

On a problem of G. O. H. Katona

By Á. VARECZA (Nyíregyháza)

Dedicated to Prof. Béla Gyires on his 80th birthday

Abstract. Let H be a finite ordered set (say, different real numbers, $|H|=n$) however the ordering is unknown for us.

G. O. H. Katona raised the following problem: If $A=\{p, q\}$ ($1 \leq p < q \leq n$) and x, y are arbitrary elements of H then find the minimal number of comparisons needed to decide whether the indexes of the elements x, y are in A or not (say in decreasing order). In the general case the answer is unknown. In this paper we prove that if $A=\{1, q\}$ ($1 < q \leq n$) then the number of comparisons is at least

$$\min \left(n + q - 3, 2n - \left\lfloor \frac{q}{2} \right\rfloor - 2 \right).$$

1. Introduction

Let H be a finite ordered set (say, different real numbers), however the ordering is unknown for us. There are many situations when we want to obtain certain information concerning H using pairwise comparisons of the elements. The simplest question of this type: Which is the largest (smallest) element in H ? It is easy to prove that any strategy for finding the largest (smallest) element needs at least $n-1$ comparisons.

If we want to determine the two largest elements then $n-2 + \lceil \log_2 n \rceil$ comparisons are needed ([3], [4, 211–212 p.]). Moreover it is proved ([8], [9]) that it is impossible to find a pair of consecutive elements with a smaller number of comparisons.

($\lfloor x \rfloor$ ($\lceil x \rceil$) denotes the smallest (largest) integer $\cong x$ ($\geq x$)). IRA POHL ([1]) has proved that at least $n + \left\lfloor \frac{n}{2} \right\rfloor - 2$ comparisons are needed if we want to determine the largest and smallest elements simultaneously (see also [5], [7], [8]). If we want to decide only whether x_1 and x_2 are the largest and smallest elements in H then we need at least $n + \left\lfloor \frac{n-1}{2} \right\rfloor - 2$ comparisons ([5]) and to decide whether x_1, x_2 are neighboring, element in H we need at least $2(n-2)$ ($n \geq 3$) comparisons ([11]).

AMS (MOS) subject classifications (1980)

Primary: 05A05

Secondary: 05A15

Key words and phrases: Finite ordered set, searching, optimal strategy.

G. O. H. KATONA suggested the following problem:

Let $A = \{r_1, r_2, \dots, r_k\}$ ($1 \leq r_1 < r_2 < \dots < r_k \leq n$) and let x be an arbitrary element of H . The question is: how many comparisons are needed if we want to decide whether the index of the element x is in A or not?

KATONA proved ([2]) that the minimal number of comparisons is $n-1$ (see also [6], [8]).

G. O. H. KATONA suggested the following problem, too:

If $A = \{p, q\}$ and x, y are arbitrary elements of H and we want to decide whether the indexes of elements x, y are in A or not then how many comparisons are needed?

He found the strategy ([2]) which gives the answer using $n+q-3$ comparisons if $1 \leq p < q \leq \left\lfloor \frac{n}{2} \right\rfloor$ and he conjectured that this strategy is optimal.

In the recent paper [6] it is proved that the conjecture of G. O. H. Katona is true (see also [8]).

In a recent paper [10] the following cases were proved: If $A = \{r_1, r_2, \dots, r_k\}$ ($1 \leq r_1 < r_2 < \dots < r_k = n$) and $r_k = k$ ($r_k \neq n$) then we need at least $n-1$ comparisons and if $r_k \neq k$, $r_k \leq \left\lfloor \frac{n}{2} \right\rfloor$ then we need at least $n+r_k-3$ comparisons.

In this paper we prove the following:

If $A = \{1, q\}$ ($1 < q \leq n$) then the number of comparisons is at least

$$\min \left(n+q-3, 2n - \left\lfloor \frac{q}{2} \right\rfloor - 2 \right) \quad (n > 2).$$

2. Notations, definitions

Let x, y be arbitrary elements of the set H and let $A = \{1, q\}$ ($1 < q \leq n$).

