Discrete Mathematics 310 (2010) 360–363

Contents lists available at ScienceDirect

Discrete Mathematics

journal homepage: www.elsevier.com/locate/disc

Note

Solution to a problem of Nicolas Lichiardopol

Adam R. Philpotts[a], Robert J. Waters[b,*]

[a] School of Mathematical Sciences, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom
[b] Department of Mathematics, University of Bristol, University Walk, Bristol, BS8 1TW, United Kingdom

ARTICLE INFO

Article history:
Received 6 August 2007
Accepted 2 September 2008
Available online 9 October 2008

Keywords:
On-line sorting
On-line algorithm

ABSTRACT

We present a solution to a problem posed by Nicolas Lichiardopol, regarding the on-line sorting of a sequence of integers.

1. The problem

The following on-line sorting problem was presented by Nicolas Lichiardopol, at the problem session of the 21st British Combinatorial Conference.

"Two integers m and n are given, with $1 \leq m \leq n$. A row of r empty boxes is available. Then m distinct random integers from $\{1, \ldots, n\}$ are announced. The task is to place each number in a box as it is announced so that, at the end, the numbers are stored in increasing order, with gaps allowed. What is the minimum value of r for which this task can always be accomplished?

Secondly, what is the minimum value needed when the m numbers are not necessarily distinct, and we require them to be in non-decreasing order?"

We will let $r(m, n)$ denote the minimum number of boxes needed in the first version of the problem, and $s(m, n)$ the minimum number in the second version. Although the problem was stated with $m \leq n$ as above, we note that this restriction is not necessary when the numbers are not distinct, and we give a solution in this case for all $m, n \in \mathbb{N} = \{1, 2, 3, \ldots\}$.

We will write $a_1, a_2, \ldots, a_m$ for the sequence of numbers in the order that they are announced.

2. Lists of distinct integers

Theorem 1. *For all $m, n \in \mathbb{N}$ with $m \leq n$,*

$$r(m, n) = \min\{n, 2^m - 1\}.$$

Proof. We begin with two simple strategies which show that $r(m, n) \leq n$ and $r(m, n) \leq 2^m - 1$ respectively. Firstly, if we start with n empty boxes, then as each number i is announced, we can just enter it into box i; thus $r(m, n) \leq n$.

The second strategy is defined inductively on m. Note that clearly $r(1, n) = 1$ for any $n \in \mathbb{N}$. Now assume that $m \geq 2$, and that $r(m-1, n) \leq 2^{m-1} - 1$. Starting with $2^m - 1$ boxes, whatever the first number a_1 is, we enter it into box 2^{m-1}.

* Corresponding author.
E-mail addresses: adam.philpotts@maths.nottingham.ac.uk (A.R. Philpotts), r.waters@bristol.ac.uk, rob@aquae.org.uk (R.J. Waters).

0012-365X/$ – see front matter
doi:10.1016/j.disc.2008.09.011

Then however many of the remaining $m-1$ numbers a_i are less than a_1, we have $2^{m-1}-1 \geq r(m-1,n)$ empty boxes to the left of a_1, so there is room to insert them. Similarly, there are $2^{m-1}-1$ empty boxes to the right of a_1, enough to contain the numbers a_i greater than a_1. Thus $r(m,n) \leq 2^m-1$.

To complete the proof we must show that $r(m,n) \geq \min\{n, 2^m-1\}$, which again we do by induction on m. Suppose that $m \geq 2$ and $a_1 = \lceil n/2 \rceil$, and that we choose to enter a_1 into box j.

Consider first the case $m \leq n < 2m$ (which implies that $n \leq 2^m-1$). Since $m-1 \geq \lceil n/2 \rceil - 1$, it may be that $\{1, \dots, \lceil n/2 \rceil - 1\} \subseteq \{a_2, \dots, a_m\}$, and so we must leave at least $\lceil n/2 \rceil - 1$ boxes to the left of box j. On the other hand, since $m-1 \geq n - \lceil n/2 \rceil = \lfloor n/2 \rfloor$, it may be that $\{\lceil n/2 \rceil + 1, \dots, n\} \subseteq \{a_2, \dots, a_m\}$, and so we must also leave at least $\lfloor n/2 \rfloor$ boxes to the right of box j. Hence

