Correction to 'On the error probability for a class of binary recursive feedback strategies'

We can recover j from $g(j)$ as follows:

$$j_n = g_n{}^j$$

$$j_k = j_{k+1} + g_k{}^j \bmod 2$$

which gives

$$j_{n-1} = g_n{}^j + g_{n-1}^j$$

$$j_{n-2} = g_n{}^j + g_{n-1}^j + g_{n-2}^j, \cdots.$$

Thus

$$j_k = \sum_{m=k}^{n} g_m{}^j \bmod 2.$$

We know that the integers i and i', where $0 \leq i < 2^{n-1}$ and $i' = i + 2^{n-1}$, differ in only one digit, i.e., $i_n = 0, i_n' = 1$, $i_k = i_k', k = 1,2,\cdots, n - 1$. Hence

$$g_k{}^i = i_k + i_{k+1} = g_k{}^{i'}, \qquad \text{if } k \leq n - 2. \qquad (2)$$

Furthermore,

$$g_n{}^i + g_{n-1}^i = i_{n-1}$$

while

$$g_n{}^{i'} + g_{n-1}^{i'} = i_{n-1}' = g_n{}^i + g_{n-1}^i. \qquad (3)$$

We have thus shown that, if we add together the first two columns of a 2^n-level Gray code and copy the remaining $n - 2$ columns, the resulting $n - 1$ columns contain two identical parts. It remains to be proved that each half is a 2^{n-1}-level Gray code. We denote the latter by $G(i)$, $0 \leq i \leq 2^{n-1} - 1$. Then

$$G_{n-1} = i_{n-1}, \qquad G_k = i_k + i_{k+1}, \qquad 0 \leq k \leq n - 2.$$

The previous are identical to the expressions (2) and (3). Thus the $n - 1$ columns do consist of two repetitions of 2^{n-1}-level Gray code. Now if we combine the first two columns again, we reduce each 2^{n-1}-level Gray code into two 2^{n-2}-level Gray codes, or, the complete array into four 2^{n-2}-level Gray codes. This can continue until we have only m columns, which would be 2^{n-m} repetitions of 2^m-level Gray code. We have thus derived the lemma.

References

[1] M. Gardner, "Mathematical games," *Sci. Amer.*, vol. 227, pp. 106–109, Aug. 1972.

Correction to "On the Error Probability for a Class of Binary Recursive Feedback Strategies"

J. PIETER M. SCHALKWIJK AND KAREL A. POST

In the above paper[1], p. 499, (2) should have read

$$p_{n+1}(\theta) = \begin{cases} \dfrac{(1 - y_{n+1})p + y_{n+1}q}{(1 - y_{n+1})[aq + (1 - a)p] + y_{n+1}[(1 - a)q + ap]} p_n(\theta), & \text{for } \theta > a_n \\[2ex] \dfrac{(1 - y_{n+1})q + y_{n+1}p}{(1 - y_{n+1})[aq + (1 - a)p] + y_{n+1}[(1 - a)q + ap]} p_n(\theta), & \text{for } \theta < a_n. \end{cases} \qquad (2)$$

Manuscript received September 14, 1973.
The authors are with the Technological University, Eindhoven, The Netherlands.
[1] J. P. M. Schalkwijk and K. A. Post, *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 498–511, July 1973.

On the right side of p. 505, the fifth and sixth line from the bottom, the lower error exponent $\bar{E}^-(R)$ is valid for the 1 output and the upper error exponent $\bar{E}^+(R)$ for the 0 output.

Acknowledgment

The authors want to thank Dr. J. C. Tiernan for pointing out the mistake in (2).

Optimal Decoding of Linear Codes for Minimizing Symbol Error Rate

L. R. BAHL, J. COCKE, F. JELINEK, AND J. RAVIV

Abstract—**The general problem of estimating the *a posteriori* probabilities of the states and transitions of a Markov source observed through a discrete memoryless channel is considered. The decoding of linear block and convolutional codes to minimize symbol error probability is shown to be a special case of this problem. An optimal decoding algorithm is derived.**

I. Introduction

The Viterbi algorithm is a maximum-likelihood decoding method which minimizes the probability of word error for convolutional codes [1], [2]. The algorithm does not, however, necessarily minimize the probability of symbol (or bit) error. In this correspondence we derive an optimal decoding method for linear codes which minimizes the symbol error probability.

We first tackle the more general problem of estimating the *a posteriori* probabilities (APP) of the states and transitions of a Markov source observed through a noisy discrete memoryless channel (DMC). The decoding algorithm for linear codes is then shown to be a special case of this problem.

The algorithm we derive is similar in concept to the method of Chang and Hancock [3] for removal of intersymbol interference. Some work by Baum and Petrie [4] is also relevant to this problem. An algorithm similar to the one described in this correspondence was also developed independently by McAdam *et al.* [5].

II. The General Problem

Consider the transmission situation of Fig. 1. The source is assumed to be a discrete-time finite-state Markov process. The M distinct states of the Markov source are indexed by the integer m, $m = 0,1,\cdots, M - 1$. The state of the source at time t is denoted by S_t and its output by X_t. A state sequence of the source extending from time t to t' is denoted by $S_t^{t'} = S_t, S_{t+1}, \cdots, S_{t'}$, and the corresponding output sequence is $X_t^{t'} = X_t, X_{t+1}, \cdots, X_{t'}$.

The state transitions of the Markov source are governed by the transition probabilities

$$p_t(m \mid m') = \Pr\{S_t = m \mid S_{t-1} = m'\}$$

and the output by the probabilities

$$q_t(X \mid m', m) = \Pr\{X_t = X \mid S_{t-1} = m'; S_t = m\}$$

where X belongs to some finite discrete alphabet.

Manuscript received July 27, 1972; revised July 23, 1973. This paper was presented at the 1972 International Symposium on Information Theory, Asilomar, Calif. January 1972.
L. R. Bahl and J. Cocke are with the IBM Thomas J. Watson Research Center, Yorktown Heights, N.Y.
F. Jelinek is with the IBM Thomas J. Watson Research Center, Yorktown Heights, N.Y., on leave from Cornell University, Ithaca, N.Y.
J. Raviv was with the IBM Thomas J. Watson Research Center., Yorktown Heights, N.Y. He is now at the IBM Scientific Center, Haifa, Israel.