Suppose that we want to decide whether the indexes of elements x, y are in the set A or not (that is the element x or y is the largest and the other the q -th)? The first pair to be compared is denoted by $S_0 = (a, b)$. If the result of the comparison is $a > b$ then the value of the variable ε_1 is 1. In the opposite case ($a < b$) $\varepsilon_1 = 0$.

The choice of the next pair $S_1(\varepsilon_1)$ depends on ε_1 . Suppose $S_1(\varepsilon_1) = (c(\varepsilon_1), d(\varepsilon_1))$. Define ε_2 to be 1 if $c(\varepsilon_1) > d(\varepsilon_1)$ and to be 0 otherwise. Continuing this procedure

$$(1) \quad S_{i-1}(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{i-1})$$

is defined for some 0, 1 sequences $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{i-1}$ with the restriction that if $S_{i-1}(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{i-1})$ is defined then $S_{i-2}(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{i-2})$ is defined, too. The value of ε_i is 1 or 0 according to whether the first or the second member is larger. A set of questions given in this way will be called a strategy suitable for deciding the questions "whether the indexes of x, y are in A or not" iff for all sequences $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_i$ when

$$(2) \quad S_{i-1}(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{i-1}) \text{ is determined, but}$$

- (3) $S_i(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_i)$ is not
- (4) $\left\{ \begin{array}{l} \text{then the answers } \varepsilon_1, \varepsilon_2, \dots, \varepsilon_i \text{ (together with the questions } S_0, S_1(\varepsilon_1), \dots \\ \dots, S_{i-1}(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{i-1}) \text{ give a unique reply to the problem: the indexes of} \\ x, y \text{ are in } A \text{ or not.} \end{array} \right.$

We use the notation $\mathcal{S}(A)$ for such a strategy. We say that the strategy $\mathcal{S}(A)$ is finished for the sequence $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_i$ if the conditions (2)—(4) are satisfied. The maximum length of the sequence $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_i$ finishing the strategy is called its length. It will be denoted by $L(\mathcal{S}(A))$.

Denote by $T_i(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_i)$ the inequality set up from the pair $S_{i-1}(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{i-1})$ on the basis of the answer ε_i . Now we can express condition (4) in a modified way: The inequalities

$$(5) \quad T_1(\varepsilon_1), T_2(\varepsilon_1, \varepsilon_2), \dots, T_i(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_i)$$

uniquely decide whether the indexes of x, y are in A or not. The situation $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_i)$ of $\mathcal{S}(A)$ is the situation after answering the question $S_{i-1}(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{i-1})$. That is, we have the inequalities $T_1(\varepsilon_1), T_2(\varepsilon_1, \varepsilon_2), \dots, T_i(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_i)$ and denote them by E_i .

We will use the following fact several times.

If a system of inequalities E_i is given, then only those inequalities are consequences of them which are deducible by transitivity.

That is let

$$H = \{h_1, h_2, \dots, h_n\}$$

be an ordered set:

$$h_1 < h_2 < \dots < h_n.$$

A set of inequalities

$$(*) \quad x_i < x_j \quad (i \neq j, i, j \in \{1, 2, \dots, n\})$$

is given and we are looking for a solution of $(*)$ in the set H . If the inequalities

$$x_{i_1} < x_{i_2}, x_{i_2} < x_{i_3}, \dots, x_{i_{k-1}} < x_{i_k} \quad (1 < k)$$

are all in $(*)$, then $x_{i_1} < x_{i_k}$ is deducible from $(*)$. The set of inequalities deducible from $(*)$ is called the extension of $(*)$. $(*)$ is non-contradictory if there is no inequality $x_i < x_i$ in the extension of $(*)$.

Lemma 0. *If $(*)$ is non-contradictory and the inequalities $x_v < x_w$ and $x_v > x_w$ are not in the extension of $(*)$, then there is a solution of $(*)$ in H satisfying $x_v > x_w$ and such a solution too for which $x_v < x_w$.*

In this paper we don't prove Lemma 0. The proof of Lemma 0, can be found in [11].