$$r(m,n) \geq (\lceil n/2 \rceil - 1) + 1 + \lfloor n/2 \rfloor = n.$$

Now we consider the case $n \geq 2m$. It may be that $a_i < \lceil n/2 \rceil$ for each $i = 2, \dots, m$, and so we must leave at least $r(m-1, \lceil n/2 \rceil - 1)$ boxes to the left of box j. On the other hand, we may have $a_i > \lceil n/2 \rceil$ for $i = 2, \dots, m$, and so we must also leave $r(m-1, \lfloor n/2 \rfloor)$ boxes to the right of box j. Hence

$$r(m,n) \geq r(m-1, \lceil n/2 \rceil - 1) + 1 + r(m-1, \lfloor n/2 \rfloor).$$

If $n \leq 2^m-1$ then $\lfloor n/2 \rfloor \leq 2^{m-1}-1$ and $\lceil n/2 \rceil - 1 \leq 2^{m-1}-1$, so we have

$$r(m,n) \geq (\lceil n/2 \rceil - 1) + 1 + \lfloor n/2 \rfloor = n;$$

and if $n \geq 2^m-1$ then $\lfloor n/2 \rfloor \geq 2^{m-1}-1$ and $\lceil n/2 \rceil - 1 \geq 2^{m-1}-1$, so

$$r(m,n) \geq (2^{m-1}-1) + 1 + (2^{m-1}-1) = 2^m-1.$$

Thus $r(m,n) \geq \min\{n, 2^m-1\}$, and this completes the proof of Theorem 1. $\square$

3. Lists with repetition

Theorem 2. *For all $m, n \in \mathbb{N}$,*

$$s(m,n) = \begin{cases} 2^b + t(m-b) - 1 & \text{if } n \leq 2^m, \\ 2^m - 1 & \text{otherwise,} \end{cases} \tag{1}$$

where $t = \lceil n/2 \rceil$ and $b = \lceil \log_2 t \rceil$.

We will need a couple of lemmas, which establish recursions for the values of $s(m,n)$, before proving Theorem 2. For these recursions it will be convenient to set $s(m,0) = 0$ for any $m \in \mathbb{N}$. Also, we note that

$$m' \leq m \quad \text{and} \quad n' \leq n \Rightarrow s(m',n') \leq s(m,n). \tag{2}$$

It is easy to see that $s(m,1) = s(m,2) = m$ for any $m \in \mathbb{N}$: we need at least m boxes to write down m numbers; then whenever a 1 is announced, we write it in the leftmost empty box, and whenever a 2 is announced, we write it in the rightmost empty box. Furthermore, it is clear that

$$s(m,n) \geq 1 + s(m-1,n) \tag{3}$$

for any $m, n \in \mathbb{N}$: irrespective of where we enter the first of a list of m numbers, in order to ensure that the remaining $m-1$ numbers are in the right order we need at least $s(m-1,n)$ additional boxes.

Lemma 3. *For all $m \in \mathbb{N}$ and $n \geq 3$,*

$$s(m,n) = \max_{1 \leq k \leq n} s(m-1, k-1) + 1 + s(m-1, n-k). \tag{4}$$

Lemma 4. *For all $m \in \mathbb{N}$ and $t \in \mathbb{N}$,*

$$s(m, 2t-1) = s(m, 2t). \tag{5}$$

Proof of Lemmas 3 and 4. Firstly, we show that the right-hand side of (4) is always a lower bound for $s(m,n)$. Suppose that $a_1 = k$, and we choose to enter it into box j. It may be the case that $a_i < k$ for all $i = 2, \dots, m$, and so we need to leave at least $s(m-1, k-1)$ boxes to the left of box j. On the other hand, we may have $a_i > k$ for $i = 2, \dots, m$, and so we also need to leave $s(m-1, n-k)$ boxes to the right of box j. Hence $s(m,n) \geq s(m-1,k-1) + 1 + s(m-1,n-k)$, for any $k = 1, 2, \dots, n$.

