AN n-DIMENSIONAL SEARCH PROBLEM WITH RESTRICTED QUESTIONS

Ervin GYŐRI

Mathematical Institute of the
Hungarian Academy of Sciences
Budapest, Hungary H-1053

Received 2 February 1981

The problem is the following: How many questions are necessary in the worst case to determine whether a point X in the n-dimensional Euclidean space $\mathbf{R}^n$ belongs to the n-dimensional unit cube Q^n, where we are allowed to ask which halfspaces of $(n-1)$-dimensional hyperplanes contain the point X? It is known that $\lceil 3n/2 \rceil$ questions are sufficient. We prove here that cn questions are necessary, where $c \approx 1.2938$ is the solution of the equation $x \log_2 x-(x-1) \log_2 (x-1)=1$.

Let $X=(x_1, x_2, ..., x_n)$ be an arbitrary point in the n-dimensional Euclidean space $\mathbf{R}^n$ and let Q^n denote the n-dimensional unit cube, that is,

$$Q^n := \{(y_1, y_2, ..., y_n)\colon\ 0 \leqq y_i \leqq 1;\ i=1, 2, ..., n\}.$$

Andrew C. Yao [2, Section 7] proposed the following

Problem. How many questions are necessary in the worst case to find out whether $X \in Q^n$ if we are allowed to ask which halfspaces of $(n-1)$-dimensional hyperplanes P_i contains the point X? (We can choose the i-th hyperplane depending on the previous answers.)

Let $f(n)$ denote the minimum number of questions necessary in the worst case, i.e.

$$f(n) := \min_{\text{srategies}} \max_{x \in \mathbf{R}^n} (\#\ \textit{questions}).$$

We have $n+1 \leqq f(n) \leqq 2n$ because the intersection of the halfspaces containing X has to be bounded at the end of the algorithm and because asking about the $2n$ hyperplanes of the $(n-1)$-dimensional lateral faces of Q^n is obviously sufficient in order to determine whether $X \in Q^n$ or not.

Notice that this problem is a generalization of the following well-known and completely solved problem: Given n distinct real numbers how many comparisons are necessary to find their maximum and minimum. In other words, let $X=(x_1, x_2, ..., x_n)$ be a point in the n-dimensional Euclidean space $\mathbf{R}^n$. If we are allowed to ask which halfspace of the $(n-1)$-dimensional hyperplanes of type $x_j-x_k=0$ contains the point X, how many questions are necessary to find the maximum and minimum coordinates of X.

AMS subject classification (1980): 05 A 05; 05 A 99

If this second problem can be solved by $g(n)$ questions and x_{max} and x_{min} are the desired coordinates then by asking about the hyperplanes $x_{min}=0$, $x_{max}=1$ we determine whether $X\in Q^n$ in $g(n)+2$ questions. I. Pohl [1] proved that $\lceil 3n/2\rceil-2$ comparisons are necessary to find the maximum and minimum numbers in the worst case. Obviously $\lceil 3n/2\rceil-2$ comparisons are sufficient: find the maximum of each of the pairs (x_1, x_2), (x_3, x_4), ... then find the maximum of the set of the greater elements (completed with x_n if n is odd) and find the minimum of the set of the smaller elements (completed with x_n if n is odd). (To find the maximum (resp. minimum) of a set of k elements we have to perform $k-1$ comparisons obviously.) Thus $f(n)\leqq\lceil 3n/2\rceil$ and the conjecture is that $f(n)=\lceil 3n/2\rceil$ i.e. we need question $\lceil 3n/2\rceil$ hyperplanes even if they can be arbitrary ones. We prove that $f(n)\geqq cn+O(\log n)$, where $c\approx 1.2938$ is the solution of the equation $x\log_2 x-(x-1)\log_2(x-1)=1$, even if the halfspaces are open.

We give an adversary strategy. This means that we do not fix the point X, we merely choose the open halfspaces of each hyperplane asked about so that the intersection of the chosen halfspaces should not be empty.

