JOURNAL OF COMBINATORIAL THEORY, Series A 59, 149–154 (1992)

Note

Solution of Ulam's Problem on Binary Search with Three Lies*

ALBERTO NEGRO

Dipartimento di Informatica ed Applicazioni, Università di Salerno, 84084 Baronissi (SA), Italy

AND

MATTEO SERENO

Dipartimento di Informatica, Università di Torino, Corso Svizzera, 185, 10149 Turin, Italy

Communicated by A. Barlotti

Received March 5, 1990

In this paper we determine the minimal number of yes–no queries needed to find an unknown integer between 1 and 1000000 if at most three of the answers may be erroneous.

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 $\log_2(1000000)$. Now suppose one were allowed to lie once, or twice, then how many questions would one need to get the right answer?

We prove that Ulam's problem has the following solution:

THEOREM. *Thirty-three is the least number of yes–no questions sufficient to find an element* $e \in \{1, 2, ..., 1000000\}$, *if up to three lies are allowed.*

* This work is supported in part by the "Ministero dell'Università e della Ricerca Scientifica" 40% funds "Analisi di Algoritmi."

582a/59/1-11

0097-3165/92 $3.00

The problem with at most one lie was solved in [4], with at most two lies was solved in [2, 3]. For our analysis we use the same terminology used in [2, 3]. A game is considered between two players: the Questioner and the Responder. The Responder chooses an element $e \in \{1, 2, ..., 1000000\}$ unknown to the Questioner who has to find it with queries of form "$e \in Q$?" for $Q \subset \{1, 2, ..., 1000000\}$. The Responder may lie at most three times. Our interest is focused on the minimal number q of queries the Questioner needs to find the unknown element. Every question modifies the state of the Questioner knowledge. Suppose the n-tuple $Q = Q_1, Q_2, ..., Q_n$ of yes–no questions has already been answered. The state of the Questioner knowledge can be summarized by the unique quadruple (A, B, C, D) of subsets of $\{1, 2, ..., 1000000\}$ with the following properties:

— $e \in A$ iff none of the answers is a lie;

— $e \in B$ iff exactly one of the answers is a lie;

— $e \in C$ iff exactly two of the answers are lies;

— $e \in D$ iff exactly three of the answers are lies.

Now assume that one more yes–no question is asked. This question has the form "$e \in Q$?" where $Q = X \cup Y \cup Z \cup K$, and $X \subseteq A$, $Y \subseteq B$, $Z \subseteq C$, $K \subseteq D$. A positive answer to Q transforms (A, B, C, D) into the quadruple:

$$(A, B, C, D)\, Q^{\text{yes}} = (A \cap Q, (B \cap Q) \cup (A \cap \bar{Q}), (C \cap Q) \cup (B \cap \bar{Q}), (D \cap Q) \cup (C \cap \bar{Q})).$$

A negative answer to Q has the same effect as a positive answer to $\bar{Q}$.

The sets in the initial quadruple $(\{1, 2, ..., 1000000\}, \varnothing, \varnothing, \varnothing)$, where $\varnothing$ is the empty set, are pairwise disjoint; furthermore, if the sets (A, B, C, D) are pairwise disjoint, then so are the sets in $(A, B, C, D)\, Q^{\text{yes}}$ and $(A, B, C, D)\, Q^{\text{no}}$.

DEFINITION 1. An *Ulam set* (with at most three lies) is a quadruple $U = (A, B, C, D)$ of pairwise disjoint finite subsets of $\mathcal{N}$. A yes–no question is a subset Q of $\mathcal{N}$. The Ulam sets UQ^{yes} and UQ^{no} are defined as

$$UQ^{\text{yes}} = (A \cap Q, (B \cap Q) \cup (A \cap \bar{Q}), (C \cap Q) \cup (B \cap \bar{Q}), (D \cap Q) \cup (C \cap \bar{Q})),$$

$$UQ^{\text{no}} = (A \cap \bar{Q}, (B \cap \bar{Q}) \cup (A \cap Q), (C \cap \bar{Q}) \cup (B \cap Q), (D \cap \bar{Q}) \cup (C \cap Q)),$$

where $\bar{Q} = \mathcal{N} \setminus Q$.

