JOURNAL OF COMBINATORIAL THEORY, Series A 52, 62–76 (1989)

Ulam's Searching Game with Lies

JUREK CZYZOWICZ

Département d'Informatique, Université du Québec à Hull, C.P. 1250, succ. "B", Hull, Québec, Canada J8X 3X7

DANIELE MUNDICI

Dipartimento Scienze Informazione, Università Studi Milano, Via Moretto da Brescia 9, 20133 Milan, Italy

AND

ANDRZEJ PELC

Département d'Informatique, Université du Québec à Hull, C.P. 1250, succ. "B", Hull, Québec, Canada J8X 3X7

Received January 21, 1988; accepted February 22, 1988

We determine the minimal number of yes–no queries sufficient to find an unknown integer between 1 and 2^m if at most two of the answers may be erroneous. © 1989 Academic Press, Inc.

1. INTRODUCTION

S. M. Ulam [6] raised the following question:

> Someone thinks of a number between one and one million (which is just less than 2^{20}). Another person is allowed to ask up to twenty questions, to each of which the first person is supposed to answer only yes or no. Obviously the number can be guessed by asking first: Is the number in the first half-million? and then again reduce the reservoir of numbers in the next question by one-half, and so on. Finally the number is obtained in less than $\log_2(1{,}000{,}000)$. Now suppose one were allowed to lie once or twice, then how many questions would one need to get the right answer? One clearly needs more than n questions for guessing one of the 2^n objects because one does not know when the lie was told. This problem is not solved in general.

Some partial answers to this question follow from results of Rivest *et al.* [4] and Spencer [5]. In both papers a search in the set $\{1, ..., n\}$ for any natural n is considered. For one lie allowed, Rivest *et al.* obtained the following:

0097-3165/89 $3.00

— it is impossible to determine the number in k questions if $2^k < n(k+1)$;

— it is possible to determine the number in k questions if $2^{k-1} \geqslant nk$.

For at most two lies allowed their result gives the following:

— it is impossible to determine the number in k questions if

$$2^k < n \frac{k^2+k+2}{2};$$

— it is possible to determine the number in k questions if

$$2^{k-2} \geqslant n \frac{k^2-3k+4}{2}.$$

It is easy to see that in the case of one lie allowed these estimates leave two possibilities for the minimal number of questions and in the case of two lies allowed they leave three possibilities. For the particular case of $n = 10^6$ the possibilities left are 25 or 26 for one lie and 29, 30, or 31 for two lies.

Spencer [5] obtained the following estimates for the case of one lie allowed:

— it is impossible to determine the unknown integer in k questions if $2^k < n(k+1)$;

— it is possible to determine the unknown integer in k questions if $2^k \geqslant \frac{8}{5} n(k+1)$.

As before this result leaves two possibilities for the minimal number of questions.

An exact solution of a generalized version of Ulam's problem for one lie was obtained in Pelc [3]:

— for even n, k questions are sufficient to determine a number in $\{1, ..., n\}$ iff $n(k+1) \leqslant 2^k$;

— for odd n, k questions are sufficient to determine a number in $\{1, ..., n\}$ iff $n(k+1)+(k-1) \leqslant 2^k$.

A particular instance of Ulam's problem for two lies and $n = 1{,}000{,}000$ was obtained in [2] where the minimal number of questions was shown to be 29.

The main objective of the present paper is to give a solution of Ulam's problem for two lies, in the case when the search space has size 2^m. We show that it is then possible to determine an unknown number in exactly

K questions where K is the lower bound provided in [4]; i.e., it is the minimal k for which

$$2^m \cdot \frac{k^2+k+2}{2} \leqslant 2^k.$$

From the above result it is easy to get the answer for $n = 10^6$, already obtained in [2] (cf. corollary in Section 3).

An important correspondence should be noted between Ulam's problem for two lies and that of finding a shortest 2-error correcting code of size n. Indeed, consider the non-interactive version of Ulam's game; that is, when the Questioner is required to state all his queries at once, then collect all answers and find the unknown number on this basis. Any optimal Questioner's strategy in this game yields a minimal length 2-error correcting code of size n and conversely. Hence, looking from the more general point of view of searching games admitting lies, the above problem of coding theory is that of finding an optimal strategy in a non-interactive game, while Ulam's problem concerns the interactive counterpart.

Two-error correcting codes of minimal length are known only for some values of the size n and the problem of finding such codes for every $n = 2^m$ is still open. It would be interesting to modify our techniques in order to solve this non-interactive version of Ulam's problem. No such modification seems apparent at this point.

2. Terminology

We begin the analysis of the case of two lies by introducing some terminology. A game is considered between two players: the Questioner and the Responder. Before the game both players agree on the size n of the search space and the number k of questions to be stated. Then the Responder chooses an element $e \in \{1, ..., n\}$ unknown to the Questioner who has to determine it in k queries of the form "$e \in T$?" for $T \subset \{1, ..., n\}$. Each move of the Questioner is such a query and the next move of the Responder is the answer yes or no. The Responder may lie at most twice. If the Questioner has a winning strategy to determine the hidden number in k queries independently of the Responder's choice and play, we say that he wins the game in k questions. Our interest is focused on the minimal k necessary for a given n.

With each stage of the game when the turn of the Questioner comes, we associate a *state* of the game which is a triple (a, b, c) of natural numbers. The first number is the size of the truth-set: the set of those elements of $\{1, ..., n\}$ which satisfy all answers given previously. The number b is the size of the one-lie-set: the set of those elements of $\{1, ..., n\}$ which satisfy all but one answer. Finally the last number in the state is the size of the two-lies-set: the set of elements of $\{1, ..., n\}$ satisfying all but two answers. The state at the beginning is clearly $(n, 0, 0)$.

Following the idea of Berlekamp [1] we define the weight of each state (a, b, c) corresponding to a stage of the game at which j questions remain to be asked. This weight is defined by

$$w_j(a, b, c) = a\left(\binom{j}{0} + \binom{j}{1} + \binom{j}{2}\right) + b\left(\binom{j}{0} + \binom{j}{1}\right) + c\binom{j}{0}$$
$$= a\,\frac{j^2 + j + 2}{2} + b(j+1) + c.$$

This can be interpreted by the fact that each number in the truth-set gives $\binom{j}{2}$ possibilities of lying twice during the remaining series of j questions, $\binom{j}{1}$ possibilities of lying once, and a single possibility of not lying at all. In the one-lie-set there are no more possibilities of lying twice and in the two-lies-set the Responder is forced to say the truth till the end, which justifies the factors $\binom{j}{0} + \binom{j}{1}$ and $\binom{j}{0}$, respectively. (See Spencer [5] for this interpretation in the case of one lie.)

Let (a, b, c) be a state and X, Y, Z subsets of the truth-set, the one-lie-set, and the two-lies-set, respectively. Let x, y, z be the cardinalities of X, Y, Z, respectively. Then the query: "is the unknown integer an element of $X \cup Y \cup Z$?" will be referred to as an $[x, y, z]$? query.

The query $[x, y, z]$? asked in the state (a, b, c) yields two states: the state $\mathrm{YES}_{a,b,c}(x, y, z)$ resulting from the answer yes and the state $\mathrm{NO}_{a,b,c}(x, y, z)$ resulting from the answer no. Obviously

$$\mathrm{YES}_{a,b,c}(x, y, z) = (x, a - x + y, b - y + z)$$

and

$$\mathrm{NO}_{a,b,c}(x, y, z) = (a - x, x + b - y, y + c - z).$$

For any $j \geqslant 1$ and any query $[x, y, z]$? asked in the state (a, b, c) we also have

$$w_j(a, b, c) = w_{j-1}(\mathrm{YES}_{a,b,c}(x, y, z)) + w_{j-1}(\mathrm{NO}_{a,b,c}(x, y, z)).$$

Definition. For any state (a, b, c), the number $\mathrm{ch}(a, b, c)$ is called the character of this state and is defined as follows

$$\mathrm{ch}(a, b, c) = \min\{k : w_k(a, b, c) \leqslant 2^k\}.$$

The Questioner's win in the remaining j questions starting at a given stage of the game depends exclusively on the state corresponding to this stage and on the number j, rather than on the particular elements of the truth-set, the one-lie-set, and the two-lies-set. This justifies the following

DEFINITION. The state (a, b, c) is called nice iff the Questioner wins in $\mathrm{ch}(a, b, c)$ questions starting from this state.

DEFINITION. The state (a, b, c) is called balanced if there exists a query $[x, y, z]$? such that

$$|w_{k-1}(\mathrm{YES}_{a,b,c}(x, y, z)) - w_{k-1}(\mathrm{NO}_{a,b,c}(x, y, z))| \leqslant 1,$$

where $k = \mathrm{ch}(a, b, c)$.