We now introduce the concept of graph — realization. Let the elements of a set H correspond to the vertices of a graph $\vec{G}$. Let a comparison be an edge of $\vec{G}$ between the corresponding vertices. Let the answer be the orientation of this edge in the following way: if we compare two elements — say c and d — in some state and the result of the comparison is $c > d$ then we direct the edge from c to d , conversely, when $c < d$ we direct the edge from d to c .

In the state $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_l)$ let $\vec{G}^i$ denote the graph derived in this way. By the above correspondence we uniquely associate an oriented graph to all states of $\mathcal{S}(A)$. It follows from the correspondence that in an arbitrary state $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_l)$ of $\mathcal{S}(A)$ the relation $a > b$ is realised if and only if an oriented path leads in $\vec{G}^i$ from a to b . Denote by G^i the graph obtained from $\vec{G}^i$ by cancelling the direction of edges.

3. The result

Theorem. *If $A = \{1, q\}$ and the set $\mathcal{C}$ is the strategy which is able to decide whether the indexes of elements x, y are in a set A or not (that is the element x or y is the largest and the other is the q -th in H), then*

$$(6) \quad \min_{\mathcal{S}(A) \in \mathcal{C}} L(\mathcal{S}(A)) = \min \left\{ n + q - 3, 2n - \left\lfloor \frac{q}{2} \right\rfloor - 2 \right\}$$

if $n > 2$.

PROOF OF THE THEOREM.

We distinguish two cases: 1. $n > \left\lfloor \frac{3q}{2} \right\rfloor - 1$, 2. $n \leq \left\lfloor \frac{3q}{2} \right\rfloor - 1$.

1. Suppose that $n > \left\lfloor \frac{3q}{2} \right\rfloor - 1$.

We can easily find a strategy $\mathcal{S}(A)$ which in at most $n + q - 3$ steps gives the answer to the question. This strategy is due to G. O. H. KATONA. Let $S_0 = (x, y)$. We can suppose that $x > y$. We compare all the elements $H \setminus \{x, y\}$ with y .

If the element y is larger in $n - q$ cases and smaller in $q - 2$ cases then y is the q -th element in H .

After this we compare the element x with the elements which are larger than y . If the element x is larger in these comparisons, then the element x is the largest in H . The number of comparisons is $n + q - 3$.

2. Suppose that $n \leq \left\lfloor \frac{3q}{2} \right\rfloor - 1$.

We determine the strategy $\mathcal{S}(A)$ which gives the answer to the question in at most $2n - \left\lfloor \frac{q}{2} \right\rfloor - 2$ steps. We compare the elements x, y with the elements from $H \setminus \{x, y\}$ until for some element from $H \setminus \{x, y\}$ — say a — $x > a > y$ or $x < a < y$ holds. Such element obviously exists if the indexes of x, y are in A .

We can suppose that $x > a > y$.

If $x > a > y$ holds and in this situation the number of those elements which are smaller than y is $n - q$, then from the elements not occurring until now we determine the largest and the smallest element and finally we compare x with the largest, y with the smallest element and we get the answer to our question.

Suppose that whenever $x > a > y$, the number of those elements which are smaller than y is smaller than $n - q$. If there exist two elements which have not occurred yet — say c, d — then we compare these two elements. We can suppose that $c > d$. We compare the element y with the element d . If $y > d$ then we compare c with the element not yet occurred — say e — and the smaller element with y etc. If there is

no such an element e then we compare c with y and in case $y < c$ we compare c with x , too. If $y < d$, then we compare c with x . We can suppose that $x > c$.

If there exist two elements which have not occurred yet — say f, g — then we compare these two elements. We can suppose that $f < g$. Compare y with f . If $y > f$, then we compare g with the element not yet occurred and the smaller with y etc. If $y < f$ then we compare x with g . If there exist two elements which have not occurred until.... In this way we give the answer to our question and we have done at most

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

comparisons.

We have proved by this that

$$\min_{\mathcal{S}(A) \in \mathcal{C}} L(\mathcal{S}(A)) \leq \min \left(n + q - 3, 2n - \left\lceil \frac{q}{2} \right\rceil - 2 \right).$$

It remains to prove that

$$(7) \quad L(\mathcal{S}(A)) \geq \min \left(n + q - 3, 2n - \left\lceil \frac{q}{2} \right\rceil - 2 \right)$$

holds for any strategy $\mathcal{S}(A)$.