For the sake of clearness sometimes we use cardinalities instead of sets. In particular, we denote an Ulam set as

$$U = (a, b, c, d),$$

where $a = |A|$, $b = |B|$, $c = |C|$, and $d = |D|$. Moreover, if $Q = X \cup Y \cup Z \cup K$ is the set involved in the yes–no question we can say

$$Q = (x, y, z, k)$$

and

$$UQ^{\text{yes}} = (x, a - x + y, b - y + z, c - z + k),$$
$$UQ^{\text{no}} = (a - x, x + b - y, y + c - z, z + d - k).$$

DEFINITION 2. An *n-solvable* Ulam set U is inductively defined as follows:

— U is 0-solvable iff $A \cup B \cup C \cup D$ contains at most one element;

— U is $(n+1)$-solvable iff there is a yes–no question such that both UQ^{yes} and UQ^{no} are n-solvable.

Following Berlekamp's idea [1] we define the weight of each state $U = (A, B, C, D)$ as follows:

DEFINITION 3. Let $U = (A, B, C, D)$ be the Ulam set when q questions remains to be asked. The *weight* of U is

$$w_q(U) = |A| \left(\left(\tbinom{q}{3}\right)\right) + |B| \left(\left(\tbinom{q}{2}\right)\right) + |C| \left(\left(\tbinom{q}{1}\right)\right) + |D| \left(\left(\tbinom{q}{0}\right)\right),$$

where $\left(\left(\tbinom{n}{m}\right)\right) = \sum_{i=0}^{m} \tbinom{n}{i}$.

PROPOSITION 1. *Let $S \neq \varnothing$, $S \subset \mathcal{N}$, and $n \in \mathcal{N}$. Then the following statements are equivalent:*

1. *n yes–no questions are sufficient to find $e \in S$, if up to three lies are allowed;*

2. *the Ulam set $(S, \varnothing, \varnothing, \varnothing)$ is n-solvable.*

Proof. By induction on n. ∎

PROPOSITION 2. *Let $U = (A, B, C, D)$ be an Ulam set and $n \in \mathcal{N}$. If U is n-solvable then:*

1. *U is $(n+1)$-solvable;*

2. $2^n \geqslant w_n(U)$;

3. *if $U' = (A', B', C', D')$ is another Ulam set, and $A' \subset A$, $B' \subset B$, $C' \subset C$, $D' \subset D$ then U' is solvable too.*

Proof. (1) By induction on n. If $n = 0$, then by definition $|A \cup B \cup C \cup D| < 1$. If we choose the question $Q = \mathcal{N}$ then $UQ^{\text{yes}} =$

(A, B, C, D), and $UQ^{\text{no}} = (\varnothing, A, B, C)$ which are both 0-solvable. For the induction step we can use the same technique.

(2) By induction on n. The case $n = 0$ is trivial. Assume that U is $(n+1)$-solvable, and let $Q \subset \mathcal{N}$ be a yes–no question such that both UQ^{yes} and UQ^{no} are n-solvable. By the inductive hypothesis we have $2^n \geqslant w_n(UQ^{\text{yes}})$ and $2^n \geqslant w_n(UQ^{\text{no}})$. Since the Ulam sets are disjoint, and using $(\!(\binom{n+1}{m})\!) = (\!(\binom{n}{m})\!) + (\!(\binom{n}{m-1})\!)$ we have $2^{n+1} \geqslant w_{n+1}(U)$.