Adversary strategy. Let P_i denote the i-th hyperplane asked about and let H_i denote the chosen open halfspace (that should contain the point X) and let $\bar{H}_i$ denote the other open halfspace of P_i. Let $V=\{V_1, V_2, ..., V_{2^n}\}$ denote the set of the vertices of Q^n. For a vertex V_j let $P_i(V_j)$ denote the set of the hyperplanes P_k $(k\leqq i)$ such that $V_j\in P_k$. We define a sequence of weight functions ω_i from V to $\mathbf{N}$. Let

$$\omega_0(V_j)=2^n$$

for $j=1, 2, ..., 2^n$. Now if P_1 is the first hyperplane asked about then choose H_1 such that

$$\sum_{V_j\in H_1}\omega_0(V_j)\geqq\sum_{V_j\in\bar{H}_1}\omega_0(V_j)$$

and

$$H_1\cap Q^n\neq\emptyset$$

should hold. Suppose that the halfspaces $H_1, H_2, ..., H_i$ have been chosen and the weight functions $\omega_0, \omega_1, ..., \omega_{i-1}$ have been defined $(i\geqq 1)$. Then let

$$\omega_i(V_j)=\begin{cases}\omega_{i-1}(V_j)/2 & \text{if } V_j\in P_i \text{ and } \dim(\cap P_i(V_j))<\dim(\cap P_{i-1}(V_j))\\ \omega_{i-1}(V_j) & \text{if } V_j\in P_i \text{ and } \dim(\cap P_i(V_j))=\dim(\cap P_{i-1}(V_j) \text{ or } V_j\in H_i\\ 0 & \text{if } V_j\in\bar{H}_i\end{cases}$$

and if P_{i+1} is the $(i+1)$-st hyperplane asked about then choose H_{i+1} such that

$$\sum_{V_j\in H_{i+1}}\omega_i(V_j)\geqq\sum_{V_j\in\bar{H}_{i+1}}\omega_i(V_j)$$

and

$$\left(\bigcap_{k=1}^{i}H_k\right)\cap Q^n=\emptyset$$

hold. It is obvious that H_{i+1} can be so chosen.

Notice that

$$\sum_{j=1}^{2^n}\omega_k(V_j)\geqq\frac{1}{2}\sum_{j=1}^{2^n}\omega_{k-1}(V_j)$$

and so

$$\sum_{j=1}^{2^n} \omega_k(V_j) \geqq 2^{2n-k} \tag{1}$$

for $k \geqq 0$.

Lemma 1. *If we know whether* $X \in Q^n$ *or not after the i-th question then* $\omega_i(V_j) \leqq 1$ *for* $j = 1, 2, 3, \dots, 2^n$.

Proof. If we know whether $X \in Q^n$ or not then $X \in Q^n$ by our strategy, so $\bigcap_{k=1}^{i} H_k \subset Q^n$. Suppose that $\omega_i(V_j) \geqq 2$ for a vertex V_j of Q^n. Then V_j is a boundary point of the polyhedron $\bigcap_{k=1}^{i} H_k$. Since $\omega_i(V_j) \geqq 2$ we thus have $\dim(\bigcap P_i(V_j)) = d \geqq 1$. Let e be a straight line such that $e \subset \bigcap P_i(V_j)$. On the other hand $V_j \in H_k$ if $P_k \notin P_i(V_j)$ and so there is a small sphere S with centre V_j such that $S \subset H_k$ if $P_k \notin P_i(V_j)$. Then the section $q = e \cap S$ belongs to the closure of the set $\bigcap_{k=1}^{i} H_k$ and this section q contains V_j in its interior. The closed cube Q^n contains the closure of the set $\bigcap_{k=1}^{i} H_k$ and so it contains also the section q. But the cube Q^n does not contain any section q that contains a vertex of Q^n in its interior, a contradiction. ∎

From now on suppose that we know whether $X \in Q^n$ or not after the i-th question. (That is $X \in Q^n$ and $\bigcap_{k=1}^{i} H_k \subset Q^n$.) Then every vertex V_j of Q^n with $\omega_i(V_j) = 1$ is contained in n independent hyperplanes P and for distinct such vertices V_j these sets of n independent hyperplanes are different. The number of vertices V_j with $\omega_i(V_j) = 1$ is at least 2^{2n-i} by (1) and Lemma 1. So we have

$$2^{2n-i} \leqq \binom{i}{n}$$

from which we can get the inequality

$$i \geqq c_0 n + o(n)$$

where $c_0 \approx 1{,}20$. We can say more about these sets of n independent hyperplanes. Let us fix such a set of n independent hyperplanes of $P_i(V_j)$ for every V_j with $\omega_i(V_j) = 1$. Let us denote this set by $P_i^n(V_j)$.