Note that if (a, b, c) is balanced and the query $[x, y, z]$? is as above then

$$\mathrm{ch}(\mathrm{YES}_{a,b,c}(x, y, z)), \qquad \mathrm{ch}(\mathrm{NO}_{a,b,c}(x, y, z)) \leqslant k-1.$$

3. THE MAIN RESULT

The main result of this paper is the following

THEOREM 1. *Let $m \geqslant 3$ be a natural number. The number of yes–no queries necessary and sufficient to determine an unknown element of $\{1, ..., 2^m\}$—if the Responder may lie at most twice—is equal to the least natural k such that*

$$k^2 + k + 2 \leqslant 2^{k-m+1}.$$

In view of the lower bound from [4] and of terminology adopted in Section 2 the above result follows from

THEOREM 2. *The state $(1, m, \binom{m}{2})$ is nice for any natural $m \geqslant 3$.*

Indeed, suppose the Questioner wins in $K = \mathrm{ch}(1, m, \binom{m}{2})$ questions starting from $(1, m, \binom{m}{2})$; K is the minimal k for which

$$w_k\left(1, m, \binom{m}{2}\right) = \frac{k^2+k+2}{2} + m(k+1) + \frac{m(m-1)}{2}$$

$$= \frac{k^2+2mk+m^2+k+m+2}{2} = \frac{(k+m)^2+(k+m)+2}{2} \leqslant 2^k.$$

It is easy to see that starting from $(2^m, 0, 0)$ one can ask m consecutive questions each time yielding equal states. After i questions the state will be

$$\left(2^{m-i}, i\cdot 2^{m-i}, \binom{i}{2}\cdot 2^{m-i}\right) \text{ and then the query}$$

$$\left[2^{m-i-1}, i\cdot 2^{m-i-1}, \binom{i}{2}\cdot 2^{m-i-1}\right]? \text{ yields two equal states}$$

$$\left(2^{m-i-1}, (i+1)\cdot 2^{m-i-1}, \binom{i+1}{2}\cdot 2^{m-i-1}\right).$$

After m questions the state $(1, m, \binom{m}{2})$ is reached. The unique element of the truth-set is then the one which satisfies all the answers, the m elements of the one-lie-set are those which do not satisfy the first, second, ..., and mth answer, respectively, and similarly for the two-lies-set. By Theorem 2 the Questioner needs K more questions to win. Hence the total number of queries he requires is $K+m$. By the definition of K, $L=K+m$ is the least natural l such that

$$\frac{l^2+l+2}{2}\cdot 2^m \leqslant 2^{l-m}\cdot 2^m = 2^l.$$

Hence $L=\mathrm{ch}(2^m, 0, 0)$ and consequently the state $(2^m, 0, 0)$ is nice, which immediately implies Theorem 1.

The rest of this section is devoted to the proof of Theorem 2. The proof is split in a series of lemmas.

LEMMA 1. *The state* $(0, 0, n)$ *is nice for any natural* n.

Proof. Straightforward.

LEMMA 2. *The state* $(0, 1, n)$ *is nice for any natural* n.

Proof. We may assume $n \geqslant 1$. Let $\mathrm{ch}(0, 1, n)=k$. We prove the lemma by induction on n. Suppose that for $m<n$ it is true. Consider two cases.

Case 1. $n<k$.

We have $\mathrm{YES}_{0,1,n}(0, 1, 0)=(0, 1, 0)$ and $\mathrm{NO}_{0,1,n}(0, 1, 0)=(0, 0, n+1)$ with $w_{k-1}(0, 1, 0) \geqslant w_{k-1}(0, 0, n+1)$. However, $(0, 1, 0)$ is already the Questioner's win and

$$w_{k-1}(0, 0, n+1) \leqslant 2^{k-1},$$

hence starting from this state the Questioner wins in $k-1$ questions. It follows that in the considered case the Questioner wins in $\mathrm{ch}(0, 1, n)$ questions starting from $(0, 1, n)$, and hence this state is nice.

Case 2. $n \geqslant k$.

In this case the state $(0, 1, n)$ is balanced. Indeed, the query $[0, 1, x]$? yields states $(0, 1, x)$ and $(0, 0, n-x+1)$. For $x=0$ we have

$$w_{k-1}(\mathrm{YES}_{0,1,n}(0, 1, x)) = k \leqslant n+1 = w_{k-1}(\mathrm{NO}_{0,1,n}(0, 1, x)).$$

For $x=n$,

$$w_{k-1}(\mathrm{YES}_{0,1,n}(0, 1, x)) = k+n > 1 = w_{k-1}(\mathrm{NO}_{0,1,n}(0, 1, x)).$$

Hence for some $0 \leqslant x < n$, we get

$$|w_{k-1}(\mathrm{YES}_{0,1,n}(0, 1, x)) - w_{k-1}(\mathrm{NO}_{0,1,n}(0, 1, x))| \leqslant 1.$$

This implies $\mathrm{ch}(\mathrm{YES}_{0,1,n}(0, 1, x))$, $\mathrm{ch}(\mathrm{NO}_{0,1,n}(0, 1, x)) \leqslant k-1$. Moreover, the state $(0, 1, x)$ is nice by the inductive hypothesis and the state $(0, 0, n-x+1)$ is nice by Lemma 1. Hence the Questioner wins in $k-1$ questions starting from each of these states, which means that he wins in k questions starting from $(0, 1, n)$.

