THE RENYI–ULAM GAME

(A 20 questions game)

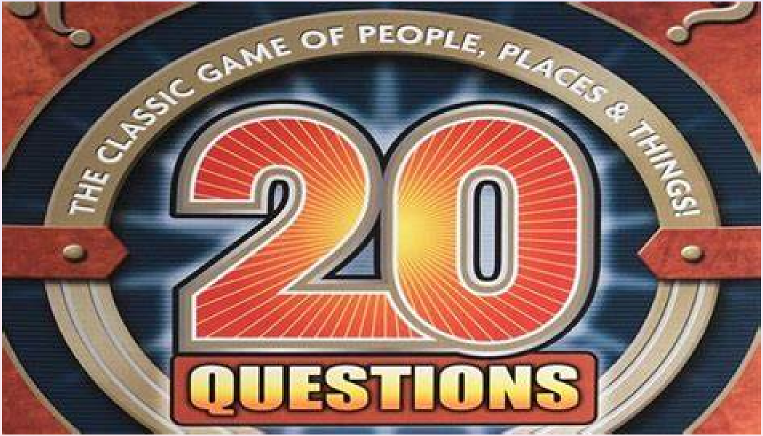
Masoud Ariannejad

Department of Mathematics, University of Zanjan

Bahar-1402

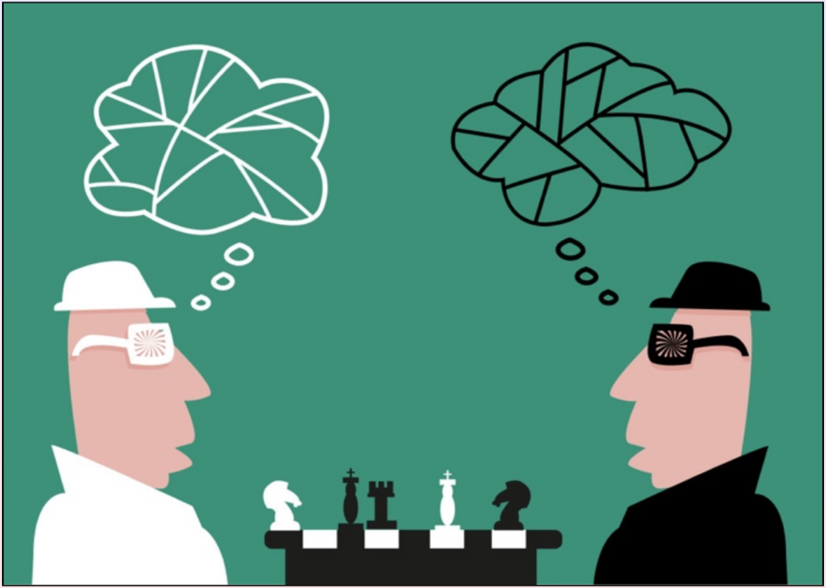
20 Questions Game

20 Questions Game



20 Questions Game

20 Questions Game



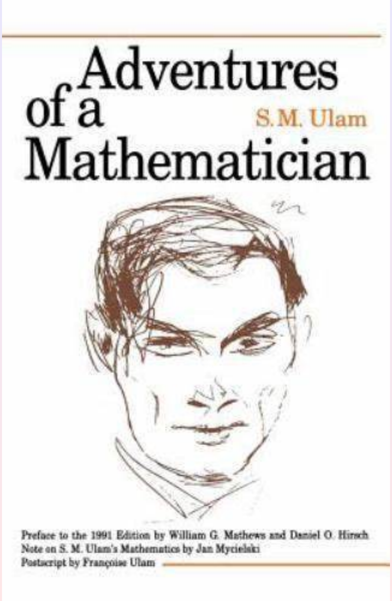
Stanislaw Ulam

Stanislaw Ulam (1909-1984), a Polish scientist, known for his works in the Manhattan Project, solving the problem of how to initiate fusion in the hydrogen bomb, inventing the Monte Carlo method of computation and deep collaboration in ergodic theory, set theory, group theory, topology, mathematical physic.

Theorem

Let $f : S^n \to \mathbb{R}^n$ be a continuous map. Then there exists $x \in S^n$ such that: $f(-x) = f(x)$.

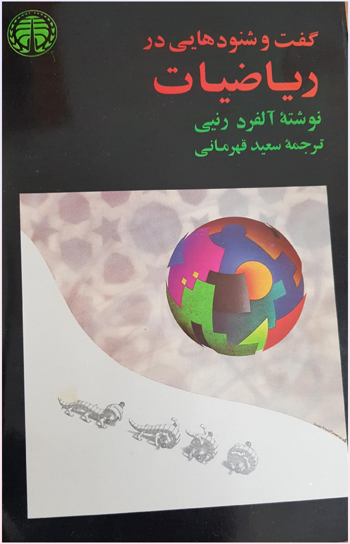
Adventures, 1976



Alfréd Rényi

Alfréd Rényi (1921-1970), a Hungarian mathematician, known for his works in probability theory, combinatorics, graph theory, and number theory.

Dialogues on Mathematics, 1965, 1373



A 20 questions game

Someone thinks of a number between one and one million. 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. Now suppose one were allowed to lie once or twice.

Obviously if there is no false answer 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 or equal to $\lceil log_2(1000000) \rceil = 20$.

Pelc contribution

JOURNAL OF COMBINATORIAL THEORY, Series A **44**, 129–140 (1987)

Solution of Ulam's Problem on Searching with a Lie

Andrzej Pelc

Département d'Informatique, Université du Québec à Hull, Case postale 1250, Succursale "B", Hull, Québec J8X 3X7, Canada

Communicated by the Managing Editors

Received March 28, 1986

ARTICLES

Coding Theory Applied to a Problem of Ulam

IVAN NIVEN
University of Oregon
Eugene, OR 97403

In this paper the theory of error-correcting codes is applied to a problem in mathematics. We add at once that no background is needed in computer science to understand the argument. The problem under consideration was stated by S. M. Ulam [1] in this way:

> 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 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?

Pelc, a survey article

ELSEVIER

Theoretical Computer Science 270 (2002) 71–109

Theoretical Computer Science

www.elsevier.com/locate/tcs

Fundamental Study

Searching games with errors—fifty years of coping with liars

Andrzej Pelc [1]

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

Received November 2000; accepted July 2001

Communicated by A.S. Fraenkel

Terminology of Pelc solution

Let the number is selected from $\{l, \cdots, n\}$. If the number has to be found after k queries of "yes" or "no" type, the game is called the $[n, k]$ game.

Two players:

1. Questioner (Q), 2. Responder (R).

• Q winning strategy is to determine provably the unknown number by least number of questions.

Terminology of Pelc solution

• R wining strategy is to give yes-no responses to Q questions in such a way that the unknown number remains uncovered up to the last possible questions.

Devil strategy The player R beside a possible lie response may adopt a devil strategy to defeat Q, that is R needs not actually think of any number at the beginning but just reply almost consistently, or in such a way that at any stage of the game there is a non-empty subset $L \subseteq \{I, \cdots, n\}$ satisfying all of his answers possibly except one.

Pelc formulation

A mathematical formulation:

State of the game

With each stage of the game, when the turn of the Questioner comes, we associate a state of the game which is a couple (a, b) of natural numbers. The first number is the size of the truth-set T, the set of those elements of $\{1, \cdots, n\}$ which satisfy all answers given previously. The second number is the size of the lie-set L, the set of those elements of $\{1, \cdots, n\}$ which satisfy all but one answer.

begining state v.s last state

$(n, 0)$ v.s. $(1, 0) or (0, 1)$ (Questioner's win)

Pelc formulation

Weight of a state

The weight of a state (a, b) at which j questions remain to be asked is defined by

$$w_j(a, b) = a(j + 1) + b.$$

This can be interpreted as:

Each number in the truth-set gives $j + 1$ possibilities of lying to each of the remaining j questions or not lying at all. In the lie-set the Responder is forced to say the truth till the end, so each number in this set yields just this one possibilty.

Pelc formulation

Asking a question

$$(a,b) \overset{\text{A question}}{\Longrightarrow} \begin{cases} (a_1, b_1) & \text{If Yes} \\ (a_2, b_2) & \text{If No} \end{cases}$$

Let T and L be the truth set and lie set of the state (a,b) and let u be the unknown number.

The question

Is the unknown number u in the subset $X \subseteq T$ of size x of the truth-set or in the subset $Y \subseteq L$ of size y of the lie-set? .

Pelc formulation

We have

$$W_j(a, b) = W_{j-1}(a_1, b_1) + W_{j-1}(a_2, b_2)$$

Why?!

If response of the qestion $u \in X$ be "yes", then $(a_1, b_1) = (x, |T \setminus X|) = (x, a_1 - x)$.

If response of the qestion $u \in X$ be "no", then $(a_2, b_2) = (|T \setminus X|, |L \cup X|) = (a_1 - x, b_1 + x)$

Pelc formulation

Devil Strategy

The Devils Strategy of the Responder consisting in always choosing the state of non-smaller weight out of the two states yielded by a query.

in other words: Devil strategy is just choosing the state (a_i, b_i) with $max\{W_{j-1}(a_1, b_1), W_{j-1}(a_2, b_2)\}$

Pelc solution

Lemma 1.

1. For even n the Responder wins the $[n, k]$ game if $n(k+1) > 2^n$.
2. For odd n the Responder wins the $[n, k]$ game if $n(k+1) + (k-1) > 2^n$.

Pelc solution

Proof(1):

Since $w_k(n,0) = n(k+1) > 2^n$ after $\leq k$ questions the weight of the resulting state will be at least 2. If there remains no possible query then the Questioner defeats and Responder wins. If there exist at least one query all we show that the last state can not be equal to (1.0). However this could occur only if the previous state was $(1,c)$, t questions remained and the question about the unique element of the truth-set yielded states $(1,0)$ and $(0,c+1)$ with $W_{t-1}(1,0) \geq W_{t-1}(0,c+1)$ or $t \geq c+1$. Since the state $(1,c)$ was reached after $k-t$ questions, we have $W_t(1,c) = t+c+1 > 2^k \cdot 2^{-(k-t)} = 2^t$. Hence $t > 2^{t-1}$, which is always false.

Pelc solution

For any state (a, b) we define its character as the number $ch(a, b) = min\{j; w_j(a, b) \leq 2^j\}$.

Lemma 2.

Let n be a natural number and $k = ch(1, n)$. The Questioner wins in at most k questions starting from the state $(1, n)$.

Pelc solution

Lemma 3.

Let (a, b) be a state such that $b \geq a - 1 \geq 1$. Then there exists a question in this state yielding states (a_1, b_1) and (a_2, b_2) such that:

1. $[\frac{a}{2}] \leq a_1 \leq [a + 12]$, $[\frac{a}{2}] \leq a_2 \leq [a + 12]$;
2. $b_1 \geq a_1 - 1$, $b_2 \geq a_2 - 1$;
3. $ch(a_1, b_1), cha(a_2, b_2) \leq ch(a, b) - 1$.

Main Result

Theorem

1. For even n the Questioner wins the $[n, k]$ game if and only if $n(k+1) \leq 2^k$.

2. For odd n the Questioner wins the $[n, k]$ game if and only if $n(k+1) + (k-1) \leq 2^k$.

Corollary

As by the case 1 above when $n = 10^6$, the minimal k for which $10^6(k+1) \leq 2^k$ is 25. Hence this is the answer to Ulams original problem: the least number k of yes-no questions sufficient to find an integer between 1 and one million, if one lie is allowed, is $k = 25$.

Niven Solution

Coding theory approach

Consider the problem with no-false-answer situation.

The questioner ask the responder to tranlates the number into binary notation which would be up to twenty binary digits. Then ask "Is the jth digit zero?" for $j = 1, 2, ..., 20$.

These questions can be asked even with decimal form of the number by asking "Is the number in the set S_j?" for $j = 1, 2, ..., 20$, considering appropriately sets: $S_1, S_2, \cdots, S_{20}$.

Niven Solution

Niven approach to this problem is that the responder does not lie but communication is through a noisy channel with up to one error in sending binary data.

.

So Niven constructs a 1-error correcting code with minimum length to determine the unknown number and to detect the lie position.

.

Niven Solution

Step One

Since $2^{19} < 10^6 < 2^{20}$, the unknown integer x in $\{1, \cdots, 10^6\}$ can be represented by up to a 20 binary digit as follows:
$x = a_1 a_2 a_3 a_4 a_5 a_6 a_7 a_8 a_9 \cdots a_{18} a_{19} a_{20}$.

Step Two

Then write respectively, these twenty digits with the following different notations:
$b_3, b_5, b_6, b_7, b_9, b_{10}, b_{11}, b_{12}, b_{13}, b_{14}, b_{15}, b_{17},$
$b_{18}, b_{19}, b_{2O}, b_{21}, b_{22}, b_{23}, b_{24}, b_{25}$.

The subscripts 1, 2, 4, 8, and 16 are not present here.

Niven Solution

Step three

Missing digits $b_1, b_2, b_4, b_8, b_{16}$ are defined by:

$$b_1 = b_3 + b_5 + b_7 + b_9 + b_{11} + b_{13} + b_{15} + b_{17} + b_{19} + b_{21} + b_{23} + b_{25}$$
$$b_2 = b_3 + b_6 + b_7 + b_{10} + b_{11} + b_{14} + b_{15} + b_{18} + b_{19} + b_{22} + b_{23}$$
$$b_4 = b_5 + b_6 + b_7 + b_{12} + b_{13} + b_{14} + b_{15} + b_{2O} + b_{20} + b_{22} + b_{23}$$
$$b_8 = b_9 + b_{10} + b_{11} + b_{12} + b_{13} + b_{14} + b_{15} + b_{24} + b_{25}$$
$$b_{16} = b_{17} + b_{18} + b_{19} + b_{20} + b_{21} + b_{22} + b_{23} + b_{24} + b_{25}$$

Niven Solution

The Questioner knows nothing about the a_i and b_j digits.
The Questioner consider a number $c = c_1c_2c_3c_4c_5 \cdots c_{23}c_{24}c_{25}$, which is constructed as follows:

Step Four

The Questions:

$$Is\ b_i = 1 \quad \overset{\text{noisy channel}}{\Longrightarrow} \quad \begin{cases} \text{if} \quad \text{Yes} & \text{let } c_i = 1 \\ \text{if} \quad \text{No} & \text{let } c_i = 0 \end{cases}$$

Niven Solution

Now the Questioner determine $b = b_1 b_2 b_3 b_4 \cdots b_{23} b_{24} b_{25}$ to find $x = a_1 a_2 a_3 \cdots a_{19} a_{20}$, but how?

First note that

• If no question is answered falsely, then $b_i = c_i$ and all five equations of page 29 satisfy for C_i too and vice versa.

• If the Responder answers one of the questions falsely, say the kth question then b and c differ only in k-position, where $b_k = c_k + 1$ and at least one of the equations of page 29 for C_i is no valid and vice versa.

Possible false position ?!

Step Five

Let S_1, S_2, S_4, S_8, and S_{16} denote respectively, the sets of subscripts in five equations of page 29, as follows:

$S_1 = \{1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25\}$,

$S_2 = \{2, 3, 6, 7, 10, 11, 14, 15, 18, 19, 22, 23\}$,

$S_4 = \{4, 5, 6, 7, 12, 13, 14, 15, 24, 25\}$,

$S_8 = \{8, 9, 10, 11, 12, 13, 14, 15, 24, 25\}$,

$S_{16} = \{16, 17, 18, 19, 20, 21, 22, 23, 24, 25\}$

Step Five

Let the binary number p_i be defined as follows:

$P_1 = \sum c_i$, where $i \in S_1$, $P_2 = \sum c_i$, where $i \in S_2$,
$P_4 = \sum c_i$, where $i \in S_4$, $P_8 = \sum c_i$, where $i \in S_8$,
$P_{16} = \sum c_i$, where $i \in S_{16}$.

Consider k as decimal representation of the binary number $p_1 p_2 p_4 p_8 p_{16}$ or $k = 16p_{16} + 8p_8 + 4P_4 + 2P_2 + P_1$.

Clearly $k \geq 0$ and its value shows the possible false position of b. This implies to determine b exactly which in turn implies to determining the unknown number a and the game is over with Questioner wining.

Thank You