This will be done in the following way.

An algorithm will be given which determines a branch of the strategy, that is, a sequence $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_l$ finishing it.

This branch will have a length

$$\geq \min \left(n + q - 3, 2n - \left\lceil \frac{q}{2} \right\rceil - 2 \right).$$

The algorithm determines the ε 's recursively.

Partitions of $H - \{x, y\}$ will be used. These partitions will also be defined recursively, for any situation $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_i)$ along the indicated branch. The branch and the partitions will be determined simultaneously. A partition has 5 classes: $B_1^i, B_2^i, B_3^i, B^i, C^i$.

The heuristic meaning of the classes is:

$B_1^i \cup B_2^i \cup B^i$: the set of elements which will be greater than the element $y(x)$ and smaller than $x(y)$;

B_3^i : the set of elements which will be smaller than both x and y ;

C^i : the set of elements yielding essential information in the given situation.

At the beginning $C^0 = H - \{x, y\}$, $B_1^0, B_2^0, B_3^0, B^0 = \emptyset$. Suppose that $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_i$ and $B_1^i, B_2^i, B_3^i, B^i, C^i$ are already defined.

The next description determines ε_{i+1} and $B_1^{i+1}, B_2^{i+1}, B_3^{i+1}, B^{i+1}, C^{i+1}$.

Let $S_i(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_i) = (g, h)$.

The new values of $\varepsilon_{i+1}, B_1^{i+1}$, etc. will depend on the classes containing g and h , resp. The classes which can be obtained by interchanging the role of g and h will not be treated separately. As regards B_1^{i+1} , etc. we will indicate the new class only for an element.

Then it will be obviously omitted from its old class. We can suppose that the element y has not occurred as x before.

Suppose that the element x first occurs in $S_j(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_j)$. Let

$$S_j(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_j) = (x, c).$$

We distinguish two cases:

Case I. $c \in B_1^i \cup B_2^i \cup B^j$. In this case $x < c$ and if $c \in B^j$ then $c \in B_1^{j+1}$.

Case II. $c \in \{y\} \cup B_3^j \cup C^j$. In this case $x > c$ and if $|B_3^j| < n - q$, $c \in C^j$ then $c \in B_3^{j+1}$ and if $|B_3^j| = n - q$, $c \in C^j$ then $c \in B_2^{j+1}$.

We define ε_{i+1} .

1. $g, h \in C^i$ $\varepsilon_{i+1} = 1$ and if $|B_1^i \cup B_2^i \cup B^i| < q - 2$,
 $|B_3^i| < n - q$ then $h \in B_3^{i+1}, g \in B^{i+1}$;
if $|B_3^i| = n - q$ then $g \in B_1^{i+1}, h \in B_2^{i+1}$;
if $|B_1^i \cup B_2^i \cup B^i| = q - 2$ then $g, h \in B_3^{i+1}$.
2. $g \in B^i, h \in C^i$ $\varepsilon_{i+1} = 1$ and
if $|B_3^i| < n - q$ then $h \in B_3^{i+1}$;
if $|B_3^i| = n - q$ then $h \in B_2^{i+1}, g \in B_1^{i+1}$.
3. $g, h \in B^i$ $\varepsilon_{i+1} = 1, g \in B_1^{i+1}, h \in B_2^{i+1}$.
4. $g \in B_r^i, h \in B_s^i$ $\varepsilon_{i+1} = 1$.
($r < s$; $r, s \in \{1, 2, 3\}$)
5. $g \in B_3^i, h \notin B_3^i$ $\varepsilon_{i+1} = 0$ and
if $h \in C^i, |B_1^i \cup B_2^i \cup B^i| < q - 2$ then $h \in B^{i+1}$;
if $h \in C^i, |B_1^i \cup B_2^i \cup B^i| = q - 2$ then $h \in B_3^{i+1}$.
6. $g \in B_1^i, h \in B^i \cup C^i$ $\varepsilon_{i+1} = 1$, and
if $|B_3^i| < n - q, h \in C^i$ then $h \in B_3^{i+1}$;
if $|B_3^i| = n - q, h \in C^i$ then $h \in B_2^{i+1}$;
if $h \in B^i$ then $h \in B_2^{i+1}$.
7. $g \in B_2^i, h \in B^i \cup C^i$ if $|B_3^i| < n - q, h \in C^i$ then $\varepsilon_{i+1} = 1, h \in B_3^{i+1}$;
if $|B_3^i| = n - q, h \in C^i$ then $\varepsilon_{i+1} = 0, h \in B_1^{i+1}$;
if $h \in B^i$ then $\varepsilon_{i+1} = 0, h \in B_1^{i+1}$.
8. $g, h \in B_r^i$ $\varepsilon_{i+1} = \text{arbitrary}$, except if it is
($r \in \{1, 2, 3\}$) determined by the extension of ε_i .
9. $g = x, h = y$ in the case I.: $x < y$;
in the case II.: $x > y$.

