On One Combinatorial Problem Concerned with the Notion of Minimal Committee[1]

M. Yu. Khachai

Institute of Mathematics and Mechanics, Ural Division, Russian Academy of Sciences,
ul. S. Kovalevskoi 16, Yekaterinburg, 620219 Russia
e-mail: mkhachay@imm.uran.ru

Abstract—A combinatorial problem concerned with the notion of minimal committee is considered. Several new results are presented.

INTRODUCTION

One of the approaches to training in pattern recognition is based on constructing so-called committee decision rules over a given class $\mathcal{F}$ of base rules. The general form of a k-element committee decision rule is

$$f(\alpha_1, \alpha_2, \ldots, \alpha_k; x) = \Theta\left(\sum_{i=1}^{\kappa} f(\alpha_i; x) - \frac{k}{2}\right). \qquad (1)$$

where, $f(\alpha_1; \cdot), f(\alpha_2; \cdot), \ldots, f(\alpha_k; \cdot)$ are elements of the base class $\mathcal{F} = \{f(\alpha; \cdot): X \longrightarrow \{0, 1\} | \alpha \in \Lambda\}$, $\Theta(t) = \begin{cases} 1, & t > 0 \\ 0, & t \le 0 \end{cases}$, and k is a fixed odd number. In the framework of this approach, a necessary condition for a "trained" rule is that this rule correctly classifies training data represented as the finite sequence

$$((x^1, \omega^1), (x^2, \omega^2), \ldots, (x^l, \omega^l)) \qquad (2)$$

of descriptions of objects from the first ($\omega = 1$) and second ($\omega = 0$) classes (patterns). Mathematically, rule (1) correctly classifies sample (2) if and only if the sequence $(\alpha_1, \alpha_2, \ldots, \alpha_k)$ of parameter values is a majority committee for the system of equations

$$f(\alpha, x^i) = \omega^i, \quad i = 1, 2, \ldots, l. \qquad (3)$$

As usual (see, e.g., [1]), by a majority committee of system (3), we understand a finite sequence $Q = (\alpha_1, \alpha_2, \ldots, \alpha_q)$ such that $|\{j: f(\alpha_j, x^i) = \omega^i\}| > \frac{q}{2}$ for each $i \in \mathbb{N}_l$ (here and in what follows, $\mathbb{N}_l$ denotes the set $\{1, 2, \ldots, l\}$). The capacity of the class of committee decision rules depends on the maximum (over the class) number of committee members and increases rapidly with this maximum number [2]; for this reason, it is important to be able to find a committee decision with minimum number of elements for a system given by Eq. (3).

In this work, one combinatorial-type problem arising in searching for a minimal committee of an arbitrary system of constraints is considered.

STATEMENT OF THE PROBLEM

Consider a more general situation. Suppose the following system of inclusions is given:

$$x \in D_i \quad (i \in \mathbb{N}_l), \qquad (4)$$

where $D_1, D_2, \ldots, D_l$ are subsets of the domain of the variable x. As is known (see, e.g., [1]), in searching for a minimal committee of this system, it suffices to consider only committees consisting of solutions to maximal consistent subsystems of this system. Let $J_1, J_2, \ldots, J_k$ be the index sets of all maximal consistent subsystems of system (4). Identifying the elements of the committees that are solutions to the same subsystem, we put each committee A in a correspondence with the sequence $A_Q = (a_1, a_2, \ldots, a_k)$, where $a_i \in Z_+$ is the number of elements in Q that are solutions to subsystem J_i. Obviously, the sum $\sum_{i=1}^{k} a_i$ coincides with the number of members in the committee Q. Let us endow $\{A_Q\}$ with the partial order defined as follows.

Definition 1. Let $A = (a_1, a_2, \ldots, a_k)$ and $B = (b_1, b_2, \ldots, b_k)$ be the finite sequences of nonnegative integers. We say that the inequality $A \ge B$ holds if

$$\sum_{i=1}^{k} a_i = q \ge p = \sum_{i=1}^{k} b_i \qquad (5)$$

and

$$(J \subseteq \mathbb{N}_k) \quad \left(\sum_{i \in J} a_i > \frac{q}{2}\right) \Leftrightarrow \left(\sum_{i \in J} b_i > \frac{p}{2}\right) \qquad (6)$$

[1] This work was financially supported by the Russian Foundation for Basic Research, project nos. 00-15-96041 and 99-01-00136.

Received July 30, 1999

Note that, if the sequences A and B are such that $A \geq B$, then, for an arbitrary system of form (4), the existence of a committee with "weights" from A is equivalent to the existence of a committee with "weights" from B; the only difference is that the number of elements in the latter may be smaller.

Assertion. *Suppose that the sequences* $A = (a_1, a_2, \ldots, a_k)$ and $C = (c_1, c_2, \ldots, c_k)$ are given such that $a_1 \geq a_2 \geq \ldots \geq a_k$, $\sum_{i=1}^{k} a_i = q$, $\sum_{i=1}^{k} c_i = p$, and

$$(J \subseteq \mathbb{N}_k) \quad \left(\sum_{i \in J} a_i > \frac{q}{2}\right) \Leftrightarrow \left(\sum_{i \in J} c_i > \frac{p}{2}\right).$$

Then, there exists a sequence $B = (b_1, b_2, \ldots, b_k)$ with $\sum_{i=1}^{k} b_i = p$ such that it satisfies Eq. (6) and $b_1 \geq b_2 \geq \ldots \geq b_k$.

This assertion allows us to consider only monotonically decreasing sequences of positive integers, which are called partitions, and extend the definition of partial order given above to sequences of different lengths.

Definition 2. A sequence A is called committee-minimal if, for any sequence B, either A and B are incomparable or $B \geq A$.

Clearly, if some committee Q is minimal for a system of form (4), then, the corresponding weight sequence A_Q is committee-minimal. On the other hand, each committee-minimal sequence can be assigned a suitable system of constraints and its minimal committee such that the weight sequence of this committee coincides with the given sequence.

Theorem. *Let A be a committee-minimal sequence. Then there exists a system of inclusions of form* (4) *such that one of its minimal committees has weight sequence A.*

Thus, by enumerating all committee-minimal sequences we approach to the solution of the problem of minimal committee. Let us give the combinatorial statement of this problem.

THE COMBINATORIAL STATEMENT OF THE PROBLEM

Let us denote the set of all partitions of an odd positive integer q by $\Lambda(q)$ and the set of partitions of q into s terms, by $\Lambda(q, s)$. On the set of all partitions of odd numbers, we introduce a binary relation κ as follows. Let $A \in \Lambda(q, s)$ and $B \in \Lambda(p, t)$. Without loss of generality, we can assume that $s \geq t$. We say that A and B are in the relation κ ($A\kappa B$) if, for any $J \subseteq \mathbb{N}_s$, the inequality $\sum_{i \in J} a_i > \frac{q}{2}$ is equivalent to $\sum_{i \in J \cap \mathbb{N}_t} b_i > \frac{p}{2}$. The relation introduced is obviously an equivalence relation.

Definition 3. A partition $A \in \Lambda(q)$ is called κ-minimal if

$$(B \in \Lambda(p)) \quad (A\kappa B) \Rightarrow (p \geq q).$$

Problem. Given an odd number q, enumerate all κ-minimal elements in the set $\Lambda(q)$.

The problem stated resembles classical enumeration problems in the theory of partitions. The main difficulty involved in this problem is that the condition on the terms essentially depends on q.

CONCLUSION

In the report, some assertions concerning conditions for a given partition A to be κ-minimal and the problem of "reducing the order" of a partition A (i.e., finding a κ-minimal partition B such that $A\kappa B$) are discussed.

REFERENCES

1. Mazurov, Vl.D. and Khachai, M.Yu., Committee Constructions, *Izv. Ural. Gos. Univ., Ser. Mat. Mech.,* 1999, vol. 2, no. 14, pp. 77–108.
2. Khachai, M.Yu., On a sufficient length of a training sample for a committee decision rule, *Artificial Intelligence*, 2000, no. 2, pp. 219–223.

SPELL: OK