THE CODING OF MESSAGES SUBJECT TO CHANCE ERRORS[1]

BY J. WOLFOWITZ

1. The transmission of messages

Throughout this paper we assume that all "alphabets" involved contain exactly two symbols, say 0 and 1. What this means will be apparent in a moment. This assumption is made only in the interest of simplicity of exposition, and the changes needed when this assumption is not fulfilled will be obvious.

Suppose that a person has a vocabulary of S words (or messages), any or all of which he may want to transmit, in any frequency and in any order, over a "noisy channel". For example, S could be the number of words in the dictionary of a language, provided that it is forbidden to coin words not in the dictionary. What a "noisy channel" is will be described in a moment. Here we want to emphasize that we do not assume anything about the frequency with which particular words are transmitted, nor do we assume that the words to be transmitted are selected by any random process (let alone that the distribution function of the random process is known). Let the words be numbered in some fixed manner. Thus transmitting a word is equivalent to transmitting one of the integers $1, 2, \cdots, S$.

We shall now explain what is meant by a "noisy channel" of memory m. A sequence of $(m + 1)$ elements, each zero or one, will be called an α-sequence. A function p, defined on the set of all α-sequences, and such that always $0 \leqq p \leqq 1$, is associated with the channel and called the channel probability function. A sequence of n elements, each of which is zero or one, will be call an x-sequence. To describe the channel, it will be sufficient to describe how it transmits any given x-sequence, say x_1. Let α_1 be the α-sequence of the first $(m + 1)$ elements of x_1. The channel "performs" a chance experiment with possible outcomes 1 and 0 and respective probabilities $p(\alpha_1)$ and $(1 - p(\alpha_1))$, and transmits the outcome of this chance experiment. It then performs another chance experiment, independently of the first, with possible outcomes 1 and 0 and respective probabilities $p(\alpha_2)$ and $(1 - p(\alpha_2))$, where α_2 is the α-sequence of the $2^{\text{nd}}, 3^{\text{rd}}, \cdots, (m + 2)^{\text{nd}}$ elements of the sequence x_1. This is repeated until $(n - m)$ independent experiments have been performed. The probability of the outcome one in the i^{th} experiment is $p(\alpha_i)$, where α_i is the α-sequence of the $i^{\text{th}}, (i + 1)^{\text{st}}, \cdots, (i + m)^{\text{th}}$ elements of x_1. The x-sequence x_1 is called the transmitted sequence. The chance sequence $Y(x_1)$ of outcomes of the experiments in consecutive order is called the received sequence. Any sequence of $(n - m)$ elements, each zero or one, will be called a y-sequence. Let y_1 be any y-sequence. If $P\{Y(x_1) = y_1\} > 0$ (the symbol

Received January 26, 1957; received in revised form May 15, 1957.

[1] Research under contract with the Office of Naval Research.