\overline{wv}_a = \overline{w'}a + \overline{v}_a + 1, \\
\overline{wv}_b = \overline{w'}b + \overline{v}_b,
\]

so

\[
\overline{wu}_a - \overline{wu}_b = \overline{w'}a - \overline{w'}b + i + 1, \\
\overline{wv}_a - \overline{wv}_b = \overline{w'}a - \overline{w'}b + i + 1.
\]

If \( w = w'x, \) where \( w' \in X^* \) and \( x \in (X \cup \{ ba \}) \setminus \{ a, b \} \), then \( \overline{wu} = \overline{w'}xau' \) and \( \overline{wv} = \overline{w'}xav' \), hence

\[
\overline{wu}_a = \overline{w'}a + \overline{u}_a, \\
\overline{wu}_b = \overline{w'}b + \overline{u}_b,
\]

and

\[
\overline{wv}_a = \overline{w'}a + \overline{v}_a, \\
\overline{wv}_b = \overline{w'}b + \overline{v}_b,
\]

so

\[
\overline{wu}_a - \overline{wu}_b = \overline{w'}a - \overline{w'}b + i, \\
\overline{wv}_a - \overline{wv}_b = \overline{w'}a - \overline{w'}b + i.
\]

Therefore, in every case, we have \( \overline{wu}_a - \overline{wu}_b = \overline{wv}_a - \overline{wv}_b \), and hence \( \overline{wu}_a = \overline{wu}_b \) if and only if \( \overline{wv}_a = \overline{wv}_b \), for any \( w \in X^* \).

2. \( u \) begins with letter \( a \) and \( v \) does not begin with letter \( a \) or vice verse. Without loss of generality, suppose that \( u = au' \), where \( u' \in X^* \). In this case, if \( w = w'b, w' \in X^* \), then for \( wu = w'bau' \), we have \( \overline{wu}_a - \overline{wu}_b = \overline{w'}a - \overline{w'}b + i - 1 \); for \( wv = w'bv \), we have \( \overline{wv}_a - \overline{wv}_b = \overline{w'}b + \overline{v}_b \).
\[ w_a - w_b + i - 1. \] If \( w = w', w' \in X^* \) and \( w_a = w' + 1 \), then for \( wu = w'A \), we have \( wu_a = w' + 1 \); for \( wv = w'B \), we have \( wv_a = w' - w_b + i - 1 \). If \( w = w'x \), where \( w' \in X^* \) and \( x \in (X \cup \{ba\}) \setminus \{a, b\} \), then for \( wu = w'x \), we have \( wu_a = w' + w_b + i \); for \( wv = w'x \), we have \( wv_a = w' - w_b + i \). Therefore, in every case, we also have \( wu_a = wv_a = w_a - w_b \), and hence \( wu_a = wv_b \) if and only if \( wu_a = wv_b \), for any \( w \in X^* \).

(3) \( u, v \) do not begin with letter \( a \). Similar to case (1), we obtain that \( wu_a - wv_b = wu_a - wv_b \), and hence \( wu_a = wv_b \) if and only if \( wu_a = wv_b \), for any \( w \in X^* \).

From the above three cases, we deduce that if \( u, v \in C_i, i = 0, 1, 2, \ldots \), then for any \( w \in X^* \),

\[ (3.2) \quad wu_a - wv_b = wu_a - wv_b, \]

and

\[ (3.3) \quad wu_a = wv_b \iff wu_a = wv_b. \]

Similarly, if \( u, v \in C_i, i = 0, 1, 2, \ldots \), then for any \( w' \in X^* \),

\[ (3.4) \quad uw_a - uw_b = uw_a - uw_b, \]

and

\[ (3.5) \quad uw_a = uw_b \iff uw_a = uw_b. \]

Then, for every \( u, v \in C_i, i = 0, 1, 2, \ldots \), for any \( w, w' \in X^* \), consider \( wuvw \) and \( wuvw' \). Since \( u, v \in C_i \), by (3.2) and by the definition of \( C_i \), we see immediately that both \( wu \) and \( wv \) are in some \( C_{i'}, i' \in \{0, 1, 2, \ldots \} \). Hence, by (3.5), we have \( wuvw_a = wuvw_b \) if and only if \( wuvw_a = wuvw_b \), that is, \( wuvw \in L_1 \) if and only if \( wuvw \in L_1 \). This result implies that \( u \equiv v(P_{L_1}) \). In view of the above facts, we have \( P_{L_1} \)-classes are

\[ C_i = \{ w \in X^* | w_a = w_b + i \}, i = 0, 1, 2, \ldots \]

It can be easily verified that each \( C_i \) is dense, \( i = 0, 1, 2, \ldots \). Hence \( L_1 \in \mathcal{D}_4 \). \( \blacksquare \)

**Proposition 3.7.** Let \( L_2 = \{ w \in X^* | w_a = 2w_b \} \). Then \( L_2 \in \mathcal{D}_4 \).

**Proof.** We first show that \( u_a - 2u_b \neq v_a - 2v_b \) implies \( u \neq v(P_{L_2}) \), for any two words \( u, v \in X^* \). In fact, if \( u_a - 2u_b = i, v_a - 2v_b = j, i \neq j \), then without loss of generality, we may let \( i \geq 0 \). Then we consider \( a^tub \) and \( a^tvb \), where \( s, t \in \mathbb{N}_0 \), by the definition of \( a^tub \), we have

\[ a^tub = s + \eta_a, \quad a^tub = s + t \eta_b. \]

Choose \( s, t \) such that \( s + i = 2t \), then by assumption, we have

\[ s + \eta_a = s + 2\eta_b + i = 2(t + \eta_b), \]

so \( a^tub = 2a^tub \), by the construction of \( L_2 \), \( a^tub \in L_2 \), while

\[ a^tvb = s + \eta_a = s + 2\eta_b + j, \]

\[ 2a^tvb = 2t + 2\eta_b, \]

clearly \( a^tub \neq 2a^tub \) since \( s + j \neq 2t \). This shows that \( a^tub \in L_2 \). Hence, \( u \neq v(P_{L_2}) \).

So let

\[ C_j = \{ w \in X^* | w_a = 2w_b + j \}, j = 0, 1, 2, \ldots \]

we have \( u \equiv v(P_{L_2}) \) implies \( u, v \in C_j \) for some \( j \in \{0, 1, 2, \ldots \} \).
Next we will show that the \( P_{L_2} \)-classes are not analogous to the \( P_{L_1} \)-classes. In fact, if \( u, v \in C_j \) for some \( j \in \{0, \pm 1, \pm 2, \ldots\} \), \( u \) begins with letter \( a \) and \( v \) does not begin with \( a \) or vice versa, then without loss of generality, we may let \( u = au', u' \in X^* \). Consider \( bu \) and \( bv \), since \( bu_a = u'_a = \overline{u}_a - 1 \), \( bu_b = u'_b = \overline{u}_b \), \( bv_a = v_a \) and \( bv_b = v_b + 1 \), we have \( \overline{bu}_a - 2\overline{bu}_b \not\equiv \overline{bv}_a - 2\overline{bv}_b \). By discussion in the previous paragraph, \( bu \not\equiv bv(P_{L_2}) \), so \( u \not\equiv v(P_{L_2}) \). Moreover, if \( u, v \in C_j \) for some \( j \in \{0, \pm 1, \pm 2, \ldots\} \), both of \( u \) and \( v \) begin with letter \( a \) and \( u \) ends at letter \( b \), \( v \) does not end at letter \( b \) or vice versa, then without loss of generality, we may let \( u = u'b \), \( u' \in X^* \). Consider \( uv \) and \( va \), then we have \( \overline{uv}_a = \overline{u}_a, \overline{uv}_b = \overline{u}_b - 1, \overline{va}_a = v_a + 1 \) and \( \overline{va}_b = v_b \), so \( \overline{uv}_a - 2\overline{uv}_b \not\equiv \overline{va}_a - 2\overline{va}_b \), by the discussion in the previous paragraph again, we have \( uv \not\equiv va(P_{L_2}) \), and so \( u \not\equiv v(P_{L_2}) \). Similarly, if \( u, v \in C_j \) for some \( j \in \{0, \pm 1, \pm 2, \ldots\} \), \( u \) and \( v \) do not begin with letter \( a \) and \( u \) ends at the letter \( b \), \( v \) does not end at the letter \( b \) or vice versa, then we also have \( u \not\equiv v(P_{L_2}) \). Hence, each \( C_j, j = 0, \pm 1, \pm 2, \ldots \), can be divided into four parts, say

\[
C_j^I = \{w \in X^* \mid \overline{w}_a = 2\overline{w}_b + j, \ w \ begins \ with \ letter \ a \ and \ ends \ at \ letter \ b \},
\]

\[
C_j^{II} = \{w \in X^* \mid \overline{w}_a = 2\overline{w}_b + j, \ w \ begins \ with \ letter \ a \ and \ does \ not \ end \ at \ letter \ b \},
\]

\[
C_j^{III} = \{w \in X^* \mid \overline{w}_a = 2\overline{w}_b + j, \ w \ does \ not \ begin \ with \ letter \ a \ and \ ends \ at \ letter \ b \},
\]

\[
C_j^{IV} = \{w \in X^* \mid \overline{w}_a = 2\overline{w}_b + j, \ w \ neither \ begins \ with \ letter \ a \ nor \ ends \ at \ letter \ b \},
\]

and we have \( u \equiv v(P_{L_2}) \) implies \( u, v \in C_j^I \) for some \( j \in \{0, \pm 1, \pm 2, \ldots\} \), \( J \in \{I, II, III, IV\} \). It is routine to check that the converse implication holds for each \( j \) and \( J \). Hence, \( P_{L_2} \)-classes are

\[
C_j^I = \{w \in X^* \mid \overline{w}_a = 2\overline{w}_b + j, \ w \ begins \ with \ letter \ a \ and \ ends \ at \ letter \ b \},
\]

\[
C_j^{II} = \{w \in X^* \mid \overline{w}_a = 2\overline{w}_b + j, \ w \ begins \ with \ letter \ a \ and \ does \ not \ end \ at \ letter \ b \},
\]

\[
C_j^{III} = \{w \in X^* \mid \overline{w}_a = 2\overline{w}_b + j, \ w \ does \ not \ begin \ with \ letter \ a \ and \ ends \ at \ letter \ b \},
\]

\[
C_j^{IV} = \{w \in X^* \mid \overline{w}_a = 2\overline{w}_b + j, \ w \ neither \ begins \ with \ letter \ a \ nor \ ends \ at \ letter \ b \},
\]

where \( j = 0, \pm 1, \pm 2, \ldots \). It is easy to check that each \( P_{L_2} \)-class is dense. Hence, \( L_2 \in \mathcal{D}_4 \).
We now construct a language in $\mathcal{D}_3$ which is a disjoint union of the above two languages $L_1$ and $L_2$ in $\mathcal{D}_4$.

Consider $L_1 \cap L_2$. If $w \in L_1 \cap L_2$, then by the constructions of $L_1$ and $L_2$, we have that $\bar{w}_a = \bar{w}_b$, $\bar{w}_a = 2\bar{w}_b$, so $\bar{w}_a = \bar{w}_b = 0$. This result implies that $w \in ((X \cup \{ba\}) \setminus \{a,b\})^*$, where $(X \cup \{ba\}) \setminus \{a,b\}$ is a finite language over $X$ since $X$ is finite, and

$$((X \cup \{ba\}) \setminus \{a,b\})^* = \{(1) \cup (X \cup \{ba\}) \setminus \{a,b\} \cup ((X \cup \{ba\}) \setminus \{a,b\})^2 \cup \cdots \}.
$$

Let $C = ((X \cup \{ba\}) \setminus \{a,b\})^*$. By the above discussion, we have $L_1 \cap L_2 \subseteq C$, and $L_1 \cap L_2 \supseteq C$ is obviously. Hence, we have $L_1 \cap L_2 = C$. By the definition of rational language [4], $C$ is a rational language over $X$. Then by Theorem 4.2.9 of [4], $C$ is regular. Hence $L_1 \setminus C \in \mathcal{D}_4$ by Corollary 2.1. Let $L'_1 = L_1 \setminus C$. Then, it is clear that $L'_1 \cap L_2 = \emptyset$. We have

$$L_1 \cup L_2 = L'_1 \cup L_2 = \{w \in X^* \mid \bar{w}_a = \bar{w}_b \} \cup \{w \in X^* \mid \bar{w}_a = 2\bar{w}_b\}.
$$

We have the following proposition.

**Proposition 3.8.** $L = L'_1 \cup L_2$ is in $\mathcal{D}_3$.

**Proof.** We first assert that if $u, v \in X^*$ and $u \equiv v(P_L)$, then $u, v \in C_{ij}$ for some $i, j \in \mathbb{N}^0$, where

$$C_{ij} = \{w \in X^* \mid \bar{w}_a = i, \bar{w}_b = j\}, \quad i, j \in \mathbb{N}^0.
$$

In fact, for any two words $u, v$ over $X$, suppose that $u, v$ are not in the same $C_{ij}$, that is, $\bar{u}_a \neq \bar{v}_a$ or $\bar{u}_b \neq \bar{v}_b$, we only discuss the case $\bar{u}_a \neq \bar{v}_a$, the case for $\bar{u}_b \neq \bar{v}_b$ can be similarly obtained. Now we divide our discussion into the following two cases:

1. $\bar{u}_a - \bar{u}_b = \bar{v}_a - \bar{v}_b$.
2. $\bar{u}_a - \bar{u}_b \neq \bar{v}_a - \bar{v}_b$.

For case (1), we consider $a^t(ba)u(ba)b'$ and $a^s(ba)v(ba)b'$, $s, t \in \mathbb{N}^0$. Choose $s, t$ such that $s, t > 0$ and $s + \bar{u}_a = 2(t + \bar{v}_b)$, then by the construction of $L_2$, we have $a^t(ba)u(ba)b' \in L_2$ and hence $a^t(ba)u(ba)b' \not\in L'_1$. Since $s, t > 0$, $a^t(ba)u(ba)b' \not\in C$. We have $a^t(ba)u(ba)b' \not\in L_1$. In view of the proof of Proposition 3.6, we have $u \equiv v(P_{L_1})$ because of the fact $\bar{u}_a - \bar{u}_b = \bar{v}_a - \bar{v}_b$, thus $a^t(ba)v(ba)b' \not\in L_1$. And we have

$$\frac{a^t(ba)v(ba)b'_a - 2a^s(ba)v(ba)b'_b}{a^t(ba)v(ba)b'_b} = s + \bar{u}_a - 2(t + \bar{v}_b)
$$

$$= \bar{v}_a - 2\bar{v}_b - (\bar{u}_a - 2\bar{v}_b)
$$

$$= \bar{v}_a - \bar{v}_b - (\bar{u}_a - \bar{u}_b) + \bar{u}_b - \bar{v}_b
$$

$$= \bar{u}_b - \bar{v}_b
$$

$$= \bar{u}_a - \bar{v}_a
$$

$$\neq 0.
$$

So $a^t(ba)v(ba)b' \not\in L_2$. Hence $a^t(ba)v(ba)b' \not\in L$. This shows $u \not\equiv v(P_L)$.

