ELSEVIER Theoretical Computer Science 137 (1995) 129–144

Theoretical Computer Science

Complexity of computing Vapnik–Chervonenkis dimension and some generalized dimensions

Ayumi Shinohara*

Research Institute of Fundamental Information Science, Kyushu University 33, Hakozaki, Higashi-ku, Fukuoka 812, Japan

Abstract

In the PAC-learning model, the Vapnik–Chervonenkis (VC) dimension plays the key role to estimate the polynomial-sample learnability of a class of $\{0,1\}$-valued functions. For a class of $\{0,\ldots,N\}$-valued functions, the notion has been generalized in various ways. This paper investigates the complexity of computing VC-dimension and generalized dimensions: VC*-dimension, Ψ_*-dimension, and Ψ_G-dimension. For each dimension, we consider a decision problem that is, for a given matrix representing a class $\mathscr{F}$ of functions and an integer K, to determine whether the dimension of $\mathscr{F}$ is greater than K or not. We prove that (1) both the VC*-dimension and Ψ_G-dimension problems are polynomial-time reducible to the satisfiability problem of length J with $O(\log^2 J)$ variables, which include the original VC-dimension problem as a special case, (2) for every constant C, the satisfiability problem in conjunctive normal form with m clauses and $C\log^2 m$ variables is polynomial-time reducible to the VC-dimension problem, while (3) Ψ_*-dimension problem is NP-complete.

1. Introduction

The PAC learnability due to Valiant [12] is to estimate the feasibility of learning a concept ($\{0,1\}$-valued function) probably approximately correctly, from a reasonable amount of examples (polynomial-sample), within a reasonable amount of time (polynomial-time). It is well known that the Vapnik–Chervonenkis dimension (VC-dimension) which is a combinatorial parameter of a concept class plays the key role to determine whether the concept class is polynomial-sample learnable or not [3, 5, 8].

This paper settles complexity issues on VC-dimension and some generalized dimensions of a class over a finite learning domain. We remark that the complexity of computing each dimension is of independent interest from the polynomial-time learnability, since it is not directly related to the running time of learning algorithms. However, there are some works on this topic. Linial et al. [5] showed that the

*Corresponding author. E-mail: ayumi@rifis.kyushu-u.a.c.jp.

SSDI 0304-3975(94)00164-2

VC-dimension of a concept class over a finite learning domain can be computed in $n^{O(\log n)}$ time, where n is the size of a given matrix which represents the concept class. Nienhuys-Cheng and Polman [9] gave another $n^{O(\log n)}$-time algorithm, although they have not analyzed its running time. On the other hand, Megiddo and Vishkin [6] defined two classes $\mathrm{SAT}_{\log^2 n}$ and $\mathrm{SAT}^{\mathrm{CNF}}_{\log^2 n}$ which are between P and NP. $\mathrm{SAT}_{\log^2 n}$ ($\mathrm{SAT}^{\mathrm{CNF}}_{\log^2 n}$) is the class of sets which are polynomial-time reducible to the satisfiability problem of a boolean formula of length J with $O(\log^2 J)$ variables (in conjunctive normal form, respectively). They showed that the problem of finding a minimum dominating set in a tournament is in $\mathrm{SAT}_{\log^2 n}$ and $\mathrm{SAT}^{\mathrm{CNF}}_{\log^2 n}$-hard. That means the problem is a kind of "complete" problem for the class of $n^{O(\log n)}$ time computable sets. Linial et al. [5] pointed out that the decision problem of computing VC-dimension may have some connection with that problem.

Along this line, we will give polynomial-time reductions between some generalized VC-dimension problems and the satisfiability problems of boolean formulae with restricted number of variables.

We consider two kinds of generalizations of the VC-dimension for the class of $\{0,\ldots,N\}$-valued functions. Recall that the VC-dimension of a class of $\{0,1\}$-valued functions is defined as the maximum cardinality of a set which is *shattered* by the class [3]. For a class of $\{0,\ldots,N\}$-valued functions, we first define *VC*-dimension* by naturally generalizing the notion of shattering. We show that the VC*-dimension problem is in $\mathrm{SAT}_{\log^2 n}$, and $\mathrm{SAT}^{\mathrm{CNF}}_{\log^2 n}$-hard. This result includes the original VC-dimension problem as a special case.

The second generalization is *Ψ-dimensions* due to the Ben-David et al. [2]. They unified various dimensions such as pseudo-dimension [4], graph dimension [7] and Natarajan dimension [7] into a general scheme by introducing a family Ψ of mappings which translate $\{0,\ldots,N\}$-valued functions into $\{0,1\}$-valued functions. They defined Ψ-dimension as the maximum VC-dimension under mappings in Ψ. Let Ψ_* be the set of all mappings from $\{0,\ldots,N\}$ to $\{0,1\}$, and let Ψ_G-dimension denote the graph dimension [7]. We show that Ψ_*-dimension problem is NP-complete, while the Ψ_G-dimension problem is in $\mathrm{SAT}_{\log^2 n}$. These results suggest that the VC-dimension gives some interesting insights not only to algorithmic learning theory, but also to computational complexity theory.

2. Preliminaries

In this paper, log denotes the logarithm to the base 2 without extra notice. For a matrix M, let M_{ij} denote the element on row i and column j of M, and the size of M is the number of elements in M. The length of a boolean formula α, denoted by $|\alpha|$, is the total number of variable occurrences in α. For a formula α and a truth assignment σ to the variables of α, let $\sigma(\alpha)$ denote the truth value of α evaluated under σ. We denote truth values by 0 and 1. For a boolean formula α and a proposition c, we denote $[\alpha;c]=\alpha$ if c is true, and $[\alpha;c]=\neg\alpha$ otherwise.

For any integers $N, t \geqslant 1$ and $1 \leqslant i \leqslant N^t$, let $digit(i,N,t)$ denote the tth digit of the number $(i-1)$ in base N, that is, $i-1=\sum_{t=1}^{\lceil \log_N i \rceil} digit\ (i,N,t)N^{t-1}$. For example, $digit(7,2,1)=0$, $digit\ (7,2,2)=1$, and $digit\ (7,2,3)=1$, since the number 6 is denoted by "110" in base 2.

Let U be a finite set called a *learning domain.* We call a subset f of U a *concept.* A concept f can be regarded as a function $f: U \to \{0,1\}$, where $f(x)=1$ if x is in the concept and $f(x)=0$ otherwise. A *concept class* is a nonempty set $\mathscr{F} \subseteq 2^U$. We represent a concept class $\mathscr{F}$ over a finite learning domain U by a $|U| \times |\mathscr{F}|$ matrix M with $M_{ij}=f_j(x_i)$. Each column represents a concept in $\mathscr{F}$. For a $\{0,1\}$-valued matrix M, let $\mathscr{F}_M$ denote the concept class which M represents.

Definition 1. We say that $\mathscr{F}$ *shatters* a set $S \subseteq U$ if for every subset $T \subseteq S$ there exists a concept $f \in \mathscr{F}$ which *cuts T out of S*, i.e., $T=S \cap f$. The *Vapnik–Chervonenkis dimension* of $\mathscr{F}$, denoted by VC-dim$(\mathscr{F})$, is the maximum cardinality of a set which is shattered by $\mathscr{F}$.

Lemma 1 (Natarajan [8]). *For any concept class* $\mathscr{F}$, VC-dim$(\mathscr{F}) \leqslant \log|\mathscr{F}|$.

By this lemma, Linial et al. [5] immediately claimed that a simple algorithm which enumerates all possible sets to be shattered shall terminate in $n^{O(\log n)}$ time, where n is the size of a given matrix.

Definition 2 (Linial et al. [5]). The *discrete VC-dimension problem* is, given a $\{0,1\}$-valued matrix M and integer $K \geqslant 1$, to determine whether VC-dim $(\mathscr{F}_M) \geqslant K$ or not.

Definition 3 (Megiddo and Vishkin [6]). The classes $\mathrm{SAT}_{\log^2 n}$ and $\mathrm{SAT}^{\mathrm{CNF}}_{\log^2 n}$ are defined as follows:

(1) A set L is in $\mathrm{SAT}_{\log^2 n}$ if there exists a Turing machine $\mathscr{M}$, a polynomial $p(n)$, and a constant C, such that for every string I of length n, $\mathscr{M}$ converts I within $p(n)$ time into a boolean formula Φ_I (whose length is necessarily less than $p(n)$) with at most $C\log^2 n$ variables, so that $I \in L$ if and only if Φ_I is satisfiable.

(2) The definition of $\mathrm{SAT}^{\mathrm{CNF}}_{\log^2 n}$ is essentially the same as that of $\mathrm{SAT}_{\log^2 n}$ except that the formula Φ_I is in conjunctive normal form.

From the definitions, it is easy to see that

$$\mathrm{P} \subseteq \mathrm{SAT}^{\mathrm{CNF}}_{\log^2 n} \subseteq \mathrm{SAT}_{\log^2 n} \subseteq \mathrm{NP}.$$

3. Complexity of the discrete VC-dimension problem

In this section, we show that the discrete VC-dimension problem is in the class $\mathrm{SAT}_{\log^2 n}$, and hard for the class $\mathrm{SAT}^{\mathrm{CNF}}_{\log^2 n}$.

Theorem 1. *The discrete VC-dimension problem is in* $\mathrm{SAT}_{\log^2 n}$.

Proof. The problem is a special case of the VC*-dimension problem, which will be shown to be in $\mathrm{SAT}_{\log^2 n}$ in Theorem 3. □

Theorem 2. *Every* $L \in \mathrm{SAT}^{\mathrm{CNF}}_{\log^2 n}$ *is polynomial-time reducible to the discrete VC-dimension problem.*

We use the following lemma in the proof of Theorem 2.

Lemma 2. *Let* $\mathscr{F}$ *be a concept class over a learning domain* U, *and* S *be a subset of* U *with* $|S| = d \geq 2$. *If* S *is shattered by* $\mathscr{F}$, *then for any two distinct* x *and* y *in* S, *the number of concepts which contain exactly one of either* x *or* y *is at least* 2^{d-1}, *i.e.,*

$$|\{f \in \mathscr{F} \mid f(x) \neq f(y)\}| \geq 2^{d-1}.$$

Proof. Let $\mathscr{F}_{\bar{x}y} = \{f \in \mathscr{F} \mid f(x) = 0, f(y) = 1\}$, and $\mathscr{F}_{x\bar{y}} = \{f \in \mathscr{F} \mid f(x) = 1, f(y) = 0\}$. Then $\{f \in \mathscr{F} \mid f(x) \neq f(y)\} = \mathscr{F}_{\bar{x}y} \cup \mathscr{F}_{x\bar{y}}$, and $\mathscr{F}_{\bar{x}y} \cap \mathscr{F}_{x\bar{y}} = \emptyset$. It is easy to see that if S is shattered by $\mathscr{F}$ then the set $S - \{x, y\}$ is shattered by both $\mathscr{F}_{\bar{x}y}$ and $\mathscr{F}_{x\bar{y}}$. By Lemma 1, $|S - \{x, y)| \leq \log|\mathscr{F}_{\bar{x}y}|$ and $|S - \{x, y\}| \leq \log|\mathscr{F}_{x\bar{y}}|$. Thus $|\mathscr{F}_{\bar{x}y}| \geq 2^{d-2}$ and $|\mathscr{F}_{x\bar{y}}| \geq 2^{d-2}$, which yield $|\{f \in \mathscr{F} \mid f(x) \neq f(y)\}| = |\mathscr{F}_{\bar{x}y}| + |\mathscr{F}_{x\bar{y}}| \geq 2^{d-2} + 2^{d-2} = 2^{d-1}$. □

Proof of Theorem 2. Let $L \in \mathrm{SAT}^{\mathrm{CNF}}_{\log^2 n}$. Then there is a constant C_L and a polynomial $p_L(n)$ such that every string I of length n can be reduced in $p_L(n)$ time to a boolean formula in conjunctive normal form with at most $C_L \log^2 n$ variables, whose satisfiability coincides with the membership $I \in L$. Therefore, we have only to show that, for any C, there is a polyomial-time reduction from the satisfiability problem in conjunctive normal form with at most $C \log^2 n$ variables to the discrete VC-dimension problem. Let $\Phi = E_1 \wedge \cdots \wedge E_m$ $(m \geq 2)$ be a boolean formula where each E_i is a disjunction and the total number of distinct variables occurring in Φ is not greater than $C \log^2 m$. Without loss of generality, we can assume that m is a power of 2. We can also assume that the number of variables is exactly $C \log^2 m$, and let us rename them, for convenience, with double indices v_{st} $(1 \leq s \leq \log m,\ 1 \leq t \leq C \log m)$. We first construct a matrix M_Φ which has $(m^C + 1)\log m$ rows and $m^2 + m(\log m - 1)$ columns, and then prove that VC-dim $(\mathscr{F}_{M_\Phi}) \geq 2 \log m$ if and only if Φ is satisfiable.

The learning domain U corresponding to Φ is defined as $U = X \cup Y$ with $X \cap Y = \emptyset$, where $X = \{x_{sl} \mid 1 \leq s \leq \log m,\ 1 \leq l \leq m^C\}$ and $Y = \{y_u \mid 1 \leq u \leq \log m\}$. Let $X_s = \{x_{sl} \in X \mid 1 \leq l \leq m^C\}$ for each $s \in \{1, \ldots, \log m\}$, and let $X^{[k]} = \bigcup_{s \in \{s \mid digit(k,2,s) = 1\}} X_s$ for each $k \in \{1, \ldots, m\}$. The ith subset $Y^{[i]}$ of Y is defined by $Y^{[i]} = \{y_u \in Y \mid digit(i, 2, u) = 1\}$ for each $i \in \{1, \ldots, m\}$.

The concept class $\mathscr{F} \subseteq 2^U$ is defined as the union of distinct subclasses $F_1, \ldots, F_m$, and G. Here, the structure of G depends only on the number m:

$$G = \{g_{ik} \mid 1 \leq i \leq m,\ 1 \leq k \leq m - 1\}, \quad \text{where } g_{ik} = Y^{[i]} \cup X^{[k]}.$$

We emphasize that k ranges only to $m-1$, but not m. On the other hand, each concept in F_i reflects the structure of the clause E_i in Φ:

$$F_i = \{f_{ij} \mid 1 \leqslant j \leqslant \log m\}, \quad \text{where } f_{ij} = Y^{[i]} \cup (X - X_j) \cup X_j^*(E_i)$$

with

$$X_j^*(E_j) = \left\{ x_{jl} \in X_j \;\middle|\; \begin{array}{l} E_i \text{ contains a positive literal } v_{jt} \text{ with } digit(l,2,t)=1 \text{ or} \\ E_i \text{ contains a negative literal } \neg v_{jt} \text{ with } digit(l,2,t)=0 \\ \text{for some } t \in \{1, \ldots, C\log m\} \end{array} \right\}.$$

Fig. 1 illustrates the structure of the matrix M_Φ.

Clearly the cardinality of learning domain, i.e., the row size of the matrix M representing $\mathscr{F}$ is

$$|U| = |X| + |Y| = m^C \cdot \log m + \log m = (m^C + 1)\log m$$

and the cardinality of the concept class $\mathscr{F}$, i.e., the column size of M is

$$|\mathscr{F}| = |G| + |F_1| + \cdots + |F_m| = m(m-1) + m \cdot \log m.$$

Moreover, it is easy to see that M_Φ can be constructed in polynomial time with respect to the length of a given formula Φ.

We now prove that if the formula Φ is satisfiable then VC-dim$(\mathscr{F}) \geqslant 2\log m$. For an assignment σ which satisfies Φ, we consider the set $S_\sigma = Y \cup X_\sigma$ with

$$X_\sigma = \{x_{s,\langle\sigma,s\rangle} \in X \mid 1 \leqslant s \leqslant \log m\}, \quad \text{where } \langle \sigma, s\rangle = \sum_{t=1}^{C\log m} 2^{t-1} \cdot \sigma(v_{st}) + 1.$$

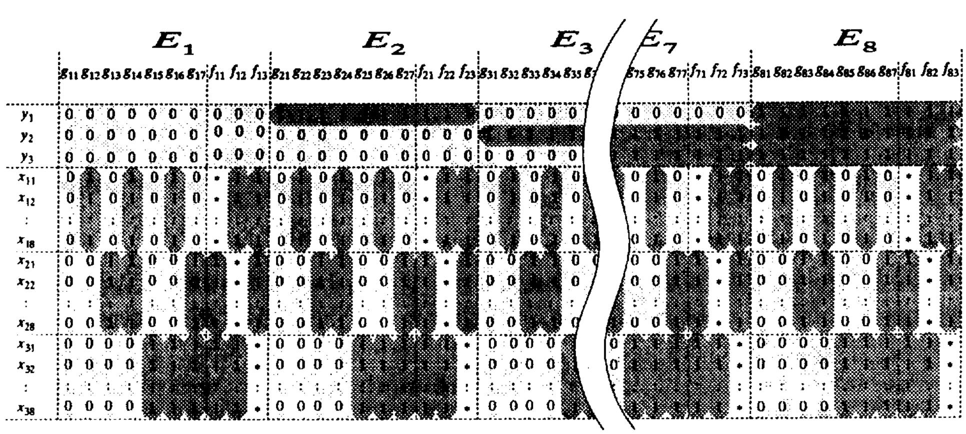

Fig. 1. Structure of the matrix M_Φ reduced from a boolean formula $\Phi = E_1 \wedge E_2 \wedge \cdots \wedge E_8$ with $C=1$. In this case, $K = 2\log 8 = 6$. The only elements marked $*$ depend on the structure of each clause E_i in Φ.

It is clear that $|S_\sigma|=|Y|+|X_\sigma|=2\log m$. We will show that S_σ is shattered by $\mathscr{F}$, i.e., for every $T\subseteq S_\sigma$, there exists an $f\in\mathscr{F}$ with $S_\sigma\cap f=T$. Let $i_T=\sum_{y_u\in T\cap Y}2^{u-1}+1$. It is easy to see that $i_T\in\{1,\ldots,m\}$ and $T\cap Y=Y^{[i_T]}$. According to $T\cap X_\sigma=X_\sigma$ or not, we have the following two cases.

(1) In case of $T\cap X_\sigma\subsetneq X_\sigma$, let $k_T=\sum_{x_{s,\langle\sigma,s\rangle}\in T\cap X_\sigma}2^{s-1}+1$. Then we can see that $k_T\in\{1,\ldots,m-1\}$ and $T\cap X_\sigma=X^{[k_T]}\cap X_\sigma$. Therefore, the concept $g_{i_T,k_T}\in G\subseteq\mathscr{F}$ cuts T out of S_σ as follows:

$$g_{i_T,k_T}\cap S_\sigma=(Y^{[i_T]}\cup X^{[k_T]})\cap(Y\cup X_\sigma)=(Y^{[i_T]}\cap Y)\cup(X^{[k_T]}\cap X_\sigma)$$

$$=(T\cap Y)\cup(T\cap X_\sigma)=T.$$

(2) In case of $T\cap X_\sigma=X_\sigma$, since σ satisfies Φ, the disjunction E_{i_T} in Φ is also satisfied by σ. That means E_{i_T} contains either positive literal v_{st} with $\sigma(v_{st})=1$, or negative literal $\neg v_{st}$ with $\sigma(v_{st})=0$, for some s and t. Let us take such an s (not necessarily unique), and let $j_T=s$. Then by the definition of $\langle\sigma,j_T\rangle$, we see $digit(\langle\sigma,j_T\rangle,2,t)=\sigma(v_{j_T,t})$ for each t. Thus, $x_{j_T,\langle\sigma,j_T\rangle}$ is included in $X^*_{j_T}(E_{i_T})$, and moreover, $X^*_{j_T}(E_{i_T})\cap X_\sigma=\{x_{j_T,\langle\sigma,j_T\rangle}\}$. Therefore, the concept $f_{i_T,j_T}\in F_{i_T}\subseteq\mathscr{F}$ cuts T out of S_σ as follows:

$$f_{i_T,j_T}\cap S_\sigma=(Y^{[i_T]}\cup(X-X_{j_T})\cup X^*_{j_T}(E_{i_T}))\cap(Y\cup X_\sigma)$$

$$=(Y^{[i_T]}\cap Y)\cup((X-X_{j_T})\cap X_\sigma)\cup(X^*_{j_T}(E_{i_T})\cap X_\sigma)$$

$$=(T\cap Y)\cup(X_\sigma-\{x_{j_T,\langle\sigma,j_T\rangle}\})\cup(x_{j_T,\langle\sigma,j_T\rangle}\}$$

$$=(T\cap Y)\cup X_\sigma=(T\cap Y)\cup(T\cap X_\sigma)=T.$$

In each case, T is shown to be cut out of S_σ by some concept in $\mathscr{F}$. Therefore, S_σ is shattered by $\mathscr{F}$.

We now show the converse. Suppose that VC-dim$(\mathscr{F})\geqslant 2\log m$. Then there is a set $S\subseteq U$ of cardinality $2\log m$ which is shattered by $\mathscr{F}$.

Claim 1. *S contains exactly one element from each X_s $(1\leqslant s\leqslant\log m)$ and all elements from Y.*

Proof. *Case $m=2$*: The learning domain is $U=\{y_1\}\cup X_1$ and the concept class is $\mathscr{F}=\{f_{11},g_{11},f_{21},g_{21}\}$. Since $|\mathscr{F}|=4$ and $g_{11}(x)=g_{21}(x)=0$ for any $x\in X_1$, no two elements from X_1 can be included in S which is to be shattered by $\mathscr{F}$. Moreover, since $|Y|=|\{y_1\}|=1$, the claim holds.

Case $m\geqslant 3$: Let $s\in\{1,\ldots,\log m\}$ be fixed arbitrarily, and x_1,x_2 be distinct elements in X_s. Suppose that S contains both x_1 and x_2. Then by Lemma 2,

$$|\{h\in\mathscr{F}\mid h(x_1)\neq h(x_2)\}|\geqslant 2^{2\log m-1}=\tfrac{1}{2}m^2.$$

On the other hand, let us consider a concept $h \in \mathscr{F}$ with $h(x_1) \neq h(x_2)$. Since $g_{ik}(x_1) = g_{ik}(x_2) = digit(k, 2, s)$ for any $g_{ik} \in G$, the concept h is not in G. Moreover, since $f_{ij}(x_1) = f_{ij}(x_2) = 1$ for any $f_{ij} \in F_1 \cup \cdots \cup F_m$ with $j \neq s$, thus h must be one of the concepts from $\{f_{1s}, f_{2s}, \ldots, f_{ms}\}$. Therefore,

$$|\{h \in \mathscr{F} \mid h(x_1) \neq h(x_2)\}| \leqslant |\{f_{1s}, f_{2s}, \ldots, f_{ms}\}| = m,$$

which yields a contradiction since $\frac{1}{2}m^2 > m$ for any $m \geqslant 3$. Thus, S can contain at most one element from X_s for each $s \in \{1, \ldots, \log m\}$. Since $|S| = 2\log m$ and $|Y| = \log m$, the set S must contain exactly one element from each X_s and all elements from Y. □

Proof of Theorem 2 (*conclusion*). Therefore, for each $s \in \{1, \ldots, \log m\}$, there is a unique $l = l(s) \in \{1, \ldots, m^C\}$ such that $x_{s,l(s)} \in S$, and we can assume that $S = Y \cup X_{(l)}$, where $X_{(l)} = \{x_{s,l(s)} \mid 1 \leqslant s \leqslant \log m\}$. Let σ_S be an assignments corresponding to S with

$$\sigma_S(v_{st}) = digit(l(s), 2, t) \quad (1 \leqslant s \leqslant \log m,\ 1 \leqslant t \leqslant C \log m).$$

We now show that σ_S satisfies all disjunctions E_i in Φ. Let $i \in \{1, \ldots, m\}$ be fixed arbitrarily. Since S is shattered by $\mathscr{F}$, for the subset $T_i = Y^{[i]} \cup X_{(l)}$ of S there is a concept $h_i \in \mathscr{F}$ with $S \cap h_i = T_i$. Since $S \cap h_i = (Y \cap h_i) \cup (X_{(l)} \cap h_i)$, the concept h_i must satisfy the following two conditions:

$$Y \cap h_i = Y^{[i]}, \tag{1}$$

$$X_{(l)} \cap h_i = X_{(l)}. \tag{2}$$

Note that no concept in G satisfies condition (2), and no concept in $F_{i'}$ with $i' \neq i$ satisfies condition (1). Therefore, such an $h_i \in \mathscr{F}$ is in F_i, and thus we can assume $h_i = f_{ij}$ for some $j \in \{1, \ldots, \log m\}$. The above condition (2) requires that f_{ij} contains all elements from $X_{(l)}$. Especially, remark that $x_{j,l(j)} \in X_{(l)}$ is included in f_{ij} for the above j. By the definition of f_{ij}, the element $x_{j,l(j)}$ is in $X_j^*(E_i)$. Thus the clause E_i satisfies either (a) or (b):

(a) E_i contains a positive literal v_{jt} with $digit(l(j), 2, t) = 1$,

(b) E_i contains a negative literal $\neg v_{jt}$ with $digit(l(j), 2, t) = 0$.

By the definition of σ_S, we see $\sigma_S(v_{jt}) = 1$ in case of (a), and $\sigma_S(v_{jt}) = 0$ in case of (b). In each case, $\sigma_S(E_i) = 1$. Therefore, σ_S satisfies every disjunction E_i in Φ. Thus Φ is satisfiable. □

4. Complexity of the VC*-dimension problem

In this section, we introduce a natural generalization of the VC-dimension, which is defined for a class of $\{0, \ldots, N\}$-valued functions. Then we show that the generalized VC-dimension problem is still in $\mathrm{SAT}_{\log^2 n}$ and $\mathrm{SAT}^{\mathrm{CNF}}_{\log^2 n}$-hard, as well as the original

VC-dimensional problem. Therefore we may interpret that the generalized VC-dimension problem is also "complete" for the class of $n^{O(\log n)}$ time computable sets.

Let $\mathbb{N}$ be the set of natural numbers. For a class $\mathscr{F}$ of functions from U to $\mathbb{N}$, we define $range(\mathscr{F}) = \bigcup_{f \in \mathscr{F}} \{f(x) \mid x \in U\}$. We represent $\mathscr{F}$ by a $|U| \times |\mathscr{F}|$ matrix M with $M_{ij} = f_j(x_i)$. For an integer matrix M, let $\mathscr{F}_M$ denote the class of functions which M represents.

The following definition seems to be one of the most natural extensions of the VC-dimension for a class of $\{0, \ldots, N\}$-valued functions.

Definition 4. Let $\mathscr{F}$ be a class of functions over U. We say that $\mathscr{F}$ *shatters* a set $S \subseteq U$ if for every function g from S to $range(\mathscr{F})$, there exists a function $f \in \mathscr{F}$ such that $f(x) = g(x)$ for all $x \in S$. The *VC*-dimension* of $\mathscr{F}$, denoted by VC*-dim$(\mathscr{F})$, is the maximum cardinality of a set which is shattered by $\mathscr{F}$.

We can easily see that VC*-dim$(\mathscr{F})$ coincides with VC-dim$(\mathscr{F})$ for any class $\mathscr{F}$ of functions with $range(\mathscr{F}) = \{0, 1\}$. It is not hard to verify the following inequality, that is a generalization of Lemma 1.

Lemma 3. *For any class $\mathscr{F}$ of functions,*

$$\text{VC*-dim}(\mathscr{F}) \leqslant \frac{\log |\mathscr{F}|}{\log |range(\mathscr{F})|}.$$

Definition 5. The *VC*-dimension problem* is, given an integer matrix M and integer $K \geqslant 1$, to determine whether VC*-dim$(\mathscr{F}_M) \geqslant K$ or not.

We now show that the VC*-dimension problem is polynomial-time reducible to the satisfiability problem of a boolean formula of length J with $O(\log^2 J)$ variables. The next theorem includes Theorem 1 as a special case.

Theorem 3. *The VC*-dimension problem is in* $\text{SAT}_{\log^2 n}$.

Proof. Let M be an $m \times r$ matrix and K be an integer, and let $N = |range(\mathscr{F})|$. We denote the elements in $range(\mathscr{F}_M)$ by $y_1, y_2, \ldots, y_N$. By Lemma 3, we can assume that $K \leqslant \log r / \log N$ without loss of generality. Moreover, we can also assume that $m = 2^l$ for some integer l; if $m < 2^l$ for $l = \lceil \log m \rceil$, then we enlarge M by duplicating the last row of M until the row size reaches 2^l. It is easy to see that the size of the enlarged matrix M' is less than twice as large as that of the original matrix M, and VC*-dim$(\mathscr{F}_{M'}) =$ VC*-dim$(\mathscr{F}_M)$.

We now construct a boolean formula Φ_M which contains $K \cdot l$ variables v_{kt} $(1 \leqslant k \leqslant K,\ 1 \leqslant t \leqslant l)$ as follows:

$$\Phi_M = \bigwedge_{s=1}^{N^K} \bigvee_{j=1}^{r} \beta_{sj},$$

$$\beta_{sj} = \bigwedge_{k=1}^{K} \alpha_{kj,digit(s,N,k)+1} \quad (1 \leqslant s \leqslant N^K,\ 1 \leqslant j \leqslant r),$$

$$\alpha_{kjq} = \bigvee_{i \in \{i \mid M_{ij} = y_q\}} \chi_{ki} \quad (1 \leqslant k \leqslant K,\ 1 \leqslant j \leqslant r,\ 1 \leqslant q \leqslant N),$$

$$\chi_{ki} = \bigwedge_{t=1}^{l} [v_{kt};\ digit(i,2,t) = 1] \quad (1 \leqslant k \leqslant K,\ 1 \leqslant i \leqslant m).$$

Note that the length of Φ_M is

$$\begin{aligned} |\Phi_M| &\leqslant l \cdot m \cdot K \cdot r \cdot N^K \\ &\leqslant \log m \cdot m \cdot \log r \cdot r \cdot r \quad (\text{since } N^K \leqslant r) \\ &< n^2 \log^2 n, \end{aligned}$$

where $n = m \cdot r$ is the size of the given matrix M. Also note that Φ_M can be constructed in polynomial time with respect to n.

Let $U = \{x_1, x_2, \ldots, x_m\}$ be the learning domain and $\mathscr{F}_M = \{f_1, f_2, \ldots, f_r\}$ be the class of functions which M represents. We will show that the formula Φ_M is satisfiable if and only if $\mathscr{F}_M$ shatters a set $S \subseteq U$ of cardinality K.

For each assignment σ, we define a set $S_\sigma \subseteq U$ as follows:

$$S_\sigma = \{x_{\langle \sigma,k \rangle} \mid 1 \leqslant k \leqslant K\}, \quad \text{where } \langle \sigma, k \rangle = \sum_{t=1}^{l} 2^{t-1} \cdot \sigma(v_{kt}) + 1.$$

It should be noted that the cardinality of S_σ is not always equal to K, since there may be two distinct k_1 and k_2 with $\langle \sigma, k_1 \rangle = \langle \sigma, k_2 \rangle$ in general.

We now show through a sequence of equivalences that an assignment σ satisfies Φ_M if and only if $|S_\sigma| = K$ and S_σ is shattered by $\mathscr{F}_M$.

First, for any $k \in \{1, \ldots, K\}$ and any $i \in \{1, \ldots, m\}$,

$$\begin{aligned} \sigma(\chi_{ki}) = 1 &\Leftrightarrow \sigma([v_{kt}; digit(i,2,t) = 1]) = 1 \ \text{ for each } t \in \{1, \ldots, l\} \\ &\Leftrightarrow \sigma(v_{kt}) = \begin{cases} 1 & \text{if } digit(i,2,t) = 1 \\ 0 & \text{if } digit(i,2,t) = 0 \end{cases} \quad \text{for each } t \in \{1, \ldots, l\} \\ &\Leftrightarrow digit(i,2,t) = \sigma(v_{kt}) \quad \text{for each } t \in \{1, \ldots, l\} \\ &\Leftrightarrow \sum_{t=1}^{l} 2^{t-1} \cdot digit(i,2,t) = \sum_{t=1}^{l} 2^{t-1} \cdot \sigma(v_{kt}) \\ &\Leftrightarrow i = \langle \sigma, k \rangle. \end{aligned}$$

Next, for any $k \in \{1, \ldots, K\}$, any $j \in \{1, \ldots, r\}$, and any $q \in \{1, \ldots, N\}$,

$$\begin{aligned} \sigma(\alpha_{kjq}) = 1 &\Leftrightarrow \sigma(\chi_{ki}) = 1 \text{ and } M_{ij} = y_q \quad \text{for some } i \in \{1, \ldots, m\} \\ &\Leftrightarrow i = \langle \sigma, k \rangle \text{ and } f_j(x_i) = y_q \\ &\Leftrightarrow f_j(x_{\langle \sigma,k \rangle}) = y_q. \end{aligned}$$

For every integer $s \in \{1, \ldots, N^K\}$, let g_s be the function from S_σ to $range(\mathscr{F}_M)$ such that $g_s(x_{\langle\sigma,k\rangle}) = y_{digit(s,N,k)+1}$ for each $k \in \{1, \ldots, K\}$. Then, for any $s \in \{1, \ldots, N^K\}$ and any $j \in \{1, \ldots, r\}$,

$$\sigma(\beta_{sj}) = 1 \Leftrightarrow \sigma(\alpha_{kj,\,digit(s,N,k)+1}) = 1 \quad \text{for each } k \in \{1, \ldots, K\}$$

$$\Leftrightarrow f_j(x_{\langle\sigma,k\rangle}) = y_{digit\,(s,N,k)+1} \quad \text{for each } k \in \{1, \ldots, K\}, \text{ and}$$

$$digit(s,N,k_1) \neq digit(s,N,k_2) \text{ implies } \langle\sigma,k_1\rangle \neq \langle\sigma,k_2\rangle$$

$$\text{for any } k_1, k_2 \in \{1, \ldots, K\}.$$

$$\Leftrightarrow f_j(x_{\langle\sigma,k\rangle}) = g_s(x_{\langle\sigma,k\rangle}) \quad \text{for all } k \in \{1, \ldots, K\}, \text{ and}$$

$$digit(s,N,k_1) \neq digit(s,N,k_2) \text{ implies } \langle\sigma,k_1\rangle \neq \langle\sigma,k_2\rangle$$

$$\text{for any } k_1, k_2 \in \{1, \ldots, K\}.$$

$$\Leftrightarrow f_j(x) = g_s(x) \quad \text{for all } x \in S_\sigma, \text{ and}$$

$$digit(s,N,k_1) \neq digit(s,N,k_2) \text{ implies } \langle\sigma,k_1\rangle \neq \langle\sigma,k_2\rangle$$

$$\text{for any } k_1, k_2 \in \{1, \ldots, K\}.$$

Finally, we get the following equivalence:

$$\sigma(\Phi_M) = 1 \Leftrightarrow \sigma\left(\bigvee_{j=1}^{r} \beta_{sj}\right) = 1 \quad \text{for any } s \in \{1, \ldots, N^K\}$$

$$\Leftrightarrow \text{for each } s \in \{1, \ldots, N^K\},$$

$$\text{there exists } f_j \in \mathscr{F}_M \text{ with } g_s(x) = f_j(x) \text{ for all } x \in S_\sigma \text{ and}$$

$$digit(s,N,k_1) \neq digit(s,N,k_2) \text{ implies } \langle\sigma,k_1\rangle \neq \langle\sigma,k_2\rangle$$

$$\Leftrightarrow k_1 \neq k_2 \text{ implies } \langle\sigma,k_1\rangle \neq \langle\sigma,k_2\rangle, \text{ and}$$

$$\text{for every function } g \text{ from } S_\sigma \text{ to } range(\mathscr{F}), \text{ there exists } f_j \in \mathscr{F}_M$$

$$\text{such that } f_j(x) = g(x) \text{ for all } x \in S_\sigma$$

$$\Leftrightarrow |S_\sigma| = K \text{ and } S_\sigma \text{ is shattered by } \mathscr{F}_M$$

Thus the formula Φ_M is satisfiable if and only if VC-dim $(\mathscr{F}_M) \geq K$. □

The next theorem shows that the VC*-dimension problem is $\mathrm{SAT}^{\mathrm{CNF}}_{\log^2 n}$-hard.

Theorem 4. *Every* $L \in \mathrm{SAT}^{\mathrm{CNF}}_{\log^2 n}$ *is polynomial-time reducible to the VC*-dimension problem.*

Proof. The generalized VC-dimension problem includes the original VC-dimension problem in which an instance is restricted to $\{0,1\}$-valued matrix. Therefore, Theorem 2 immediately derives this theorem. □

5. Complexity of the Ψ-dimension problems

The VC*-dimension introduced in the previous section seems to be one of the most natural extension of the VC-dimension to the class of $\{0,\ldots,N\}$-valued functions. However, it has not been used actually in the literatures. The reason is that the cardinality of the largest class $\mathscr{F}$ of functions over U of a given dimension grows exponentially in $|U|$ for all $|range(\mathscr{F})|>2$ [1,2], whereas polynomial growth is desirable for the PAC-learning model. As alternative definitions, a variety of notions of dimension to classes of $\{0,\ldots,N\}$-valued functions had been proposed [4,7], and Ben-David et al. [2] gave a general scheme which unified them. They introduced Ψ-dimension, where Ψ is a family of mappings which translate $\{0,\ldots,N\}$-valued functions into $\{0,1\}$-valued ones. In this section, we investigate the complexity of computing Ψ-dimension for two special families Ψ_* and Ψ_G. We show that the Ψ_*-dimension problem is NP-complete, while the Ψ_G-dimension is still in $\mathrm{SAT}_{\log^2 n}$.

Definition 6. Let Ψ be a family of the mappings ψ from $\mathbb{N}$ to $\{0,1,*\}$, where $*$ will be thought of as a null element. Let $\mathscr{F}$ be a class of functions over U. We say that $\mathscr{F}$ *Ψ-shatters* a set $S\subseteq U$ if there exists a mapping $\psi\in\Psi$ which satisfies the following condition: for every subset $T\subseteq S$, there exists a function $f\in\mathscr{F}$ with $\psi(f(x))=1$ for any $x\in T$ and $\psi(f(x))=0$ for any $x\in S-T$. That is, Ψ-shattering requires that under some mapping $\psi\in\Psi$, $\mathscr{F}$ contains all functions from U to $\{0,1\}$. The *Ψ-dimension* of $\mathscr{F}$, denoted by $\Psi\text{-dim}(\mathscr{F})$, is the maximum cardinality of a set which is Ψ-shattered by $\mathscr{F}$.

Remark 1. In [2], they introduced more general notions of Ψ-shatter and Ψ-dimension. Our definition of the Ψ-dimension corresponds to the *uniform Ψ-dimension* they call.

Definition 7. For a family Ψ of mappings from $\mathbb{N}$ to $\{0,1,*\}$, we define the *Ψ-dimension problem* as the decision problem to determine whether $\Psi\text{-dim}(\mathscr{F}_M)\geqslant K$ or not for given integer matrix M and an integer $K\geqslant 1$.

Let Ψ_* be the family of all mappings from $\mathbb{N}$ to $\{0,1*\}$. Therefore, the Ψ_*-dimension problem is the most general one in the family of Ψ-dimension problems.

Theorem 5. *The Ψ_*-dimension problem is* NP-*complete.*

Proof. It is easy to see that the Ψ_*-dimension problem is in NP: guess a set $S\subseteq U$ of size K and a mapping $\psi\in\Psi_*$ nondetereministically, and verify that $\mathscr{F}_M$ Ψ_*-shatters

S by consecutively guessing appropriate f's in $\mathscr{F}_M$. This procedure terminates in polynomial time.

We now give a polynomial-time reduction from 3SAT to the problem. Let $\Phi = E_1 \cdots E_m$ be a formula in 3-CNF with n variables. Without loss of generality, we can assume that m is a power of 2, and variables are indexed as $v_2, v_3, \ldots, v_n, v_{n+1}$. We first construct a matrix M_Φ which has $\log m + 2$ rows and $6m$ columns, and then prove that $\Psi_*\text{-dim}(\mathscr{F}_{M\Phi}) \geqslant \log m + 2$ if and only if Φ is satisfiable.

The learning domain U corresponding to Φ is defined as $U = \{x_0, x_1, y_1, y_2, \ldots, y_{\log m}\}$. The class $\mathscr{F}_{M_\Phi}$ of functions from U to $\{0, 1, \ldots, n+1\}$ is

$$\mathscr{F}_{M_\Phi} = \bigcup_{i=1}^{m} \{f_{i1}, f_{i2}, f_{i3}, g_{i1}, g_{i2}, g_{i3}\},$$

where each function in $\mathscr{F}_{M_\Phi}$ is defined as follows. For each $i \in \{1, \ldots, m\}$ and $u \in \{1, \ldots, \log m\}$,

$$f_{ik}(y_u) = g_{ik}(y_u) = \mathit{digit}(i, 2, u) \quad \text{for all } k \in \{1, 2, 3\},$$

that is, the value of f_{ik} and g_{ik} on y_u corresponds with the uth binary digit of the number $i-1$. For each $i \in \{1, \ldots, m\}$,

$$g_{i1}(x_0) = 0, \qquad g_{i2}(x_0) = 1, \qquad g_{i3}(x_0) = 1,$$

$$g_{i1}(x_1) = 0, \qquad g_{i2}(x_1) = 0, \qquad g_{i3}(x_1) = 1,$$

which are independent of the structure of Φ. On the other hand, the values of f_{i1}, f_{i2} and f_{i3} on x_0 and x_1 reflect the structure of the clause $E_i = (l_{i1} \vee l_{i2} \vee l_{i3})$ in Φ. For each $k \in \{1, 2, 3\}$, if the literal l_{ik} is a negative literal $\neg v_q$, then we define

$$f_{ik}(x_0) = q, \qquad f_{ik}(x_1) = 1.$$

Otherwise (l_{ik} is a positive literal v_q),

$$f_{ik}(x_0) = 0, \qquad f_{ik}(x_1) = q.$$

It should be noted that q ranges from 2 to $n+1$. Fig. 2 illustrates the structure of the matrix M_Φ.

It is easy to see that M_Φ can be constructed in polynomial time with respect to the length of given formula Φ.

We now prove that if the formula Φ is satisfiable then $\Psi_*\text{-dim}(\mathscr{F}) = \log m + 2$. For an assignment σ which satisfies Φ, we consider a mapping $\psi_\sigma \in \Psi_*$ with

$$\psi_\sigma(0) = 0, \qquad \psi_\sigma(1) = 1,$$

$$\psi_\sigma(q) = \sigma(v_q) \quad \text{for each } q \in \{2, \ldots, n+1\}.$$

Let $i \in \{1, \ldots, m\}$ be fixed arbitrarily. Since the clause E_i is satisfied by the assignment σ, E_i contains either positive literal v_q with $\sigma(v_q) = 1$ or negative literal $\neg v_q$ with $\sigma(v_q) = 0$.

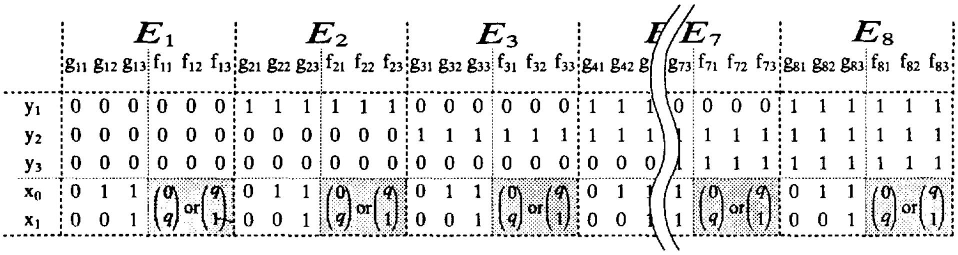

Fig. 2. Structure of the matrix M_Φ reduced from a boolean formula $\Phi = E_1 \wedge E_2 \wedge \cdots \wedge E_8$. The only shadowed elements depend on the structure of each clause E_i in Φ.

In each case, we have $f_{ik} \in \mathscr{F}_{M_\Phi}$ such that

$$\psi_\sigma(f_{ik}(x_0)) = 0, \qquad \psi_\sigma(f_{ik}(x_1)) = 1,$$

for some $k \in \{1,2,3\}$. Thus we can see that, under the mapping ψ_σ, the class $\mathscr{F}_{M_\Phi}$ contains all functions from U to $\{0,1\}$. Therefore, U is Ψ_*-shattered by $\mathscr{F}_{M_\Phi}$.

We now show the converse. Suppose that $\Psi_*\text{-dim}(\mathscr{F}_{M_\Phi}) = \log m + 2$. Then there is a subset of U with cardinality $\log m + 2$ which is Ψ_*-shattered by $\mathscr{F}_{M_\Phi}$. Since $|U| = \log m + 2$, U itself is Ψ_*-shattered by $\mathscr{F}_{M_\Phi}$. That is, there exists a mapping $\psi \in \Psi_*$, under which $\mathscr{F}_{M_\Phi}$ contains all functions from U to $\{0,1\}$. We can assume that $\psi(0) = 0$ and $\psi(1) = 1$ without loss of generality. For such a ψ, we define the assignment σ_ψ with $\sigma_\psi(v_q) = \psi(q)$ for each $q \in \{2, \ldots, n+1\}$. Then we can verify that σ_ψ satisfies every disjunction E_i in Φ. Thus Φ is satisfiable. □

Natarajan [7] introduced the *graph dimension* in order to characterize the learnability of a class of $\{0, \ldots, N\}$-valued functions.

Definition 8. The *graph dimension* is the Ψ_G-dimension, where $\Psi_G = \{\psi_{G,\tau} \mid \tau \in \mathbb{N}\}$, and $\psi_{G,\tau}$ is defined by

$$\psi_{G,\tau}(a) = \begin{cases} 1 & \text{if } a = \tau, \\ 0 & \text{otherwise.} \end{cases}$$

From the definition, Ψ_G is a subset of Ψ_*. The following theorem gives an interesting contrast with the Theorem 5.

Theorem 6. *The Ψ_G-dimension problem is in* $\mathrm{SAT}_{\log^2 n}$.

Proof. The outline and notations follow the proof of Theorem 3. Let M be an $m \times r$ matrix and K be an integer. We should remark that in this time, the inequality

$K \leqslant \log r/\log N$ does not necessarily hold. Instead, we can only assume that $K \leqslant \log r$. We construct a boolean formula Φ'_M as

$$\Phi'_M = \bigvee_{q=1}^{N} \Phi'_q,$$

$$\Phi'_q = \bigwedge_{s=1}^{2^K} \bigvee_{j=1}^{r} \beta'_{sjq} \quad (1 \leqslant q \leqslant N),$$

$$\beta'_{sjq} = \bigwedge_{k=1}^{K} [\alpha_{kjq}; digit(s,2,k)=1] \quad (1 \leqslant s \leqslant 2^K,\ 1 \leqslant j \leqslant r,\ 1 \leqslant q \leqslant N),$$

where α_{kjq} is the same as that in the proof of Theorem 3. Then the length of Φ'_M is

$$|\Phi'_M| \leqslant l \cdot m \cdot K \cdot r \cdot 2^K \cdot N \leqslant n^3 \log^2 n,$$

since $2^K \leqslant r$. We now show that an assignment σ satisfies Φ'_M if and only if $|S_\sigma| = K$ and S_σ is Ψ_G-shattered by $\mathscr{F}_M$.

For every integer $s \in \{1, \ldots, 2^K\}$, let g_s be the function from S_σ to $\{0,1\}$ such that $g_s(x_{\langle\sigma,k\rangle}) = digit(s,2,k)$ for each $k \in \{1, \ldots, K\}$. Then for any $s \in \{1, \ldots, 2^K\}$, any $j \in \{1, \ldots, r\}$, and any $q \in \{1, \ldots, N\}$,

$$\begin{aligned}
\sigma(\beta'_{sjq}) = 1 \iff & \sigma([\alpha_{kjq}; digit(s,2,k)=1]) = 1 \quad \text{for each } k \in \{1, \ldots, K\} \\
\iff & \begin{cases} f_j(x_{\langle\sigma,k\rangle}) = y_q & \text{if } digit(s,2,k)=1 \\ f_j(x_{\langle\sigma,k\rangle}) \neq y_q & \text{if } digit(s,2,k)=0 \end{cases} \quad \text{for each } k \in \{1, \ldots, K\} \\
\iff & \psi_{G,y_q}(f(x_{\langle\sigma,k\rangle})) = digit(s,2,k) \quad \text{for each } k \in \{1, \ldots, K\}, \text{ and} \\
& digit(s,2,k_1) \neq digit(s,2,k_2) \text{ implies } \langle\sigma,k_1\rangle \neq \langle\sigma,k_2\rangle \\
& \text{for any } k_1, k_2 \in \{1, \ldots, K\}. \\
\iff & \psi_{G,y_q}(f_j(x_{\langle\sigma,k\rangle})) = g_s(x_{\langle\sigma,k\rangle}) \quad \text{for all } k \in \{1, \ldots, K\}, \text{ and} \\
& digit(s,2,k_1) \neq digit(s,2,k_2) \text{ implies } \langle\sigma,k_1\rangle \neq \langle\sigma,k_2\rangle \\
& \text{for any } k_1, k_2 \in \{1, \ldots, K\}. \\
\iff & \psi_{G,y_q}(f_j(x)) = g_s(x) \quad \text{for all } x \in S_\sigma, \text{ and} \\
& digit(s,2,k_1) \neq digit(s,2,k_2) \text{ implies } \langle\sigma,k_1\rangle \neq \langle\sigma,k_2\rangle \\
& \text{for any } k_1, k_2 \in \{1, \ldots, K\}.
\end{aligned}$$

For each $q \in \{1, \ldots, N\}$,

$$\begin{aligned}
\sigma(\Phi'_q) = 1 \iff & digit(s,2,k_1) \neq digit(s,2,k_2) \text{ implies } \langle\sigma,k_1\rangle \neq \langle\sigma,k_2\rangle, \text{ and} \\
& \text{for every function } g \text{ from } S_\sigma \text{ to } \{0,1\}, \text{ there exists } f_j \in \mathscr{F}_M \\
& \text{such that } \psi_{G,y_q}(f_j(x)) = g(x) \quad \text{for all } x \in S_\sigma \\
\iff & |S_\sigma| = K \text{ and } S_\sigma \text{ is shattered by } \mathscr{F}_M \text{ under the mapping } \psi_{G,y_q}.
\end{aligned}$$

Thus Ψ'_M is satisfiable if and only if Ψ_G-dim$(\mathscr{F}) \geqslant K$. □

According to the hardness result, the following corollary is immediately derived from the Theorem 2.

Corollary 1. *Every* $L \in \mathrm{SAT}^{\mathrm{CNF}}_{\log^2 n}$ *is polynomial-time reducible to the* Ψ_G*-dimension problem.*

6. Conclusion

We showed that the VC-dimension problem, VC*-dimension problem and Ψ_G-dimension problem are all in $\mathrm{SAT}_{\log^2 n}$ and $\mathrm{SAT}^{\mathrm{CNF}}_{\log^2 n}$-hard, while the Ψ_*-dimension problem is NP-complete. It remains open that these problems are in $\mathrm{SAT}^{\mathrm{CNF}}_{\log^2 n}$, or $\mathrm{SAT}_{\log^2 n}$-hard.

As a dual to the VC-dimension, Romanik [11] defined the *testing dimension* of a concept class $\mathscr{F}$ as the *minimum* cardinality of a set $S \subseteq U$ which is *not* shattered by $\mathscr{F}$. We can see that the testing dimension problem is also in $\mathrm{SAT}_{\log^2 n}$, by a similar reduction in the proof of Theorem 1. It is also open whether the problem is $\mathrm{SAT}^{\mathrm{CNF}}_{\log^2 n}$-hard or not.

As another crucial characterization of the complexity of computing VC-dimension, recently Papadimitriou and Yannakakis [10] defined a new complexity class LOGNP, for which the (original) VC-dimension problem becomes complete. Their results imply the polynomial-time reducibility from the satisfiability problem with $O(\log^2 n)$ variables in conjunctive normal form to the VC-dimenson problem, which we gave explicitly in the proof of Theorem 2.

Acknowledgement

The author would like to thank H. Arimura, S. Shimozono, and T. Uchida for helpful discussions.

References

[1] N. Alon, On the density of sets of vectors, *Discrete Math.* **46** (1983) 199–202.

[2] S. Ben-David, N. Cesa-Bianchi and P.M. Long, Characterizations of learnability for classes of $\{0, \ldots, n\}$-valued functions, in: *Proc. 5th Ann. Workshop on Computational Learning Theory* (1992) 333–340.

[3] A. Blumer, A. Ehrenfeucht, D. Haussler and M.K. Warmuth, Learnability and the Vapnik–Chervonenkis dimension, *J. ACM* **36** (1989) 929–965.

[4] D. Haussler, Decision theoretic generalizations of the PAC model for neural net and other learning applications, *Inform. and Comput.* **100** (1992) 78–150.

[5] N. Linial, Y. Mansour and R.L. Rivest, Results on learnability and the Vapnik–Chervonenkis dimension, *Inform. and Comput.* **90** (1991) 33–49.

[6] N. Megiddo and U. Vishkin, On finding a minimum dominating set in a tournament, *Theoret. Comput. Sci.* **61** (1988) 307–316.
[7] B.K. Natarajan, On learning sets and functions, *Machine Learning* **4**(1) (1989) 67–97.
[8] B.K. Natarajan, *Machine Learning – A Theoretical Approach* (Morgan Kaufmann, Los Altos, CA, 1991).
[9] S.H. Nienhuys-Cheng and M. Polman, Complexity dimensions and learnability, in: *Proc. European Conf. on Machine Learning*, Lecture Notes in Artificial Intelligence, Vol. 667 (Springer, Berlin, 1993) 348–353.
[10] C.H. Papadimitriou and M. Yannakakis, On limited nondeterminism and the complexity of the V-C dimension, in: *Proc. 8th Ann. Conf. on Structure in Complexity Theory* (1993) 12–18.
[11] K. Romanik, Approximate testing and learnability, in: *Proc. 5th Ann. Workshop on Computational Learning Theory* (1992) 327–332.
[12] L.G. Valiant, A theory of the learnable, *Comm. ACM* **27** (1984) 1134–1142.