ON THE SMALLEST AND LARGEST ELEMENTS

By

Á. VARECZA

(Received 15 May, 1981)

1. Introduction

Let $H = \{z_1, z_2, \ldots, z_n\}$ be a finite ordered set (say, different real numbers). However, the ordering is unknown for us. There are many cases when we want to obtain certain information concerning $H$ using pairwise comparisons of its elements.

The simplest question of this type: Which is the largest (smallest) element in $H$? It is easy to prove that any strategy finding the largest element needs at least $n - 1$ comparisons.

Pohl [4] proved that at least $n + \left[ \frac{n}{2} \right] - 2$ ([x] denotes the smallest integer $\leq x$) comparisons are needed if we want to determine the largest and smallest elements simultaneously (see also [3], [5]).

To find the two largest elements, $n + \left[ \log_2 n \right] - 2$ comparisons are needed [2] (for similar results see also [1], [3], [5]). Moreover it is proved [6] that it is impossible to find a pair of consecutive elements with a smaller number of comparisons. In a recent paper [7] it is proved that we need $2(n-2)$ comparisons if we want to decide whether $z_i$ and $z_j$ are neighbouring elements in $H$ ($z_i, z_j$ are arbitrary elements of $H$).

In this paper we shall solve the following problem: What is the minimal number of comparisons needed to decide whether $z_i$ and $z_j$ are the largest and the smallest elements (in this order) in $H$. The answer is $n + \left[ \frac{n}{2} \right] - 3$ ($n \geq 3$). We also solve the modified problem, when the question is if the pair $z_i, z_j$ coincides with the pair of the largest and smallest elements, regardless of their order.

The result is, surprisingly, slightly more than $n + \left[ \frac{n}{2} \right] - 3$, namely, $n + \left[ \frac{n-1}{2} \right] - 2$ ($n \geq 3$). Our method gives a new proof for Pohl's result as well.
2. Notations, definitions

The first pair to be compared is denoted by \( S_0 = (c, d) \). If the result of the comparison is \( c > d \) then the value of the variable \( e_1 \) is 1, and in the opposite case \( c < d \) \( e_1 = 0 \). The choice of the next pair \( S_1 (e_1) \) depends on \( e_1 \), say \( S_1 (e_1) = (e(e_1), f(e_1)) \). Define \( e_2 \) to be 1 if \( e(e_1) > f(e_1) \) and to be 0 otherwise. Continuing this procedure in the same way,

\[
S_{i-1} (e_1, e_2, \ldots, e_{i-1})
\]

is defined for some 0—1 sequences \( e_1, e_2, \ldots, e_{i-1} \) with the restriction that if \( i \geq 2 \), and \( S_{i-1} (e_1, e_2, \ldots, e_{i-1}) \) is defined then \( S_{i-2} (e_1, e_2, \ldots, e_{i-2}) \) is defined too. The value of \( e_i \) is 1 or 0 according to whether the first or the second member of \( S_{i-1} \) is larger. A set of questions put in this way will be called a strategy suitable for deciding the problem "whether or not \( z_i \) is the largest and \( z_j \) is the smallest element in \( H \)" iff for all sequences \( e_1, e_2, \ldots, e_i \) if

\[
S_{i-1} (e_1, e_2, \ldots, e_{i-1})
\]

is defined, but

\[
S_i (e_1, e_2, \ldots, e_i)
\]

is not, then

\[
\begin{align*}
& \text{the answers } e_1, e_2, \ldots, e_i (\text{to the questions } S_0, S_1 (e_1), \\
& \ldots, S_{i-1} (e_1, \ldots, e_{i-1}) ) \text{ decide the problem, whether} \\
& \text{or not } z_i \text{ is the largest and } z_j \text{ is the smallest element in } H.
\end{align*}
\]

We use the notation \( \mathcal{S} \) for such a strategy. We say that a strategy \( \mathcal{S} \) is finished for the sequence \( e_1, e_2, \ldots, e_i \) if conditions (2)—(4) are satisfied. The maximum length of the sequence \( e_1, \ldots, e_i \) finishing the strategy is called its length. It will be denoted by \( L (\mathcal{S}) \).

