ELSEVIER

Information Processing Letters 56 (1995) 89–93

Information Processing Letters

An algorithm for "Ulam's Game" and its application to error correcting codes

Eugene L. Lawler [1], Sergei Sarkissian *

Department of Electrical Engineering and Computer Sciences, Computer Science Division, University of California, Berkeley, CA 94720, USA

Received 13 June 1994; revised 23 January 1995
Communicated by F.B. Schneider

Abstract

A near-optimal algorithm for "Ulam's Game" is presented. The relationship between the game and multiple error correcting codes is discussed. For many cases, codes derived from winning strategies of the game are optimal for the communication scheme with noisy forward and noiseless feedback channels.

Keywords: Algorithms; Game tree; Error correcting codes

1. Introduction

The existence of an optimal or near-optimal strategy for playing a cooperative game known to us as "Ulam's Game", or the game of "Twenty Questions with a Liar", has occupied researchers in the past two–three decades [1,7,10,11]. Also, it has been known for some time that "Ulam's Game" can form a basis for the analysis of adaptive error correcting codes [5,8].

One of the latest papers by Spencer [15] gives a comprehensive theoretical analysis of "Ulam's Game" in terms of what are the players' necessary and sufficient conditions to win. However, the major disadvantage of his paper is that it does not provide a practical algorithm of generating the winning strategy for any particular case of interest. Thus, the difficulties in obtaining a reliable procedure for calculating a winning strategy for this game – a method that could be extended to the determining of an optimal code length q in the presence of k errors – has left adaptive error correcting schemes behind the major scene of modern communication techniques.

In the first part of our paper, we present a simple heuristic that gives optimal and near-optimal strategies for playing "Ulam's Game". In the second part, we discuss the correspondence between this game and error correcting codes for a binary channel with noiseless, delayless feedback. We have shown that such a scheme greatly increases the error correcting capability of the communication channel.

* Corresponding author. Mailing address: S. Sarkissian, 4120 Brooklyn Avenue, N.E., Apt. 205, Seattle, WA 98105, USA.

[1] Research supported in part by NSF Grant No. IRI-9120074.

SSDI 0020-0190(95)00127-1

2. A strategy for playing "Ulam's Game"

In the game of "Twenty Questions with a Liar", one person (the Responder) thinks of a "target" object and the other person (the Questioner) attempts to identify the target by asking questions that can be answered "yes" or "no". If the Responder stipulates that the target is an integer in the interval $[1,\dots,N]$, then it is easy to see that a bisection search enables the Questioner to identify the object in $[\log_2 N]$ questions. Furthermore, an adversary argument shows that no smaller number of questions will suffice in the worst case.

The Responder is permitted to answer as many as k questions with lies, where k is a number agreed upon in advance by the players. S.M. Ulam posed a special case of this game in his autobiography, *Adventures of a Mathematician* [17]. Hence the present name of the problem – "Ulam's Game".

The Questioner can ask any question of the type: "Is the target a member of the set S?", where $S \subseteq \{1,\dots,N\}$. The state of knowledge of the Questioner at any point in the game is given by a collection of disjoint subsets, $A_0, A_1,\dots,A_k$, where $A_0 \cup A_1 \cup \cdots \cup A_k \subseteq \{1,\dots,N\}$ and where A_i is the subset of possible targets under the condition that the Responder has told exactly i lies ($i=0,\dots,k$). For example, suppose $N=4$ and $k=1$. If the Questioner initially asks "Is the target in the set {1,2}?" and the Responder answers "yes", then $A_0=\{1,2\}$ and $A_1=\{3,4\}$. A complete game tree, in which the target can be identified in five questions (a worst-case optimal number) is given in Fig. 1.

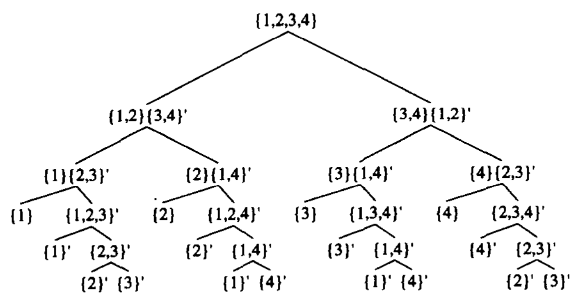

Fig. 1. All numbers in the sets A_0 are denoted as $\{*\}$ and in the sets A_1 as $\{*\}'$.

It is not hard to verify that in order for the Questioner to win, N, q and k must satisfy the inequality

$$N\sum_{j=0}^{k}\binom{q}{j} \leqslant 2^q. \tag{1}$$

Spencer [11] has shown that, when q is sufficiently large, the Questioner can always win the game in q questions, if

$$N\sum_{j=0}^{k}\binom{q}{k} + c_k q^k \leqslant 2^q, \tag{2}$$

where c_k is a constant depending on k. However, since this is an asymptotic result, it is not of much help in verifying the existence of winning strategies for specific values of N, q and k. However, we have found that a simple heuristic yields a nearly optimal strategy for the Questioner. We have used his heuristic to confirm the existence of winning strategies, as described below.

Suppose the state of knowledge of the Questioner is $(A_0, A_1,\dots,A_k)$, with r questions remaining to be asked. Let us define the "weight" of the state $(A_0, A_1,\dots,A_k)$ as

$$\begin{aligned} w_r(A_0, A_1,\dots,A_k) &= |A_0|\binom{r}{\leqslant k} + |A_1|\binom{r}{\leqslant k-1} \\ &\quad + \cdots + |A_k|\binom{r}{\leqslant 0}, \end{aligned}$$

where for notational convenience,

$$\binom{r}{\leqslant j} = \binom{r}{0} + \binom{r}{1} + \cdots + \binom{r}{j}.$$

By generalization of (1), it must be the case that $w_r(A_0, A_1,\dots,A_k) \leqslant 2^r$, in order for the Questioner to be able to win the game. Hence the Questioner is constrained to pose a question for which the states $(Y_0, Y_1,\dots,Y_k)$ and $(N_0, N_1,\dots,N_k)$ resulting from "yes" and "no" answers are such that

$$w_{r-1}(Y_0, Y_1,\dots,Y_k) \leqslant 2^{r-1},$$

$$w_{r-1}(N_0, N_1,\dots,N_k) \leqslant 2^{r-1},$$

and

$$w_r(A_0, A_1, \dots, A_k)$$
$$= w_{r-1}(Y_0, Y_1, \dots, Y_k) + w_{r-1}(N_0, N_1, \dots, N_k),$$

for any possible question that may be posed. Hence the Questioner would seem to be well advised to pose a question for which $w_{r-1}(Y_0, Y_1, \dots, Y_k)$ and $w_{r-1}(N_0, N_1, \dots, N_k)$ are as nearly equal as possible. Let $S = S_0 \cup S_1 \cup \cdots \cup S_k$, where $S_j \subseteq A_j$, be a set about which a question is asked, and let

$$y_S = \binom{r-1}{k} |S_0| + \cdots + \binom{r-1}{0} |S_k|$$

and

$$n_S = \binom{r-1}{k} |A_0 \setminus S_0| + \cdots + \binom{r-1}{0} |A_k \setminus S_k|.$$

As Spencer observed,

$$|w_{r-1}(Y_0, Y_1, \dots, Y_k) - w_{r-1}(N_0, N_1, \dots, N_k)|$$
$$= |y_S - n_S|.$$

Accordingly, our heuristic is very simple:

Choose_S $(A_0, A_1, \dots, A_k, r)$
  $y_S := n_S := 0$;
  for $j = 0$ to k
    Choose S_j to minimize

$$\left| \left(y_S + \binom{r-1}{k-j} |S_j| \right) - \left(n_S + \binom{r-1}{k-j} |A_j \setminus S_j| \right) \right|;$$

    $y_S := y_S + \binom{r-1}{k-j} |S_j|$;
    $n_S := n_S + \binom{r-1}{k-j} |A_j \setminus S_j|$;
  endfor
  return $(S_0, S_1, \dots, S_k)$;
end Choose_S

It is necessary to note that Eq. (1) is analogous to the well-known sphere-packing inequality for (nonadaptive) k-error correcting codes [4]. Except that here the sum represents the number of root-to-leaf paths that must exist for each of the N possible target objects in a game tree of height q, rather than the number of code words in a sphere of radius k.

Therefore, to verify how close the results of our heuristic are to the actual optimal solutions, we have generated a number of instances for which the algorithm is able to find winning strategies for the Questioner in the smallest number of questions satisfying the sphere-packing inequality. In Table 1 we have indicated, for given k and i, where $N = 2^i$, the smallest number of questions for which the heuristic found a winning strategy for the Questioner. When this number is greater than the minimum number satisfying the sphere-packing inequality, the latter number is indicated in parentheses.

3. "Ulam's Game" vs. error correcting codes

Suppose the Transmitter is to encode i information bits for transmission to the Receiver, with provision for the correction of at most k errors. The Transmitter assumes the role of Responder and the Receiver the role of Questioner, the two parties playing the game cooperatively, making reference to the same game tree. The Receiver, in the role of Questioner, must guess the correct target out of a set of size $N = 2^i$. The Transmitter, in the role of Responder, anticipates the questions that the Receiver wishes to ask. Each error that occurs in the transmission amounts to a lie told by the Responder, and the Transmitter knows this by comparing the signal sent in the forward channel to the signal received from the feedback channel.

The target is one of the $N = 2^i$ bit patterns, $0\dots0$ to $1\dots1$. We propose to predetermine the first i questions that are asked. Question j, $1 \leqslant j \leqslant i$, is: "Is bit position j of the target pattern a 1?". Thus the Transmitter simply transmits the i information bits in answer to these first i questions, thereby making the code systematic, i.e., one having distinct information bits and check bits. After the i information bits have been transmitted, the state of the knowledge of the Receiver is $(A_0, A_1, \dots, A_k)$, where

$$|A_j| = \binom{i}{j}, \quad j = 0, \dots, k.$$

By reference to the feedback channel, the Transmitter records the sequence of bits actually received by the Receiver. The Transmitter exclusive-or's this sequence with the sequence of bits actually sent, thereby obtaining an error pattern. There are exactly $|A_j|$ possible error patterns for j transmission errors, for $j=0,\dots,k$, for a total of $\binom{i}{\leqslant k}$ patterns. The Receiver is able to determine the target pattern if and only if the Transmitter can inform the Receiver of the correct error pattern with the $q-i$ check bits that remain to be transmitted.

Remarkably, our heuristic gives results that very much correspond to the relation established by Berlekamp [2]. According to his "translation bound theorem", the optimal code length for the transmission of a message containing i information bits through the channel where k errors may occur differs from that for the transmission of the same message through the noisy channel with $k-1$ possible errors in three additional check bits. Table 1 shows that only in three cases, namely, for 8, 11, and 14 information bits, the violation is observed. For example, in the case of 8 information bits and 5 errors, the algorithm has generated the code length of 25 bits. However, the translation rule predicts only 24 bits for the optimal code length. In all other instances, the heuristic has produced optimal solutions.

Most importantly, the above results provide good example of the usefulness of the feedback channel in communication schemes. The fact that a sequential recursive coding strategy increases transmission efficiency is well documented in the literature [3,5]. Schalkwijk and Post [14] also showed that in a binary symmetric channel with noiseless feedback code digits can be decoded in such a way that the error probability vanishes exponentially even for a fixed coding delay. More recently Spencer and Winkler [16] observed that without a feedback channel no block code can correct for more than $\lceil q/4 \rceil$ errors, when the number of information bits is sufficiently large. With a feedback channel, it is possible to correct as many as $\lfloor q/3 \rfloor$ errors for any number of information bits i. (Note, however, that some coding schemes for two binary forward channels allow one to achieve essentially the same efficiency in communication as it would be in the presence of feedback channel [14].) Also, in comparison with the best known multiple-error correcting linear codes (e.g., BCH codes) [6,9], the improvements in error correcting capability are achieved for $i \geqslant 3$ and $k \geqslant 2$ with a few exceptions.

Table 1

i	k							
	1	2	3	4	5	6	7	8
1	3	5	7	9	11	13	15	17
2	5	8(7)	11(10)	14(12)	17(14)	20(16)	23(18)	26(21)
3	6	9	12(11)	15(14)	18(16)	21(18)	24(21)	27(23)
4	7	10	13	16(15)	19(18)	22(20)	25(23)	28(25)
5	9	12	15(14)	18(17)	21(20)	24(22)	27(25)	30(27)
6	10	13	16	19(18)	22(21)	25(24)	28(26)	31(29)
7	11	14	17	20	23	26(25)	29(28)	32(31)
8	12	15	18	21	25(24)	28(27)	31(30)	34(32)
9	13	17	20	23	26	29(28)	32(31)	35(34)
10	14	18	21	24	27	30	33	36(35)
11	15	19	22	25	28	32(31)	35(34)	38(37)
12	17	20	23	27	30	33	36	39(38)
13	18	21	25	28	31	34	37	40
14	19	22	26	29	32	35	39(38)	42(41)
15	20	24	27	30	34	37	40	43
16	21	25	28	32	35	38	41	44

4. Implementation

As noted in the previous section, the first phase of transmission is a straightforward, non-adaptive transmission of i information bits. In particular, this means that unlike multiple repetition coding [2] our scheme allows bit subsequences of any form. In the second phase of the transmission, during which check bits are transmitted, the Transmitter makes reference to a fully worked out game tree. This tree has height $q-i$. A bit vector of length $\binom{i}{\leqslant k}$ is stored at each node of the game tree with each component of the bit vector corresponding to a different error pattern. Suppose component h is identified with the error pattern to be communicated to the Receiver. When the Transmitter has arrived at a given node of the tree, it transmits the bit found in component h of the vector at that node. The Transmitter observes the bit returned by the feedback channel and proceeds to the child of the current node, as specified by the bit returned. Thus, only a few machine language instructions need to be executed for each successive check bit transmitted.

One of the advantages of implementing left-to-right decoding strategy presented here is that in practice it is very likely that the Receiver will be able to decode the transmitted message before the last bit of the error pattern is received. In the contrast to that, simple block coding strategies developed by Schalkwijk [12] lead to the reverse (right-to-left) direction of decoding: this implies that the whole codeword must be scanned every time.

In addition, it is observed that for a given number of information bits the actual size of a game tree as a percent of the upper bound decreases as the number of errors to be corrected increases. For instance, an 8-error correcting code with 6 information bits requires only 21.2 megabytes of storage, which corresponds to about 8.3% of the upper bound of 256 megabytes.

References

[1] J.A. Aslam and A. Dhagat, Searching in the presence of linearly bounded errors, in: *Proc. 23rd Ann. ACM Symp. on the Theory of Computing* (1991).

[2] D.W. Becker, Multiple-repetition feedback coding, Ph.D. Thesis, University of California, San Diego, 1973.

[3] E. Berlekamp, Block coding for the binary symmetric channel with noiseless, delayless feedback, in: *Error Correcting Codes* (1968) 61–85.

[4] R.E. Blahut, *Theory and Practice of Error Control Codes* (Addison-Wesley, Reading, MA, 1983).

[5] M. Horstein, Sequential transmission using noiseless feedback, *IEEE Trans. Inform. Theory* **9** (1963) 136–143.

[6] F.J. MacWilliams and N.J.A. Sloane, *The Theory of Error-Correcting Codes*, Part II (North-Holland,Amsterdam, 1977).

[7] A. Pelc, Solution of Ulam's problem on searching with a lie, *J. Combin. Theory Ser. A* **44** (1987) 129–140.

[8] A. Pelc, Detecting errors in searching games, *J. Combin. Theory Ser. A* **51** (1989) 43–54.

[9] W.W. Peterson, Encoding and error-correction procedures for the Bose–Chaudhuri codes, *IRE Trans. Inform. Theory* **6** (1960) 459–470.

[10] B. Ravikumar and K.B. Lakshamanan, Coping with known patterns of lies in a search game, *Theoret. Comput. Sci.* **33** (1984) 85–94.

[11] R.L. Rivest, A.R. Meyer, D.J. Kleitman and K. Winklmann, Coping with errors in binary search procedures, *J. Comput. System. Sci.* **20** (1980) 396–404.

[12] J.P.M. Schalkwijk, A class of simple and optimal strategies for block coding on the binary symmetric channel with noiseless feedback, *IEEE Trans. Inform. Theory* **17** (1971) 283–287.

[13] J.P.M. Schalkwijk, A coding scheme for duplex channels, *IEEE Trans. Comm.* **22** (1974) 1369–1373.

[14] J.P.M. Schalkwijk and K.A. Post, On the error probability for a class of binary recursive feedback strategies, *IEEE Trans. Inform. Theory* **19** (1973) 498–511.

[15] J. Spencer, Ulam's searching game with a fixed number of lies, *Theoret. Comput. Sci.* **95** (1992) 307–321.

[16] J. Spencer and P. Winkler, Three thresholds for a liar, *Combin. Probab. Comput.* **1** (1992) 81–93.

[17] S. Ulam, *Adventures of a Mathematician* (Charles Scribner's Sons, 1976).