We will prove the opposite inequality and Lemma 4 simultaneously, by induction on m. As before, we note that $s(1,n) = 1$ for any $n \in \mathbb{N}$; note also that we have already shown that Lemma 4 holds when $t = 1$. Now fix $m \geq 2$, and assume that Eqs. (4) and (5) hold for smaller values of m. We need to show that, for each $t \geq 2$,

$$s(m, 2t-1) \leq \max_{1\leq k\leq 2t-1} s(m-1, k-1) + 1 + s(m-1, 2t-k-1), \tag{6a}$$

$$s(m, 2t) \leq \max_{1\leq k\leq 2t} s(m-1, k-1) + 1 + s(m-1, 2t-k). \tag{6b}$$

We first show that the right-hand sides of (6a) and (6b) are the same. Write $P_k = s(m-1, k-1) + 1 + s(m-1, 2t-k-1)$ and $Q_k = s(m-1, k-1) + 1 + s(m-1, 2t-k)$. The inductive hypothesis $s(m-1, 2t'-1) = s(m-1, 2t')$, together with (2), tells us that for even $k \in \{2, 4, \ldots, 2t-2\}$, $P_k = Q_k = Q_{k+1}$, and $P_{k\pm 1} \leq P_k$. This means that

$$\max_{1\leq k\leq 2t-1} P_k = \max_{2\leq k\leq 2t-1} Q_k,$$

and we just need to check that Q_1 is not larger than this maximum (note that $Q_{2t} = Q_1$). However $Q_1 = 1 + s(m-1, 2t-1)$ (since $s(m-1, 0) = 0$), which by the inductive hypothesis equals $s(m-2, k-1) + 2 + s(m-2, 2t-k-1)$ for some k, which by (3) is less than $s(m-1, k-1) + 1 + s(m-1, 2t-k-1) = P_k$.

Next we prove (6b): given m numbers announced from $\{1, \ldots, 2t\}$, we want to show that $\max\{Q_k : 1 \leq k \leq 2t\}$ boxes are enough in all cases. Whatever number a_1 is announced first, we place it in box $j = s(m-1, a_1) + 1$. If a_1 is even, we will place any of the remaining numbers a_i that are less than *or equal to* a_1 to the left of box j, and any a_i that are greater than a_1 to the right of box j. To do this we need $s(m-1, a_1) + 1 + s(m-1, 2t-a_1)$ boxes; but this equals Q_{a_1} since, by induction, $s(m-1, a_1) = s(m-1, a_1-1)$. On the other hand, if a_1 is odd, we will place any a_i that are less than a_1 to the left of box j, and any a_i that are greater than *or equal to* a_i to the right of box j. In this case we need $s(m-1, a_1-1) + 1 + s(m-1, 2t-a_1+1)$ boxes; but this equals Q_{a_1} since, by induction, $s(m-1, 2t-a_1) = s(m-1, 2t-a_1+1)$.

Finally, we observe that (6a) follows from (6b), since $s(m, 2t-1) \leq s(m, 2t)$ by (2), and the right-hand sides are the same; this in turn implies Eq. (5). Thus we have completed the proof of both Lemmas 3 and 4. □

Proof of Theorem 2. Since relation (5) is satisfied by (1), it is only necessary to prove (1) for even $n = 2t$. We will prove (1) by induction on m. Notice that the two expressions for $s(m, n)$ are equal when $b = m$, which is the case when $2^m + 1 \leq n \leq 2^{m+1}$; this will simplify the inductive step.

It is easily verified that (1) correctly gives $s(1, n) = 1$ for all $n \in \mathbb{N}$. Now assume that $m \geq 2$, and that

$$s(m-1, 2t) = 2^{\lceil \log_2 t\rceil} + t(m-1-\lceil \log_2 t\rceil) - 1 \tag{7}$$

for all $t \leq 2^{m-1}$ (i.e. $2t \leq 2^{(m-1)+1}$, using the comment in the previous paragraph). Using the fact that equality holds in (6b), and the subsequent comments about the values Q_k, we can deduce that

$$s(m, 2t) = \max_{1\leq j\leq t-1} s(m-1, 2j) + 1 + s(m-1, 2t-2j). \tag{8}$$

In particular, this means that $s(m, 2t) \geq s(m-1, 2j) + 1 + s(m-1, 2t-2j)$ for $j = \lfloor t/2 \rfloor$. We show that this gives the right-hand side of (1) as a lower bound for $s(m, 2t)$, by considering four cases.

(i) If $t > 2^{m-1}$, then $s(m, 2t) \geq (2^{m-1} - 1) + 1 + (2^{m-1} - 1) = 2^m - 1$.

In the remaining three cases we assume that $t \leq 2^{m-1}$, and set $j = \lfloor t/2 \rfloor$ and $b = \lceil \log_2 t\rceil$. We also make use of (7) in each case.