Denote by \( T_i (e_1, e_2, \ldots, e_j) \) the inequality corresponding to \( e_i \). Now we can express condition (4) in a modified way:

The inequalities

\[
T_1 (e_1), T_2 (e_1, e_2), \ldots, T_i (e_1, e_2, \ldots, e_i)
\]

decide whether or not \( z_i \) is the largest and \( z_j \) is the smallest element in \( H \). Let the situation after answering the question \( S_{i-1} (e_1, e_2, \ldots, e_{i-1}) \) be called the situation \( (e_1, e_2, \ldots, e_i) \) of \( \mathcal{S} \), that is, we have then the inequalities \( T_1 (e_1), T_2 (e_1, e_2), \ldots, T_i (e_1, e_2, \ldots, e_i) \). This system of inequalities will be denoted by \( \mathcal{E}_i \). The extension of \( \mathcal{E}_i \) consists of all the inequalities \( z_r < z_t \) which can be deduced from \( \mathcal{E}_i \). It is proved (Lemma 0 in [7]) that if \( z_r < z_t \) follows from \( \mathcal{E}_i \) then there is a chain of inequalities

\[
z_r = z_{r_1} < \ldots < z_{r_k} = z_t
\]
where $z_{r,v} = z_{r,v+1}$ (1 ≤ v ≤ k) are in $\mathcal{E}_v$. We now introduce the concept of graph-realization. Let us regard the set $H$ as the vertex set of a directed graph $G$. Let the comparison of any two elements of $H$ be an arc in $G$, directed from the greater element (vertex) to the smaller one. In the state $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_i)$ let $G^i$ denote the graph derived in this way. $H$ is totally ordered, so $G^i$ contains no directed cycle. By the above correspondence, with every state of $\mathcal{S}$ we associate an oriented graph. It follows from the correspondence that in an arbitrary state $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_i)$ of $\mathcal{S}$ the relation $e \rightarrow f$ is realized if and only if a directed path leads in $G^i$ from $e$ to $f$. If $\mathcal{S}$ is finished and $z_i$ is the largest and $z_j$ is the smallest element, then (by Lemma 0 of [7]) there exists a directed path from $z_i$ to all the other elements of $H$ and there exists a directed path to $z_j$ from all the other elements of $H$.

We can define in a similar way a strategy which can determine the largest and smallest elements simultaneously, and another strategy which can decide whether $z_i$ and $z_j$ are the largest and smallest elements in $H$ (regardless of their order). For this purpose it will be better to use the notations $z_i = z_1 = x$ and $z_j = z_2 = y$, i.e. $H = \{x, y, z_3, \ldots, z_n\}$.

Let $\mathcal{S}_1$ be a strategy which can define the largest and smallest elements of $H$ simultaneously; let $\mathcal{S}_2$ be a strategy which can decide whether $x$ and $y$ are the largest and smallest elements in $H$ (regardless of their order); and let $\mathcal{S}_3$ be a strategy which can decide whether $x$ is the largest and $y$ is the smallest element in $H$.

### 3. The results

We shall prove the following theorems.

**Theorem 1.** (Pohl [4])

\[
\min_{\mathcal{S}_1} L(\mathcal{S}_1) = n + \left\lfloor \frac{n}{2} \right\rfloor - 2 \quad (n \geq 2).
\]

**Theorem 2.**

\[
\min_{\mathcal{S}_2} L(\mathcal{S}_2) = n + \left\lfloor \frac{n-1}{2} \right\rfloor - 2 \quad (n \geq 3).
\]

**Theorem 3.**

\[
\min_{\mathcal{S}_3} L(\mathcal{S}_3) = n + \left\lfloor \frac{n}{2} \right\rfloor - 3 \quad (n \geq 3).
\]

**Proof of Theorem 1.** It is easy to find a strategy satisfying (6) (see [3], [4], [5]).

It remains to prove

\[
L(\mathcal{S}_1) \equiv n + \left\lfloor \frac{n}{2} \right\rfloor - 2
\]

or any strategy $\mathcal{S}_1$. 

This will be done in the following way. An algorithm will be given which determines a branch of the strategy, that is, a sequence $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_l$ finishing it.

This branch will have a length $\geq n + \left\lfloor \frac{n}{2} \right\rfloor - 2$. The algorithm determines the $\varepsilon$'s recursively. Partitions of $H$ will be used. The partitions will also be defined recursively, for any situation $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_i)$ along the indicated branch. The branch and the partitions will be determined simultaneously. A partition has three classes: $H = N^i \cup K^i \cup A^i$.

At the beginning $A^0 = H$, $N^0 = K^0 = \emptyset$.

Suppose that $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_i$ and $A^i$, $N^i$, $K^i$ are already defined. Then the next description determines $\varepsilon_{i+1}$ and $A^{i+1}$, $N^{i+1}$, $K^{i+1}$.

Let $S_i(\varepsilon_1, \ldots, \varepsilon_i) = (g, h)$. The new values of $\varepsilon_{i+1}$, $A^{i+1}$, etc. will depend on the classes containing $g$ and $h$, respectively. The cases obtained by interchanging the role of $g$ and $h$ will not be treated separately. As regards $A^{i+1}$ etc., we shall only indicate the new class for an element. Then it will be obviously omitted from its old class. The system of inequalities

$$T_1(\varepsilon_1), T_2(\varepsilon_1, \varepsilon_2), \ldots, T_i(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_i)$$

is denoted by $\mathcal{C}_i$ where $\varepsilon_1, \ldots, \varepsilon_i$ go along this branch.

Definition of the $(i+1)$-st point of the branch:

\begin{itemize}
  \item case 1: $g, h \in A^i \quad \varepsilon_{i+1} = 1, \quad g \in N^{i+1}, \quad h \in K^{i+1}$
  \item case 2: $g, h \in N^i \quad (K^i) \quad \varepsilon_{i+1} = 1, \quad g \in K^{i+1}$
  \item case 3: $g \in N^i, \quad h \in A^i \cup K^i \quad \varepsilon_{i+1} = 1, \quad h \in K^{i+1}$
  \item case 4: $g \in A^i, \quad h \in K^i \quad \varepsilon_{i+1} = 1, \quad g \in N^{i+1}$.
\end{itemize}

In this way we defined a branch $\varepsilon_1, \ldots, \varepsilon_i$ of the strategy $\mathcal{S}_i$. It will be denoted by $P_i$. The length $|P_i|$ of $P_i$ is $l$. We shall prove $l \equiv n + \left\lfloor \frac{n}{2} \right\rfloor - 2$ with a sequence of lemmas, valid for this special branch.

**Lemma 1.** $a \in N^i(K^i)$ implies $a \in N^r(K^r)$ for any $r > i$.

**Proof.** It can be easily seen by checking the cases of the definition of $P_i$.

**Lemma 2.** Suppose that $a < b$ is an inequality in $\mathcal{E}_i$ and $a \in N^i \quad (b \in K^i)$. Then $b \in N^i \quad (a \in K^i)$ follows.

**Proof.** It is sufficient to prove the first statement. The other one follows analogously.

Thus, suppose that $a < b$ is in $\mathcal{E}_i$ and $a \in N^i$. It follows that $S_j(\varepsilon_1, \ldots, \varepsilon_j) = (a, b)$ (or $(b, a)$) for some $j < i$ and $\varepsilon_{j+1} = 0$ (or $= 1$).

It follows from Lemma 1 that $a \in N^i \cup A^i$. If $a \in N^i$ then the statement follows from case 2 evidently.

If $a \in A^i$, then $b \in A^i \cup N^i$ and $b \in N^{i+1}$, $a \in K^{i+1}$ follows from cases 1 and 3 and $a \in K^i$ follows from Lemma 1, and this is a contradiction.

The lemma is proved.
Lemma 3. If $a \in N^i$ then there is an inequality $a \rightarrow b$ in $E_i$ with $b \in K^i$. Analogously, if $a \in K^i$ then there is an inequality $a \leftarrow b$ in $E_i$ with $b \in N^i$.

Proof. We prove the first half of the statement only, the other half can be proved in the same way. Thus suppose that $a \in N^i$.