Lemma 2. *We have the inequality*

$$\left| \bigcap_{k=1}^{2^l+1} P_i^n(V_{j_k}) \right| < n - l,$$

for any integer $l \geqq 0$ *and for any indices* $1 \leqq j_1 < j_2 < \dots < j_{2^l+1} \leqq 2^n$ *with* $\omega_i(V_{jk}) = 1$.

Proof. Assume by way of contradiction that $P_{i_1}, P_{i_2}, \dots, P_{i_{n-l}} \in \bigcap_{k=1}^{2^l+1} P_i^n(V_{jk})$. The hyperplanes $P_{i_1}, P_{i_2}, \dots, P_{i_{n-l}}$ imply $n-l$ independent equations for the coordi nates $x_1, x_2, \ldots, x_n$ of the points of $\bigcap_{m=1}^{n-l} P_{i_m}$. Then there are l free variables in the general solution of the system of equations. If we fix these l coordinates as 0 or 1 then the other $n-l$ coordinates are determined so the number of the $0-1$ solutions is at most 2^l. But the vertices $V_{j_1}, V_{j_2}, \ldots, V_{j2^l+1}$ imply different $0-1$ solutions of the system of equations, a contradiction.

These at most $\binom{i}{n-l}$ sets of $n-l$ independent hyperplanes can be completed to sets $P_i^n(V_j)$ of n independent sets in at most $\binom{i}{n-l}2^l$ different ways. Any set $P_i^n(V_j)$ of n independent hyperplanes is obtained $\binom{n}{n-l}$ times. Thus the number of the sets $P_i^n(V_j)$ is at most $\binom{i}{n-l}2^l \Big/ \binom{n}{n-l}$. On the other hand we have at least 2^{2n-i} sets $P_i^n(V_j)$ by (1) and Lemma 1. Thus we have

$$2^{2n-i} \leqq \binom{i}{n-l} 2^l \Big/ \binom{n}{n-l} \tag{2}$$

for $l=0, 1, 2, \ldots, n$. It is easy to see that (2) gives the best estimate if $l=i-n$ so we get

$$2^{3n-2i} \leqq \binom{i}{2n-i} \Big/ \binom{n}{2n-i} \tag{3}$$

Using Stirling's formula and taking logarithms to the base 2 we get

$$3n-2i \leqq i \log_2 i + (i-n)\log_2(i-n) - (2i-2n)\log_2(2i-2n) - n\log_2 n + O(\log_2 n).$$

Hence

$$n \leqq i\log_2 i - (i-n)\log_2(i-n) - n\log_2 n + O(\log_2 n).$$

Let $i=cn$. Dividing by n we get

$$1 \leqq c\log_2 cn - (c-1)\log_2(c-1)n - \log_2 n + O\left(\frac{\log_2 n}{n}\right)$$

$$1 \leqq c\log_2 c - (c-1)\log_2(c-1) + O\left(\frac{\log_2 n}{n}\right).$$

Finally since the derivative of the function $x\log_2 x - (x-1)\log_2(x-1)$ has a positive lower bound in the interval $(1; 1.5)$ we find that $f(n) \geqq cn + O(\log n)$, where $c \approx 1{,}2938$ and $c\log_2 c - (c-1)\log_2(c-1) = 1$.

References

[1] I. Pohl, A sorting problem and its complexity, *Comm. of the ACM* **15** (1972), 462—464.

[2] Andrew C. Yao, On the complexity of comparison problems using linear functions, *Proc. 16th Ann. IEEE FOCS Symp.*, Berkeley 1975, 85—89.