For case (2), if $\bar{u}_a - 2\bar{u}_b = \bar{v}_a - 2\bar{v}_b$, then similar to the discussion in the case (1), we choose $s, t$ such that $s, t > 0$ and $s + \bar{u}_a = t + \bar{u}_b$. Then by the construction of $L'_1$, $a^t(ba)u(ba)b' \in L'_1$ and hence we deduce that $a^t(ba)u(ba)b' \not\in L_2$. This implies that $s + \bar{u}_a - 2t - 2\bar{b} \neq 0$. This shows that $a^t(ba)v(ba)b'_a - 2a^s(ba)v(ba)b'_b = s + \bar{u}_a - 2t - 2\bar{v}_b = s - 2t + \bar{u}_a - 2\bar{b} \neq 0$, $a^t(ba)v(ba)b' \not\in L_2$. Also $a^t(ba)v(ba)b'_a - a^s(ba)v(ba)b'_b = s + \bar{u}_a - 2t - 2\bar{v}_b = \bar{u}_a - 2t - 2\bar{b} \neq 0$. Therefore, we deduce that $a^t(ba)v(ba)b' \not\in L_2$. Hence $a^t(ba)v(ba)b' \not\in L$. This shows $u \not\equiv v(P_L)$. 


we have $u \not\equiv v(P_L)$. If $\overline{v}_a - 2\overline{v}_b \neq \overline{v}_a - 2\overline{v}_b$, then choose $s, t$ such that $s, t > 0$ and $s + \overline{v}_a = t + \overline{v}_b$. Then by the discussion in the previous paragraph, we conclude that $a^s(ba)v(ba)b' \in L_1'$ and $a^s(ba)v(ba)b' \not\in L_1$. Consider

$$a^s(ba)v(ba)b'_a - 2a^s(ba)v(ba)b'_b = s + \overline{v}_a - 2t - 2\overline{v}_b,$$

if $s + \overline{v}_a - 2t - 2\overline{v}_b \neq 0$, then $a^s(ba)v(ba)b' \not\in L_2$. So we have $a^s(ba)v(ba)b' \not\in L$. This shows that $u \not\equiv v(P_L)$. If $s + \overline{v}_a - 2t - 2\overline{v}_b = 0$, then, we choose another $t' \in \mathbb{N}^0$ such that $s + \overline{v}_a = t' + \overline{v}_b$. Then $t' = 2t + \overline{v}_b$, and $t' > 0$ by $t > 0$, and so $a^s(ba)v(ba)b' \in L_1'$. On the other hand, we have $a^s(ba)(ua)(ba)b'_a - a^s(ba)(ua)(ba)b'_b = s + \overline{v}_a - t' + \overline{v}_b$. Since $s + \overline{v}_a + t' - \overline{v}_b = t' - \overline{v}_b$, so $s + \overline{v}_a - t' + \overline{v}_b \neq 0$ by $t' = 2t + \overline{v}_b$ and $t > 0$. Hence, $a^s(ba)(ua)(ba)b' \not\in L_1$. And $a^s(ba)(ua)(ba)b'_a - 2a^s(ba)(ua)(ba)b'_b = s + \overline{v}_a - 2t' - 2\overline{v}_b = -3t - 2\overline{v}_b - \overline{v}_b \neq 0$. Clearly, we see that $a^s(ba)(ua)(ba)b' \not\in L_2$. Hence, $a^s(ba)(ua)(ba)b' \not\in L$. We also have $u \not\equiv v(P_L)$. Thus, we have shown that our assertion holds.

By the above assertion, we see immediately that each $P_L$-class is contained in some $C_{ij}$, $i, j \in \mathbb{N}^0$. Since $C_{ij}$ is thin (if otherwise, we let $w = (ba)^{i+1}b^{j+1}(ba)$). Then for any $u, v \in X^*$, we have $uv \not\in C_{ij}$, a contradiction) for $i, j = 0, 1, 2, \ldots$, we easily see that each $P_L$-class is thin.

Next we continue to show that there are infinite $P_L$-classes. Consider

$$C_{00} = \{ w \in X^* | \overline{w}_a = 0, \overline{w}_b = 0 \}.$$

Clearly, $1 \in C_{00}$. For any word $w \in C_{00} \setminus \{1\}$, we shall show that $1 \not\equiv w(P_L)$. Take $x = b^2, y = a^3$. Then

$$xyy = b^2a^3 = b^2a^3, xy = b^2wa^3,$$

by the definition of $xy$, we have $xy = ba^2$, hence $x_1y = 2$, $x_1y_b = 1$ and $xy \in L_2$. While $xwyr = b^2wa^3$ because $w \neq 1$ and $w$ neither begins with $a$ nor ends at $b$. This leads to $xwyr = 2$, $xwyr_b = 2$. Hence, we have $xyw \not\in L_1$, $xyw \not\in L_2$, and so $xyw \not\in L$. Hence, we have $1 \not\equiv w(P_L)$. On the other hand, for any two words $u, v \in C_{00} \setminus \{1\}$, by the definition of $C_{00}$, we have $\overline{u}_a = \overline{u}_b = 0, \overline{v}_a = \overline{v}_b = 0$, and so for any $x, y \in X^*$,

$$xy = x_a + y_a, xy = x_b + y_b,$$

This shows that

$$xy \in L_1 \text{ if and only if } xy \in L',$$

and

$$xy \in L \text{ if and only if } xy \in L_2.$$

Thus, we have

$$xy \in L \text{ if and only if } xy \in L.$$

Hence, we have proved that $u \equiv v(P_L)$.

Now, $C_{00}$ can be divided into two parts, namely, $C_{00} \setminus \{1\}$ and $\{1\}$. In view of the above facts, we see that $C_{00} \setminus \{1\}$ and $\{1\}$ are both $P_L$-classes. Notice that $(ba)^+ \subseteq C_{00} \setminus \{1\}$, where

$$(ba)^+ = (ba)^+ \setminus \{1\} = \{ba, (ba)^2, (ba)^3, \ldots\},$$

so $C_{00} \setminus \{1\}$ is infinite. Therefore, $L \in \mathcal{D}_3$. \[\qed\]
4. The decomposition from \( D_1 \) to \( D_2 \)

In [12], Shyr and Yu have shown the existence of a disjunctive language which can be partitioned into two parts such that both of them are midst-languages. The disjunctive language can be constructed on \( X \) with \(|X| \geq 3\). In this section, we will improve their construction on \( X \) with \(|X| \geq 2\) and simplify the construction of the disjunctive language as well. We will show that the disjunctive language is not only a disjoint union of midst-languages but is also a disjoint union of languages in \( D_2 \).

In our study, the free monoid \( X^* \) sometimes needs to be equipped with a total order. In this paper, we only adopt the standard total order \( \leq \) which is defined on \( X^* \) as follows [10]: For any \( u, v \in X^* \), if \( lg(u) < lg(v) \), then \( u < v \); if \( lg(u) = lg(v) \), then \( \leq \) is the lexicographical order on \( X^n \) for all \( n \geq 1 \). For a word \( x \in X^* \), we write \( \xi x = m \) if \( x \) stands at the \( m \)th position in this order.

Recall that \( C_L(w) \) is the set of \( L \)-coefficients of \( w \), for any word \( w \) over \( X \) and any nonempty language \( L \) in \( X^+ \).

We begin with the following definition.

**Definition 4.1.** Let \( L \subseteq X^+, L \neq \emptyset \). Then, we define a binary relation \( \sigma_L \) on \( X^* \) as follows:

\[
(w_1, w_2) \in \sigma_L \iff C_L(w_1) \cap C_L(w_2) \neq \emptyset.
\]

If \( (w_1, w_2) \in \sigma_L \), then \( w_1 \) and \( w_2 \) are said to be \( L \)-related.

**Corollary 4.1.** \( \sigma_L \) is left compatible and right compatible with the operation on free monoid \( X^* \).

*Proof.* We only consider the left compatibility. For the case of right compatibility, it can be proved analogously. Suppose that \( w_1 \sigma_L w_2 \) holds. Then, by the definition of \( \sigma_L \), there exist \((x_1, x_2, \ldots, x_{n+1})\) such that \( w_1 \) has an \( L \)-representation

\[
w_1 = x_1y_1x_2y_2 \cdots x_ny_nx_{n+1},
\]

and \( w_2 \) has \( L \)-representation

\[
w_2 = x_1y_1'x_2y_2' \cdots x_ny_n'x_{n+1}.
\]

Now, for any \( x \in X^* \), we have

\[
xw_1 = xx_1y_1x_2y_2 \cdots x_ny_nx_{n+1},
\]

\[
xw_2 = xx_1y_1x_2y_2' \cdots x_ny_n'x_{n+1}.
\]

If \( I(xx_1) \cap L = \emptyset \), then, we let \( xx_1 = x_1' \). Now, we see that

\[
xw_1 = x_1'y_1x_2y_2 \cdots x_ny_nx_{n+1},
\]

\[
xw_2 = x_1'y_1x_2y_2' \cdots x_ny_n'x_{n+1}
\]

are the \( L \)-representations of \( xw_1 \) and \( xw_2 \) respectively, so \( xx_1 \sigma_L xw_2 \). If \( I(xx_1) \cap L \neq \emptyset \), then \( xx_1 \) has an \( L \)-representation

\[
xx_1 = u_1v_1 \cdots u_mv_mu_{m+1},
\]

where \( v_i \in L, I(u_j) \cap L = \emptyset, i = 1, 2, \ldots, m, j = 1, 2, \ldots, m + 1 \). Clearly,

\[
xw_1 = u_1v_1 \cdots u_mv_mu_{m+1}y_1x_2y_2 \cdots x_ny_nx_{n+1},
\]

\[
xw_2 = u_1v_1 \cdots u_mv_mu_{m+1}y_1x_2y_2' \cdots x_ny_n'x_{n+1}
\]
are $L$-representations of $xw_1$ and $xw_2$ respectively. Hence, we have proved that $xw_1 \sigma_L xw_2$.

**Lemma 4.1.** If $L$ is a solid code, then $\sigma_L$ is a congruence on $X^*$.

**Proof.** By the definition of $\sigma_L$, $\sigma_L$ is clearly reflexive and symmetric. If $L$ is a solid code, then $\sigma_L$ is transitive by Proposition 3.5 since, at this point, $(w_1, w_2) \in \sigma_L$ if and only if $c_L(w_1) = c_L(w_2)$. Hence, $\sigma_L$ is an equivalence relation on $X^*$. By Corollary 4.1, we have proved that $\sigma_L$ is a congruence on $X^*$.

For the solid codes, we have the following lemma.

**Lemma 4.2.** Let $L$ be a solid code. Then the following statements are equivalent:

1. $L$ is finite.
2. every $\sigma_L$-class is finite.

**Proof.** (1) $\Rightarrow$ (2). Suppose that $|L| = m$, $m \in \mathbb{N}$. For any $w \in X^*$, if $I(w) \cap L \neq \emptyset$ and $w$ has the unique $L$-representation $w = x_1 y_1 x_2 y_2 \cdots x_n y_n x_{n+1}$, then $|w|_\sigma_L = 2^n$. Moreover, if $I(w) \cap L = \emptyset$ and $w$ has the trivial $L$-representation $w = x_1$, then $|w|_\sigma_L = 1$.

(2) $\Rightarrow$ (1). Observe that $L$ is a $\sigma_L$-class. Hence, every $\sigma_L$-class is finite implies $L$ is finite.

Now, let $|X| \geq 2$ and $X = \{a, b, \ldots\}$. Consider $u_1 = a^3b^3, u_2 = a^2bab^2, v = a^2b^2ab$. We have

\[
\begin{align*}
P_m(u_1) &= \{a, a^2, a^3, a^3b, a^3b^2\}, \quad S_m(u_1) = \{b, b^3, ab, a^2b^3\}, \\
P_m(u_2) &= \{a, a^2, a^2b, a^2ba, a^2bab\}, \quad S_m(u_2) = \{b, b^2, bab, bab^2, abab^2\}, \\
P_m(v) &= \{a, a^2, a^2b, a^2b^2, a^2b^2a\}, \quad S_m(v) = \{b, ab, bab, ba^2b, abab\}.
\end{align*}
\]

Clearly $\{u_1, u_2, v\}$ is a solid code. By Corollary 3.2, we obtain that $\{u_i, v\}$-representation of every $w \in X^+$ is unique, for $i = 1, 2$.

We define the following sets, where $s \in \mathbb{N}$:

\[
A_{u_1, v}(s) = \{w \in X^+ \mid I_{\{u_1, v\}}(w) = s\}.
\]

\[
\overline{A}_{u_1, v}(s) = \{ab^6wda^6+q_1ba^6+q_2\cdots ba^6+q_{s+1} \mid w \in A_{u_1, v}(s) \text{ and } \#x_i = q_i, i = 1, 2, \ldots, s + 1, \text{ when the } \{u_1, v\} - \text{representation of } w \text{ is } w = x_1 y_1 x_2 y_2 \cdots x_n y_n x_{n+1}\}.
\]

\[
A_{u_1, v} = \bigcup_{s \geq 1} \overline{A}_{u_1, v}(s).
\]

The following lemma is a crucial lemma. The proof can be found in [13]. However, we notice that the proof given in [13] has a possible gap (that is, let $\{u\}$ be a solid code with $u \in \{aX^+b\}$ and $u \notin \{a^+b, ab^+\}$, $lg(u) = n, n \geq 3$. If $u \in I(a^kb^n w)$ for any $k \in \mathbb{N}^0$, and any $w \in X^+$, then $u \notin I(a^kb^n)$. But we notice that there exist such solid codes $\{u\}$ such that $u \in I(a^kb^n)$, for example, let $n = 6, k \geq 3$. Then by the above discussion, $\{u = a^3b^3\}$ is a solid code, and clearly $u \in I(a^kb^n)$. Therefore, in the following revised proof of this lemma, it contains the consideration for $u \in I(a^kb^n)$.

**Lemma 4.3.** [13, Lemma 3.11] Let $\{u, v\}$ be a solid code, $\{u, v\} \subseteq \{aX^+b \cap X^n\}$ for some $n \geq 3$, and $\{u, v\} \cap \{a^+b, ab^+\} = \emptyset, w_1, w_2 \in X^*$. If $(w_1, w_2) \notin \sigma_{\{u, v\}}$, then for any $i, j, k \in \mathbb{N}^0$,

\[
(u^i v^j a^k b^n w_1, u^i v^j a^k b^n w_2) \notin \sigma_{\{u, v\}}.
\]
Proof. (revised) Let $z = u^iv^j a^kb^n w_1$ and $z' = u^iv^j a^kb^n w_2$. We first suppose that $(z, z') \in \sigma_{\{u, v\}}$. Then, the $\{u, v\}$-representations of $z$ and $z'$ are

$$z = x_1y_1x_2y_2 \cdots x_my_mx_{m+1},$$

and

$$z' = x_1y_1'x_2y_2' \cdots x_my'_mx_{m+1}'.$$

Thus, $x_1 = x_2 = \cdots = x_{i+j} = 1$, and the $\{u, v\}$-representations of $a^kb^n w_1$ and $a^kb^n w_2$ are

$$a^kb^n w_1 = x_{i+j+1}y_{i+j+1} \cdots x_my_mx_{m+1},$$

and

$$a^kb^n w_2 = x_{i+j+1}y_{i+j+1}' \cdots x_my'_mx_{m+1}'.$$

On the one hand, if $I(a^kb^n) \cap \{u, v\} \neq \emptyset$, then by our hypothesis and the $\{u, v\}$-representations of $a^kb^n w_1$ and $a^kb^n w_2$, we have $a^kb^n = x_{i+j+1}y_{i+j+1}b^p$ with $1 < p < n - 1$, and $b^p$ is a prefix of $x_{i+j+2}$ (that is, $x_{i+j+2} = b^p, x \in X^\ast$). Hence, the $\{u, v\}$-representations of $w_1$ and $w_2$ have the forms