Let $(\epsilon_1, \ldots, \epsilon_j)$ be the situation for which $a \in N^j$ but $a \notin N^{j+1}$. It follows from cases 1 and 4 that $a$ occurs in the question $S_j(\epsilon_1, \ldots, \epsilon_j)$, say $S_j(\epsilon_1, \ldots, \epsilon_j) = (a, b)$. It follows from cases 1 and 4 that $a \in A^j$ and $b \in A^j \cup K^j$ and $a \notin N^{j+1}$, $b \in K^{j+1}$. Lemma 1 implies $b \in K^i$.

The lemma is proved.

Lemma 4. If the strategy is finished for the sequence $\epsilon_1, \epsilon_2, \ldots, \epsilon_i$ then $A^i = \emptyset$.

Proof. Suppose on the contrary that $A^i \neq \emptyset$. If $a \in A^i$, then $a$ is not in $E_i$ and $a$ can be the smallest and largest element of $H$. This contradicts the supposition that the strategy is finished.

Let us turn back to the proof of (9).

Suppose that our strategy $\delta$ along branch $P_1$ is finished after the $l$-th comparison. It follows from Lemma 4 that $A^l = \emptyset$. Consequently,

$$|N^l \cup K^l| = n$$

holds. Let $|N^i| = i$, $|K^i| = j$ where $i + j = n$. Consider the graph $G^i$. Denote the largest element by $x$ and the smallest one by $y$. It follows from the definition of $P_1$ that $x \in N^i$, $y \in K^i$, and to all elements ($\neq x$) of $H$ ($N^i$) there is a directed path from $x$, and from all elements ($\neq y$) of $H$ ($K^i$) there is a directed path to $y$. First we consider those inequalities $a \leftarrow b$ in $E_i$ where $a \in N^i$, $b \in N^i$. Take the corresponding edges in $G^i$. There is a path from $x$ to any element of $N^i$. This path cannot go through an element of $K^i$ by Lemma 2.

Therefore the subgraph induced by $N^i$ is connected. Consequently there are at least $i - 1$ edges among the vertices in $N^i$. That is, the number of inequalities $a \leftarrow b$ in $E_i$ where $a, b \in N^i$ is at least $i - 1$. Similarly, there are at least $j - 1$ inequalities with $a, b \in K^i$.

Consider now those inequalities $a \leftarrow b$ in $E_i$ where $a \in K^i$, $b \in N^i$. One of $i$ and $j$ is at least $\left\lfloor \frac{n}{2} \right\rfloor$. Suppose that $|N^i| \geq \left\lfloor \frac{n}{2} \right\rfloor$ holds. Then, by Lemma 3, there are at least as many such inequalities in $E_i$. Summing up our results:

$$l \geq i - 1 + j - 1 + \left\lfloor \frac{n}{2} \right\rfloor = n + \left\lfloor \frac{n}{2} \right\rfloor - 2.$$ 

Theorem 1 is proved.

Proof of Theorem 2. It is easy to find a strategy satisfying (7).

Let $x$ and $y$ be these elements of $H$, we want only to decide whether $x$ and $y$ are the largest and smallest elements.

Let $S_0 = (x, z_2), S_1(\epsilon_1) = (y, z_3)$.

If $\epsilon_1 = 1$ ($\epsilon_1 = 0$) and $\epsilon_2 = 1$ ($\epsilon_2 = 0$) then $\delta_2$ is finished and the answer is no.
Suppose that $\varepsilon_1 = 1$ ($\varepsilon_2 = 0$) and $\varepsilon_2 = 0$ ($\varepsilon_2 = 1$), that is, $x > z_3$ ($x < z_3$) and $y < z_3$ ($y > z_3$). We determine the largest and smallest elements of $H - \{x, y, z_3\}$ simultaneously. It follows from Theorem 1 that we can do it by making $n - 3 + \left\lceil \frac{n - 3}{2} \right\rceil - 2$ comparisons.

Denote by $a$ ($b$) the largest (smallest) element of $H - \{x, y, z_3\}$.

Now we compare elements $x$ and $a$ ($y$ and $a$) and elements $y$ and $b$ ($x$ and $b$). If $x > a$ ($y > a$) and $y < b$ ($x < b$), the answer is yes, otherwise the answer is no.

The number of comparisons is

$$2 + n - 3 + \left\lceil \frac{n - 3}{2} \right\rceil - 2 + 2 = n + \left\lceil \frac{n - 1}{2} \right\rceil - 2.$$

This proves:

$$\min \limits_{\mathcal{S}_2} L(\mathcal{S}_2) \leq n + \left\lceil \frac{n - 1}{2} \right\rceil - 2.$$

It remains to prove

$$L(\mathcal{S}_2) \leq n + \left\lceil \frac{n - 1}{2} \right\rceil - 2$$

for any strategy $\mathcal{S}_2$.

This will be done similarly as we have proved Theorem 1. Consider branch $P_1$ of strategy $\mathcal{S}_2$. We shall use the subsets $A$, $N$, $K$ of $H$, similarly as we have used them in the proof of Theorem 1. Now we shall use a modification of $P_1$.

Let $S_i(\varepsilon_1, \ldots, \varepsilon_i)$ be the first comparison involving the element $x$ or $y$, that is, the comparisons $S_j(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_j)$ with $j < i$ do not involve $x$ and $y$.

We can suppose that $x$ occurs in $S_i(\varepsilon_1, \ldots, \varepsilon_i)$. If $S_i(\varepsilon_1, \ldots, \varepsilon_i) = (g, h)$ then we define the $(i+1)$-st point of the branch:

case 5: $g = x$, $h = y$, $\varepsilon_{i+1} = 1$, $x \in N^{i+1}$, $y \in K^{i+1}$ and if $S_r(\varepsilon_1, \ldots, \varepsilon_r) = (x, b)$ ($b \in H$, $r > i$) then $\varepsilon_{r+1} = 1$, and if $S_r(\varepsilon_1, \ldots, \varepsilon_r) = (y, b)$ ($b \in H$, $r > i$) then $\varepsilon_{r+1} = 0$.

case 6: $g = x$, $h \in K^i \cup A^i$, $\varepsilon_{i+1} = 1$, $x \in N^{i+1}$, $h \in K^{i+1}$, $y \in K^{i+1}$ ($h \neq y$) and if $S_r(\varepsilon_1, \ldots, \varepsilon_r) = (x, b)$ ($y, b)$ ($r > i$), then $\varepsilon_{r+1} = 1$ ($0$) ($b \in H$).

case 7: $g = x$, $h \in N^i$, $\varepsilon_{i+1} = 0$, $x \in K^{i+1}$, $y \in N^{i+1}$, and ($h \neq y$) if $S_r(\varepsilon_1, \ldots, \varepsilon_r) = (x, b)$ ($y, b)$ ($r > i$, $b \in H$) then $\varepsilon_{r+1} = 0$ ($1$).

Denote by $P_2$ the modification of branch $P_1$. Obviously, if strategy $\mathcal{S}_2$ is finished along branch $P_2$, then the largest element is $x$ ($y$) and the smallest one is $y$ ($x$) in $H$. We can see easily that Lemmas 1, 2, and 4 apply here too, and the setting of Lemma 3 is not realized for element $y$ only.

Suppose that our strategy $\mathcal{S}$ along branch $P_2$ is finished after the $l$-th comparison.

Let $|N^i| = i$, $|K^l| = j$ ($i + j = n$).
Similarly to the proof of Theorem 1, we can determine the number of comparisons easily. The number of inequalities \(a < b\) in \(E_i\) where \(a, b \in N^j\) (\(a, b \in K^l\)) is at least \(i - 1\) \((j - 1)\), and the number of inequalities \(a < b\) in \(E_i\) where \(a \in K^l\), \(b \in N^l\) is at least \(\left\lfloor \frac{n-1}{2} \right\rfloor\).

Summing up our results:

\[
I = i - 1 + j - 1 + \left\lfloor \frac{n-1}{2} \right\rfloor = n - 2 + \left\lfloor \frac{n-1}{2} \right\rfloor.
\]

Theorem 2 is proved.

