# INCLUSION PROBLEMS OF LANGUAGES GENERATED BY REGULAR PATTERNS AND CO-REGULAR PATTERNS

Mikiharu TERADA

Received October 14, 2004; revised December 1, 2004

Abstract. A *pattern* is a finite string consisting of constant symbols and variables. A pattern is *regular* if each variable appears in the pattern at most once. The language generated by a pattern is the set of constant strings obtained from the pattern by substituting nonempty strings for variables in the pattern. This paper deals with inclusion problems of unions or intersections of languages defined by regular patterns and co-regular patterns. The semantics of a co-pattern is defined by a particular subset of the complement of the original pattern language. We show the equivalence between the semantic inclusion and the syntactic inclusion.

**1 Introduction.** A *pattern* is a nonempty finite string consisting of constant symbols and variables. A pattern is *regular* if each variable appears in the pattern at most once. The language $L(p)$ generated by a pattern $p$ is the set of constant strings obtained from $p$ by substituting *nonempty* constant strings for variables in $p$. The inclusion problem for pattern languages is shown to be undecidable (Jiang et al.[6]) and the membership problem is NP complete (Angluin [2]), while both problems for regular pattern languages are polynomial time computable (Shinohara [13]). The class $\mathcal{PL}$ of pattern languages has been introduced by Angluin [2] as a learnable class from positive examples in the Gold's framework [5]. Pattern languages merely are not used for applications because of their simplicity. From a practical point of view, various kinds of languages generated by patterns have been investigated in Gold's framework, PAC learning and so on. Languages generated by decision trees over regular patterns are paid much attention in some practical applications such as genome informatics (Arikawa et al.[3]).

The previous paper [17] due to the present author dealt with learning problem of decision trees over regular patterns with bounded depth from positive example. For each regular pattern $p$, we introduced a particular type of string $p^c$ called *co-pattern* of $p$, and defined its semantics as the subset of the complement $L(p)^c$ consisting of strings with lengths more than or equal to that of $p$. In terms of languages $L(p)$s and $L(p^c)$s, we gave expressions for languages generated by decision trees, and showed the learnability of such decision trees from positive examples. In designing an efficient learning algorithm for decision trees over regular patterns, it is an important key to solve the inclusion problem for unions or intersections of regular pattern languages and co-regular pattern languages.

In the present paper, we investigate such inclusion problems. We introduce two kinds of partially ordered syntactic relations on generalizations and instances of finite sets of patterns, and show the equivalence between the semantic inclusion and the syntactic inclusion, under some assumption of the cardinality of the alphabet. We obtain some results about relations between the semantic inclusion $L(\pi_1) \cap L(\pi_2) \subseteq L(\tau_1) \cup L(\tau_2)$ and the syntactic inclusion for the pairs $(\pi_1, \pi_2)$ and $(\tau_1, \tau_2)$, where $\pi_1, \pi_2, \tau_1$ and $\tau_2$ are patterns or co-patterns.

2000 *Mathematics Subject Classification.* 68Q70, 68R15.
*Key words and phrases.* pattern, co-pattern, pattern language, inclusion problem.

**2 Regular Pattern and Co-Regular Pattern Languages.** Let $\Sigma$ be a fixed alphabet of *constant* symbols and let $X = \{x, y, z, x_1, x_2, \cdots\}$ be a countable set of *variables*. Assume that $\Sigma \cap X = \phi$. For a positive integer $l$, $\Sigma^l$ denotes the set of strings over $\Sigma$ with length $l$. We define $\Sigma^{\leq l} = \bigcup_{i=1}^{l} \Sigma^i$ and $\Sigma^{[l_1,l_2]} = \bigcup_{i=l_1}^{l_2} \Sigma^i$ for positive integers $l, l_1$ and $l_2$ with $l_1 \leq l_2$. Similarly, we also define $\Sigma^{<l}$ and $\Sigma^{\geq l}$ for a positive integer $l$. By $X^+$, we denote the set of strings consisting of variables only. We put $\chi_l = x_1 x_2 \cdots x_l$ for distinct variables $x_i$s and for a positive integer $l$.

For a finite set $S$, $\sharp S$ denotes the number of elements in $S$.

A *pattern* over $\Sigma$ is a nonempty finite string over $\Sigma \cup X$. A pattern $p$ is *regular* if each variable occurs at most once in $p$. For instance, $p = axbyz$ and $q = axbyy$ are patterns over $\{a, b, c\}$, and $p$ is regular but not $q$. And the string $\chi_l$ is also a regular pattern with length $l$. By $\mathcal{RP}$, we denote the class of regular patterns. In what follows, we confine ourselves to regular patterns, and call a regular pattern a *pattern* simply. The *length*, denoted by $l_p$, of a pattern $p$ is the number of symbols in $p$.

A *substitution* is a homomorphism from patterns to patterns that maps each constant symbol in $\Sigma$ to itself. For a substitution $\theta$, we denote the image of $p$ by $p\theta$. We assume that $x\theta$ is not empty for any variable $x$.

A pattern $p$ is an *instance* of a pattern $q$ (or $q$ is a *generalization* of $p$), denoted by $p \preceq q$, if $q\theta = p$ for some substitution $\theta$. For a pattern $p$, we define the language $L(p) = \{w \in \Sigma^+ \mid w \preceq p\}$. Clearly $p \preceq q$ implies $L(p) \subseteq L(q)$ and $l_q \leq l_p$. A pattern $p$ is *equivalent* to a pattern $q$, denoted by $p \equiv q$, if $p \preceq q$ and $q \preceq p$. If $p \equiv q$, these patterns are identical except naming variables, and thus we identify these patterns in this paper.

Let $P$ be a set of patterns. A pattern $q$ is an *instance* of $P$, if $q \preceq p$ for any $p \in P$. Similarly, $q$ is a *generalization* of $P$, if $p \preceq q$ for any $p \in P$. A pattern $p$ is a *maximal* instance (mi for short) of $P$, if $p$ is an instance of $P$ and there is no instance $q (\neq p)$ of $P$ such that $q \preceq p$. Similarly, we define the notions of a *minimal* generalization (mg for short). By mi $P$ (resp., mg $P$), we denote the sets of maximal instances (resp., minimal generalizations) of $P$.

**Lemma 2.1 (Mukouchi [10])** *The following two statements are valid:*

*(1) If $p$ is an instance of a set $P$ of patterns, then there is a maximal instance $q \in$mi $P$ such that $p \preceq q$.*

*(2)* $\mathrm{mi}\ \{p_1, p_2, \cdots, p_n\} \subseteq \bigcup_{p \in \mathrm{mi}\{p_1, p_2, \cdots, p_{n-1}\}} \mathrm{mi}\ \{p, p_n\}$.

**Lemma 2.2** *Let $P$ be a nonempty finite set of patterns. Then*

$$\bigcap_{p \in P} L(p) = \bigcup_{r \in \mathrm{mi} P} L(r).$$

For a pattern $p$ and a positive integer $n$, we consider the set mi $\{p, \chi_n\}$. If $l_p \leq n$ then mi $\{p, \chi_n\} = \{p\}$. Otherwise it is given by the set of patterns with just length $n$ obtained from $p$ by substituting strings consisting of distinct variables for variables in $p$. For example, let $p = x_1 a x_2 b$. Then we have mi $\{p, \chi_n\} = \{x_1 a x_2 b\}$ for $n = 1, 2, 3, 4$. On the other hand, mi $\{p, \chi_5\} = \{x_1 x_2 a x_3 b, x_1 a x_2 x_3 b\}$ holds.

For a pattern $p$ and a positive integer $i$, we introduce a particular finite subset of $L(p)$ as follows:

$$S_i(p) = L(p) \cap \Sigma^{\leq i \times l_p}.$$

Clearly the set $S_1(p)$ is the set of shortest strings of $L(p)$.