10. $g = x, h \in C^i$

in the case I.:

if $|B_1^i \cup B_2^i \cup B^i| < q-2$ then $\varepsilon_{i+1} = 0, h \in B_1^{i+1}$;

if $|B_1^i \cup B_2^i \cup B^i| = q-2$ then $\varepsilon_{i+1} = 1, h \in B_3^{i+1}$;

in the case II.: $\varepsilon_{i+1} = 1$ and

if $|B_3^i| < n-q$ then $h \in B_3^{i+1}$;

if $|B_3^i| = n-q$ then $h \in B_2^{i+1}$.

11. $g = x, h \in B_1^i \cup B_2^i \cup B^i$

in the case I.: $\varepsilon_{i+1} = 0$ and if $h \in B^i$

then $h \in B_1^{i+1}$;

in the case II.: $\varepsilon_{i+1} = 1$ and if $h \in B^i$

then $h \in B_2^{i+1}$.

12. $g = y, h \in C^i$

in the case I.: $\varepsilon_{i+1} = 1$ and

if $|B_3^i| < n-q$ then $h \in B_3^{i+1}$;

if $|B_3^i| = n-q$ then $h \in B_2^{i+1}$.

in the case II.:

if $|B_1^i \cup B_2^i \cup B^i| < q-2$ then $\varepsilon_{i+1} = 0, h \in B_1^{i+1}$;

if $|B_1^i \cup B_2^i \cup B^i| = q-2$ then $\varepsilon_{i+1} = 1, h \in B_3^{i+1}$.

13. $g = y, h \in B_1^i \cup B_2^i \cup B^i$

in the case I.: $\varepsilon_{i+1} = 1$ and if $h \in B^i$ then $h \in B_2^{i+1}$ in
the case II.: $\varepsilon_{i+1} = 0$ and if $h \in B^i$ then $h \in B_1^{i+1}$.

In this way we have defined the value of ε_{i+1} and the sets $B_1^{i+1}, B_2^{i+1}, B_3^{i+1}, B^{i+1}, C^{i+1}$.

It is worth-while to show in a figure the possible changes of elements among the classes (see Fig. 1).

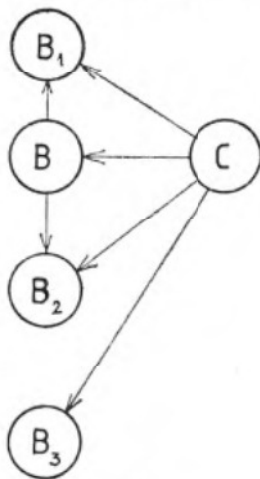


Fig. 1.

Suppose that the strategy $\mathcal{S}(A)$ is finished for the sequence $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_l$. It will be denoted by $P(A)$. The length $|P(A)|$ of $P(A)$ is l . We shall prove

$$l \cong \min \left(n+q-3, 2n - \left\lfloor \frac{q}{2} \right\rfloor - 2 \right).$$

We can easily see that if the strategy $\mathcal{S}(A)$ is finished for the sequence $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_l$ then case I. or case II. holds. If case I. has resulted then largest element in H is y and the q -th is x , and if case II. has resulted the largest element in H is x and the q -th is y .