(ii) If t is even, then $\lceil \log_2 j\rceil = \lceil \log_2 (t-j)\rceil = b-1$, and so

$$s(m, 2t) \geq 2\left(2^{b-1} + \frac{t}{2}[(m-1)-(b-1)] - 1\right) + 1 = 2^b + t(m-b) - 1.$$

(iii) If t is odd, and $t \neq 2^{b-1} + 1$, then again $\lceil \log_2 j\rceil = \lceil \log_2 (t-j)\rceil = b-1$, and so

$$\begin{aligned} s(m, 2t) &\geq \left(2^{b-1} + \frac{t-1}{2}[(m-1)-(b-1)] - 1\right) + 1 + \left(2^{b-1} + \frac{t+1}{2}[(m-1)-(b-1)] - 1\right) \\ &= 2^b + t(m-b) - 1. \end{aligned}$$

(iv) If $t = 2^{b-1} + 1$, then $j = 2^{b-2}$ and $\lceil \log_2 j\rceil = b-2$, but $t - j = 2^{b-2} + 1$ and $\lceil \log_2 (t-j)\rceil = b-1$; so we have

$$\begin{aligned} s(m, 2t) &\geq \left(2^{b-2} + 2^{b-2}[(m-1)-(b-2)] - 1\right) + 1 + \left(2^{b-1} + (2^{b-2}+1)[(m-1)-(b-1)] - 1\right) \\ &= 2^b + (2^{b-1}+1)(m-b) - 1. \end{aligned}$$

We now need to prove the opposite inequality, that the right-hand side of (1) is an upper bound for $s(m, 2t)$. Observe that the second strategy in the proof of Theorem 1 works equally well when the a_i are not all distinct, and so $s(m, n) \leq 2^m - 1$ always. This completes the proof for case (i) above, so we turn to the cases where $t \leq 2^{m-1}$. Given (8), we need to show that

$$s(m-1, 2j) + 1 + s(m-1, 2t-2j) \leq 2^b + t(m-b) - 1 \tag{9}$$

for each $j = 1, \ldots, t-1$. Write $k = t - j$, $c = \lceil \log_2 j \rceil$, and $d = \lceil \log_2 k \rceil$; substituting (7) into (9), and rearranging, give

$$2^c + 2^d + t(b-1) - jc - kd \leq 2^b. \tag{10}$$

For any $x \in \mathbb{Z}$, $x \leq 2^{x-1}$; now let $x = b - c$. This gives $2^c(b-c) \leq 2^{b-1}$, and since $j \leq 2^c$, we have $2^c + j(b-c-1) \leq 2^{b-1}$. In the same way, setting $x = b - d$ gives $2^d + k(b-d-1) \leq 2^{b-1}$. However, adding these last two inequalities together, we obtain exactly (10). Thus the proof of Theorem 2 is complete. $\square$

Note added in proof

A referee has pointed out that Lichiardopol has added a third version of the problem, the same as the second version but "with the supplementary constraint that if the list contains more than one instance of a given number, the identical numbers must be written from left to right (with gaps allowed) in the order they are announced".

Writing $t(m, n)$ for the minimum number of boxes needed in this third version, we give the following solution.

Theorem 5. *For all $m, n \in \mathbb{N}$,*

$$t(m,n) = \begin{cases} 2^b + n(m-b) - 1 & \text{if } n \leq 2^m, \\ 2^m - 1 & \text{otherwise}, \end{cases} \tag{11}$$

where $b = \lceil \log_2 n \rceil$.

Proof. We omit a full proof, since it is largely analogous to that of Theorem 2, but mention some differences. The result corresponding to Lemma 3 is that, for all $m \in \mathbb{N}$ and $n \geq 2$,

$$t(m,n) = \max_{1 \leq k \leq n} t(m-1, k-1) + 1 + t(m-1, n-k+1), \tag{12}$$

with the starting condition that $t(m, 1) = m$ for any $m \in \mathbb{N}$. The extra '+1' at the end of (12) takes into account the fact that whenever $a_i = a_1$ for $i \geq 2$, we must now write a_i to the right of a_1. There is no analogue to Lemma 4; in fact this considerably simplifies the proof of (12). The proof of Theorem 5 follows just as above, by showing that recursion (12) is satisfied by (11). $\square$