PRINCIPLES OF CODING, FILTERING, AND INFORMATION THEORY

BY
LEONARD S. SCHWARTZ

Principles of CODING, FILTERING, *and* INFORMATION THEORY

BY LEONARD S. SCHWARTZ

New York University

1963

SPARTAN BOOKS, INC.

BALTIMORE

CLEAVER-HUME PRESS, LTD.

LONDON

Library of Congress Catalog Card No. 63-194[illegible]

Sole Distributors in Great Britain, the British Commonwealth and the Continent of Europe:

CLEAVER-HUME PRESS
10-15 ST. MARTINS STREET
LONDON W.C. 2

Preface

IN THIS ERA OF shrinking distances and growing mutual dependence among nations, the success or failure of civilization may hinge upon the expansion of world-wide communications. The media of communications—the newspapers, books, radio, and television are technologically based upon communications engineering. A general, unifying theory of communications was initiated by Wiener and Shannon, whose significant contributions in the 1940's gave the impetus to what we now designate information theory.

The initial successes of pioneers in a new field are rarely followed by sustained advances on a broad front until a basic theory has been developed. This general observation, true of most fields of science and technology, has in fact been fulfilled in communications engineering. With the advent of the new theory, advances in understanding and progress in application have been appreciable. If advances have not been as fast as some impatient observers and even practitioners of information theory might have wished, the delays are nonetheless normal. It takes time until the implications of a theory can fully permeate to the working levels of practice. Considering the fact that the theory to be presented in this book is barely twenty years old, we may be justly pleased with the current level of accomplishment.

The evolution of information theory since the late forties has proceeded along two fairly distinct lines that are complementary rather than competitive. One line of development —emphasizing the importance of concepts of space and dimensionality and of notions of structural and metrical information—has been primarily concerned with the problem of the information contained in the representation of a physical situation. European contributors have been the leading expounders of this approach. The other avenue of development—concerning itself with the concept of statistical ensembles of messages and noise in a channel—has developed in a direction of more specific interest to communications. Because this book grew from the author's work and teaching in a number of communication engineering areas, it is primarily devoted to an exploration of the latter.

Presenting an account of information theory in order to display its power to unify and explain a number of diverse phenomena of communications is one objective of this book. Another is the illustration of some ways in which insights into the theory may suggest improvements in the performance of communication systems.

Information theory, because of its depth and generality, can be applied to all forms of communication, not only to those that come within the province of communication engineering. Information theory contributes to the basic concepts of cryptography—a form of coding—and to the problems of translation from one language to another (although meaning as well as information must be considered in this field) and of the logical design of computing machines. New potentialities for information theory are being explored in the study of sensory information and biological models, learning mechanisms, pattern recognition, information retrieval, and in a general approach to classification, syntactics and semantics.

The basic ideas of information theory were developed originally for use in communication systems involving a transmitter communicating with a receiver over a noisy channel. Although some of the recent extensions of information theory are of interest in their own right, we are restricted, in a book of limited size to a discussion of the basic elements within the original setting of the theory. In this way we may avoid the danger of accomplishing too little by attempting too much; and we are on much firmer ground, since some of the proposed extensions may be of doubtful validity. Moreover, all the essential elements of the theory can be brought into clear focus in a discussion of the applications of the theory to communication systems.

Although much has already been published about information theory in book form, the treatment in each case is primarily confined to certain specific areas, so that the student does not have the essential unity of different aspects of the subject matter presented to him. In this book—a survey of the range of information theory—coding, filtering, detection of signals in noise, and communication feedback are included.

In some of the current literature a highly detailed and specific approach is accompanied by an extremely rigorous mathematical presentation. This is an approach that undeniably instills confidence in the reader who appreciates the validity of conclusions based on mathematical reasoning. But it falls somewhat short of the urgent requirements not only of the beginning student and practicing engineer, but of sophisticated and technically well-trained members of other disciplines with an interest in communications and information theory. For them a presentation is needed that is sufficiently mathematical to be convincing and which, at the same time, can be accepted intuitively. The present book attempts to fill this need. A prior knowledge of the calculus and the elements of probability theory is desirable in order to derive full benefit from the discussions, but even without this training the reader should be able to follow essentials.

The introductory chapter is concerned with the need for information theory, its origins, and the topics entailed in the theory. The next three chapters are devoted to the properties of information at the source, the discrete noisy channel, and the maximum rate at which error-free information can be transmitted over a noisy channel, respectively. A discussion of various techniques (methods of coding) to achieve a near-maximum reliable information rate follows. The sixth chapter contains a discussion of the continuous-information channel in which an explicit expression is derived for maximum error-free information rate over a noisy channel. Certain useful inferences regarding trade-offs between signal-to-noise ratio and bandwidth are based on this discussion, and some concepts of importance to modulation systems are touched upon.

So far, we have been concerned with the problem of the transmitter, i.e., the organization of the signal against noise. We have now to take up the receiver problem, i.e., the recovery of a signal from a background of random noise, the achievement of which depends on the technique of mathematics known as Generalized Harmonic Analysis, discussed in Chapter seven. The application of this technique leads to optimum filtering systems whose properties are described in chapter eight. After optimum filtering, follows the problem of making decisions about the nature of the received waveform. This problem is studied with the aid of statistical decision theory in chapter nine. The aim is to develop methods for improving system performance through the optimization of decision mechanisms at the receiver. This material leads to a description of the general probability detector and its effect on information rate. Subsequently, there is an analysis of a simple practical form of probability detector, called a "null" detector, that withholds decision when interpretation is excessively ambiguous.

Although there are many advantages in a discussion of the ideas connected with null detection, much more can be gained from linking null-detection schemes and other methods of withholding and correcting received information with "feedback." Hence, feedback and its effects on performance are discussed in chapter ten. In order to bring to the fore the potentialities of feedback to improve the reliability of communication systems, the discussion is—at times—more specifically oriented than in preceding chapters; but the intent is to go beyond the objective of merely exhibiting the advantages of feedback. It is hoped that the reader will gain from this chapter a greater feeling for the applicability of the concepts and principles of information theory than the more generalized approach of the earlier chapters could have given him.

It is characteristic of any problem involving both theory and practice that in the course of its elaboration associated problems clamoring for solution, or at least for definition, will be engendered. The relative claims of syntactics, semantics, and pragmatics certainly come to mind in any thorough-going consideration of information theory. Here, however, we find ourselves in an area in which science and philosophy meet, perhaps to their mutual advantage. The future may hold an answer in store for this fascinating question.

In acknowledging the assistance given me in the preparation of this book, I wish to express my grateful appreciation to Miss Alice Mary Hilton for the initial stimulus to write the book and to my colleagues, Dr. John Metzner and Professors Sheldon Chang, Bernard Harris, Arthur Hauptschein and Kenneth Morgan of the Electrical Engineering Department of New York University for many helpful discussions and insights and the permission to use some graphs resulting from their calculations. My thanks are due also to Professor Max I. Baym of Polytechnic Institute of Brooklyn with whom I have had many fruitful discussions on the rôle of information theory in the study of the Humanities. Moreover, appreciation is extended to my students whose comments and criticisms of the text material have proved helpful and to Mrs. Eleanor Gilmore and Mrs. Beatrice Schwartz for typing assistance. Thanks also are due my wife, Jeanne Schwartz, who has given me constant encouragement in the writing of the book and valuable assistance in the preparation of a portion of the manuscript Finally, grateful acknowledgement is made to the Communications Theory Branch, Office of Aerospace Research (USAF), Air Force Cambridge Research Laboratories, Laurence G. Hanscom Field, Bedford, Massachusetts under contract AF 19(604)–1964 for support of the work leading to some of the results mentioned in chapter ten.

Leonard S. Schwartz

New York
February 1963

Contents

CHAPTER ONE

Introduction

COMMUNICATION EXPERIENCE in recent years has stimulated the development of Information Theory. Although FM radio has demonstrated better signal-to-noise ratio than AM, and pulse-code modulation (PCM) proved to be better than any of its predecessors, ultimately limitations are imposed on reliable message rate by disturbances which are usually beyond the control of the transmitter or the receiver. These disturbances take the form of tube and resistor noise and of atmospherics in radio reception, of motor-car ignition noise in television, or of cross-talk in telephone circuits. The engineer is interested in ways to eliminate or reduce the limiting effects of such disturbances. His measures include changing the position of signal-carrier frequency in the spectrum, of increasing the signal-to-noise ratio, of modifying the modulation system, or of coding. The purpose of a Theory of Information is to establish a quantitative relationship between these measures and their influence on reliable message rates.

That communication reliability can be improved by increasing the signal-to-noise ratio is well known. It is also well known that shifting to another frequency channel is often helpful. Thus, parameters such as signal-to-noise ratio, bandwidth, and time play key roles in limiting communication rate and reliability, and these parameters are interrelated. The first attempt to formulate the underlying relationship of communication systems was made by Hartley in 1928 by considering the amplitudes of signal waveforms as quantized into N levels.[1] The different levels represent possible states of the waveform source, and successive sample ordinates may be regarded as selections from these states. The information contributed by selection of a state from such a source is proportional to log N. In a time interval of T seconds and a signal bandwidth of W, there are $2WT$ independent sample ordinates and thus N^{2WT} possible different waveforms. Hartley defined the information rate of such a source as

$$R = 2WT \log N \tag{1-1}$$

Hartley's theory is inadequate. He failed to consider limitations on the fineness of quantization, did not involve the effects of noise, and did not include the probabilities of the various states of a message source. Shannon's Theory of Communications, published in 1948, is built on the foundations of Hartley's theory, but the concept of a determinate source with an output unperturbed by noise has been replaced by that of a probabilistic source which "feeds" a noisy channel.[2] Shannon's theory deals with the question of how to organize (i.e., code) the message at the transmitter—within certain constraints on the parameters of power, bandwidth and time—in order to resist most effectively the corrupting effect of channel noise. It takes into account prior information (regarding the signals to be transmitted) that is mainly of a statistical character, so that—from the mathematical point of view—modern communication theory is the application of probability theory to communication problems.

The Shannon theory of information is of great importance for the communications engineer because his primary interest centers on the rate of correct transmission and reception, i.e., on a reliable message rate and on potential rather than actual information content. As far as the engineering aspect of communications is concerned, the content of a message is immaterial. It may contain meaning or nonsense—the engineer is interested in the relative frequency of occurrence of individual symbols or groups of symbols in messages and in the fact that these frequencies and not the symbol meanings determine the data rate, form of modulation, etc., that he must design into his communication system. In other words, the semantic aspects of communication are irrelevant to the engineering aspects, although the engineering aspects are not necessarily irrelevant to the semantic aspects.

Preceding the statistical approach to communication problems, the term "filtering" was used to designate the removal of unwanted signals by the use of frequency-selective networks. For such removal, it is necessary for the frequency components of the unwanted signals to lie outside the passband of the desired signal. If unwanted frequency components fall within this band, there is unavoidable interference. With the advent of the statistical approach, however, a new dimension is added to the process of filtering because it is discovered that interfering signals, or noise within the channel, can be removed, or at least reduced by statistical processing, without a corresponding reduction in the signal level. Classical methods of signal analysis are unsuited to this application because noise and other forms of interference are random and not simply periodic or a-periodic.

The statistical approach makes use of the method of Generalized Harmonic Analysis developed by Wiener.[3] With this method it is possible to analyze random functions and to discover their time and frequency-domain characterizations.

A further difficulty remains after all possible steps have been taken at the transmitter in respect to coding and at the receiver in respect to filtering: the detection of signals in noise due to the fact that a threshold is always present in receivers because of natural or artificial biases on tubes or semi-conductor devices. Thus, in the last analysis a receiver is a decision device. It must decide whether a signal is present or not, and it must make this decision with a minimum probability of error. Detection problems are statistical in nature because of the random properties of signals and noise. It was fortunate that years ago workers in mathematical statistics evolved techniques for their own purposes that have recently proved useful in developing a theory of detection of signals in noise.[4,5] An outgrowth of advances in the theory of the testing of hypotheses in mathematical statistics, these techniques and the consequent theory have led to detection systems of greater reliability.[6] An interesting aspect of this study is that it permits us to approach the problem of optimum filter design alternatively as a filtering problem or as a decision problem.[7]

A further big step can be made by the introduction of a feedback path from the receiver to the transmitter for the purposes of error control. Considerable study of feedback of this kind has been undertaken and has been shown to have marked advantages over coded uni-directional operation.[5] It must be emphasized that coding, filtering, and feedback are not mutually exclusive and that the best results accrue from their combined use. The techniques for the employment of feedback are also statistical, and, in part, are suggested by the Shannon theory of coding. The introduction of feedback gives a communication channel an ability to adjust its transmitted information rate to the signal and noise conditions on the channel, speeding up when the noise level is low and slowing down when it is high. Thus, systems equipped with a feedback link are *adaptive communication systems*.

As we can surmise from this introduction, Information Theory is a powerful tool for the evaluation and comparison of the performance of communication and radar systems. More important even, the theory is a fertile source for methods to improve these systems.

References

1. R. V. L. Hartley, "Transmission of Information" BSTJ Vol. 7 (1928) pp. 535–563.
2. C. E. Shannon, "A Mathematical Theory of Communication" BSTJ Vol. 27 (1948) pp. 379–423, pp. 623–658.
3. N. Wiener. *The Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications*, John Wiley and Sons, Inc., New York (1949).
4. A. M. Mood, *Introduction to the Theory of Statistics*, Chaps. 12, 15, McGraw-Hill Book Company, Inc., New York (1950).