From definition $P(A)$ it follows that if some element — say b — $\in B_r^i$ ($r \in \{1, 2, 3\}$) then $b \in B_r^j$ ($j > i$) and if $b \in B^i$ then $b \in B_1^i \cup B_2^i$.

From definition $P(A)$ it follows, too, that if $b > a \in \mathcal{E}_i$ and $b \in B_r^i$, $a \in B_s^i$ ($r \neq s$) then $r < s$ ($r, s \in \{1, 2, 3\}$) and the elements of $B_1^i \cup B_2^i$ are smaller than $y(x)$ and larger than $x(y)$ and the elements of B_3^i are smaller than x, y .

Suppose that $|B_1^i \cup B_2^i \cup B^i| < q-2$, $|B_3^i| < n-q$ and that either $|B_1^{i+1} \cup B_2^{i+1} \cup B^{i+1}| = q-2$ or $|B_3^{i+1}| = n-q$ holds. We prove the following Lemma:

Lemma 1. *If $|B_3^{i+1}| = n-q$, then the elements of B_3^{i+1} occur at least twice as smaller ones in $\mathcal{E}_i$.*

If $|B_1^{i+1} \cup B_2^{i+1} \cup B^{i+1}| = q-2$ then in case I. (in case II.) the elements of $B_1^i \cup B_2^i$ occur with at least one element from $\{x\} \cup B_3^i$ ($\{y\} \cup B_3^i$) in $\mathcal{E}_i$.

PROOF. We prove the first half of the statement. Suppose that $|B_3^i| < n-q$, $|B_1^i \cup B_2^i \cup B^i| < q-2$ and $|B_3^{i+1}| = n-q$ hold.

Let — say — a be an arbitrary element in B_3^{i+1} . Suppose that the element a first occurs in the $S_j(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_j)$. Let

$$S_j(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_j) = (a, b) \quad (j < i+1).$$

From definition $P(A)$ it follows that $a \in C^j$, $\varepsilon_{j+1} = 0$ (because $a \in B_3^{j+1}$). Thus we used 1., 2., 6., 7., 10. or 12.

If we used 1., 2., 6. or 7. then $b \in B_1^{j+1} \cup B_2^{j+1} \cup B^{j+1}$ and $b \in B_1^i \cup B_2^i$. It follows from this that the element a occurs in $\mathcal{E}_i$ as a smaller one with an element from $\{x\} \cup B_3^i$ (in case I.) or with an element from $\{y\} \cup B_3^i$ (in case II.), that is the statement holds. If we use 10., then x has already occurred and in its first occurrence with some element, say e , x is greater than e . On the basis of $\mathcal{E}_i$ $a < y$ because the element a occurs in $\mathcal{E}_i$ with an element from $\{y\} \cup B_3^i$ and the element a is smaller in this inequality.

The statement follows from this.

If we use 12. then — as we can easily see — the statement again holds.

With this we proved the first half of the Lemma.

Now we prove the other half.

Suppose that $|B_3^i| < n-q$, $|B_1^i \cup B_2^i \cup B^i| < q-2$ and $|B_1^{i+1} \cup B_2^{i+1} \cup B^{i+1}| = q-2$.

Let — say — a be an arbitrary element in $B_1^i \cup B_2^i$. Suppose that the element a first occurs in the $S_j(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_j)$. Let

$$S_j(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_j) = (a, b).$$

From the definition of $P(A)$ it follows that $a \in C^j$, $\varepsilon_{j+1}=1$ (because $a \in B_1^{i+1} \cup B_2^{i+1} \cup B^{i+1}$). Now we use 1., 2., 5., 6., 7., 10. or 12.

If we use 1. then $a > b$, $b \in B_3^{j+1}$, $b \in B_3^i$. We can easily see that the use of cases 2., 6., 7. is impossible. If we use 5. then $b \in B_3^j$, $b \in B_3^i$. If we use 10. then $b = x$, $\varepsilon_{j+1}=1$ (because $a \in B_1^i \cup B_2^i$).

