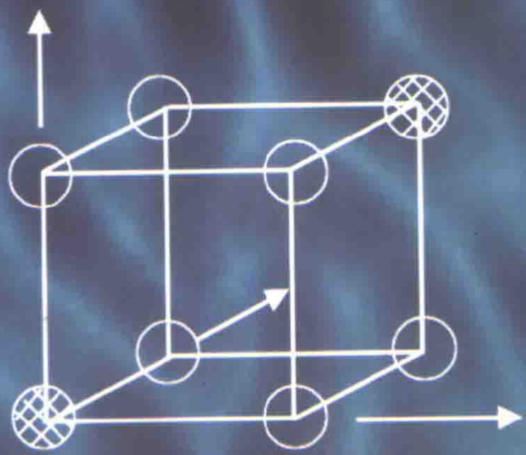
信息科学经典原版教材

Prentice Hall

工程应用编码与信息理论

（英文版）

Applied Coding and Information Theory for Engineers



（美） Richard B. Wells 著

Pearson Education

机械工业出版社
CHINA MACHINE PRESS

信息科学经典原版教材

工程应用编码与信息理论

（英文版）

APPLIED CODING AND INFORMATION THEORY FOR ENGINEERS

（美）Richard B. Wells　著

机械工业出版社
China Machine Press

English reprint copyright © 2002 by Pearson Education North Asia Limited and China Machine Press.

Original English language title: **Applied Coding and Information Theory for Engineers**, by Richard B. Wells
Copyright © 1999 by Prentice-Hall, Inc.
All rights reserved.

Published by arrangement with the original publisher, Pearson Education, Inc., publishing as Prentice-Hall, Inc.

This edition is authorized for sale only in the People's Republic of China (excluding the Special Administrative Region of Hong Kong and Macau).

本书英文影印版由美国 Pearson Education（培生教育出版集团）授权机械工业出版社在中国大陆境内独家出版发行，未经出版者许可，不得以任何方式抄袭、复制或节录本书中的任何部分。

本书封面贴有 Pearson Education（培生教育出版集团）激光防伪标签，无标签者不得销售。

北京市版权局著作权合同登记号：图字：08-2002-3638

图书在版编目（CIP）数据

工程应用编码与信息理论/（美）韦尔斯（Wells, R. B.）著.—北京：机械工业出版社，2002.9

书名原文：Applied Coding and Information Theory for Engineers

信息科学经典原版教材

ISBN 7-111-10983-X

Ⅰ.工…　Ⅱ.韦…　Ⅲ.①通信工程-编码理论-高等学校-教材-英文 ②通信工程-信息论-高等学校-教材-英文　Ⅳ.TN911.2

中国版本图书馆 CIP 数据核字（2002）第 072465 号

机械工业出版社（北京市百万庄大街 22 号　邮政编码 100037）
责任编辑：刘　青　封面设计：饶　薇　责任印制：付方敏
北京忠信诚胶印厂印刷·新华书店北京发行所发行
2002 年 10 月第 1 版·第 1 次印刷
787mm × 1092mm $^1/_{16}$·20 印张·493 千字
0 001—3 000 册
定价：36.00 元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换
本社购书热线电话：（010）68993821、68326677-2527
封面无防伪标均为盗版

出版说明

随着我国加入 WTO，我们将全方位地参与国际竞争。而国际间的竞争最根本的就是人才的竞争、教育的竞争。引进国外先进的教学思想和教学方法，同时引进国外优秀的原版教材，在有条件的学校推动英语教学和双语教学，对提高我国自编教材的水平，加强学生英语实际应用能力，加快我国教育改革步伐，使我国的高等教育尽快与国际接轨，加快培养具有国际竞争力的高水平技术人才有着重要的意义。因此，国家教育部近来出台一系列政策，大力倡导各高校积极引进原版教材，开展英语或双语教学。

以此为契机，机械工业出版社近期将推出一系列国外引进版教材，涉及到高校公共基础课，以及机、电、信息领域的专业基础课和专业课。

为了做好教材的引进工作，机械工业出版社成立了由著名专家组成的国外高校优秀教材审定委员会，通过对双语教学的深入研究，从原版教材的教学思想和教学方法的先进性、系统性、经济性等方面对引进国外原版教材提出了许多建设性意见。同时，结合我国高校教学课程体系的设置和要求，对每一本要引进的原版教材进行精选，以保证这些教材能够满足编写质量高、有较强的权威性和系统性，能够适应我国学生的外语水平和学习特点等多方面的要求。可以肯定地讲，目前机械工业出版社所引进的原版教材，都是经过精心挑选，由国际知名出版公司出版，在各个学科领域和教学实践中具有很高的学术价值和很强的代表性。相信这些教材的出版，对我国的双语教学工作将会起到积极的作用。