For a set $P$ of patterns, we put $L(P) = \bigcup_{p \in P} L(p)$ and $S_i(P) = \bigcup_{p \in P} S_i(p)$ for each positive integer $i$.

For nonempty finite sets $P$ and $Q$ of patterns, we introduce a relation $\sqsubseteq$ defined as follows:

$$P \sqsubseteq Q \quad \Longleftrightarrow \quad \text{for any pattern } p \in P \text{, there is a pattern } q \in Q \text{ such that } p \preceq q.$$

**Theorem 2.3 (Sato et al.[11])** *Let $k$ be a positive integer and suppose $\sharp\Sigma \geq 2k+1$. Let $P$ be a nonempty finite set of patterns and let $Q$ be a set of $k$ patterns. Then the following equivalences are valid:*

$$S_1(P) \subseteq L(Q) \quad \Longleftrightarrow \quad L(P) \subseteq L(Q) \quad \Longleftrightarrow \quad P \sqsubseteq Q.$$

**Theorem 2.4 (Sato et al.[11])** *Suppose $\sharp\Sigma \geq 4$. Let $p, q_1, q_2$ be patterns. Then the following equivalence is valid:*

$$L(p) \subseteq L(q_1) \cup L(q_2) \quad \Longleftrightarrow \quad p \preceq q_1 \text{ or } p \preceq q_2.$$

**Theorem 2.5 (Sato and Mukouchi [12])** *Let $p, q \in \mathcal{RP}$. Then for any pattern $r \in$mi $\{p, q\}$,*

$$\max\{l_p, l_q\} \leq l_r \leq l_p + l_q.$$

By the above, for any pattern $r \in$mi $P$, the following inequality is valid:

$$l_r \leq \sum_{p \in P} l_p.$$

By Lemma 2.2 and Theorem 2.3, it follows that:

**Theorem 2.6** *Let $k$ be a positive integer and suppose $\sharp\Sigma \geq 2k+1$. Let $P$ be a nonempty finite set of patterns and let $Q$ be a set of $k$ patterns. Then the following equivalences are valid:*

$$S_i(\text{mi } P) \subseteq L(Q) \quad \Longleftrightarrow \quad \bigcap_{p \in P} L(p) \subseteq L(Q) \quad \Longleftrightarrow \quad \text{mi } P \sqsubseteq Q,$$

*where $i$ is the least integer such that $i \times l \geq \sum_{p \in P} l_p$ with $l = \max\{l_p \mid p \in P\}$.*

**Lemma 2.7** *Suppose $\sharp\Sigma \geq k+1$. Let $l$ be a positive integer and let $p_i \in \mathcal{RP}$ with $l_{p_i} \leq l$ for $i = 1, 2, \cdots, k$. If $\Sigma^l \subseteq L(p_1) \cup \cdots \cup L(p_k)$, then $p_i \in X^+$ holds for some $i$.*

*Proof.* We assume that $\Sigma^l \subseteq L(p_1) \cup \cdots \cup L(p_k)$, but $p_i \notin X^+$ for any $i$. Then for each $i$, $p_i$ contains at least one constant symbol, say $a_i \in \Sigma$. Since $\sharp\Sigma \geq k+1$, there exists a symbol $b \in \Sigma$ different from $a_i$ $(i = 1, \cdots, k)$. Consider the string $w = b \cdots b$ with length $l$. As easily seen, $w \in \Sigma^l$ but $w \notin L(p_1) \cup \cdots \cup L(p_k)$. It is a contradiction. ∎

Hereafter, we assume that $\sharp\Sigma \geq 4$. Thus by Theorem 2.3 and Theorem 2.4, $S_1(p) \subseteq L(q)$ iff $p \preceq q$, and $L(p) \subseteq L(q_1) \cup L(q_2)$ iff $p \preceq q_1$ or $p \preceq q_2$.

For each pattern $p \notin X^+$, we introduce a particular type of string called a *co-pattern* of $p$, denoted by $p^c$. We define the semantics (language) of a co-pattern $p^c$ as follows:

$$L(p^c) = L(p)^c \cap \Sigma^{\geq l_p},$$

where $L(p)^c$ is the complement of the language $L(p)$, i.e., $L(p)^c = \Sigma^+ - L(p)$. By co-$\mathcal{RP}$, we denote the set of co-patterns, and put $\mathcal{JRP} = \mathcal{RP} \cup$ co-$\mathcal{RP}$. For a co-pattern $p^c$ of a pattern $p$, let us define its length $l_{p^c}$ as the length $l_p$ of $p$, that is, we put $l_{p^c} = l_p$.

**Lemma 2.8** *Let $L_1$ and $L_2$ be languages, and let $p, q \in \mathcal{RP}$. Then the following equivalences are valid:*

(1) $L_1 \cap L(p^c) \subseteq L_2 \iff L_1 \cap \Sigma^{\geq l_p} \subseteq L_2 \cup L(p)$,

(2) $L_1 \subseteq L_2 \cup L(q^c) \iff L_1 \cap (L(q) \cup \Sigma^{<l_q}) \subseteq L_2$.

*Proof.* By the definition of $L(p^c)$, the assertion (1) can be easily proved. We only prove that the assertion (2) is valid. The implication $\Rightarrow$ is clear from the definition of $L(p^c)$. Assume $L_1 \cap (L(q) \cup \Sigma^{<l_q}) \subseteq L_2$. Let $w \in L_1$. By our assumption, clearly if $w \in L(q)$ then $w \in L_2$. Otherwise $w \in L(q^c)$ or $w \in \Sigma^{<l_q}$. By our assumption, the latter yields $w \in L_2$. In any case, $w \in L_2 \cup L(q^c)$ holds. Hence the assertion (2) is valid. ∎

For a co-pattern $p$ and for a positive integer $i$, we define

$$S_i(p^c) = L(p^c) \cap \Sigma^{\leq i \times l_p}.$$

Clearly $S_i(p) \cup S_i(p^c) = \Sigma^{[l_p, i \times l_p]}$ and $S_i(p) \cap S_i(p^c) = \phi$.

**Lemma 2.9** *Let $p, q \in \mathcal{RP}$. Then the following equivalences are valid:*

(1) $S_2(p) \subseteq L(q^c) \iff L(p) \subseteq L(q^c) \iff \text{mi}\,\{p, q\} = \phi,\ l_q \leq l_p$,

(2) $S_1(p^c) \subseteq L(q) \iff L(p^c) \subseteq L(q) \iff q \in X^+,\ l_q \leq l_p$,

(3) $S_1(p^c) \subseteq L(q^c) \iff L(p^c) \subseteq L(q^c) \iff \text{mi}\,\{q, \chi_{l_p}\} \sqsubseteq \{p\},\ l_q \leq l_p$.

*Proof.* We only prove that the assertion (1) is valid. The other assertions can be proved similarly.

Assume that mi $\{p, q\} = \phi$ and $l_q \leq l_p$. Then by Lemma 2.2, the former implies $L(p) \cap L(q) = \phi$. By $l_q \leq l_p$, it means that $L(p) \subseteq L(q^c)$.

Clearly $L(p) \subseteq L(q^c)$ implies $S_2(p) \subseteq L(q^c)$.

Now we prove $S_2(p) \subseteq L(q^c)$ implies mi $\{p, q\} = \phi$ and $l_q \leq l_p$. Clearly $l_q \leq l_p$ holds. If there is a pattern $r \in$mi $\{p, q\}$, then by Theorem 2.5, $l_r \leq 2 \times l_p$ holds. By the definition of $S_2$, it implies $S_1(r) \subseteq S_2(p)$, and so $S_1(r) \subseteq L(q^c)$. It contradicts that $r \preceq q$. ∎

## 3 Inclusion Problem (1).

This section considers the inclusion problem of the form

$$L(\pi) \subseteq L(\tau_1) \cup L(\tau_2),$$

where $\pi, \tau_1, \tau_2 \in \mathcal{JRP}$.

**Theorem 3.1** *Let $\pi, \tau_1, \tau_2 \in \mathcal{JRP}$ with $l_{\tau_1} \leq l_{\tau_2}$. If $2 \times l_\pi < l_{\tau_2}$ holds for a case of $\pi \in \mathcal{RP}$ and $\tau_1 \in$ co-$\mathcal{RP}$, and $l_\pi < l_{\tau_2}$ for the other cases, then the following equivalence is valid:*

$$L(\pi) \subseteq L(\tau_1) \cup L(\tau_2) \iff L(\pi) \subseteq L(\tau_1).$$

*Proof.* We prove only for the following two cases. The other cases can be proved similarly.

A case of $(\pi, \tau_1, \tau_2) = (p, q_1, \tau_2)$. Assume that $l_p < l_{\tau_2}$ and $l_{q_1} \leq l_{\tau_2}$. If $L(p) \subseteq L(q_1) \cup L(\tau_2)$, then $S_1(p) \subseteq L(q_1) \cup L(\tau_2)$. By $l_p < l_{\tau_2}$, it follows that $S_1(p) \subseteq L(q_1)$. By Theorem 2.3, it leads to $L(p) \subseteq L(q_1)$.

A case of $(\pi, \tau_1, \tau_2) = (p, q_1^c, \tau_2)$. Assume that $2 \times l_p < l_{\tau_2}$ and $l_{q_1} \leq l_{\tau_2}$. If $L(p) \subseteq L(q_1^c) \cup L(\tau_2)$, then $S_2(p) \subseteq L(q_1^c) \cup L(\tau_2)$. By $2 \times l_p < l_{\tau_2}$ and the definition of $S_2(p)$, $S_2(p) \cap L(\tau_2) = \phi$, and thus $S_2(p) \subseteq L(q_1^c)$. By Lemma 2.9(1), it leads to $L(p) \subseteq L(q_1^c)$. ∎

On the other hand, the conditions for the equivalence in the theorem above is not valid, then the following theorem holds:

**Theorem 3.2** *Let $\pi, \tau_1, \tau_2 \in \mathcal{JRP}$ with $l_{\tau_1} \leq l_{\tau_2}$. If $2 \times l_\pi \geq l_{\tau_2}$ holds for a case of $\pi \in \mathcal{RP}$ and $\tau_1 \in$ co-$\mathcal{RP}$, and $l_\pi \geq l_{\tau_2}$ for the other cases, then the following statements are equivalent:*

*(1) $S_4(\pi) \subseteq L(\tau_1) \cup L(\tau_2)$, (2) $L(\pi) \subseteq L(\tau_1) \cup L(\tau_2)$,*

*(3) These patterns satisfy the syntactic inclusions shown in Table 1.*

| $\pi$ | $\tau_1$ | $\tau_2$ | Syntactic Inclusions |
|---|---|---|---|
| | $q_1$ | $q_2$ | $p \preceq q_1$ or $p \preceq q_2$ |
| $p$ | $q_1$ | $q_2^c$ | mi $\{p, q_2\} \sqsubseteq \{q_1\}$ |
| | $q_1^c$ | $q_2$ | mi $\{p, q_1\} \sqsubseteq \{q_2\}$, $l_{q_1} \leq l_p$ |
| | $q_1^c$ | $q_2^c$ | mi $\{p, q_1, q_2\} = \phi$, $l_{q_1} \leq l_p$ |
| | $q_1$ | $q_2$ | $q_1 \in X^+$ or $q_2 \in X^+$ |
| $p^c$ | $q_1$ | $q_2^c$ | mi $\{q_2, \chi_{l_p}\} \sqsubseteq \{q_1, p\}$ |
| | $q_1^c$ | $q_2$ | mi $\{q_1, \chi_{l_p}\} \sqsubseteq \{q_2, p\}$ |
| | $q_1^c$ | $q_2^c$ | mi $\{q_1, q_2, \chi_{l_p}\} \sqsubseteq \{p\}$ |

Table 1: Syntactic Inclusions

The proof will be shown by Theorem 2.3 for a case of $\pi, \tau_1, \tau_2 \in \mathcal{RP}$ and six lemmas for the other cases given below.

**Lemma 3.3** *Let $l_{q_2} \leq l_p$. Then the following statements are equivalent:*

*(1) $S_2(p) \subseteq L(q_1) \cup L(q_2^c)$, (2) $L(p) \subseteq L(q_1) \cup L(q_2^c)$, (3)* mi $\{p, q_2\} \sqsubseteq \{q_1\}$.

*Proof.* By the definition of $S_2(p)$, clearly the assertion (2) implies the assertion (1). Since the assertion (3) implies $L(p) \cap L(q_2) \subseteq L(q_1)$, by the assumption of $l_{q_2} \leq l_p$, it gives the assertion (2).

Now we prove that (1) implies (3). We assume that the assertion (1) is valid. Then by $l_{q_2} \leq l_p$, the assertion (1) implies that $S_2(p) \cap L(q_2) \subseteq L(q_1)$. Let $r$ be an arbitrary pattern in mi $\{p, q_2\}$. Then by $l_{q_2} \leq l_p$ and Theorem 2.5, $l_r \leq l_p + l_{q_2} \leq 2 \times l_p$ holds. Hence we have $S_1(r) \subseteq S_2(p)$. It implies together with $r \preceq q_2$ that $S_1(r) \subseteq S_2(p) \cap L(q_2)$. By the assertion (1), it means that $S_1(r) \subseteq L(q_1)$, i.e., $r \preceq q_1$. Therefore the assertion (3) is valid. ∎

**Lemma 3.4** *Let $l_{q_1} \leq l_{q_2}$. Then the following statements are equivalent:*

*(1) $S_2(p) \subseteq L(q_1^c) \cup L(q_2)$, (2) $L(p) \subseteq L(q_1^c) \cup L(q_2)$, (3)* mi $\{p, q_1\} \sqsubseteq \{q_2\}$, $l_{q_1} \leq l_p$.

*Proof.* By Lemma 3.4, (3) implies (1) and (2).

Now we prove that (1) implies (3). We assume that the assertion (1) is valid. Then, as easily seen, $\min\{l_{q_1}, l_{q_2}\} \leq l_p$ holds. This implies $l_{q_1} \leq l_p$, because $l_{q_1} \leq l_{q_2}$ by the assumption. Thus, by Lemma 3.4, the assertion (3) is valid.

Similarly, we can show that (2) implies (3). ∎

**Lemma 3.5** *Let $l_{q_1} \leq l_{q_2} \leq 2 \times l_p$. Then the following statements are equivalent:*

*(1) $S_4(p) \subseteq L(q_1^c) \cup L(q_2^c)$, (2) $L(p) \subseteq L(q_1^c) \cup L(q_2^c)$, (3)* mi $\{p, q_1, q_2\} = \phi$, $l_{q_1} \leq l_p$.

*Proof.* We only show that (1) implies (3).

Assume that the assertion (1) is valid. Then clearly $l_{q_1} \leq l_p$ holds. Suppose there is a pattern $r \in$mi $\{p, q_1, q_2\}$. By $l_{q_1} \leq l_{q_2} \leq 2 \times l_p$ and Theorem 2.5, $l_r \leq l_p + l_{q_1} + l_{q_2} \leq 4 \times l_p$

holds. Hence by the definition of $S_4(p)$, we have $S_1(r) \subseteq S_4(p)$. By the assertion (1), it implies $S_1(r) \subseteq L(q_1^c) \cup L(q_2^c)$. It is a contradiction because of $r \preceq q_1$ and $r \preceq q_2$. ∎

**Lemma 3.6** *Let $l_{q_1} \leq l_p$ and $l_{q_2} \leq l_p$. Then the following statements are equivalent:*
*(1) $S_1(p^c) \subseteq L(q_1) \cup L(q_2)$, (2) $L(p^c) \subseteq L(q_1) \cup L(q_2)$, (3) $q_1 \in X^+$ or $q_2 \in X^+$.*

*Proof.* We only show that (1) implies (3).

Assume that the assertion (1) is valid. Then we have $\Sigma^{l_p} \subseteq L(p) \cup L(q_1) \cup L(q_2)$. By Lemma 2.7 and $p \notin X^+$, the assertion (3) is valid. ∎

**Lemma 3.7** *Let $l_{q_2} \leq l_p$. Then the following statements are equivalent:*
*(1) $S_1(p^c) \subseteq L(q_1) \cup L(q_2^c)$, (2) $L(p^c) \subseteq L(q_1) \cup L(q_2^c)$, (3)* $\mathrm{mi}\ \{q_2, \chi_{l_p}\} \sqsubseteq \{p, q_1\}$.

*Proof.* We only show that (1) implies (3).

Assume that the assertion (1) is valid. Then by $l_{q_2} \leq l_p$, the assertion (1) implies that $\Sigma^{l_p} \cap L(q_2) \subseteq L(p) \cup L(q_1)$. Suppose that there is a pattern $r \in \mathrm{mi}\ \{q_2, \chi_{l_p}\}$ such that $r \not\preceq p$ and $r \not\preceq q_1$. Then by $l_{q_2} \leq l_p$, $l_r = l_p$ holds. It means that $S_1(r) \subseteq L(p) \cup L(q_1)$. By Theorem 2.3, it follows that $r \preceq p$ or $r \preceq q_1$, and a contradiction. ∎

**Lemma 3.8** *Let $l_{q_1} \leq l_p$ and $l_{q_2} \leq l_p$. Then the following statements are equivalent:*
*(1) $S_2(p^c) \subseteq L(q_1^c) \cup L(q_2^c)$, (2) $L(p^c) \subseteq L(q_1^c) \cup L(q_2^c)$, (3)* $\mathrm{mi}\ \{q_1, q_2, \chi_{l_p}\} \sqsubseteq \{p\}$.

*Proof.* We only show that (1)implies (3).

Assume that the assertion (1) is valid. Then by $l_{q_1} \leq l_p$ and $l_{q_2} \leq l_p$, the assertion (1) implies that $\Sigma^{[l_p, 2\times l_p]} \cap L(q_1) \cap L(q_2) \subseteq L(p)$. Suppose that there is a pattern $r \in \mathrm{mi}\ \{q_1, q_2, \chi_{l_p}\}$ such that $r \not\preceq p$. By $l_{q_1} \leq l_p$ and $l_{q_2} \leq l_p$, $r \in \mathrm{mi}\ \{q_1, q_2\}$ or $r \in \mathrm{mi}\ \{r', \chi_{l_p}\}$ holds for some $r' \in \mathrm{mi}\ \{q_1, q_2\}$. Hence by Theorem 2.5, we have $l_r \leq l_{q_1} + l_{q_2} \leq 2 \times l_p$. It means that $S_1(r) \subseteq \Sigma^{[l_p, 2\times l_p]} \cap L(q_1) \cap L(q_2)$. By the assertion (1), it implies that $S_1(r) \subseteq L(p)$, and thus $r \preceq p$. It is a contradiction. ∎

Let $\pi_1$ and $\pi_2$ be regular patterns or co-regular patterns. We put

$$S_i(\pi_1, \pi_2) = S_i(\pi_1) \cup S_i(\pi_2) \quad (i \geq 1).$$

Then by Theorem 3.2, the next results immediately follow:

**Corollary 3.9** *Let $\pi_1, \pi_2, \tau_1, \tau_2 \in \mathcal{JRP}$. Then the following two statements are equivalent:*
*(1) $S_4(\pi_1, \pi_2) \subseteq L(\tau_1) \cup L(\tau_2)$, (2) $L(\pi_1) \cup L(\pi_2) \subseteq L(\tau_1) \cup L(\tau_2)$.*

**4 Inclusion Problem (2).** This section considers the inclusion problem of the form

$$L(\pi_1) \cap L(\pi_2) \subseteq L(\tau),$$

where $\pi_1, \pi_2, \tau \in \mathcal{JRP}$.

We first consider the length of shortest strings in $L(\pi_1) \cap L(\pi_2)$. For $\pi_1, \pi_2 \in \mathcal{RP}$, say $\pi_1 = p_1$ and $\pi_2 = p_2$, the length of shortest strings in the language $L(p_1) \cap L(p_2)$ is given by $l_{p_1,p_2}$, where $l_{p_1,p_2}$ denotes the length of shortest patterns in mi $\{p_1, p_2\}$. On the other hand, if $\pi_1 \in$ co-$\mathcal{RP}$ or $\pi_2 \in$ co-$\mathcal{RP}$, then the length of shortest strings in $L(\pi_1) \cap L(\pi_2)$ is given by $\max\{l_{\pi_1}, l_{\pi_2}\}$.

Then we introduce particular finite sets of strings as follows:

$$T_i(\pi_1,\pi_2) = \begin{cases} S_i(\mathrm{mi}\ \{\pi_1,\pi_2\}), & \text{if } \pi_1,\pi_2 \in \mathcal{RP}, \\ L(\pi_1)\cap L(\pi_2)\cap \Sigma^{\le i\times l}, & \text{otherwise,} \end{cases}$$

where $i \ge 1$, and $l = \max\{l_{\pi_1}, l_{\pi_2}\}$.

By the above definition, for any language $L$, we have the followings:

$$\begin{array}{rcl} T_i(p_1,p_2^c) \subseteq L & \iff & S_i(\mathrm{mi}\ \{p_1,\chi_{l_{p_2}}\}) \subseteq L(p_2)\cup L, \\ T_i(p_1^c,p_2^c) \subseteq L & \iff & \Sigma^{[l,i\times l]} \subseteq L(p_1)\cup L(p_2)\cup L. \end{array}$$

For patterns $p_1,\cdots,p_k$ and a positive integer $l$, $\mathrm{mi}_{<l}\ \{p_1,\cdots,p_k\}$ denotes the set of patterns in $\mathrm{mi}\ \{p_1,\cdots,p_k\}$ with length less than $l$. Clearly if $l < \max\{l_{p_1},\cdots,l_{p_k}\}$, then $\mathrm{mi}_{<l}\ \{p_1,\cdots,p_k\} = \phi$. By definition, $r \in\mathrm{mi}_{<l}\ \{p_1,\cdots,p_k\}$ is equivalent to the relation $L(r)\cap\Sigma^{<l} \subseteq L(p_1)\cap\cdots\cap L(p_k)\cap\Sigma^{<l}$.

**Theorem 4.1** *Let $\pi_1,\pi_2,\tau \in \mathcal{JRP}$. Then the following statements are equivalent:*
*(1) $T_2(\pi_1,\pi_2) \subseteq L(\tau)$, (2) $L(\pi_1)\cap L(\pi_2) \subseteq L(\tau)$,*
*(3) These patterns satisfy syntactic conditions in Table 2, where $l = \max\{l_{p_1}, l_{p_2}\}$.*

| $\pi_1$ | $\pi_2$ | $\tau$ | Syntactic Inclusions |
|---|---|---|---|
| $p_1$ | $p_2$ | | $\mathrm{mi}\ \{p_1,p_2\} \sqsubseteq \{q\}$ |
| $p_1$ | $p_2^c$ | $q$ | $\mathrm{mi}\ \{p_1,\chi_{l_{p_2}}\} \sqsubseteq \{p_2,q\}$ |
| $p_1^c$ | $p_2^c$ | | $q \in X^+$, $l_q \le l$ |
| $p_1$ | $p_2$ | | $\mathrm{mi}\ \{p_1,p_2,q\} = \phi$, $l_q \le l_{p_1,p_2}$ |
| $p_1$ | $p_2^c$ | $q^c$ | $\mathrm{mi}\ \{p_1,q,\chi_{l_{p_2}}\} \sqsubseteq \{p_2\}$ |
| $p_1^c$ | $p_2^c$ | | $\mathrm{mi}\ \{q,\chi_l\} \sqsubseteq \{p_1,p_2\}$ |

Table 2: Syntactic Inclusions

The proof will be shown by Theorem 2.6 and lemmas given below.

**Lemma 4.2** *The following statements are equivalent:*
*(1) $T_1(p_1,p_2^c) \subseteq L(q)$, (2) $L(p_1)\cap L(p_2^c) \subseteq L(q)$, (3) $\mathrm{mi}\ \{p_1,\chi_{l_{p_2}}\} \sqsubseteq \{p_2,q\}$.*

*Proof.* If $p_1 \preceq p_2$, then clearly $T_1(p_1,p_2^c) = L(p_1)\cap L(p_2^c)\cap\Sigma^{l_{p_1}} = \phi$ and $\mathrm{mi}\ \{p_1,\chi_{l_{p_2}}\} = \{p_1\}$ hold. Thus our lemma is valid. We only show that (1) implies (3) under the assumption $p_1 \not\preceq p_2$. We assume that $T_1(p_1,p_2^c) \subseteq L(q)$, i.e., $S_1(\mathrm{mi}\ \{p_1,\chi_{l_{p_2}}\}) \subseteq L(p_2)\cup L(q)$. Then for every $r \in\mathrm{mi}\ \{p_1,\chi_{l_{p_2}}\}$, we have $S_1(r) \subseteq L(p_2)\cup L(q)$, and so $r \preceq p_2$ or $r \preceq q$. Thus the assertion (3) is valid. ∎

**Lemma 4.3** *Let $l_{p_1} \le l_{p_2}$. Then the following statements are equivalent:*
*(1) $T_1(p_1^c,p_2^c) \subseteq L(q)$, (2) $L(p_1^c)\cap L(p_2^c) \subseteq L(q)$, (3) $q \in X^+$, $l_q \le l_{p_2}$.*

*Proof.* We only show that (1) implies (3).

Assume that the assertion (1) is valid. Then by $l_{p_1} \le l_{p_2}$, we have $l_q \le l_{p_2}$ and $\Sigma^{l_{p_2}} \subseteq L(p_1)\cup L(p_2)\cup L(q)$. Therefore by Lemma 2.7 and $\Sigma \ge 4$, we get $q \in X^+$. ∎

**Lemma 4.4** *Let $l_{p_1} \le l_{p_2}$. Then the following statements are equivalent:*
*(1)* $T_2(p_1, p_2) \subseteq L(q^c)$*, (2)* $L(p_1) \cap L(p_2) \subseteq L(q^c)$*, (3)* $\mathrm{mi}\ \{p_1, p_2, q\} = \phi$*,* $l_q \le l_{p_1,p_2}$*.*

*Proof.* We only show that (1) implies (3).

We assume that $T_2(p_1, p_2) = S_2(\mathrm{mi}\ \{p_1, p_2\}) \subseteq L(q^c)$. Then clearly $l_q \le l_{p_1,p_2}$ holds. Suppose that $\mathrm{mi}\ \{p_1, p_2, q\} \neq \phi$, i.e., there is a pattern $r \in \mathrm{mi}\ \{p_1, p_2, q\}$. Then by Lemma 2.1, there is a pattern $r' \in \mathrm{mi}\ \{p_1, p_2\}$ such that $r \in \mathrm{mi}\ \{r', q\}$. Since $l_q \le l_{p_1,p_2}$ and $l_{p_1,p_2} \le l_{r'}$, we have $l_q \le l_{r'}$. By $r \in \mathrm{mi}\ \{r', q\}$, $l_q \le l_{r'}$ and Theorem 2.5, we have $l_r \le l_{r'} + l_q \le 2 \times l_{r'}$. On the other hand, by the assertion (1), we have $S_2(r') \subseteq T_2(p_1, p_2) \subseteq L(q^c)$. It means that $S_1(r) \subseteq S_2(r')$. Hence we have $S_1(r) \subseteq L(q^c)$. It contradicts that $r \preceq q$. ∎

**Lemma 4.5** *The following statements are equivalent:*
*(1)* $T_2(p_1, p_2^c) \subseteq L(q^c)$*, (2)* $L(p_1) \cap L(p_2^c) \subseteq L(q^c)$*, (3)* $\mathrm{mi}\ \{p_1, q, \chi_{l_{p_2}}\} \sqsubseteq \{p_2\}$*.*

*Proof.* We only show that (1) implies (3).

We assume that the assertion (1) is valid, but there is a pattern $r \in \mathrm{mi}\ \{p_1, q, \chi_{l_{p_2}}\}$ such that $r \not\preceq p_2$. Then there is a pattern $r' \in \mathrm{mi}\ \{p_1, \chi_{l_{p_2}}\}$ such that $r \in \mathrm{mi}\ \{r', q\}$. Clearly if $l_{p_1} \le l_{p_2}$ then $l_{r'} = l_{p_2}$. Otherwise $r' = p_1$. In any case, since $l_r \le l_{r'} + l_q \le 2 \times l_{r'}$ by Theorem 2.5, we have $S_1(r) \subseteq S_2(r')$. On the other hand, since the assertion (1) is equivalent to $S_2(\mathrm{mi}\ \{p_1, \chi_{l_{p_2}}\}) \subseteq L(p_2) \cup L(q^c)$, we have $S_1(r) \subseteq S_2(r') \subseteq L(p_2) \cup L(q^c)$. Since $r \preceq q$, it means that $S_1(r) \subseteq L(p_2)$, and so $r \preceq p_2$. It is a contradiction. ∎

**Lemma 4.6** *Let $l_{p_1} \le l_{p_2}$. Then the following statements are equivalent:*
*(1)* $T_1(p_1^c, p_2^c) \subseteq L(q^c)$*, (2)* $L(p_1^c) \cap L(p_2^c) \subseteq L(q^c)$*, (3)* $\mathrm{mi}\ \{q, \chi_{l_{p_2}}\} \sqsubseteq \{p_1, p_2\}$*.*

*Proof.* We only show that (1) implies (3).

Assume that the assertion (1) is valid. Then by $l_{p_1} \le l_{p_2}$, clearly $l_q \le l_{p_2}$ must hold. It implies that $L(q) \cap \Sigma^{l_{p_2}} \subseteq L(p_1) \cup L(p_2)$, and so $S_1(\mathrm{mi}\ \{q, \chi_{l_{p_2}}\}) \subseteq L(p_1) \cup L(p_2)$. By Theorem 2.3, it means that (3) is valid. ∎

**Corollary 4.7** *Let $\pi_1, \pi_2, \tau_1, \tau_2 \in \mathcal{JRP}$. Then the following statements are equivalent:*
*(1)* $T_2(\pi_1, \pi_2) \subseteq L(\tau_1) \cap L(\tau_2)$*, (2)* $L(\pi_1) \cap L(\pi_2) \subseteq L(\tau_1) \cap L(\tau_2)$*.*

## 5 Inclusion Problem (3).

Finally, this section considers the inclusion problem of the form

$$L(\pi_1) \cap L(\pi_2) \subseteq L(\tau_1) \cup L(\tau_2),$$

where $\pi_1, \pi_2, \tau_1, \tau_2 \in \mathcal{JRP}$.

**Theorem 5.1** *Let $\pi_1, \pi_2, \tau_1, \tau_2 \in \mathcal{JRP}$ with $l_{\tau_1} \le l_{\tau_2}$. If $\max\{l_{\pi_1}, l_{\pi_2}\} < l_{\tau_2}$ holds for a case of $\pi_1 \in$ co-$\mathcal{RP}$ or $\pi_2 \in$ co-$\mathcal{RP}$, and $2 \times \max\{l_{\pi_1}, l_{\pi_2}\} < l_{\tau_2}$ for the other cases, then the following statements are equivalent:*

$$L(\pi_1) \cap L(\pi_2) \subseteq L(\tau_1) \cup L(\tau_2) \quad \Longleftrightarrow \quad L(\pi_1) \cap L(\pi_2) \subseteq L(\tau_1).$$

*Proof.* We prove only for the following two cases. The other cases can be proved similarly.

A case of $(\pi_1, \pi_2, \tau_1, \tau_2) = (p_1, p_2, q_1, q_2^c)$ with $l_{p_1} \le l_{p_2}$, $l_{q_1} \le l_{q_2}$, $2 \times l_{p_2} < l_{q_2}$. If $L(p_1) \cap L(p_2) \subseteq L(q_1) \cup L(q_2^c)$, then $S_1(\mathrm{mi}\ \{p_1, p_2\}) \subseteq L(q_1) \cup L(q_2^c)$. There is a pattern $r \in \mathrm{mi}\ \{p_1, p_2\}$. By Theorem 2.5, $l_r \le 2 \times l_{p_2}$. Therefore $S_1(r) \cap L(q_2^c) = \phi$, it leads to $L(p_1) \cap L(p_2) \subseteq L(q_1)$.

A case of $(\pi_1, \pi_2, \tau_1, \tau_2) = (p_1, p_2^c, q_1, q_2^c)$ with $l_{p_1} \leq l_{p_2}$, $l_{q_1} \leq l_{q_2}$, $l_{p_2} < l_{q_2}$. If $L(p_1) \cap L(p_2^c) \subseteq L(q_1) \cup L(q_2^c)$, then $S_1(\text{mi } \{p_1, \chi_{l_{p_2}}\}) \subseteq L(p_2) \cup L(q_1) \cup L(q_2^c)$. By $l_{p_2} < l_{q_2}$, it follows that $S_1(\text{mi } \{p_1, \chi_{l_{p_2}}\}) \subseteq L(p_2) \cup L(q_1)$, and thus $T_1(p_1, p_2^c) \subseteq L(q_1)$. By Lemma 4.2, it leads to $L(p_1) \cap L(p_2^c) \subseteq L(q_1)$. ∎

**Theorem 5.2** *Let $\pi_1, \pi_2, \tau_1, \tau_2 \in \mathcal{JRP}$ and $\sharp\Sigma \geq 5$. Then the following statements are equivalent provided that $l_{\tau_1}, l_{\tau_2} \leq \max\{l_{\pi_1}, l_{\pi_2}\}$ if $\pi_1 \in$ co-$\mathcal{RP}$ or $\pi_2 \in$ co-$\mathcal{RP}$.*
*(1) $T_4(\pi_1, \pi_2) \subseteq L(\tau_1) \cup L(\tau_2)$, (2) $L(\pi_1) \cap L(\pi_2) \subseteq L(\tau_1) \cup L(\tau_2)$,*
*(3) These patterns satisfy the syntactic conditions in Table 3.*

Note that if $\pi_i \in$ co-$\mathcal{RP}$ and $l_{\tau_j} > \max\{l_{\pi_1}, l_{\pi_2}\}$ for some $i, j$, by Theorem 4.1 and Theorem 5.1, the assertion (2) in the above is equivalent to the assertions (1) and (3) in Theorem 4.1 for $\tau = \tau_{j'}$ with $j' \neq j$.

| $\pi_1$ | $\pi_2$ | $\tau_1$ | $\tau_2$ | Syntactic Conditions |
|---|---|---|---|---|
| $p_1$ | $p_2$ | $q_1$ | $q_2$ | mi $\{p_1, p_2\} \sqsubseteq \{q_1, q_2\}$ |
| | | $q_1$ | $q_2^c$ | mi $\{p_1, p_2, q_2\} \sqsubseteq \{q_1\}$, mi$_{<l_{q_2}}$ $\{p_1, p_2\} \sqsubseteq \{q_1\}$, $\min\{l_{q_1}, l_{q_2}\} \leq l_{p_1,p_2}$ |
| | | $q_1^c$ | $q_2^c$ | mi $\{p_1, p_2, q_1, q_2\} = \phi$, mi$_{<l_{q_2}}$ $\{p_1, p_2, q_1\} = \phi$, $l_{q_1} \leq l_{p_1,p_2}$, where $l_{q_1} \leq l_{q_2}$ |
| $p_1$ | $p_2^c$ | $q_1$ | $q_2$ | mi $\{p_1, \chi_{l_{p_2}}\} \sqsubseteq \{p_2, q_1, q_2\}$ |
| | | $q_1$ | $q_2^c$ | mi $\{p_1, q_2, \chi_{l_{p_2}}\} \sqsubseteq \{p_2, q_1\}$ |
| | | $q_1^c$ | $q_2^c$ | mi $\{p_1, q_1, q_2, \chi_{p_2}\} \sqsubseteq \{p_2\}$ |
| $p_1^c$ | $p_2^c$ | $q_1$ | $q_2$ | $q_1 \in X^+$ or $q_2 \in X^+$ |
| | | $q_1$ | $q_2^c$ | mi $\{q_2, \chi_{l_{p_2}}\} \sqsubseteq \{p_1, p_2, q_1\}$, where $l_{p_1} \leq l_{p_2}$ |
| | | $q_1^c$ | $q_2^c$ | mi $\{q_1, q_2, \chi_{l_{p_2}}\} \sqsubseteq \{p_1, p_2\}$, where $l_{p_1} \leq l_{p_2}$ |

Table 3: Syntactic Inclusions

The proof will be shown by lemmas given below.

**Lemma 5.3** *The following statements are equivalent:*
*(1) $T_2(p_1, p_2) \subseteq L(q_1) \cup L(q_2^c)$, (2) $L(p_1) \cap L(p_2) \subseteq L(q_1) \cup L(q_2^c)$,*
*(3) mi $\{p_1, p_2, q_2\} \sqsubseteq \{q_1\}$, mi$_{<l_{q_2}}$ $\{p_1, p_2\} \sqsubseteq \{q_1\}$, $\min\{l_{q_1}, l_{q_2}\} \leq l_{p_1,p_2}$.*

*Proof.* We only show that (1) implies (3).

Assume that the assertion (1) is valid. Then clearly $l = \min\{l_{q_1}, l_{q_2}\} \leq l_{p_1,p_2}$ holds. Since $S_1(\text{mi}_{<l_{q_2}} \{p_1, p_2\}) \subseteq S_2(\text{mi } \{p_1, p_2\}) = T_2(p_1, p_2)$, we have $S_1(\text{mi}_{<l_{q_2}} \{p_1, p_2\}) \subseteq L(q_1)$, and so mi$_{<l_{q_2}}$ $\{p_1, p_2\} \sqsubseteq \{q_1\}$.

Assume that mi $\{p_1, p_2, q_2\} \not\sqsubseteq \{q_1\}$, and there is a pattern $r \in$mi $\{p_1, p_2, q_2\}$ satisfying $r \not\preceq q_1$. By the choice of $r$, there is a pattern $r' \in$mi $\{p_1, p_2\}$ such that $r \in$mi $\{r', q_2\}$. By the assertion (1), it implies that

$$S_2(r') \subseteq S_2(\text{mi } \{p_1, p_2\}) \subseteq L(q_1) \cup L(q_2^c), \quad L(r) \cap L(q_2^c) = \phi.$$

A case of $l_{q_2} \leq l_{r'}$. In this case, by $r \in$mi $\{r', q_2\}$ and Theorem 2.5, we obtain $l_r \leq 2 \times l_{r'}$ and so, $S_1(r) \subseteq S_2(r') \subseteq L(q_1) \cup L(q_2^c)$. By the above, it means that $S_1(r) \subseteq L(q_1)$, i.e., $r \preceq q_1$ and a contradiction.

