LEARNING OF LANGUAGES GENERATED BY PATTERNS FROM POSITIVE EXAMPLES

MASAKO SATO AND YASUHITO MUKOUCHI

Received March 26, 2003

ABSTRACT. The paper investigates knowledge discovery based on learning of languages generated by patterns from positive examples. A pattern p is a finite string of constant symbols and variables, and the language defined by p is the set of constant strings obtained from p by substituting nonempty and constant strings for variables. We consider the class ${\mathcal{PL}_*}^k$ of unions of at most k intersections of finitely many pattern languages. It is well known that the class of unions of finitely many pattern languages is not inferable from positive examples. Every intersection of finitely many pattern languages can be represented as a union of pattern languages, but it may be a union of infinitely many ones. The class $\mathcal{PL}_*$ of intersections, however, is shown to have finite thickness. Using the result, we show that the class ${\mathcal{PL}_*}^k$ is refutably inferable from complete examples as well as inferable from positive examples.

In order to study efficient learning algorithm, we introduce two kind of syntactic ordered relations for finite sets of *regular* patterns, and show that the semantic containment of unions of intersections of regular pattern languages is equivalent to the syntactic containment of some sets of regular patterns. In terms of this result, the class of intersections of regular pattern languages is polynomial time (refutably) inferable from positive (complete) examples, under some assumption.

1 Introduction.

Inductive learning or inductive inference is a process to find general rules from their concrete examples. The present paper deals with inductive/refutable learning of languages generated by patterns from examples. A *pattern* is a nonempty finite string consisting of constant symbols and variables. The language $L(p)$ generated by a pattern p is the set of constant strings obtained from p substituting nonempty constant strings for variables in p.

Learnability of pattern languages has been extensively investigated in the framework of *identification in the limit* due to Gold[8] ([20], [9], [21]). Angluin[2] proved a theorem characterizing inferable classes from positive examples in Gold's framework, and gave a useful sufficient condition, called *finite thickness*, for the inferability. The class $\mathcal{PL}$ of pattern languages has finite thickness, and thus is inferable from positive examples ([2]). The class $\mathcal{PL}$ is one of the most basic class in the framework of elementary formal systems which was introduced by Smullyan[22] to develop a new theory of recursive functions, and was proposed as a unifying framework for language learning by Arikawa et al.([3], [5]). That is, an elementary formal system consisting of only one definite clause defines a pattern language. From practical point of view, pattern languages were investigated in another framework of learning such as PAC learning ([10], [19]) and learning from neighbor systems ([14], [15]).

Pattern languages merely are not used for some applications because of their simplicity. Various kinds of languages generated by patterns have been investigated in Gold's

2000 *Mathematics Subject Classification.* 68T05, 68T25.

Key words and phrases. Pattern language, inductive inference, refutable inference, identification in the limit.

framework such as languages generated by decision trees over patterns and so on ([4]). In particular, unions of pattern languages were extensively studied by Wright, Shinohara and Arimura ([23], [6], [7], [21]). Wright[23] introduced a more general sufficient condition called *finite elasticity* for the inferability than finite thickness, and showed that the property is closed under union operation. Moreover, Moriyama and Sato[11] and Sato[16] showed that the property is closed under various operations such as intersection, concatenation and so on. As a result, the classes of bounded unions of intersections of pattern languages are shown to have finite elasticity, and thus are inferable from positive examples. Note that an intersection of finite number of pattern languages can be expressed as a union of pattern languages, but not always finite number. However, the class of intersections of finitely many pattern languages is shown to have finite thickness, and thus is inferable from positive examples. It is well known that the class of finite unions of pattern languages is super-finite and thus is not inferable from positive examples. In some practical applications such as genome informatics, such languages generated by patterns are paid much attentions (Arikawa et al.[4]).

On the other hand, Mukouchi and Arikawa[13] proposed refutably inductive inference as a framework of machine discovery, and showed that the class $\mathcal{PL}$ is refutably inferable from complete examples.

In this paper, we investigate inductive learnability and refutability of bounded unions of intersections of pattern languages from examples. The class ${\mathcal{PL}_*}^k$ of unions of at most k intersections of finitely many pattern languages is shown to be inferable from positive examples as well as refutably inferable from complete examples.

In order to construct an efficient learning algorithm, the complexity of the membership problem or finding one of compact expressions generating a given sample is an important key. Concerning the membership problem, it is well known that the problem of regular pattern languages is polynomial time computable although that of pattern languages is NP-complete. A pattern is *regular* if each variable occurs at most once in the pattern.

We introduce two kinds of syntactic partially ordered relations on generalizations and instances of finite sets of patterns, and show an equivalence between the semantic containment and the syntactic containment, under some assumption of the cardinality of the alphabet. In terms of the result, if the set of minimal generalizations of a given sample is polynomial time computable, then the class $\mathcal{RPL}_*$ of intersections of finitely many regular pattern languages is polynomial time inferable from positive examples.

2 Preliminaries.

Let Σ be a fixed alphabet of *constant* symbols, and $X = \{x, y, z, x_1, x_2, \cdots\}$ be a countable set of *variables*. Assume $\Sigma \cap X = \phi$. A *pattern* 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, $p = axbyy$ and $q = axbyz$ are patterns over $\{a, b, c\}$ and p is regular but not q. By $\mathcal{P}$ and $\mathcal{RP}$, we mean the classes of patterns and regular patterns, respectively. The *length* of a pattern p, denoted by $|p|$, is the number of symbols in p.

A *substitution* θ is a homomorphism from patterns to patterns that maps each constant to itself. By $p\theta$, we denote the image of a pattern p by a substitution θ. We assume that $x\theta$ is not empty for any variable x. Thus $|p| \le |p\theta|$ for any θ.

We define the language generated by a pattern p as follows:

$$L(p) = \{w \in \Sigma^+ \mid \exists\theta \text{ s.t. } w = p\theta\}.$$

We denote $\mathcal{PL}$ and $\mathcal{RPL}$ the classes of pattern languages and regular pattern languages, respectively.

The present paper deals with various classes of languages generated by pattern languages.

We give the notions of *identification in the limit* from positive examples (Gold[8], Angluin[2]), and *refutable inferability* from complete examples (Mukouchi and Arikawa[13]).

Let N be the set of nonnegative integers. A language class $\mathcal{L} = L_0, L_1, \cdots$ over Σ is an *indexed class of recursive languages* if there is a computable function $f : N \times \Sigma^* \to \{0, 1\}$ such that $f(i, w) = 1$ if $w \in L_i$, otherwise 0. Hereafter we confine ourselves to indexed classes of recursive languages, and identify a class with a hypothesis space.

A *positive presentation*, or a *text*, of a nonempty language L is an infinite sequence of strings $w_1, w_2, \cdots$ in Σ^* such that $\{w_n \mid n \geq 1\} = L$. A *complete presentation*, or an *informant*, of L is an infinite sequence $(w_1, t_1), (w_2, t_2), \cdots$ of elements in $\Sigma^* \times \{+, -\}$ such that $\{w_n \mid n \geq 1,\ t_n = +\} = L$ and $\{w_n \mid n \geq 1,\ t_n = -\} = L^c (= \Sigma^+ \setminus L)$. In what follows, σ denotes a positive or complete presentation, and $\sigma[n]$ denotes the σ's initial segment of length n.

An *inductive inference machine* (IIM, for short) is an effective procedure M that requests inputs from time to time and produces nonnegative integers, called *guesses*, from time to time. Let M be an IIM and $M(\sigma[n])$ be the last guess of M which is successively presented $\sigma[n]$ on its input request.

An IIM M *converges* to $h \in N$ for a positive presentation, if there is an $n \in N$ such that for any $m \geq n$, $M(\sigma[m]) = h$. An IIM M *infers* a class $\mathcal{L}$ *in the limit from positive examples*, if for any $L_i \in \mathcal{L}$ and for any positive presentation σ of L_i, M converges to an index j for σ such that $L_j = L_i$. A class $\mathcal{L}$ is *inferable from positive examples*, if there is an IIM which infers $\mathcal{L}$ from positive examples.