$$w_1 = x_{i+j+2}y_{i+j+2} \cdots x_my_mx_{m+1},$$

and

$$w_2 = x_{i+j+2}y_{i+j+2}' \cdots x_my'_mx_{m+1}.$$  

with $b^p x_{i+j+2} = x_{i+j+2}, x_{i+j+2}' \in X^\ast$. On the other hand, if $I(a^kb^n) \cap \{u, v\} = \emptyset$, then from the $\{u, v\}$-representations of $a^kb^n w_1$ and $a^kb^n w_2$ again, $a^kb^n$ is a prefix of $x_{i+j+1}$ or $x_{i+j+1}$ is a proper prefix of $a^kb^n$ (that is, $x_{i+j+1} = a^kb^n, x \in X^\ast$). We now claim that $x_{i+j+1}$ is not the proper prefix of $a^kb^n$, for otherwise, if $x_{i+j+1}a^kb^n = a^kb^n, k' \geq 1$, then by the hypothesis $lg(u) = lg(v) = n$, we have $y_{i+j+1}x = a^k b^n, x \in X^\ast$. This result contradicts to $I(a^kb^n) \cap \{u, v\} = \emptyset$; if $x_{i+j+1}b^n = a^kb^n, 1 \leq n' \leq n$, then $y_{i+j+1}$ begins with letter $b$, this contradicts to $\{u, v\} \subseteq aX^\ast b$. This result hence shows that $a^kb^n$ is the prefix of $x_{i+j+1}$. Hence, the $\{u, v\}$-representations of $w_1$ and $w_2$ are

$$w_1 = x_{i+j+1}y_{i+j+1} \cdots x_my_mx_{m+1} \text{ and } w_2 = x_{i+j+1}y_{i+j+1}' \cdots x_my'_mx_{m+1},$$

where $a^kb^n x_{i+j+1} = x_{i+j+1}, x_{i+j+1}' \in X^\ast$.

Both of the above two cases imply that $(w_1, w_2) \in \sigma_{\{u, v\}}$. Thus, we arrive at a contradiction and our proof is completed.

We state the following proposition.

**Proposition 4.1.** $P_{\lambda_{u_1}, v} = \sigma_{\{u_1, v\}}$.

**Proof.** Let $w_1$ and $w_2$ be two words over $X$, $(w_1, w_2) \notin \sigma_{\{u_1, v\}}$. Then, we will show that $w_1 \neq w_2(P_{\lambda_{u_1}, v})$. Consider $z_1 = u_1^i v^j a^kb^nw_1$ and $z_2 = u_1^i v^j a^kb^nw_2$. By Lemma 4.3, we have $(z_1, z_2) \notin \sigma_{\{u_1, v\}}$. Now, we choose $i$ and $j$ such that $l_{\{u_1, v\}}(z_1) = 2^i$ for some $i \in \mathbb{N}$. Let the $\{u_1, v\}$-representation of $z_1$ be

$$z_1 = x_1y_1x_2y_2 \cdots x_ny_nx_{n+1},$$

$n = 2^i$ and let the $\{u_1, v\}$-representation of $z_2$ be

$$z_2 = x_1'y_1'x_2'y_2' \cdots x_m'y_m'x_{m+1}'.$$
for some $m \in \mathbb{N}^0$. Since $(z_1, z_2) \not\in \sigma_{\{u_1, v\}}$, we have $n \neq m$, or $n = m$ but $x_h \neq x'_h$ for some $h$, $1 \leq h \leq n + 1$. Let $z_x = q_p$, $p = 1, 2, \ldots, n + 1$. Then by the construction of $A_{u_1, v}$, we have

$$\overline{z_1} = ab^6z_1a^{6+q_1}ba^{6+q_2}b\cdots a^{6+q_{n+1}} \in \overline{A}_{u_1, v}(2'),$$

and

$$\overline{z_2} = ab^6z_2a^{6+q_1}ba^{6+q_2}b\cdots a^{6+q_{n+1}} \not\in \overline{A}_{u_1, v}(2'),$$

for each $s \in \mathbb{N}$. Hence $\overline{z_1} \not\equiv \overline{z_2}(P_{A_{u_1, v}})$. It follows that $z_1 \not\equiv z_2(P_{A_{u_1, v}})$, and hence $w_1 \neq w_2(P_{A_{u_1, v}})$. So

$$P_{A_{u_1, v}} \subseteq \sigma_{\{u_1, v\}}.$$

We now proceed to prove the converse statement of the above proposition. We first suppose that $w_1$ and $w_2$ are two different words over $X$, $(w_1, w_2) \in \sigma_{\{u_1, v\}}$. Since $\sigma_{\{u_1, v\}}$ is a congruence, $(xw_1y, xw_2y) \in \sigma_{\{u_1, v\}}$ for any $x, y \in X^*$. From the definition of $\sigma_{\{u_1, v\}}$, we have

$$xw_1y = x_1y_1x_2y_2\cdots x_ny_nx_{n+1},$$

$$xw_2y = x'_1y'_1x'_2y'_2\cdots x'_n'y'_nx'_{n+1}.$$  

If $xw_1y$ is in some $\overline{A}_{u_1, v}(2')$, $t \in \mathbb{N}$, then $xw_1y = x_1y_1x_2y_2\cdots x_ny_nx_{n+1}$ has the form $ab^6w'a^{6+q_1}ba^{6+q_2}b\cdots a^{6+q_{2'i}+1}$, where $w' \in X^*$, $q_i \in \mathbb{N}$, $i = 1, 2, \ldots, 2' + 1$. Since $I(ab^6) \cap \{u_1, v\} = \emptyset$ by the definitions of $u_1$ and $v$, similar to the proof of Lemma 4.3, we have $ab^6x'_i = x_1, x'_i \in X^*$. Similarly, we have $x'_i+1a^{6+q_1}ba^{6+q_2}b\cdots a^{6+q_{2'i}+1} = x_{n+1}, x'_{n+1} \in X^*$. Thus, $xw_1y$ and $xw_2y$ can be written as

$$(4.1) \quad \quad ab^6x'_1y'_1x'_2y'_2\cdots x'_n'y'_nx'_{n+1}a^{6+q_1}ba^{6+q_2}b\cdots a^{6+q_{2'i}+1},$$

$$(4.2) \quad \quad ab^6x'_1y'_1x'_2y'_2\cdots x'_n'y'_nx'_{n+1}a^{6+q_1}ba^{6+q_2}b\cdots a^{6+q_{2'i}+1},$$

respectively. By our assumption that $xw_1y \in \overline{A}_{u_1, v}(2')$, we have $n = 2'$ in (4.1), and $x'_i = q_1, x_i = q_i, i = 2, \ldots, 2'$, $x'_{2'i} = q_{2'i} = q_{2'i+1}$. This means that $xw_2y \in \overline{A}_{u_1, v}(2')$. Dually, it can be proved that $xw_1y \in \overline{A}_{u_1, v}(2')$ if and only if $xw_2y \in \overline{A}_{u_1, v}(2')$. Hence, $w_1 \equiv w_2(P_{A_{u_1, v}})$ and so

$$\sigma_{\{u_1, v\}} \subseteq P_{A_{u_1, v}}. \quad \blacksquare$$

**Proposition 4.2.** The following statements always hold.

1. $u_1^{v_i}v_j \equiv v_i^{u_1^i}(P_{A_{u_1, v}})$, for all $i, j \in \mathbb{N}^0$.
2. $A_{u_1, v}$ is an f-disjunctive language.

**Proof.** (1) Clearly, $(u_1^{v_i^j}, v_j^{u_1^i}) \in \sigma_{\{u_1, v\}}$, for all $i, j \in \mathbb{N}^0$, by Proposition 4.1, the result holds.

(2) By Proposition 4.1, $P_{A_{u_1, v}} = \sigma_{\{u_1, v\}}$ and by Lemma 4.2, every $\sigma_{\{u_1, v\}}$-class contains only finite elements. This shows that $A_{u_1, v}$ is f-disjunctive. \quad \blacksquare

By Proposition 4.2, we have $A_{u_1, v} \in \mathcal{D}_2$, it is clear that $A_{u_1, v}$ is a midst-language. Similarly, by replacing $u_1$ with $u_2$, the languages $A_{u_2, v}(s), \overline{A}_{u_2, v}(s), A_{u_2, v}$ are defined respectively. Clearly, Proposition 4.2 is also valid for $A_{u_2, v}$. Thus, $A_{u_2, v} \in \mathcal{D}_2$.

Let $B_1 = A_{u_1, v}(b), B_2 = \{b\}A_{u_2, v}$. Then, by Lemma 3.1 and Lemma 4.3 of [12], $B_1$ and $B_2$ are both in $\mathcal{D}_2$. Clearly, $B_1$ and $B_2$ are disjoint. We now show that $B_1 \cup B_2$ is a disjunctive language.

**Proposition 4.3.** The language $B_1 \cup B_2$ is disjunctive.
5. Some questions related to the decompositions of r-disjunctive languages

We first observe that Proposition 1.4 can be modified into the following form.

Proposition 1.4′. Let L be a disjunctive language over X, L = L₁ ∪ L₂. Then the following statements hold.

1. |{L₁, L₂} ∩ ∅| = 1, or
2. |{L₁, L₂} ∩ ∅| = 2, or
3. {L₁, L₂} ⊆ ∅\ ∅.

By the main result in [6], we see immediately that any disjunctive language L has a decomposition of L = L₁ ∪ L₂ such that L₁ ∈ ∅, L₂ ∈ ∅. At this time, we have |{L₁, L₂} ∩ ∅| = 1; and any disjunctive language has the decomposition of case (2), in fact, the disjunctive language L is dense and so by [3] (a dense language can be divided into two disjoint disjunctive languages), L has the decomposition L = L₁ ∪ L₂ such that |{L₁, L₂} ∩ ∅| = 2; for the case (3), not every disjunctive language has this decomposition, but it has been shown that there exist such languages, see Section 4. Meanwhile, for those r-disjunctive languages, we see that any f-disjunctive (t-disjunctive, r-disjunctive) language L has the decomposition of L = L₁ ∪ L₂ and |{L₁, L₂} ∩ ∅| = 1 (|{L₁, L₂} ∩ ∅| = 1, |{L₁, L₂} ∩ ∅| = 1) by Proposition 2.1. Now, we can also see that there exist f-disjunctive (t-disjunctive, r-disjunctive) languages L such that L = L₁ ∪ L₂ and {L₁, L₂} ⊆ ∅\ {f|} or {L₁, L₂} ⊆ ∅\ {r|}, see Section 3. But we still do not know whether every f-disjunctive (t-disjunctive, r-disjunctive) language L has the decomposition L = L₁ ∪ L₂ with |{L₁, L₂} ∩ ∅| = 2 (|{L₁, L₂} ∩ ∅| = 2, |{L₁, L₂} ∩ ∅| = 2)?

In closing this paper, we point out that the above question also leads to a more special question.

Does every f-disjunctive (t-disjunctive, r-disjunctive) language L can be decomposed into L = L₁ ∪ L₂ such that P₁ = P₁ = P₁ = P₁?

We remark that for disjunctive languages, the above two questions are actually the same question, but for f-disjunctive (t-disjunctive, r-disjunctive) languages, they are different
questions.

Acknowledgement. Yuqi Guo is the corresponding author of the article. The research was supported by the National Natural Science Foundation of China (No. 11371177). The authors would like to thank the referees for their revision suggestions.

References