This implies that the case I. holds and $a \in B_1^{j+1}$, $a \in B_1^i$. We can similarly prove the statement when we use 12. With this the Lemma is proved.

Finally we prove (7).

We distinguish two cases: 1. $|B_3^{i+1}| = n - q$; 2. $|B_1^{i+1} \cup B_2^{i+1} \cup B^{i+1}| = q - 2$.

1. If $|B_3^{i+1}| = n - q$ ($|B_3^i| < n - q$, $|B_1^i \cup B_2^i \cup B^i| < q - 2$), then $B_3^{i+1} = B_3^i$ and — according to Lemma 1. — the elements of B_3^i occur in $\mathcal{E}_i$ at least $2(n - q)$ times as a smaller one.

We can suppose that the element x is the largest and y is the q -th in H . (If the element y is the largest and x is the q -th then we can prove the statement similarly.) From definition $P(A)$ it follows that $B^i \cup C^i = \emptyset$. Consider the inequalities between the elements of the set $\{x, y\} \cup B_1^i \cup B_2^i$.

We can easily see that if $a \in B_1^i(B_2^i)$ then there exists an inequality $a > b$ ($a < b$) in $\mathcal{E}_i$ with $b \in \{y\} \cup B_2^i$ ($b \in \{x\} \cup B_1^i$).

Suppose that $a \in B_1^i$ and the element a first occurs in $S_j(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_j)$. Let $S_j(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_j) = (a, b)$. By the definition of $P(A)$ we have $a \in C^j$, $\varepsilon_{j+1}=1$. If $j < i+1$, $b \neq y$ then we use 1., 5. (supposing $x > y$ and $a \in B_1^{j+1}$, $b \in B_3^{j+1}$).

Later the element a occurs with the element e from $\{y\} \cup B \cup B_2$: in S_j as a larger one because $a \in B_1^i$ and $e \in \{y\} \cup B_2^i$.

If $j < i+1$ and $b = y$ then we use 12. and $a \in B_1^{j+1}$, $a \in B_1^i$.

If $j \geq i+1$ then we use 1., 5., 7. or 12. If we used 1., 7., 12. then $a \in B_1^{j+1}$, $b \in \{y\} \cup B_2^{j+1}$. If we used 5. then $b \in B_3^j$, $a \in C^j$, $a \in B_1^{j+1}$ and later the element a occurs with the element e from $\{y\} \cup B_2 \cup B$ and $a > e$ because we supposed that $a \in B_1^i$ therefore the element e is in $\{y\} \cup B_2^i$.

With this we have proved that if $a \in B_1^i$ then there exists an inequality $a > b$ in $\mathcal{E}_i$ with $b \in \{y\} \cup B_2^i$. We can prove similarly that if $a \in B_2^i$ then there exists an inequality $a < b$ in $\mathcal{E}_i$ with $b \in \{x\} \cup B_1^i$. With this the proof of the statement is finished. We prove that the element x occurs with an element from $\{y\} \cup B_2^i$ or the element y occurs with an element from $\{x\} \cup B_1^i$ in $\mathcal{E}_i$ (we supposed $x > y$). Suppose the element y does not occur with an element from $\{x\} \cup B_1^i$. We prove that in this case the element x occurs with an element from $\{y\} \cup B_2^i$ in $\mathcal{E}_i$. Let $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{j+1})$ be the first statement in which the element x occurs with an element from $B_1^{j+1} \cup B_2^{j+1} \cup B^{j+1} \cup \{y\}$ in $\mathcal{E}_{j+1}$. Let

$$S_j(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_j) = (x, b)$$

and — because we supposed the element x to be the largest — $x > b$. From the definition of $P(A)$ it follows that

$$b \in B_1^j \cup B_2^j \cup B^j \cup C^j \cup \{y\}, b \in B_2^{j+1} \cup \{y\}.$$

With this the statement is proved.

From these it follows that the number of inequalities $a > b$ in $\mathcal{E}_i$ in which $a \in B_1^i \cup \{x\}$, $b \in B_2^i \cup \{y\}$ is at least $\left\lfloor \frac{q-1}{2} \right\rfloor$ because $|B_1^i \cup \{x\}| + |B_2^i| = q - 1$.

Consider the graph G^l . We can easily see that the subgraph induced by $\{x\} \cup B_1^l$ ($\{y\} \cup B_2^l$) is connected.

Consequently there are at least $|B_1^l|$ ($|B_2^l|$) edges among the vertices in $\{x\} \cup B_1^l$ ($\{y\} \cup B_2^l$).

From these it follows that the number of inequalities in $\mathcal{E}_l$ is at least

$$2(n-q) + \left\lceil \frac{q-1}{2} \right\rceil + q - 2 = 2n - \left\lfloor \frac{q}{2} \right\rfloor - 2.$$

2. Suppose that $|B_1^l \cup B_2^l \cup B^l| < q - 2$, $|B_3^l| < n - q$ and $|B_1^{l+1} \cup B_2^{l+1} \cup B^{l+1}| = q - 2$. (We can suppose that x is the largest element in H .)

Suppose that the element x first occurs in $S_j(e_1, e_2, \dots, e_j)$. Let

$$S_j(e_1, e_2, \dots, e_j) = (x, c).$$

Because $x > c$ it follows that the case II. holds and $c \in \{y\} \cup B_3^{l+1}$.

It follows from Lemma 1. that if $a \in B_1^l \cup B_2^l$ then there exists an inequality $a > b$ in $\mathcal{E}_l$ with $b \in B_3^l \cup \{y\}$. This implies that the number of inequalities $a > b$ in $\mathcal{E}_l$ with $a \in \{x\} \cup B_1^l \cup B_2^l$, $b \in \{y\} \cup B_3^l$ is at least $q - 1$. We can easily see that the subgraph induced by $\{x\} \cup B_1^l \cup B_2^l$ ($\{y\} \cup B_3^l$) is connected.

Consequently there are at least $|B_1^l \cup B_2^l|$ ($|B_3^l|$) edges among the vertices in $\{x\} \cup B_1^l \cup B_2^l$ ($\{y\} \cup B_3^l$). From these it follows that the number of inequalities in $\mathcal{E}_l$ is at least

$$q - 1 + |B_1^l \cup B_2^l| + |B_3^l| = n + q - 3$$

since $|B_1^l \cup B_2^l| + |B_3^l| = n - 2$. From the cases 1. and 2. it follows, that

$$l \geq \min \left(n + q - 3, 2n - \left\lfloor \frac{q}{2} \right\rfloor - 2 \right).$$

With this the proof of our theorem complete.

Acknowledgement. Finally the author would like to thank professor G. O. H. KATONA for his valuable remarks.

References

- [1] IRA POHL, A sorting problem and its complexity *Com. of the ACM* **15**, 6 (1972), 462—464.
- [2] G. O. H. KATONA, Personal communication.
- [3] S. S. KISZLICYN, On the Selection of k -th Element of an Ordered Set by Pairwise Comparisons (russian) *Sib. Math. Z.*, **5** (1964), 557—564.
- [4] D. E. KNUTH, The Art of Computer Programming vol. 3., Sorting and Searching, Addison — Wesley, New-York, 1975.
- [5] Á. VARECZA, On the smallest and the largest elements, *Annales Univ. Sci. Budapest* **IV**. (1983), 1—10.
- [6] Á. VARECZA, On a conjecture of G. O. H. Katona, *Studia Sci. Math. Hung.* **23** (1988), 41—52.
- [7] Á. VARECZA, Methods for determination of principle bounds of ordering algorithms (in Hungarian) *Alk. Mat. Lapok* **5** (1979), 191—202.
- [8] Á. VARECZA, Optimális rendezési algoritmusok. *Kandidátusi értekezés* (1981).
- [9] Á. VARECZA, Finding two consecutive elements, *Stud. Sci. Math. Hung.* **17** (1982), 291—302.
- [10] Á. VARECZA, Katona G. O. H. egy problémájának általánosításáról. *Alk. Mat. Lapok* **10** (1984), 359—372.
- [11] Á. VARECZA, Are two given elements neighbouring? *Disc. Math.* **42** (1982), 107—117.

(Received June 2, 1984, in a reviewed form on February 20, 1989.)