This proves the lemma in the second case, and hence it concludes the entire proof.

LEMMA 3. *The state* $(1, 0, n)$ *is nice for any natural* n.

Proof. We may assume $n \geqslant 1$. Let $\mathrm{ch}(1, 0, n) = k$. We prove the lemma by induction on n. Suppose that for $m < n$ it is true. We have

$$\mathrm{YES}_{1,0,n}(1, 0, 0) = (1, 0, 0)$$

and

$$\mathrm{NO}_{1,0,n}(1, 0, 0) = (0, 1, n).$$

Two cases are possible:

1. If $w_{k-1}(1, 0, 0) > w_{k-1}(0, 1, n)$ then $w_{k-1}(0, 1, n) \leqslant 2^{k-1}$. The state $(1, 0, 0)$ is the Questioner's win and the state $(0, 1, n)$ is nice by the previous lemma. It follows that the Questioner can win in $k-1$ questions starting from it. Hence he can also win in k questions starting from $(1, 0, n)$, which proves that this state is nice in the case we considered.

2. If $w_{k-1}(1, 0, 0) \leqslant w_{k-1}(0, 1, n)$ then $(1, 0, n)$ is balanced: the query $[1, 0, x]$? yields states $(1, 0, x)$ and $(0, 1, n-x)$; for $x=0$ we have

$$w_{k-1}(\mathrm{YES}_{1,0,n}(1, 0, n)) \leqslant w_{k-1}(\mathrm{NO}_{1,0,n}(1, 0, x))$$

and for $x=n$,

$$w_{k-1}(\mathrm{YES}_{1,0,n}(1,0,x))=w_{k-1}(1,0,n)>w_{k-1}(0,1,0)$$
$$=w_{k-1}(\mathrm{NO}_{1,0,n}(1,0,x)).$$

Hence for some $0\leqslant x<n$, we get

$$|w_{k-1}(\mathrm{YES}_{1,0,n}(1,0,x))-w_{k-1}(\mathrm{NO}_{1,0,n}(1,0,x))|\leqslant 1,$$

which implies $\mathrm{ch}(\mathrm{YES}_{1,0,n}(1,0,x))$, $\mathrm{ch}(\mathrm{NO}_{1,0,n}(1,0,x))\leqslant k-1$.

Moreover, the state $(1, 0, x)$ is nice by the inductive hypothesis and the state $(0, 1, n-x)$ is nice by Lemma 2. Hence the Questioner wins in $k-1$ questions starting from each of these states, which means that he wins in k questions starting from $(0, 1, n)$.

This completes the analysis of the second case and concludes the proof of the lemma.

LEMMA 4. *The state* $(0, a, n)$ *is nice for any natural n and any even a.*

Proof. It was proved in Pelc [3] that for $b\geqslant a-1$ any state $(0, a, b)$ is nice. (Note that in the abovementioned paper the case of one lie was considered and hence states were defined as couples of natural numbers, the state (a, b) corresponding to $(0, a, b)$ in our present terminology). Let $a=2c$ be any even number, n any natural number, and $k=\mathrm{ch}(0, a, n)$. We have

$$\mathrm{YES}_{0,a,n}\left(0,c,\left[\frac{n}{2}\right]\right)=\left(0,c,c+\left[\frac{n}{2}\right]\right)$$

and

$$\mathrm{NO}_{0,a,n}\left(0,c,\left[\frac{n}{2}\right]\right)=\left(0,c,c+n-\left[\frac{n}{2}\right]\right).$$

($[\frac{n}{2}]$ denotes the integer part of $\frac{n}{2}$.) Hence the state $(0, a, n)$ is balanced:

$$\left|w_{k-1}\left(\mathrm{YES}_{0,a,n}\left(0,c,\left[\frac{n}{2}\right]\right)\right)-w_{k-1}\Delta\left(\mathrm{NO}_{0,a,n}\left(0,c,\left[\frac{n}{2}\right]\right)\right)\right|\leqslant 1.$$

This implies

$$\mathrm{ch}\left(\mathrm{YES}_{0,a,n}\left(0,c,\left[\frac{n}{2}\right]\right)\right),\ \mathrm{ch}\left(\mathrm{NO}_{0,a,n}\left(0,c,\left[\frac{n}{2}\right]\right)\right)\leqslant k-1.$$

Since both obtained states are of the form $(0, c, d)$ with $d\geqslant c-1$, they are nice and hence the Questioner wins in $k-1$ queries starting from any of them. Consequently he wins in k queries starting from $(0, a, n)$, which implies that this state is nice.