**Remark.** It is easy to see that if \(n = 2\) then the minimal number of comparisons is 0.

**Proof of Theorem 3.** It is easy to find a strategy \(S_3\) satisfying (8). For \(n = 3\) let \(S_0 = (x, z_3)\), \(S_1(e_1) = (y, z_3)\), then \(S_3\) is finished and (8) holds. Suppose \(n \geq 4\). Let \(S_0 = (z_3, z_4)\) and suppose that \(e_1 = 1\) holds. Let \(S_1(e_1) = (x, z_3)\) and \(S_2(e_1, e_2) = (y, z_4)\). If \(e_2 = 0\) or \(e_3 = 1\) then \(S_3\) is finished and the answer is no.

Let \(e_2 = 1\) and \(e_3 = 0\), that is, \(x > z_3\) and \(y < z_4\). Now we determine the largest and smallest element of \(H - \{y, x, z_3, z_4\}\) simultaneously. It follows from Theorem 1 that we can do it by making \(n - 6 + \left\lfloor \frac{n-4}{2} \right\rfloor\) comparisons.

Denote a (b) the largest (smallest) element of the set \(H - \{x, y, z_3, z_4\}\). Now we compare element \(x\) with element \(a\), and element \(y\) with element \(b\). If the results are \(x > a\) and \(y < b\) then the answer is yes, otherwise the answer is no.

The number of comparisons is

\[
3 + n - 6 + \left\lfloor \frac{n-4}{2} \right\rfloor + 2 = n + \left\lfloor \frac{n}{2} \right\rfloor - 3.
\]

This proves:

\[
\min_{S_3} S_3 \leq n + \left\lfloor \frac{n}{2} \right\rfloor - 3 \quad (n \geq 3).
\]

It remains to prove

\[
L(S_3) \geq n + \left\lfloor \frac{n}{2} \right\rfloor - 3
\]

for any strategy \(S_3\) \((n > 3)\).

Consider branch \(P_1\) of strategy \(S_3\).

We will use subsets \(A, N, K\) of \(H\), similarly as we have used them in the proof of Theorem 1, and now we shall use a modification of \(P_1\).

Let \(N^0 = \{x\}, K^0 = \{y\}\) and if for any \(j\) \(S_j(\epsilon_1, \epsilon_2, \ldots, \epsilon_j) = (x, h)\) \((y, h)\) holds then \(\epsilon_{j+1} = 1\) \((0)\). Denote by \(P_3\) the modification of branch \(P_1\).

Obviously, if strategy \(S_3\) is finished along branch \(P_3\), and the answer is yes, then the largest element is \(x\) and the smallest one is \(y\).
We shall prove that the length of branch $P_3$ is at least $n + \left\lfloor \frac{n}{3} \right\rfloor - 3$.

We can see easily that Lemmas 1, 2, and 4 apply here too, and the setting of Lemma 3 is not realized for elements $x$ and $y$ only.

Suppose that our strategy $S_3$ is finished along branch $P_3$, after the $l$-th comparison.

Let $|N^l| = i$, $|K^l| = j$ ($i + j = n$).

Similarly to the proof of Theorem 1, we can determine the number of comparisons easily.

The number of inequalities $a < b$ in $E_l$ where $a, b \in N^l$ ($a, b \in K^l$) is at least $i - 1$ ($j - 1$), and the number of inequalities $a < b$ in $E_l$ where $a \in K^l$, $b \in N^l$ is at least $\left\lfloor \frac{n - 2}{2} \right\rfloor$.

That is, the length of branch $P_3$ is at least

$$i - 1 + j - 1 + \left\lfloor \frac{n - 2}{2} \right\rfloor = n + \left\lfloor \frac{n}{2} \right\rfloor - 3.$$ 

Theorem 3 is proved.

**Remark.** It is easy to see that if $n = 2$ then the minimal number of comparisons is 1.

**Acknowledgement**

Finally, the author would like to thank G. Katona for his valuable remarks.

**REFERENCES**


Bessenyei György Tanárképző Főiskola
Matematikai Tanszék
4400 Nyíregyháza, Sóstói út 31/b.