(3) The proof can be found in [2]. ∎

PROPOSITION 3. *Let $U = (A, B, C, D)$ be an Ulam set with $A = \phi$, $|B| = 1$, $|C| = m$, $|D| = \binom{m}{2}$ with $m \geqslant 3$. Then U is ch-solvable, where $ch = \min_i \{i : w_i(U) \leqslant 2^i\}$.*

Proof. This proof can be found in [3] substituting the triplet of subsets (B, C, D) with the quadruplet $(\varnothing, B, C, D)$. ∎

PROPOSITION 4. *Let $U_n = (A_n, B_n, C_n, D_n)$ be an Ulam set, with A_n, B_n, C_n, D_n pairwise disjoint, where $A_n = \phi$, $|B_n| = 2^n$, $|C_n| = (8-n)\,2^n$, $|D_n| = \binom{8-n}{2}\,2^n$ with $n \geqslant 0$. Then U_n is $(7+n)$-solvable.*

Proof. By induction on n. For $n = 0$ $U_0 = (\varnothing, \{b\}, \{c_1, ..., c_8\}, \{d_1, ..., d_{\binom{s}{2}}\})$ this is 7-solvable by Proposition 3, because $ch = 7$ with $m = 8$.

Induction step. Let $n+1 \leqslant 8$, $s = 2^{n+1}$, $t = (8-(n+1))\,2^{n+1}$, and $v = \binom{8-(n+1)}{2}\,2^{n+1}$ then $U_{n+1} = (\phi, \{b_1, ..., b_s\}, \{c_1, ..., b_t\}, \{d_1, ..., d_v\})$.

If $Q = \{b_1, ..., b_{s/2}, c_1, ..., c_{t/2}, d_1, ..., d_{v/2}\}$ then $U_{n+1}Q^{\text{yes}} = (\varnothing, \{b_1, ..., b_{s/2}\}, \{c_1, ..., c_{t/2}, b_{s/2+1}, ..., b_s\}, \{d_1, ..., d_{v/2}, c_{t/2+1}, ..., c_t\}) = (A_n, B_n, C_n, D_n)$ and $A_n = \phi$, $|B_n| = s/2$, $|C_n| = (8-n)\,2^n$, $|D_n| = \binom{8-n}{2}\,2^n$.

Symmetrically we can write $U_{n+1}Q^{\text{no}} = (A'_n, B'_n, C'_n, D'_n)$, where $A'_n = A_n = \phi$, $|B'_n| = |B_n|$, $|C'_n| = |C_n|$, $|D'_n| = |D_n|$. $U_{n+1}Q^{\text{no}}$ is then $(7+n)$-solvable. Therefore, U_{n+1} is $(7+(n+1))$-solvable, as required. ∎

PROPOSITION 5. *Let $U_n = (A_n, B_n, C_n, D_n)$ be an Ulam set, with A_n, B_n, C_n, D_n pairwise disjoint, where $A = \phi$, $|B| = 2^n$, $|C| = (14-n)\,2^n$, $|D| = \binom{14-n}{2}\,2^n$ with $n \geqslant 0$. Then U_n is $(8+n)$-solvable.*

Proof. The proof is the same as in Proposition 4. In this case $U_0 = (\varnothing, \{b\}, \{c_1, ..., c_{14}\}, \{d_1, ..., d_{\binom{14}{2}}\})$ that is 8-solvable by Proposition 3 because $ch = 8$ with $m = 14$. ∎

PROPOSITION 6. *Let $U_n = (A_n, B_n, C_n, D_n)$ be an Ulam set, with A'_n, B_n, C_n, D_n pairwise disjoint, where $A = \varnothing$, $|B| = 2^n$, $|C| = (11-n)\,2^n$, $|D| = \binom{22-n}{2}\,2^n$ with $n \geqslant 0$. Then U_n is $(9+n)$-solvable.*

Proof. The proof is the same as in Propositions 4 and 5. In this case $U_0 = (\varnothing, \{b\}, \{c_1, \ldots, c_{22}\}, \{d_1, \ldots, c_{\binom{22}{2}}\})$ that is 9-solvable by Proposition 3 because $ch = 9$ with $m = 22$. ∎

PROPOSITION 7. *Let* $U = (A, B, C, D)$ *be where* A, B, C, D *are pairwise disjoint subsets of* $\mathcal{N}$ *of cardinalities* $1, 20, \binom{20}{2}, \binom{20}{3}$, *respectively.* U *is* 13*-solvable.*

Proof. The complete analysis of a Questioner's strategy is

$(1, 20, \binom{20}{2}, \binom{20}{3})$
$Q_1 = (0, 7, 139, 1140)$ —yes→ (0, 8, 152, 1191) 12-solvable by Prop. 6 ($n = 3$) and Prop. 2 (case 3)
no ↓

$(1, 13, 58, 139)$
$Q_2 = (0, 7, 58, 139)$ —yes→ (0, 8, 64, 139) 11-solvable by Prop. 5 ($n = 3$) and Prop. 2 (case 3)
no ↓

$(1, 6, 7, 58)$
$Q_3 = (0, 3, 7, 58)$ —yes→ (0, 4, 10, 58) 10-solvable by Prop. 5 ($n = 2$) and Prop. 2 (case 3)
no ↓

$(1, 3, 3, 7)$
$Q_4 = (0, 3, 3, 7)$ —yes→ (0, 4, 3, 7) 9-solvable by Prop. 4 ($n = 2$) and Prop. 2 (case 3)
no ↓

$(1, 0, 3, 3)$
$Q_5 = (0, 0, 3, 3)$ —yes→ (0, 1, 3, 3) 6-solvable by Prop. 3 (if $m = 3$ then $ch = 6$) by Prop. 2 (case 1) is 8-solvable
no ↓

$(1, 0, 0, 3)$
$Q_6 = (0, 0, 0, 3)$ —yes→ (0, 2, 0, 3) 6-solvable by Prop. 3 ($m = 3$) by Prop. 2 (case 1) is 8-solvable, too
no ↓

$(1, 0, 0, 0)$
0-solvable by definition of Ulam set
by Prop. 2 (case 1) is 7-solvable, too. ∎

Proof of theorem. By Proposition 3 (case 2) and Proposition 1, $q < 33$ questions are not sufficient to find an integer in $\{1 \cdots 1000000\}$ when up to three answers may be erroneous. The theorem can be proved using a stronger result. In fact, we will show that 33 questions are sufficient to find a number $e \in \{0, 1, \ldots, 2^{20} - 1\}$.

Suppose that $U=(2^{20}, 0, 0, 0)$. The first 20 questions can be asked in such a way that the Ulam set after the $i<20$ question $Q=(2^{20-i}, (i-1)\,2^{20-i}, \binom{i-1}{2}2^{20-i}, \binom{i-1}{3}2^{20-i})$ will be $U_i=(2^{20-i}, i2^{20-i}, \binom{i}{2}2^{20-i}, \binom{i}{3}2^{20-i})$. (Note that in this phase of the algorithm $U_{i-1}Q_i^{\text{yes}}=U_{i-1}Q_i^{\text{no}}$). After 20 questions the state $U=(1, 20, \binom{20}{2}, \binom{20}{3})$, 13-solvable by Proposition 7, is reached. ∎

ACKNOWLEDGMENT

Thanks to Professor Daniele Mundici for his encouragement.

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. J. CZYZOWICZ, D. MUNDICI, AND A. PELC, Ulam's searching game with lies, *J. Combin. Theory Ser. A* **52** (1989), 62–76.
4. A. PELC, Solution of Ulam's problem on searching with lie, *J. Combin. Theory Ser. A* **44** (1987), 129–140.
5. 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), 396–406.
6. S. M. ULAM, "Adventures of a Mathematician," p. 281, Scribner, New York, 1976.