LEMMA 5. *The state $(1, 1, n)$ is nice for any natural n.*

Proof. We first observe two simple facts:

$$\mathrm{ch}(1, 1, n) \geqslant 4 \qquad \text{for any } n. \tag{$*$}$$

Indeed,

$$w_k(1, 1, n) = \frac{k^2+k+2}{2} + k + 1 + n$$

and for $k < 4$, we have

$$\frac{k^2+k+2}{2} + k + 1 > 2^k.$$

$$w_{l-1}(0, 2, 0) \leqslant w_{l-1}(1, 0, n+1) \qquad \text{for } l = \mathrm{ch}(1, 1, n) \text{ and any } n. \tag{$**$}$$

Indeed, we have $l \geqslant 4$, hence $5l - 4 \leqslant l^2$, which gives

$$4l \leqslant l^2 - l + 4;$$

i.e.,

$$\begin{aligned} w_{l-1}(0, 2, 0) = 2l &\leqslant \frac{(l-1)^2 + (l-1) + 2}{2} + 1 \\ &\leqslant \frac{(l-1)^2 + (l-1) + 2}{2} + (n+1) = w_{l-1}(1, 0, n+1). \end{aligned}$$

The lemma will be proved by induction on n. Suppose it is true for $m < n$.

Consider two cases.

Case 1. $w_{l-1}(1, 0, 1) > w_{l-1}(0, 2, n)$.

The query $[1, 0, 0]$? asked in the state $(1, 1, n)$ yields states $(1, 0, 1)$ and $(0, 2, n)$. Since $w_{l-1}(0, 2, n) \leqslant 2^{l-1}$ and the state $(0, 2, n)$ is nice by Lemma 4, the Questioner wins in $l-1$ questions starting from this state.

Next we turn attention to the state $(1, 0, 1)$. The following holds:

$\mathrm{YES}_{1,0,1}(1, 0, 0) = (1, 0, 0)$—an immediate win,

$\mathrm{NO}_{1,0,1}(1, 0, 0) = (0, 1, 1)$ and then

$\mathrm{YES}_{0,1,1}(0, 1, 0) = (0, 1, 0)$—an immediate win,

$\mathrm{NO}_{0,1,1}(0, 1, 0) = (0, 0, 2)$—a win after the next question.

This shows that starting from $(1, 0, 1)$ the Questioner wins in at most 3 questions. Hence starting from $(1, 1, n)$ he always wins in $\max(l, 4) = l$ questions, which shows that $(1, 1, n)$ is nice in Case 1.

Case 2. $w_{l-1}(1, 0, 1) \leqslant w_{l-1}(0, 2, n)$.

The query $[1, 0, x]$? asked in the state $(1, 1, n)$ yields states $(1, 0, x+1)$ and $(0, 2, n-x)$. The inequalities

$$w_{l-1}(1, 0, 1) \leqslant w_{l-1}(0, 2, n)$$

and

$$w_{l-1}(0, 2, 0) \leqslant w_{l-1}(1, 0, n+1)$$

imply that for $x = 0$ we have

$$w_{l-1}(\mathrm{YES}_{1,1,n}(1, 0, x)) \leqslant w_{l-1}(\mathrm{NO}_{1,1,n}(1, 0, x))$$

and for $x = n$,

$$w_{l-1}(\mathrm{YES}_{1,1,n}(1, 0, x)) \geqslant w_{l-1}(\mathrm{NO}_{1,1,n}(1, 0, x)).$$

It follows that for some $0 \leqslant x \leqslant n$,

$$|w_{l-1}(\mathrm{YES}_{1,1,n}(1, 0, x)) - w_{l-1}(\mathrm{NO}_{1,1,n}(1, 0, x))| \leqslant 1.$$

Hence $\mathrm{ch}(\mathrm{YES}_{1,1,n}(1, 0, x))$, $\mathrm{ch}(\mathrm{NO}_{1,1,n}(1, 0, x)) \leqslant l-1$.

Since the state $(1, 0, x+1)$ is nice by Lemma 3 and the state $(0, 2, n-x)$ is nice by Lemma 4, it follows that the Questioner wins in $l-1$ questions starting from any of them, hence he wins in l questions starting from $(1, 1, n)$, which proves that this state is nice in Case 2 as well.

Now we are able to present the final part of the proof of Theorem 2.

LEMMA 6. *The state* $(1, m, \binom{m}{2})$ *is nice for any* $3 \leqslant m \leqslant 29$.

Proof. First note that $\mathrm{ch}(1, m, \binom{m}{2})$ equals: 6 for $3 \leqslant m \leqslant 4$, 7 for $5 \leqslant m \leqslant 8$, 8 for $9 \leqslant m \leqslant 14$, 9 for $15 \leqslant m \leqslant 22$, and 10 for $23 \leqslant m \leqslant 29$. Clearly, if $a \leqslant a_1$, $b \leqslant b_1$, and $c \leqslant c_1$, the minimal number of questions needed for the Questioner's win starting from state (a, b, c) cannot exceed that starting from state (a_1, b_1, c_1). Hence it suffices to show that the states $(1, m, \binom{m}{2})$ are nice for $m = 4$, 8, 14, 22, 29.

$m = 4$. $\mathrm{ch}(1, 4, 6) = 6$. The question $[1, 1, 0]$? in this state yields states $\mathrm{YES}_{1,4,6}(1, 1, 0) = (1, 1, 3)$ and $\mathrm{NO}_{1,4,6}(1, 1, 0) = (0, 4, 7)$, both of character 5. The first is nice by Lemma 5 and the second by Lemma 4. Hence $(1, 4, 6)$ is nice.

$m=8$. $\mathrm{ch}(1, 8, 28)=7$. The question $[1, 1, 28]$? in this state yields states $\mathrm{YES}_{1,8,28}(1, 1, 28)=(1, 1, 35)$ and $\mathrm{NO}_{1,8,28}(1, 1, 28)=(0, 8, 1)$, both of character 6. The first is nice by Lemma 5 and the second by Lemma 4. Hence $(1, 8, 28)$ is nice.

$m=14$. $\mathrm{ch}(1, 14, 91)=8$. The question $[1, 1, 78]$? in this state yields states $\mathrm{YES}_{1,14,91}(1, 1, 78)=(1, 1, 91)$ and $\mathrm{NO}_{1,14,91}(1, 1, 78)=(0, 14, 14)$, both of character 7. The first is nice by Lemma 5 and the second by Lemma 4. Hence $(1, 14, 91)$ is nice.

$m=22$. $\mathrm{ch}(1, 22, 231)=9$. The question $[1, 1, 189]$? in this state yields states $\mathrm{YES}_{1,22,231}(1, 1, 189)=(1, 1, 210)$ and $\mathrm{NO}_{1,22,231}(1, 1, 189)=(0, 22, 43)$, both of character 8. The first is nice by Lemma 5 and the second by Lemma 4. Hence $(1, 22, 231)$ is nice.

$m=29$. $\mathrm{ch}(1, 29, 406)=10$. The question $[1, 0, 406]$? in this state yields states $\mathrm{YES}_{1,29,406}(1, 0, 406)=(1, 0, 435)$ and $\mathrm{NO}_{1,29,406}(1, 0, 406)=(0, 30, 0)$, both of character 9. The first is nice by Lemma 3 and the second by Lemma 4. Hence $(1, 29, 406)$ is nice. This concludes the proof of the lemma.

LEMMA 7. *The state $(1, m, \binom{m}{2})$ is nice for any $m \geqslant 30$.*

First observe that for $m \geqslant 30$ we have

$$\frac{16m^2}{9}+\frac{4m}{3}+2 \leqslant 2^{m/3+1},$$

which implies

$$\frac{(m/3)^2+m/3+2}{2}+m\left(\frac{m}{3}+1\right)+\binom{m}{2} \leqslant 2^{m/3}$$

and this in turn gives

$$\mathrm{ch}\left(1, m, \binom{m}{2}\right) \leqslant \frac{m}{3}.$$

Now let $\mathrm{ch}(1, m, \binom{m}{2})=k+1$. We prove two inequalities:

(a) $(k^2+k+2)/2+(k+1) \leqslant m(k+1)$

(b) $(k^2+k+2)/2+\binom{m}{2}+m \geqslant (m+1)(k+1)$.

For (a) we have: $k \geqslant 3$; hence $k+3 \leqslant 2k \leqslant 2m$, which gives

$$k^2+3k+4 \leqslant k^2+4k+3=(k+1)(k+3) \leqslant (k+1)\,2m.$$

This implies

$$k^2+k+2 \leqslant 2mk+2m-2k-2=2(m-1)(k+1)$$

and, finally,

$$\frac{k^2+k+2}{2}+(k+1) \leqslant m(k+1).$$

For (b) we have: $k \leqslant m/3$; hence

$$(2m+1)k \leqslant (2m+1)\frac{m}{3}=\frac{2m^2}{3}+\frac{m}{3} \leqslant m^2-m \leqslant k^2+m^2-m,$$

which implies

$$k^2+k+2+m(m-1)+2m \geqslant 2mk+2k+2m+2$$

and, finally,

$$\frac{k^2+k+2}{2}+\binom{m}{2}+m \geqslant (m+1)(k+1).$$

In order to prove that $(1, m, \binom{m}{2})$ is nice, consider two cases:

Case 1. m is even. The query $[1, 1, x]$? asked in the state $(1, m, \binom{m}{2})$ yields states $(1, 1, x+m-1)$ and $(0, m, \binom{m}{2}-x+1)$. In view of inequality (a) we get

$$\begin{aligned} w_k(1, 1, m-1) &= \frac{k^2+k+2}{2}+(k+1)+(m-1) \\ &\leqslant m(k+1)+(m-1) \\ &\leqslant m(k+1)+\binom{m}{2}+1 = w_k\left(0, m, \binom{m}{2}+1\right) \end{aligned}$$

and inequality (b) implies

$$\begin{aligned} w_k\left(1, 1, \binom{m}{2}+m-1\right) &= \frac{k^2+k+2}{2}+(k+1)+\binom{m}{2}+m-1 \\ &\geqslant m(k+1)+1 = w_k(0, m, 1). \end{aligned}$$

Hence for $x=0$ we have

$$w_k(\mathrm{YES}_{1,m,\binom{m}{2}}(1, 1, x)) \leqslant w_k(\mathrm{NO}_{1,m,\binom{m}{2}}(1, 1, x))$$

and for $x=\binom{m}{2}$

$$w_k(\mathrm{YES}_{1,m,\binom{m}{2}}(1,1,x)) \geqslant w_k(\mathrm{NO}_{1,m,\binom{m}{2}}(1,1,x)).$$

It follows that the state $(1, m, \binom{m}{2})$ is balanced. Indeed, for some $0 \leqslant x \leqslant \binom{m}{2}$, we have

$$|w_k(\mathrm{YES}_{1,m,\binom{m}{2}}(1,1,x)) - w_k(\mathrm{NO}_{1,m,\binom{m}{2}}(1,1,x))| \leqslant 1.$$

Hence $\mathrm{ch}(\mathrm{YES}_{1,m,\binom{m}{2}}(1,1,x))$, $\mathrm{ch}(\mathrm{NO}_{1,m,\binom{m}{2}}(1,1,x)) \leqslant k$.

Since the state $(1, 1, x+m-1)$ is nice by Lemma 5 and the state $(0, m, \binom{m}{2}-x+1)$ is nice by Lemma 4 for even m, we conclude that $(1, m, \binom{m}{2})$ is also nice in this case.

Case 2. m is odd. The query $[1, 0, x]$? asked in the state $(1, m, \binom{m}{2})$ yields states $(1, 0, x+m)$ and $(0, m+1, \binom{m}{2}-x)$. In view of inequality (a) we have

$$w_k(1,0,m) = \frac{k^2+k+2}{2} + m \leqslant (m+1)(k+1) + \binom{m}{2} = w_k\left(0, m+1, \binom{m}{2}\right)$$

(note that for $m \geqslant 30$ the inequality $m \leqslant \binom{m}{2}$ holds) and inequality (b) implies

$$\begin{aligned} w_k\left(1, 0, \binom{m}{2}+m\right) &= \frac{k^2+k+2}{2} + \binom{m}{2} + m \geqslant (m+1)(k+1) \\ &= w_k(0, m+1, 0). \end{aligned}$$

Hence for $x=0$ we have

$$w_k(\mathrm{YES}_{1,m,\binom{m}{2}}(1,0,x)) \leqslant w_k(\mathrm{NO}_{1,m,\binom{m}{2}}(1,0,x))$$

and for $x=\binom{m}{2}$

$$w_k(\mathrm{YES}_{1,m,\binom{m}{2}}(1,0,x)) \geqslant w_k(\mathrm{NO}_{1,m,\binom{m}{2}}(1,0,x)).$$

It follows that the state $(1, m, \binom{m}{2})$ is balanced. Indeed, for some $0 \leqslant x \leqslant \binom{m}{2}$ we have

$$|w_k(\mathrm{YES}_{1,m,\binom{m}{2}}(1,0,x)) - w_k(\mathrm{NO}_{1,m,\binom{m}{2}}(1,0,x))| \leqslant 1.$$

Hence $\mathrm{ch}(\mathrm{YES}_{1,m,\binom{m}{2}}(1,0,x))$, $\mathrm{ch}(\mathrm{NO}_{1,m,\binom{m}{2}}(1,0,x)) \leqslant k$.

Now the state $(1, 0, x+m)$ is nice by Lemma 3 and the state $(0, m+1, \binom{m}{2}-x)$ is nice by Lemma 4 ($m+1$ being even in this case), hence $(1, m, \binom{m}{2})$ is nice for odd m as well. This completes the proof of Lemma 7 and the proof of Theorem 2.

The following corollary yields the solution of the second part of Ulam's original question, obtained via a different argument in [2].