In the above definition, the behavior of an IIM is not specified, when we feed a positive presentation of a target language not in the hypothesis space. Mukouchi and Arikawa[13] proposed an inference machine that can refute the entire space of hypothesis.

An *inductive inference machine that can refute hypothesis space* (RIIM, for short) is an effective procedure that works like an IIM and, moreover, has an ability to refute the class. A class $\mathcal{L}$ is *refutably inferable from complete examples*, if there is an RIIM M which satisfies the following condition: If L is contained in the hypothesis space, then M infers each target language L in the limit from complete examples, otherwise M refutes and stops at some stage.

3 Bounded unions of intersections of pattern languages.

A notion of finite thickness for a language class due to Angluin[2] is a very useful sufficient condition for inferability and defined as follows: A class $\mathcal{L}$ has *finite thickness*, if for any nonempty subset $S \subseteq \Sigma^+$, $\{L \in \mathcal{L} \mid S \subseteq L\}$ is finite. Since $\mathcal{PL}$ has finite thickness, the class is inferable from positive examples (cf. Angluin[2]).

We denote by $\mathcal{P}^*$ the class of nonempty finite subsets of $\mathcal{P}$. For the class $\mathcal{P}^*$, we define two language classes as follows:

$$\mathcal{PL}^* = \{L(P) \mid P \in \mathcal{P}^*\}, \qquad \mathcal{PL}_* = \{L(\{P\}) \mid P \in \mathcal{P}^*\},$$

where $L(P) = \bigcup_{p \in P} L(p)$ and $L(\{P\}) = \bigcap_{p \in P} L(p)$ for each $P \in \mathcal{P}^*$. Since $\mathcal{PL}$ has finite thickness, it immediately follows that:

Theorem 1 *The class $\mathcal{PL}_*$ of unbounded intersections of pattern languages has finite thickness. Thus it is inferable from positive examples.*

Note that the class $\mathcal{PL}^*$ of *unbounded* unions of pattern languages is not inferable from positive examples since it is a super-finite class (cf. Gold[8]). For a positive integer k, by $\mathcal{PL}^k$ we denote the class of unions of at most k pattern languages, i.e.,

$$\mathcal{PL}^k = \{L(P) \mid P \in \mathcal{P}^k\}, \qquad \text{where } \mathcal{P}^k = \{P \in \mathcal{P}^* \mid \sharp P \leq k\}.$$

A notion of finite elasticity due to Wright[23] is defined as follows (cf. Wright[23] and Motoki et al.[12]): A class $\mathcal{L}$ has *finite elasticity*, if there does not exist an infinite sequence $w_0, w_1, w_2, \cdots \in \Sigma^+$ and an infinite sequence $L_1, L_2, \cdots \in \mathcal{L}$ such that

$$\{w_0, w_1, \cdots, w_{k-1}\} \subseteq L_k, \quad \text{but} \quad w_k \notin L_k \quad (k \geq 1).$$

Finite elasticity is a good property in a sense that (1) it is a more general sufficient condition for inferability than finite thickness and (2) it is closed under various class operations such as union, intersection and so on (Wright[23], Moriyama and Sato[11], Sato[16]). Wright showed that the class $\mathcal{PL}^k$ $(k \geq 1)$ has finite elasticity and thus is inferable from positive examples.

For $P_1, \cdots, P_n \in \mathcal{P}^*$, we define

$$L(\{P_1, \cdots, P_n\}) = \bigcup_{i=1}^{n} L(\{P_i\}) = \bigcup_{i=1}^{n} \bigcap_{p \in P_i} L(p).$$

Clearly $L(P) = L(\{\{p_1\}, \cdots, \{p_n\}\})$ holds if $P = \{p_1, \cdots, p_n\}$ for patterns p_i's. Moreover, we define

$$\mathcal{PL}_*{}^k = \{L(\{P_1, \cdots, P_n\}) \mid 1 \leq n \leq k,\ P_1, \cdots, P_n \in \mathcal{P}^*\}.$$

Clearly $\mathcal{PL}^k \subsetneq \mathcal{PL}_*{}^k$ $(k \geq 1)$ holds. Since $\mathcal{PL}_*$ has finite thickness, we obtain the following result:

Theorem 2 *Let $k \geq 1$. The class $\mathcal{PL}_*{}^k$ has finite elasticity. Thus it is inferable from positive examples.*

Now we consider refutable inferability of the class $\mathcal{PL}_*{}^k$ from complete examples.

M-finite thickness introduced by Sato[16] is another generalized notion of finite thickness: A class $\mathcal{L}$ has *M-finite thickness*, if for any nonempty finite set $T \subseteq \Sigma^+$, (1) for any $L \in \mathcal{L}$ containing T, there is a minimal language $L' \in \mathcal{L}$ of S satisfying $L' \subseteq L$, and (2) $\{L \in \mathcal{L} \mid L \text{ is a minimal language of } S\}$ is finite.

Let us define the econ function for a class $\mathcal{L}$ as follows: For finite sets $T, F \subseteq \Sigma^+$,

$$\text{econ}(T, F) = \begin{cases} 1, & \text{if } \exists L \in \mathcal{L} \text{ s.t. } T \subseteq L \text{ and } F \subseteq L^c, \\ 0, & \text{o.w. .} \end{cases}$$

If a class $\mathcal{L}$ has finite elasticity and M-finite thickness and the econ function is computable, then the class is refutably inferable from complete examples (Sato[16]).

Theorem 3 *The class $\mathcal{PL}_*{}^k$ has M-finite thickness, and the* econ *function is computable. Thus the class is refutably inferable from complete examples.*

For regular patterns and regular pattern languages, the classes $\mathcal{RP}^*$, $\mathcal{RP}^k$, $\mathcal{RPL}^*$, $\mathcal{RPL}^k$, $\mathcal{RPL}_*$, and $\mathcal{RPL}_*{}^k$ are defined similarly to those for patterns.

We note that the theorems obtained in this section are valid for the subclasses $\mathcal{RPL}_*$ and $\mathcal{RPL}_*{}^k$.

4 Generalizations of sets of patterns and efficient learning. An inference machine M is *polynomial time updating* if after receiving w_n, M produces a guess h_n within a polynomial time in the sum of lengths of examples so far received. A class $\mathcal{L}$ is polynomial time inferable from positive examples if there is a polynomial time updating inference machine that infers the class from positive examples.

If a class has finite elasticity, and both of the membership and MINL problems are polynomial time computable, then the class is polynomial time inferable from positive examples (Angluin[1], Arimura et al.[6]). Here the MINL problem for a class $\mathcal{L}$ is finding one of minimal languages in $\mathcal{L}$ of a given finite set $S \subseteq \Sigma^+$, and the membership problem is to decide whether $w \in L$ or not for given $w \in \Sigma^+$ and $L \in \mathcal{L}$.

In this paragraph, we consider efficient learning algorithms for the classes $\mathcal{RPL}_*$ and ${\mathcal{RPL}_*}^k$ $(k \geq 1)$ from positive examples.

A pattern q is a *generalization* of a pattern p (or p is an *instance* of q), denoted by $p \preceq q$, if there is a substitution θ such that $p = q\theta$. If $p \preceq q$ and $q \preceq p$, p and q are equal except renaming variables. In this paper, we identify such patterns, and thus $\mathcal{P}$ is a partially ordered set under the relation $\preceq$.

Let $P \in \mathcal{P}^*$. A pattern q is a generalization of P if $p \preceq q$ for every $p \in P$. A pattern q is an instance of P if $q \preceq p$ for every $p \in P$. A pattern p is a *minimal generalization* (*mg*) if p is a generalization of P and there is no generalization p' of P such that $p' \prec p$. We define a *maximal instance* (*mi*) of P similarly. We denote by mg P and mi P the sets of minimal generalizations and maximal instances of P, respectively. Clearly mg P is finite since lengths of patterns in mg P are less than or equal to the shortest length of patterns in a finite set P. However mi P is not always finite. In fact, mi $\{axxa, xx\} = \{a^{2n} \mid n \geq 1\}$ over $\Sigma = \{a\}$. As easily seen, $L(\{P\}) = L(\text{mi } P)$, i.e., $\bigcap_{p \in P} L(p) = \bigcup_{q \in \text{mi } P} L(q)$ holds. Hence $L(axxa) \cap L(xx) = \{a^{2n} \mid n \geq 1\}$.

For regular patterns, the following result is given:

Lemma 1 *Let P be a nonempty finite set of regular patterns. Then lengths of regular patterns in* mi P *are less than or equal to* $\sum_{p \in P} |p|$. *Thus the set* mi P *is finite.*

We define relations on $\mathcal{P}^*$ as follows:

$$\begin{array}{rcl} P \sqsubseteq Q & \iff & \forall p \in P, \exists q \in Q \text{ s.t. } p \preceq q, \\ P \sqsubseteq' Q & \iff & \forall q \in Q, \exists p \in P \text{ s.t. } p \preceq q. \end{array}$$

Obviously $P \sqsubseteq Q$ implies $L(P) \subseteq L(Q)$, and $P \sqsubseteq' Q$ implies $L(\{P\}) \subseteq L(\{Q\})$. The converses, however, are not valid in general. The relation $\preceq$ is polynomial time computable for *regular* patterns ([20]) but NP-complete for patterns ([1]). Thus the relations $\sqsubseteq$ and $\sqsubseteq'$ are polynomial time computable for sets in $\mathcal{RP}^*$ but NP-complete for sets in $\mathcal{P}^*$. Hence the membership problems for the classes $\mathcal{RPL}_*$ and ${\mathcal{RPL}_*}^k$ are easily shown to be polynomial time computable.

Note that since $\mathcal{PL}_*$ and ${\mathcal{PL}_*}^k$ have finite elasticity, so have these classes considered.

We first consider the MINL problem for the class $\mathcal{RPL}_*$.

By the definition $\sqsubseteq'$, it is easily shown that mg mi $P \sqsubseteq' P$ holds for any $P \in \mathcal{RP}^*$, but not always the converse. A set $P \in \mathcal{RP}^*$ is *reduced* if mg mi $P = P$ holds.

Lemma 2 *For any $P \in \mathcal{RP}^*$, there is a reduced set Q in $\mathcal{RP}^*$ satisfying $L(\{P\}) = L(\{Q\})$.*

Let *reduced-*$\mathcal{RP}^*$ be the set of all reduced sets in $\mathcal{RP}^*$. Then the next result follows:

Lemma 3 *Let $P, Q \in$reduced-$\mathcal{RP}^*$. Then*

$$L(\{P\}) \subseteq L(\{Q\}) \iff P \sqsubseteq' Q \iff \text{mi } P \sqsubseteq Q.$$

By Lemma 3, the next result is given.

Theorem 4 *If for any finite set $S \subseteq \Sigma^+$,* mg S *is polynomial time computable, then the class $\mathcal{RPL}_*$ is polynomial time inferable from positive examples.*

Note that the problem finding one in mg S is polynomial time computable ([20]).

Next, we consider the MINL problem for the class $\mathcal{RPL}_*{}^k$.

Arimura et al.[7] presented an efficient algorithm for the MINL problem of the class $\mathcal{RPL}^k$ using a framework of generalization systems, and the class is polynomial time inferable from positive examples, provided that the class has compactness.

Here the class $\mathcal{RP}^k$ has *compactness w.r.t. containment* if for any $P, Q \in \mathcal{RP}^k$, the syntactic containment $P \sqsubseteq Q$ is equivalent to the semantic containment $L(P) \subseteq L(Q)$.

Theorem 5 (Sato et al.[17]) *The class $\mathcal{RP}^k$ has compactness w.r.t. containment if and only if $\sharp\Sigma \geq 2k-1$ for $k \geq 3$ and $\sharp\Sigma \geq 4$ for $k = 2$.*

For a pattern p, by $S_n(p)$ the set of strings over Σ obtained from p by substituting strings with length at most n to each variable. For $P \in \mathcal{P}^*$, put $S_n(P) = \bigcup_{p \in P} S_n(p)$.

Theorem 6 *Let $k \geq 3$, $\sharp\Sigma \geq 2k-1$, and let $P_1, \cdots, P_n, Q_1, \cdots, Q_m \in$ reduced-$\mathcal{RP}^*$ for $n \geq 1$ and $1 \leq m \leq k$. Then*

$$S_2\left(\bigcup_{i=1}^{n} \mathrm{mi}\ P_i\right) \subseteq L(\{Q_1, \cdots, Q_m\})$$
$$\iff \quad L(\{P_1, \cdots, P_n\}) \subseteq L(\{Q_1, \cdots, Q_m\}) \quad \iff \quad \bigcup_{i=1}^{n} \mathrm{mi}\ P_i \sqsubseteq \bigcup_{j=1}^{m} \mathrm{mi}\ Q_j.$$

Note that the former equivalence means that the set $S_2\left(\bigcup_{i=1}^{n} \mathrm{mi}\ P_i\right)$ is a *characteristic set* of $\bigcup_{i=1}^{n} L(\{P_i\})$ within $\mathcal{RPL}_*{}^k$, and the latter shows the equivalence between the semantic containment and the syntactic containment.

References

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

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

[3] S. Arikawa, T. Shinohara and A. Yamamoto: *Elementary formal systems as a unifying framework for language learning*, in Proceedings of the 2nd Annual ACM Workshop on Computational Learning Theory, pp. 312–327 (1989).

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

[5] S. Arikawa, M. Sato, A. Shinohara, and T. Shinohara: *Developments in computational learning and discovery*, Machine Intelligence, **15**, pp. 227–247 (1999).

[6] H. Arimura, T. Shinohara and S. Otsuki: *Polynomial time inference of unions of two tree pattern languages*, IEICE Transactions on Information and Systems, **E75-D**(4), pp. 426–434 (1992).

[7] H. Arimura, T. Shinohara and S. Otsuki: *Finding minimal generalizations for unions of pattern languages and its application to inductive inference from positive data*, Lecture Notes in Computer Science, **775**, pp. 649–660 (1994).

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

[9] S. Lange and R. Wiehagen: *Polynomial-time inference of arbitrary pattern languages*, New Generation Computing, **8**, pp. 361–370 (1991).

[10] A. Marron: *Learning pattern languages from a single initial example and from queries*, in Proceedings of the 1st Workshop on Computational Learning Theory, pp. 345–358 (1988).

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

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

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

[14] Y. Mukouchi and M. Sato: *Language Learning with a Neighbor System*, Lecture Notes in Artificial Intelligence, **1967**, pp. 183–196 (2000).

[15] Y. Mukouchi and M. Sato: *Refutable language learning with a neighbor system*, to appear in Theoretical Computer Science.

[16] M. Sato: *Inductive inference of formal languages*, Bulletin of Informatics and Cybernetics, **27**(1), pp. 85–106 (1995).

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

[18] M. Sato, Y. Mukouchi and M. Terada: *Refutable/inductive learning from neighbor examples and its application to decision trees over patterns*, Lecture Notes in Artificial Intelligence, **1967**, pp. 183–196 (2002).

[19] R.E. Schapire: *Pattern languages are not learnable*, in Proceedings of the 3rd Annual Workshop on Computational Learning Theory, pp. 122–129 (1990).

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

[21] T. Shinohara and H. Arimura: *Inductive inference of unbounded unions of pattern languages from positive data*, Lecture Notes in Artificial Intelligence, **1160**, pp. 256–271 (1996).

[22] R.M. Smullyan: "Theory of formal systems," Princeton University Press (1961).

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

Department of Mathematics and Information Sciences, College of Integrated Arts and Sciences, Osaka Prefecture University, Sakai, Osaka 599-8531, Japan