这套教材出版后，我们将根据各高校的双语教学计划，适时开展原版教材培训活动，并及时将其推荐给各高校使用。希望高校师生提出宝贵意见和建议，帮助我们不断提高教材出版水平，更好地为教学改革服务。

机械工业出版社
2002 年 8 月

Preface

Welcome to the study of information and coding theory. We are living, as the saying goes, in the dawn of the information era, a time which many have likened to the next industrial revolution. Although this observation has been made so often as to become a cliche, the underlying significance of this new age for business, industry, and society in general is nonetheless difficult to overstate. When the industrial revolution occurred, it brought with it a new need for people to become skilled in the technical sciences, arts, and crafts. In a similar fashion, the information revolution brings with it the need for a greater number of people with understanding of and skills in the crafting of information for a variety of uses.

This book has been written for the beginner. It is the result of class notes from an undergraduate-level course I teach to students in electrical and computer engineering, computer science, and mathematics. The level of exposition in this book has been aimed at undergraduate students in their junior or senior years of study and for the practicing engineer who has little or no previous exposure to this subject. The goal of this book is to help you get started in the *practice* of information engineering.

In recent years, introductory textbooks on this subject have become virtually extinct. Some very good graduate-level textbooks do exist, but these texts are often a bit too theory laden and a bit light on practice for the eager new student motivated by the need to develop a marketable skill or for the busy practitioner looking for an introductory-level treatment so she can get started on that hot new project. With these readers in mind, I have deliberately abandoned the traditional "theorem-proof" format found in most books on this subject. The material in this text comes equipped with theory but I have tried to structure the book in such a way that the required mathematical developments immediately precede the material on "how to" apply the methods and theory. Thus, there is no grand chapter where all of the mathematical theorems are condensed. Topics are introduced as they are needed and in a "just in time" fashion. The text is liberally sprinkled with examples (with numbers) to illustrate the concepts.

The material in this textbook is adequate for a one-semester course in the junior or senior year. The text assumes the reader has previously acquired a background in elementary linear algebra and in introductory probability. A previous course in digital logic design or in introductory communication systems is helpful but is not vital.

Chapter 1 begins with an overview of digital communication systems and an introduction to the concept of information. I have found that students are frequently

此为试读，需要完整PDF请访问： www.ertongbook.com

surprised to learn that "information" and "data" are not the same thing. Chapter 1 introduces discrete information sources and the fundamental concepts of entropy and joint entropy. This leads to an introduction to source coding for data compression where we apply the theory to Huffman codes, Lempel–Ziv codes, and arithmetic codes. The topics of source modeling and adaptive coding are lightly touched upon and references are provided for the reader wishing to delve deeper into this important topic.

Our study of information theory continues in Chapter 2 with the introduction of discrete memoryless channels. After describing and defining these channels, we introduce mutual information and channel capacity. The Arimoto—Blahut algorithm for calculating channel capacity for discrete memoryless channels is described. We also are introduced to the all-important binary symmetric channel, which is described in some detail. This leads us to the idea of block coding and Shannon's famous second theorem. This chapter also introduces Markov processes and channels with memory which leads us to a number of important concepts. Constrained channels are then introduced, along with the important notions of the autocorrelation and power spectrum of a sequence. We close the chapter with applications of the theory to data translation codes and introduce run-length-limited (d,k) codes.

Chapter 3 is all application and the instructor may skip this chapter without loss of continuity if he feels the pinch of the clock and calendar. This chapter is about the particular class of data translation codes known variously as line codes, modulation codes, or run-length limited codes. It is a survey of prefix block coding techniques including state-independent fixed-rate/fixed-block codes, state-dependent coding for fixed-length block codes, variable-length/fixed-rate block codes, look-ahead codes, and concludes with a few words about dc-free codes.

Chapter 4 introduces the general theory of linear-block error correcting codes. It begins with a discussion of the coding problem and the calculation of error probabilities for noisy channels. Error correction using binary repetition codes is next discussed along with some important bounds and constraints which any linear block code must obey. We then provide some brief background on binary fields and binary vector spaces in preparation for the more theoretical development of algebraic codes. The fundamental ideas of Hamming distance, Hamming weight, and the Hamming cube are introduced along with some other important mathematical definitions and concepts. Decoding is introduced using the standard array, and systematic block codes are defined. This leads us to a very in-depth discussion of the Hamming codes, our first "important" practical code. Along with the basic Hamming codes, we also discuss some useful "variations on a theme" including the dual codes and the expanded Hamming code. Codes for correction and detection are discussed. The chapter concludes with a discussion of error rates for linear-block codes and code performance for error-correcting codes and for automatic repeat-request systems.

Chapter 5 continues the discussion of linear-block codes with the introduction of cyclic block codes. Following basic definitions and properties, the polynomial representation of cyclic codes is introduced and we discuss polynomial modulo arithmetic for polynomials constructed from the binary field. Attention is then turned to efficient methods for the generation and decoding of cyclic codes. A number of practical circuits for implementing encoders and decoders are given. We briefly discuss error trapping and pipelined error-trapping decoders. We finish up this chapter by providing several useful and important standard codes including the Hamming codes (again!), some of the simpler BCH codes, and some good burst-correcting codes. Error detection using cyclic redundancy check (CRC) codes is also discussed along with some useful "variations on a theme," including interleaving and shortened codes.

In Chapter 6, we turn away from block codes and introduce linear convolutional codes. After discussing the basic encoder, we examine some structural properties of convolutional codes and the representation of these codes using state diagrams and trellis diagrams. The notions of the transfer function of a code and its uses are discussed. We discuss the Viterbi algorithm in depth. The presentation here differs from that of most texts and papers in that we first describe what the algorithm is and how it works before describing why it works. This reversal of the usual presentation is reported by my students to be easier to follow than the traditional pedagogy. We discuss both hard- and soft-decision Viterbi decoding and compare and discuss the performance differences between these two methods. Some of the known good convolutional codes are then presented in tabular form. We return the Viterbi algorithm to discuss some practical implementation matters including the traceback method of decoding and the use of punctured convolutional codes to obtain higher coding rates.

Chapter 7 is a brief introduction to trellis-coded modulation. We introduce two-dimensional I-Q channels and transmitters and receivers for these channels. The error rate properties of encoded channels is discussed for phase modulation and for quadrature amplitude modulation systems. This is followed by an introduction to systematic recursive convolutional encoders and their representation using trellis diagrams. Ungerboeck's canonical encoder is presented and octal description of TCM codes using parity check polynomials is given. This is followed by a discussion of set partitioning and how this is used to construct TCM codes using Ungerboeck encoders. The chapter concludes with a summary of some known good codes for phase modulation and quadrature amplitude modulation.

Chapter 8 is a brief introduction to the application of information theory to cryptography. Some simple cryptosystems, based on ciphers, are introduced along with a brief description of some of the methods by which cryptosystems may be attacked. Shannon's definition of perfect secrecy is then introduced and conditions for attaining perfect secrecy are derived. The entropy rate of a natural language and how the redundancy of natural languages can be exploited in cryptoanalysis is then discussed. This leads us to the important notions of spurious keys and unicity distance. Following this, the issue of computational security is discussed. Shannon's techniques of diffusion and confusion are described, leading to the important technique of product cipher systems. Finally, the chapter concludes with brief descriptions of codes, public-key cryptosystems, and certain other issues. Public-key cryptosystems, while important, are not described in depth since, the theory of public key cryptosystems is more involved with number theory than with information theory; but the reader is provided with several good references for following up in more detail on the theory and practice of public key cryption.

The text concludes with Chapter 9. This brief chapter provides a proof of Shannon's second theorem for the special case of the binary symmetric channel. It introduces the notion of the random-coding argument and provides a discussion of what the theorem does and does *not* tell us. We then turn to the derivation of Shannon's noiseless coding theorem and the existence of prefix codes for performing source compression. The text then concludes with a few final words on information theory and where the reader may go from here.

Although a certain amount of formalism is necessary in presenting coding-and-information theory, I have tried to keep the presentation a little informal whenever possible. The reader will run across, from time to time, some light-hearted commentary from the author who feels, "If the writing of a textbook has its lighter moments, shouldn't the reading of it have some too?" One of the more important aspects of coding and information theory that is often lost in the classical graduate-level textbooks on this subject is that it is *fun.* I have tried not to lose sight of this aspect during the swirl of mathematical presentation.

While it is true that an author is a character of some importance in the writing of a textbook, the merit of a textbook lies not with its author but, rather, with the audience that book is meant to serve. Furthermore, any textbook (and this one in particular) owes its existence to a number of people who make important contributions. With this in mind, I would like to acknowledge my appreciation to my students who provided me with plenty of feedback about the manuscript, the homework exercises, and the solution manual. I would also like to thank Mr. Aaron Brennan for his assistance in the design of some of the homework exercises. I would especially like to thank Dr. George Freeman of the University of Waterloo for his insightful comments and suggestions which did much to improve the quality of this book. Finally, although they are seldom mentioned by an author, I would like to thank Alice, Tom, and the rest of the merry band at Prentice-Hall whose craftsmanship makes the difference between a manuscript and a book.

Richard B. Wells
Moscow, Idaho

Contents

此为试读，需要完整PDF请访问：www.ertongbook.com

CHAPTER 1

Discrete Sources and Entropy

1.1 OVERVIEW OF DIGITAL COMMUNICATION AND STORAGE SYSTEMS

Systems dedicated to the communication or storage of information are commonplace in everyday life. