Mathematical Problems of Computer Science 22, 2001, 113–117.

On One Problem of Finding Optimal Test for Permutations

Vahagn R.Tonoyan

Abstract

The article defines the concept of d-test for a set. For the set of permutations $\pi:\{1,2,...,n\}\to\{1,2,...,n\}$ it constructs an optimal d-test for $\pi(i)=j$ $(1\le i,j\le n)$ type of set of questions, when $\alpha=0$ or 1.

Let's define some concepts used in the problem discussed below ([1]-[3].)

We will define the concept of a 2-tree $T_u=(V,X)$ through induction.

1. $T_u=(\{u\},\varnothing)$ is a 2-tree. It has a marked vertex u, which will be called leaf.
2. If $T_{u_1}=(V_1,X_1)$ and $T_{u_2}=(V_2,X_2)$ are two distinct 2-trees without common vertex, then the tree $T_u=(V,X)$, $V=\{u\}\cup V_1\cup V_2$, $X=X_1\cup X_2\cup\{uu_1\}\cup\{u_1u_2\}$, obtained through adding new vertex u and sides $\{u,u_1\}$, $\{u_1,u_2\}$, is a 2-tree.

The vertex u will be considered a marked vertex of a tree – root, and u_1 and u_2 will be correspondingly considered as left and right descendants of the root.

If $T_u=(V,X)$ is a leaf of a 2-tree, then it is clear that there is only one sequence $u=u_0,u_1,u_2,...,u_k=v$, so that $\{u_i,u_{i+1}\}\in X$. We will call that sequence a branch, connecting root u to leaf v. The number k will be called the length of the branch.

We will call the length of the longest branch of a 2-tree $T_u=(V,X)$ height.

Suppose M is some finite set and the distance $d(m,n)$ is defined between the elements of that set, which satisfies to the axioms of metrical space.

We will call the reflection $h:M\to\{0,1\}$ a question regarding the elements of M.

The set of questions of M is a set of different from each other questions $H=\{h_1,h_2,...,h_s\}$, through which it is possible to separate the elements of M. For any elements $m_1,m_2\in M$ there is $h\in H$, so that $h(m_1)\ne h(m_2)$.

Suppose the set H of questions and the 2-tree $T_u=(V,X)$ are given and a rule φ is indicated, which corresponds each inner vertex u of the 2-tree to a question $h_u\in H$, in a way, that the questions corresponding to vertices located on the same branch are different from each other. In that case we will say that a search tree $(T_u;H;\varphi)$ is given in the H set of questions.

Suppose v is some leaf of the search tree $(T_u;H;\varphi)$ and $u=u_0,u_1,u_2,...,u_k=v$ is the branch connecting the root to the leaf.

We will say that the element $m\in M$ corresponds to the leaf u_k, if for the values $i=0,1,2,...,k-1$ the following conditions are satisfied:

if the vertex u_{i+1} is the left descendant of the vertex u_i, then $h_{u_i}(m)=0$, if the vertex u_{l+1} is the right descendant of the vertex u_l, then $h_{u_l}(m)=1$.

It is obvious that through the above-mentioned rule each element $m \in M$ will correspond to some leaf and each leaf u will correspond to some subset (maybe empty) $M(u)$.

Let's fix the number $d \geq 0$.

We will say that the search tree $(T_n;H;\varphi)$ is a d-test for the set M, if for any leaf u of that set the following condition is true:

If $m_1 \in M(u)$ and $m_2 \in M(u)$ $(m_1 \neq m_2)$, then $d(m_1,m_2) \leq d$.

This is the same as the search tree is d-test for the set M, if the elements $M(u)$ corresponding to any leaf u are within the sphere of radius d in the metrical space M.

We will call the height of the 2-tree corresponding to the d-test of the set M the complexity of the test. d-test, having the smallest complexity, will be called optimal d-test.

The problem is the following:

For the given set M and H set of questions, it is necessary to find the optimal d-test.

It is known, that even in the case $d=0$ the above-mentioned problem is NP-complete. Thus, it is natural to discuss the possibility of solving the problem in particular cases.

Let's observe the set of all permutations $\pi=\{1,2,...,n\} \to \{1,2,...,n\}$ of the set $\{1,2,...,n\}$ - S_n.

Let's define the distance $d(\pi 1,\pi 2)=\max\limits_{1\leq i\leq n}|\pi_1(i)-\pi_2(i)|$ of permutations $\pi_1,\pi_2 \in S_n$ and $H=\{h_{ij}/1\leq i,j\leq n\}$ set of questions, where $h_{ij}^{(\pi)}=1$, if $\pi(i)=j$ and $h_{ij}(\pi)=0$, if $\pi(i)\neq j$.

We will discuss the problem of finding optimal d-test for the set S_n.

It has the following meaning. Suppose an unknown permutation $\pi \in S_n$ is given. It is allowed to ask $\pi(i)=j$ types of questions and obtain answers. It is necessary to find the given permutation $\pi \in S_n$ (0-test) or at least find some $\pi' \in S_n$, so that $d(\pi,\pi')\leq d$ (d-test), through smallest possible number of questions.

The complexity of the optimal d-test in S_n set of questions will be denoted by $g(n,d)$.

The following statements are proven in the work.

Theorem 1: $g(n,0)=\frac{n(n-1)}{2}$.

Theorem 2: $g(n,1)=\left\lceil \frac{n^2}{2} \right\rceil - n$, where $\lceil a \rceil$ is the smallest natural number, to which a does not exceed.

Proof of the theorem 1.

Let's discuss the algorithm of finding the permutation $\pi \in S_n$. We find the value of $\pi(1)$ through questions $\pi(1)=1,\pi(2)=2,...$ (asking $n-1$ questions in the worst case). Then, using the same method we sequentially decide the values $\pi(2),\pi(3),...,\pi(n-1)$ ($\pi(n)$ is decided exclusively using the answers from previous questions).

The complexity of the mentioned algorithm is $(n-1)+(n-2)+...+2+1$.

As $g(n,0)$ is the complexity of the optimal algorithm finding the permutation $\pi \in S_n$, then $g(n,0)\leq \frac{n(n-1)}{2}$.

In order to prove the inequality $g(n,0) \geq \frac{n(n-1)}{2}$, we will compose "guesser" – rules of selecting branches in the search tree. It will indicate a branch in any search tree, solving the problem, which will have a length of at least $\frac{n(n-1)}{2}$.

Let's define the following strategy:

In the first step the "guesser" observes the two-sided graph $G_1 = (V, X_1)$, where

$V = \{1,2,...,n\} \cup \{l(1), l(2),...,l(n)\}$,

$X_1 = \{(i, l(j)) / 1 \leq i \leq n, 1 \leq j \leq n\}$.

In the step k it observes the graph $G_k = (V, X_k)$ $(k = 1,2,3,...)$ and in the case of asking question $\pi(i_k) = j_k$, selects:

a) the branch 0, if there is a matching in the graph G_k, which does not contain the side $(i_k, l(j_k))$ and composes the graph $G_{k+1} = (V, X_{k+1})$, where $X_{k+1} = X_k \setminus \{(i_k, l(j_k))\}$,

b) the branch 1, if any matching of the graph G_k contains the side $(i_k, l(j_k))$ and $G_{k+1} = G_k$.

It is clear that the "guesser" with such strategy indicates a branch, connecting root to some leaf in each search tree finding 0-test in the set S_n, and the length of the branch is at least $n^2 - z$. Here z is the number of sides of the graph G^* answering to the last question.

Let's note that the graph G^* has the only matching $\{1, l(i_1)\}, \{2, l(i_2)\},...,\{n, l(i_n)\}$ and the permutation $\pi(k) = l(i_k)$, $k = 1,2,...,n$ is the one that corresponds to the case selected by the "guesser".

We will denote the maximum number of sides in above-mentioned types of graphs by z_n.

It is easy to check that in the graph G^*, containing the only matching, there is a vertex, which has a degree of 1. Suppose $\{j, l(j)\}$ is the side of the matching passing through that vertex. If the vertices $(j, l(j))$ and the sides passing through that vertices are removed from the graph, whose number does not exceed n, then the new graph also contains the only matching.

Thus, $z_n \leq n + z_{n-1}$, $z_1 = 1$ and therefore $g(n,0) \geq \frac{n(n-1)}{2}$.

Proof of the theorem 2.

Let's observe the following method of composing 1-test for the set S_n.

We find the approximate value of $\pi(1)$ (by accuracy of 1) through sequential questions $\pi(1) = 1, \pi(1) = 2,...$. Then, using the same method, we decide the values $\pi(2), \pi(3),...$. It is easy to see, that the number of questions will be the maximum (let's denote by $h(n)$), when $n-2$ questions are asked for finding each of the values $\pi(1)$ and $\pi(2)$, and therefore $h(n) = h(n-2) + 2n - 4$, $n \geq 4$, and $h(2) = 0$, $h(3) = 2$.

Through the method of induction it is proved that $h(n) = \left\lceil \frac{n^2}{2} \right\rceil - n$, and therefore $g(n,1) \leq \left\lceil \frac{n^2}{2} \right\rceil - n$.

Let's show that the "guesser" composed for proving the previous theorem finds a branch in the search tree corresponding to each 1-test of the set S_n, which has a length of at least h_n.

Suppose G^* is the graph correspondingly answering to the last question selected by the "guesser".

Let's note that the graph G^* satisfies to the following conditions:

1. It is a two-sided graph with $2n$ vertices and it contains a matching.
2. That matching is the only one or, as the same, for any two matchings $\{1,l(i_1)\},\{2,l(i_2)\},...,\{n,l(i_n)\}$ and $\{1,l(j_1)\},\{2,l(j_2)\},...,\{n,l(j_n)\}$ the condition $|i_k - j_k| \le 1$, $k=1,2,...,n$ is met.

Suppose $G_0 = (V, X_0^*)$ is a graph satisfying the conditions 1 and 2, and having maximum number of sides. If $|x^*| = y_n$, then it is clear that the length of the branch selected by the "guesser" is at least $n^2 - y_n$.

Let's observe some matching $\{k_1,l(1)\},\{k_2(l2)\},...,\{k_n,l(n)\}$ of the graph $G_0 = (V, X_0)$. There are two possibilities:

b) Each matching of the graph G_0 contains the side $\{k_1,l(1)\}$. In that case the number of vertices of the graph G_0, that are neighboring to one of the vertices k_1, $l(1)$ and are different from each other, does not exceed $(n-1)$ and, therefore, $y_n \le n + y_{n-1}$.

b) There is a matching $\{u_1,l(1)\},\{u_2,l(2)\},...,\{u_n,l(n)\}$ in the graph G_0 so that $u_1 \ne k_1$. It is easy to check that $u_2 = k_1$ and $u_1 = k_2$. It can be drawn from the condition 2 that the number of vertices of the graph G_0 that are neighboring to any of the vertices k_1, k_2, $l(1)$ and $l(2)$ and are different from each other, does not exceed $(2n-4)$, and therefore $y_n \le 2n + y_{n-2}$.

Using the obtained recurrences, it can be checked that the number of sides of the graph G_0 is $\left\lfloor \frac{n^2}{2} \right\rfloor + n$ (here $\lfloor a \rfloor$ is the largest natural number not exceeding a).

Thus, actually $g(n,1) = \left\lceil \frac{n^2}{2} \right\rceil - n$.

References

[1] F. Harary, *Graph Theory*, Addison-Wesley Publishing Company, 1969.

[2] D. E. Knuth, "The Art of Computer Programming", vol. 1, Fundamental Algorithms, Addison-Wesley Publishing Company, 1968.

[3] E. Reingold, J. Nievergelt, D. Narsing, *Combinatorial Algorithms – Theory and Practice*, Prentice-Hall, Inc., New Jersey, 1977.

Տեղադրությունների համար օպտիմալ թեստ գտնելու մի խնդրի մասին

Վ.Տոնոյան

Ամփոփագիր

Աշխատանքում սահմանվում է բազմության d թեստի գաղափարը: Տեղադրությունների $\pi:\{1,2,...,n\}\rightarrow\{1,2,...,n\}$ բազմության համար $\pi(i)=j$ $1\leq i,j\leq n$ հարցաշարում կառուցված է լավագուն d թեստ, երբ $\alpha=0$ կամ 1: