INTERNATIONAL CENTRE FOR MECHANICAL SCIENCES

COURSES AND LECTURES No. 229

THE INFORMATION THEORY APPROACH TO COMMUNICATIONS

EDITED BY
G. LONGO

INTERNATIONAL CENTRE FOR MECHANICAL SCIENCES

COURSES AND LECTURES No. 229

THE INFORMATION THEORY APPROACH TO COMMUNICATIONS

EDITED BY

G. LONGO

UNIVERSITY OF TRIESTE

SPRINGER VERLAG WIEN NEW YORK

ISBN 3-211-81484-1 Springer Verlag Wien - New York

ISBN 0-387-81484-1 Springer Verlag New York - Wien

FOREWORD

Thirty years ago, Claude E. Shannon published the first paper on information theory, whose impact and consequences have been far-reaching. One of these consequences is the series of summer schools on information theory and coding theory which I have been organizing for several years at the International Centre for Mechanical Sciences.

If any merit has to be attributed to these schools, it is mainly because they made people from different parts of the world come together and have the chance to meet and discuss both scientific and non-scientific matters. Several of these people are now friends and I myself met some of my best friends during these schools.

I wish to take the opportunity to thank all those who have contributed to the success of this school: first of all the lecturers, for their skill and knowledge, and the audience, for their active participation and curiosity. I am grateful, in particular, to Dr. Andrew J. Viterbi, who undertook the job of reading all the contributions and wrote a clear and penetrating preface.

Finally, it gives me great pleasure to acknowledge the continuous and consistent help that before, during and after the school has been patiently and cheerfully given to me by Enzo Ceschia, Ezio Cum, Claudio di Giorgio, Fabio Ferdani, Enny Rovere, Pierina Mattiussi, Manuela Miconi, Antonio Sfiligoi, Elsa Venir.

Giuseppe Longo

Udine, April 1978

PREFACE

This volume contains a collection of eight sets of lectures presented at CISM Summer Schools of 1976 and 1977. The common thread of all eight lecture series is the development of information and communication theory, from classical results through the most current research and applications. Taken together they present a broad summary of the state of knowledge in the fields of channel coding and source coding for both point-to-point and multiterminal links, intersymbol interference, optical communication and image compression.

The lectures by Omura represent a mini-textbook on information theory, with emphasis on error coding bounds and rate-distortion bounds, intended for the initiated researcher, but of value also to the graduate student. Restriction to the binary symmetric channel and discrete memoryless source limits the length but not the usefulness of the presentation. While the development follows the generally accepted evolution from block coding to convolutional coding to rate-distortion results, the emphasis on commonality of technique and duality of results is unusual and enlightening. Particularly valuable and unique is the idea of starting with a strong converse error bound for block codes and then showing that the better-known random coding bound takes on the same form over the complementary parameter range.

This volume contains two welcome contributions on multiterminal coding for networks of both channels and sources. The paper by Bergmans on the "Gaussian Network" deals with the broadcast channel introduced by Cover, which revived interest in multiliterminal information theory, and the interference channel; the first involves a single trasmitter accessing several receivers dissimilarly disadvantaged by Gaussian noise, and the second consists of N transmitters accessing N receivers, each receiving a linear combination of all transmitted signals in the presence of Gaussian noise which is independent of that received by all others.

The paper by Berger, after a brief but thorough review of classical rate-distortion theory, treats networks of sources communicating with each other over noiseless channels. This begins with the basic Slepian-Wolf results concerning correlated sources and the coding rates required for each so that the receiver can achieve almost distortionless decoding of both sources. This development is extended through use of a new result which the author calls the Markov Lemma, *which he applies to distortionless source coding with side information, and rate-distortion theory with side information, involving several multiterminal source configurations.*

The contribution by Schalkwijk concentrates on convolutional (linear finite-state) codes, presenting a new approach to the maximum likelihood decoder, which reduces the state requirements and hence the implementation complexity when these codes are used on the binary symmetric channel.

Also related to the theory of convolutional codes are the lectures on error bounds for intersymbol interference channels. Presented by Viterbi and based on work done in collaboration with Omura, these lectures begin with the earlier results of Forney on optimum demodulation for the Gaussian channel with intersymbol interferences, develop more detailed error bounds based on the error-state diagram and, finally, extend these results to error bounds for convolutional codes on the Gaussian intersymbol interference channel.

The lectures by Biglieri, based on work in collaboration with Elia, concerning geometrically oriented block group codes for the Gaussian channel, review numerous results spanning two decades of the work of Slepian and others. Recent resuls of the authors are included.

The contribution on Optical Communication by Ephremides differs from the others in that it is oriented more toward detection and estimation (referred to in a footnote as det-estable) than toward coding theory. It is a well-organized, readable presentation of the fundamental models, including both the point-process approach and the quantum mechanical theory, and the main results in quantum detection and estimation theory, with a final lecture on probable future directions of research.

The contributions by Huang and by Huang and Burnett represent applications of information theory, including some of the recent results presented by other authors in this series, to the very practical problems of image mensuration and efficient coding.

This collection accurately portrays the state of progress in key areas of information and communication theory at the time of its publication.

Andrew J. Viterbi

La Jolla, California
March 1978.

G. LONGO (ed.): The Information Theory Approach to Communications
CISM Courses and Lectures No. 229 — Springer-Verlag Wien-New York 1978

ERRATUM

p. 249, top: the two equations should read

$$Y_1 = X_1 + X_2 \sqrt{a_{21}} + Z_1$$

$$Y_2 = X_2 + X_1 \sqrt{a_{12}} + Z_2$$

CONTENTS

SOURCE AND CHANNEL CODING WITH BLOCK AND CONVOLUTIONAL CODES

Jim K. Omura
System Science Department
University of California
Los Angeles, California

PREFACE

I am grateful to CISM and Professor Giuseppe Longo for the opportunity to present these lectures during the 1976 CISM summer school session. My purpose in these lectures was to develop the basic concepts of channel coding theory and rate distortion theory in a unifying manner. The interest and enthusiasm of the participants in this course made this a stimulating endeavor.

Jim K. Omura

SOURCE AND CHANNEL CODING WITH BLOCK AND CONVOLUTIONAL CODES*

Jim K. Omura

System Science Department

University of California

Los Angeles, California

1. INTRODUCTION

In these lectures we examine the performance of block codes and convolutional codes when these codes are used for source coding and channel coding. We also explore the natural duality between channel coding theory and rate distortion theory as well as discuss some more direct relationships between these theories.

In order to facilitate our presentation without excessive technical details we restrict our proofs of error bounds in channel coding theory to the binary symmetric channel (BSC). For rate distortion theory we

*This work was supported in part by the National Science Foundation under Grant ENG75-03224.

limit our discussion to discrete memoryless sources (DMS) and single-letter bounded distortion measures. Generalizations and references will be given at the end of each section.

2. CHANNEL CODING THEORY

Throughout this discussion we assume that information consists of independent equally likely binary symbols. We call such an information source a binary symmetric source (BSS). Let $\mathscr{U} = \{0,1\}$ be the source alphabet and $u \in \mathscr{U}$ an information bit. For the set of sequences, we use subscripts. The set of sequences of K information bits is denoted $\mathscr{U}_K$. Our channel is assumed to be a BSC with crossover probability ε. Let $\mathscr{X} = \{0,1\}$ and $\mathscr{Y} = \{0,1\}$ denote input and output alphabets for the BSC with $\mathscr{X}_N$ and $\mathscr{Y}_N$ denoting corresponding sets of sequences of length N. For the channel input symbol $x \in \mathscr{X}$ and channel output symbol $y \in \mathscr{Y}$ we have channel conditional probability

$$P(y|x) = \begin{cases} \varepsilon & ; \; y \neq x \\ 1-\varepsilon & ; \; y = x \end{cases} \tag{2.1}$$

For channel input sequence $\underset{\sim}{x} = (x_1, x_2, \ldots, x_N) \in \mathscr{X}_N$ and channel output sequence $\underset{\sim}{y} = (y_1, y_2, \ldots, y_N) \in \mathscr{Y}_N$ we have channel conditional probability

$$\begin{aligned} P_N(\underset{\sim}{y}|\underset{\sim}{x}) &= \prod_{n=1}^{N} P(y_n|x_n) \\ &= \varepsilon^{\omega(\underset{\sim}{y},\underset{\sim}{x})} (1-\varepsilon)^{N-\omega(\underset{\sim}{y},\underset{\sim}{x})} \end{aligned} \tag{2.2}$$

where $\omega(\underset{\sim}{y},\underset{\sim}{x})$ is the number of components where $y_n \neq x_n$; $n = 1,2,\ldots,N$. $\omega(\underset{\sim}{y},\underset{\sim}{x})$ is called the Hamming distance between binary sequences $\underset{\sim}{y}$ and $\underset{\sim}{x}$. The BSC channel is sketched in Figure 1.

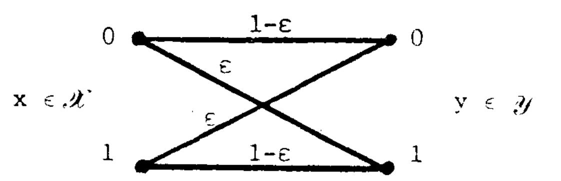

Figure 1. Binary Symmetric Channel

In order to send information bits reliably over the BSC, some redundancy must be added to the information bits to allow correction of channel errors. Suppose for K information bits we send N channel input bits as shown below in Figure 2. The amount of redundancy is measured in the information rate

$$R = \frac{K}{N} \text{ information bits/channel use} \tag{2.3}$$

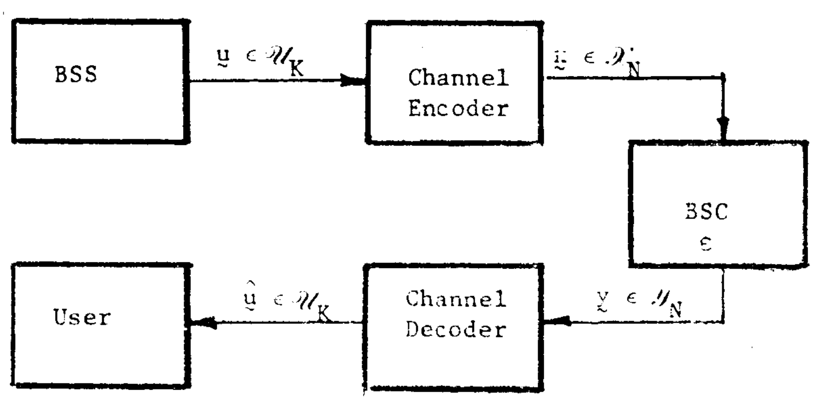

Figure 2. Basic Channel Coding System

The channel encoder merely assigns a code word to each of the 2^K possible information bit sequence. Let $\{\underset{\sim}{x}_1, \underset{\sim}{x}_2, \ldots, \underset{\sim}{x}_M\}$ where $M = 2^K = 2^{RN}$ be the code words. Hence there are only $M = 2^K$ possible channel input sequences. Because of possible channel errors, however, any of the 2^N possible channel output sequences can occur. The channel decoder must take $\underset{\sim}{y} \in \mathcal{Y}_N$ and decide which code word in $\{\underset{\sim}{x}_1, \underset{\sim}{x}_2, \ldots, \underset{\sim}{x}_M\}$ was sent over the channel. Once it decides the transmitted code word is $\hat{\underset{\sim}{x}}$ then the decoder can output the corresponding information bit sequence $\hat{\underset{\sim}{u}} \in \mathcal{U}_K$.

Any deterministic decoder decision rule can be characterized by decision functions

$$\delta_m(\underset{\sim}{y}) = \begin{cases} 1 \; ; & \text{If the decoder decides } \underset{\sim}{x}_m \text{ given } \underset{\sim}{y} \in \mathcal{Y}_N. \\ 0 \; ; & \text{otherwise} \end{cases} \qquad (2.4)$$

$$m = 1, 2, \ldots, M.$$

Clearly since only one decision is assumed for each channel output sequence, for any $\underset{\sim}{y} \in \mathcal{Y}_N$ we have

$$\sum_{m=1}^{M} \delta_m(\underset{\sim}{y}) = 1 \; . \qquad (2.5)$$

In channel coding the usual criterion is the minimization of the error

$$P_e = \Pr\{\hat{\underset{\sim}{u}} \neq \underset{\sim}{u}\} \qquad (2.6)$$

or, equivalently, maximization of the probability of a correct decision

$$\begin{aligned} P_c &= 1 - P_e \\ &= \Pr\{\hat{\underset{\sim}{u}} = \underset{\sim}{u}\} \end{aligned} \qquad (2.7)$$

For any set of distinct code words $\{\underset{\sim}{x}_1, \underset{\sim}{x}_2, \ldots, \underset{\sim}{x}_M\}$ and any decoder decision rule we have

$$\begin{aligned} P_c &= \Pr\{\hat{\underset{\sim}{u}} = \underset{\sim}{u}\} \\ &= \sum_{m=1}^{M} \Pr\{\hat{\underset{\sim}{u}} = \underset{\sim}{u}_m | \underset{\sim}{u}_m\} P_N(\underset{\sim}{u}_m) \\ &= \frac{1}{M} \sum_{m=1}^{M} \Pr\{\hat{\underset{\sim}{x}} = \underset{\sim}{x}_m | \underset{\sim}{x}_m\} \\ &= \frac{1}{M} \sum_{m=1}^{M} \sum_{\underset{\sim}{y}} \delta_m(\underset{\sim}{y}) P_N(\underset{\sim}{y}|\underset{\sim}{x}_m) \\ &= \frac{1}{M} \sum_{\underset{\sim}{y}} \left(\sum_{m=1}^{M} \delta_m(\underset{\sim}{y}) P_N(\underset{\sim}{y}|\underset{\sim}{x}_m) \right) \end{aligned} \tag{2.8}$$

Note that P_c is maximized over the decision rules if for each $\underset{\sim}{y} \in \mathcal{Y}_N$, $\delta_m(\underset{\sim}{y})$; $m = 1,2,\ldots,M$ is chosen to maximize

$$\sum_{m=1}^{M} \delta_m(\underset{\sim}{y}) P_N(\underset{\sim}{y}|\underset{\sim}{x}_m) \ . \tag{2.9}$$

Clearly the optimum decoder is the maximum likelihood decoder which has the decision rule*

$$\delta_m(\underset{\sim}{y}) = \begin{cases} 1 & ; \text{ If } P_N(\underset{\sim}{y}|\underset{\sim}{x}_m) > P_N(\underset{\sim}{y}|\underset{\sim}{x}_{m'}) \\ & \quad \text{for all } m' \neq m. \\ 0 & ; \text{ otherwise.} \end{cases} \tag{2.10}$$

For the BSC since

$$P_N(\underset{\sim}{y}|\underset{\sim}{x}) = \left(\frac{\varepsilon}{1-\varepsilon}\right)^{\omega(\underset{\sim}{y},\underset{\sim}{x})} (1-\varepsilon)^N \tag{2.11}$$

*Assume that when more than one code has the maximum likelihood value then any choice among them is made.