A case of $l_{q_2} > l_{r'}$. In this case, by $S_1(r') \subseteq S_2(r')$ and $S_1(r') \cap L(q_2^c) = \phi$, it implies together with the above that $S_1(r') \subseteq L(q_1)$, and so $r' \preceq q_1$. It means that $r \preceq q_1$, and a contradiction. ∎

**Lemma 5.4** *Let $l_{q_1} \le l_{q_2}$. Then the following statements are equivalent:*
*(1)* $T_4(p_1, p_2) \subseteq L(q_1^c) \cup L(q_2^c)$, *(2)* $L(p_1) \cap L(p_2) \subseteq L(q_1^c) \cup L(q_2^c)$,
*(3)* $\mathrm{mi}\ \{p_1, p_2, q_1, q_2\} = \phi$, $\mathrm{mi}_{<l_{q_2}}\ \{p_1, p_2, q_1\} = \phi$, $l_{q_1} \le l_{p_1,p_2}$.

*Proof.* We only show that (1) implies (3).

We assume that the assertion (1) holds but not the assertion (3). Since $l_{q_1} \le l_{q_2}$, clearly $l_{q_1} \le l_{p_1,p_2}$ holds.

Suppose that there is a pattern $r \in \mathrm{mi}\ \{p_1, p_2, q_1, q_2\}$. Then there is a pattern $r' \in \mathrm{mi}\ \{p_1, p_2\}$ satisfying $r \in \mathrm{mi}\ \{r', q_1, q_2\}$. By $l_{q_1} \le l_{p_1,p_2}$ and $l_{p_1,p_2} \le l_{r'}$, we have $l_{q_1} \le l_{r'}$. Moreover, by Theorem 2.5, $l_r \le l_{r'} + l_{q_1} + l_{q_2} \le 2 \times l_{r'} + l_{q_2}$.

A case of $l_{q_2} \le 2 \times l_{r'}$. In this case, we have $l_r \le 4 \times l_{r'}$ and so $S_1(r) \subseteq S_4(r')$. By the assertion (1), it means that $S_1(r) \subseteq L(q_1^c) \cup L(q_2^c)$. Since $r \preceq q_1$ and $r \preceq q_2$, it is a contradiction.

A case of $2 \times l_{r'} < l_{q_2}$. In this case, since $S_2(r') \cap L(q_2^c) = \phi$, we have $S_2(r') \subseteq L(q_1^c)$. By Lemma 2.1, it implies that $\mathrm{mi}\ \{r', q_1\} = \phi$. It contradicts that $\mathrm{mi}\ \{r', q_1, q_2\} \ne \phi$. Therefore $\mathrm{mi}\ \{p_1, p_2, q_1, q_2\} = \phi$ is valid.

Similarly, we can prove $\mathrm{mi}_{<l_{q_2}}\ \{p_1, p_2, q_1\} = \phi$. ∎

**Lemma 5.5** *Let $l_{q_1} \le l_{q_2}$. Then the following statements are equivalent:*
*(1)* $T_1(p_1, p_2^c) \subseteq L(q_1) \cup L(q_2)$, *(2)* $L(p_1) \cap L(p_2^c) \subseteq L(q_1) \cup L(q_2)$,
*(3)* $\mathrm{mi}\ \{p_1, \chi_{l_{p_2}}\} \sqsubseteq \{p_2, q_1, q_2\}$.

*Proof.* We only show that (1) implies (3).

By the assertion (1), we have $S_1(\mathrm{mi}\ \{p_1, \chi_{l_{p_2}}\}) \subseteq L(p_2) \cup L(q_1) \cup L(q_2)$. It implies that $\mathrm{mi}\ \{p_1, \chi_{l_{p_2}}\} \sqsubseteq \{p_2, q_1, q_2\}$. ∎

**Lemma 5.6** *Let $l_{q_1} \le l$, $l_{q_2} \le l$, where $l = \max\{l_{p_1}, l_{p_2}\}$. Then the following statements are equivalent:*
*(1)* $T_2(p_1, p_2^c) \subseteq L(q_1) \cup L(q_2^c)$, *(2)* $L(p_1) \cap L(p_2^c) \subseteq L(q_1) \cup L(q_2^c)$,
*(3)* $\mathrm{mi}\ \{p_1, q_2, \chi_{l_{p_2}}\} \sqsubseteq \{p_2, q_1\}$.

*Proof.* We only show that (1) implies (3) under the assumption of $p_1 \not\preceq p_2$.

By assumptions on lengths of $l_{q_1}$, $l_{q_2}$ and $l$, the assertion (1) is equivalent to

$$S_2(\mathrm{mi}\ \{p_1, \chi_{p_2}\}) \cap L(q_2) \subseteq L(p_2) \cup L(q_1).$$

Suppose that there is a pattern $r \in \mathrm{mi}\ \{p_1, q_2, \chi_{l_{p_2}}\}$ such that $r \not\preceq p_2$ and $r \not\preceq q_1$. Then there is a pattern $r' \in \mathrm{mi}\ \{p_1, \chi_{l_{p_2}}\}$ satisfying $r \in \mathrm{mi}\ \{r', q_2\}$. Clearly $l_{r'} = l$ and $l_r \le l + l_{q_2}$ hold. By $l_{q_2} \le l$, it follows that $l_r \le 2 \times l = 2 \times l_{r'}$, and so $S_1(r) \subseteq S_2(r')$. Hence by the assertion (1) we get $S_1(r) \subseteq L(p_2) \cup L(q_1) \cup L(q_2^c)$. Since $r \preceq q_2$ holds, and so we have $S_1(r) \cap L(q_2^c) = \phi$. It means that $r \preceq p_2$ or $r \preceq q_1$. It is a contradiction. ∎

**Lemma 5.7** *Let $l_{q_1} \le l_{q_2} \le l$, where $l = \max\{l_{p_1}, l_{p_2}\}$. Then the following statements are equivalent:*
*(1)* $T_3(p_1, p_2^c) \subseteq L(q_1^c) \cup L(q_2^c)$, *(2)* $L(p_1) \cap L(p_2^c) \subseteq L(q_1^c) \cup L(q_2^c)$,
*(3)* $\mathrm{mi}\ \{p_1, q_1, q_2, \chi_l\} \sqsubseteq \{p_2\}$.

*Proof.* We only show that (1) implies (3).

Assume that the assertion (1) is valid. Then by the assumption on lengths of $l_{q_1}$, $l_{q_2}$ and $l$, the assertion (1) is equivalent to

$$S_3(\mathrm{mi}\ \{p_1, \chi_{l_{p_2}}\}) \subseteq L(p_2) \cup L(q_1^c) \cup L(q_2^c).$$

Suppose that there is a pattern $r \in \mathrm{mi}\ \{p_1, q_1, q_2, \chi_l\}$ such that $r \not\preceq p_2$. Then there is a pattern $r' \in \mathrm{mi}\ \{p_1, \chi_{l_{p_2}}\}$ satisfying $r \in \{r', q_1, q_2\}$. As easily seen, $l_{r'} = l$ holds. By Theorem 2.5 and the assumptions on lengths, it means that $l_r \leq l_{r'} + l_{q_1} + l_{q_2} \leq 3 \times l$, and thus $S_1(r) \subseteq S_3(\mathrm{mi}\ \{p_1, \chi_{l_{p_2}}\})$. Since $r \preceq q_1$ and $r \preceq q_2$, the assertion (1) gives $S_1(r) \subseteq S_3(\mathrm{mi}\ \{p_1, \chi_{l_{p_2}}\}) \cap L(q_1) \cap L(q_2) \subseteq L(p_2)$. It implies that $r \preceq p_2$, and a contradiction. ∎

**Lemma 5.8** *Let $l_{p_1} \leq l_{p_2}$, $l_{q_1} \leq l_{q_2}$ and $l_{q_1} \leq l_{p_2}$. If $\sharp\Sigma \geq 5$, then the following statements are equivalent:*

*(1) $T_1(p_1^c, p_2^c) \subseteq L(q_1) \cup L(q_2)$, (2) $L(p_1^c) \cap L(p_2^c) \subseteq L(q_1) \cup L(q_2)$,*

*(3) $q_1 \in X^+$ or $q_2 \in X^+$.*

*Proof.* We only show that (1) implies (3).

Assume that the assertion (1) is valid. Then by $l_{p_1} \leq l_{p_2}$, the assertion (1) is equivalent to $\Sigma^{l_{p_2}} \subseteq L(p_1) \cup L(p_2) \cup L(q_1) \cup L(q_2)$. By the assumption on $\sharp\Sigma \geq 5$ and Lemma 2.7, since $p_1, p_2 \notin X^+$ holds, it follows that $q_1 \in X^+$ or $q_2 \in X^+$. ∎

**Lemma 5.9** *Let $l_{p_1} \leq l_{p_2}$ and $\max\{l_{q_1}, l_{q_2}\} \leq l_{p_2}$. Then the following statements are equivalent:*

*(1) $T_1(p_1^c, p_2^c) \subseteq L(q_1) \cup L(q_2^c)$, (2) $L(p_1^c) \cap L(p_2^c) \subseteq L(q_1) \cup L(q_2^c)$,*

*(3) $\mathrm{mi}\ \{q_2, \chi_{l_{p_2}}\} \sqsubseteq \{p_1, p_2, q_1\}$.*

*Proof.* We only show that (1) implies (3).

Assume that the assertion (1) is valid. Then by $l_{p_1} \leq l_{p_2}$ and $l_{q_2} \leq l_{p_2}$, the assertion (1) is equivalent to $\Sigma^{l_{p_2}} \cap L(q_2) \subseteq L(p_1) \cup L(p_2) \cup L(q_1)$. It follows that $\mathrm{mi}\ \{q_2, \chi_{l_{p_2}}\} \sqsubseteq \{p_1, p_2, q_1\}$. ∎

**Lemma 5.10** *Let $l_{p_1} \leq l_{p_2}$ and $l_{q_1} \leq l_{q_2} \leq l_{p_2}$. Then the following statements are equivalent:*

*(1) $T_2(p_1^c, p_2^c) \subseteq L(q_1^c) \cup L(q_2^c)$, (2) $L(p_1^c) \cap L(p_2^c) \subseteq L(q_1^c) \cup L(q_2^c)$,*

*(3) $\mathrm{mi}\ \{q_1, q_2, \chi_{l_{p_2}}\} \sqsubseteq \{p_1, p_2\}$.*

*Proof.* We only show that (1) implies (3).

Assume that the assertion (1) is valid. Then by $l_{p_1} \leq l_{p_2}$ and $l_{q_1} \leq l_{q_2} \leq l_{p_2}$, the assertion (1) implies that

$$\Sigma^{[l_{p_2}, 2\times l_{p_2}]} \cap L(q_1) \cap L(q_2) \subseteq L(p_1) \cup L(p_2).$$

Suppose that there is a pattern $r \in \mathrm{mi}\ \{q_1, q_2, \chi_{l_{p_2}}\}$ such that $r \not\preceq p_1$ and $r \not\preceq p_2$. Then there is a pattern $r' \in \mathrm{mi}\ \{q_1, q_2\}$ satisfying $r \in \mathrm{mi}\ \{r', \chi_{l_{p_2}}\}$. Hence by Theorem 2.5, we have $l_{r'} \leq l_{q_1} + l_{q_2} \leq 2 \times l_{p_2}$. We show that $l_r \leq 2 \times l_{p_2}$.

A case of $l_{r'} \leq l_{p_2}$. In this case, by $r \in \mathrm{mi}\ \{r', \chi_{l_{p_2}}\}$, clearly $l_r = l_{p_2}$ holds.

A case of $l_{r'} > l_{p_2}$. In the case, clearly $r' = r$ must hold. Hence $l_r = l_{r'} \leq 2 \times l_{p_2}$.

By the above, it implies together with $r \preceq q_1$ and $r \preceq q_2$ that $S_1(r) \subseteq \Sigma^{[l_{p_2}, 2\times l_{p_2}]} \cap L(q_1) \cap L(q_2) \subseteq L(p_1) \cup L(p_2)$. It contradicts that $r \not\preceq p_2$ and $r \not\preceq p_2$. ∎

**Acknowledgement**

The author wishes to thank Prof. Masako Sato, Osaka Prefecture University, for many suggestions and encouragement.

## References

[1] D. Angluin, *Finding patterns common to a set of strings*, in Proceedings of the 11th Annual Symposium on Theory of Computing, (1979) 130–141.

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

[3] S. Arikawa, S. Kuhara, S. Miyano, Y. Mukouchi, A. Shinohara, and T. Shinohara, *A machine discovery from amino acid sequences by decision trees over regular patterns*, New Generation Computing, **11**(3,4) (1993) 361–375.

[4] H. Arimura, T. Shinohara, and S. Otsuki, *Finding minimal generalizations of pattern languages and its application to inductive inference from positive data*, in Proceedings of the 11th Symposium on Theoretical Aspects Computer Science, Lecture Notes in Computer Science, **775** (1994) 646–660.

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

[6] T. Jiang, A. Salomaa, K. Salomaa, and S. Yu, *Decision problems for patterns*, Journal of Computer and System Sciences, **50**(1) (1995) 53–63.

[7] S. Miyano, *Learning theory towards genome informatics*, in Proceedings of the 4th Workshop on Algorithmic Learning Theory, Lecture Notes in Artificial Intelligence, **744** (1993) 19–36.

[8] T. Moriyama and M. Sato, *Properties of language classes with finite elasticity*, IEICE Transactions on Information and Systems, **E78-D**(5) (1995) 532–538.

[9] T. Motoki and T. Shinohara, *The correct definition of finite elasticity: corrigendum to identification of unions*, in Proceedings of the 4th Annual ACM Workshop on Computational Learning Theory, (1991) 375–375.

[10] Y. Mukouchi, *Characterization of pattern languages*, IEICE Transactions on Information and Systems, **E75-D**(4) (1992) 420–425.

[11] M. Sato, Y. Mukouchi, and D. Zheng, *Characteristic sets for unions of regular pattern languages and compactness*, Lecture Notes in Artificial Intelligence, **1501** (1998) 220–233.

[12] M. Sato and Y. Mukouchi, *Learning of languages generated by patterns from positive examples*, Scientiae Mathematicae Japonicae, **8** (2003) 479–485.

[13] T. Shinohara, *Polynomial time inference of pattern languages and its applications*, in Proceedings of the 7th IBM Symposium on Mathematical Foundations of Computer Science, (1982) 191–209.

[14] T. Shinohara, *Inductive inference from negative data*, Bulletin of Informatics and Cybernetics, **21**(3, 4) (1985) 67–70.

[15] T. Shinohara and H. Arimura, *Inductive inference of unbounded unions of pattern languages from positive data*, in Proceedings of the 7th International Workshop on Algorithmic Learning Theory, Lecture Notes in Artificial Intelligence, **1160** (1996) 256–271.

[16] R.M. Smullyan, "Theory of formal Systems," Princeton University Press, Princeton, New Jersey, (1961).

[17] M. Terada, Y. Mukouchi, and M. Sato, *Inductive Inference of Decision Trees over Patterns from Positive Example*, IEICE Transactions on Information and Systems, **J83-D-I**(1) (2000) 60–67.

[18] K. Wright, *Identification of unions of language drawn from an identifiable class*, in Proceedings of the 2nd Workshop on Computational Learning Theory, (1989) 328–333.

International Buddhist University, Habikino, Osaka, 583-8501, Japan