COROLLARY. *The number of yes–no queries necessary and sufficient to find an unknown number in the set* $\{1, ..., 10^6\}$*—if the Responder may lie at most twice—is* 29.

Proof. The least k such that

$$\frac{k^2+k+2}{2}\cdot 2^n \leqslant 2^k$$

is 29 for $n = 10^6$ and for $n = 2^{20}$.

By Rivest's lower bound the Questioner does not win in less than 29 questions when searching $\{1, ..., 10^6\}$. By Theorem 1, 29 questions are sufficient for the set $\{1, ..., 2^{20}\}$. Since $10^6 < 2^{20}$, it follows that 29 is the least number of questions for the set $\{1, ..., 10^6\}$.

4. MISCELLANEOUS REMARKS

1. For $m = 1$ both Theorems 1 and 2 remain true but for $m = 2$ both become false. Indeed, for $m = 1$ we have to consider the state $(1, 1, 0)$ which is nice by Lemma 5. On the other hand, for $m = 2$ we have the state $(1, 2, 1)$ whose character is 5. The following proposition shows that this state is not nice.

PROPOSITION. *The minimal number of questions yielding the Questioner's win starting from the state* $(1, 2, 1)$ *is* $k = 6$.

Proof. We first show that five questions are not sufficient starting from state $(1, 2, 1)$. Without loss of generality we may assume that the first question is $[0, y, z]$? with $y \leqslant 2$ and $z \leqslant 1$. Hence $\mathrm{YES}_{1,2,1}(0, y, z) = (0, 1 + y, 2 - y + z)$ and $\mathrm{NO}_{1,2,1}(0, y, z) = (1, 2 - y, y + 1 - z)$. We have

$$w_4(0, 1 + y, 2 - y + z) = 5(1 + y) + 2 - y + z = 4y + z + 7$$

and

$$w_4(1, 2 - y, y + 1 - z) = 11 + 5(2 - y) + y + 1 - z = 22 - 4y - z.$$

If $y \leqslant 1$ then $22 - 4y - z \geqslant 17$, hence four questions are not sufficient starting from state $\mathrm{NO}_{1,2,1}(0, y, z)$. If $y > 1$ (that is, $y = 2$) then $\mathrm{YES}_{1,2,1}(0, y, z) = (0, 3, z)$. It follows from [3] that the least number of questions yielding the Questioner's win starting from $(0, 3, 0)$ is five. Hence four questions cannot suffice starting from $(0, 3, z)$.

This shows that five questions are not sufficient starting from state $(1, 2, 1)$. On the other hand, it is easy to show that six questions suffice. Consider the question $[1, 0, 1]$? We get $\mathrm{YES}_{1,2,1}(1, 0, 1) = (1, 0, 3)$ and $\mathrm{NO}_{1,2,1}(1, 0, 1) = (0, 3, 0)$. By Lemma 3 the state $(1, 0, 3)$ is nice. Since $\mathrm{ch}(1, 0, 3) = 4$, four questions suffice starting from this state. As remarked before, five questions suffice starting from state $(0, 3, 0)$, hence six questions suffice starting from state $(1, 2, 1)$. This concludes the proof.

It follows that the number of yes–no queries necessary and sufficient to find an unknown number in the set $\{1, 2, 3, 4\}$—if the Responder may lie at most twice—is 8.

2. It should be noted that in Rivest *et al.* [4] and Spencer [5] the lower bound for the minimal number of questions is computed allowing arbitrary yes–no queries (i.e., queries of the type "$x \in T$?" for any subset T of $\{1, ..., n\}$) whereas the upper bound is computed using only comparison queries (i.e., queries of the type "$x < a$?" for $a \in \{1, ..., n\}$). We worked with arbitrary yes–no queries following Ulam's original problem.

REFERENCES

1. E. R. BERLEKAMP, Block coding for the binary symmetric channel with noiseless, delayless feedback, *in* "Error-Correcting Codes," pp. 61–85, Wiley, New York, 1968.
2. J. CZYZOWICZ, D. MUNDICI, AND A. PELC, Solution of Ulam's problem on binary search with two lies, *J. Combin. Theory Ser. A* **49** (1988), 384–388.
3. A. PELC, Solution of Ulam's problem on searching with a lie, *J. Combin. Theory Ser. A* **44** (1987), 129–140.
4. R. L. RIVEST, A. R. MEYER, D. J. KLEITMAN, K. WINKLMANN, AND J. SPENCER, Coping with errors in binary search procedures, *J. Comput. System Sci.* **20** (1980), 396–404.
5. J. SPENCER, Guess a Number—with Lying, *Math. Mag.* **57**, No. 2 (1984), 105–108.
6. S. M. ULAM, "Adventures of a Mathematician," p. 281, Scribner, New York, 1976.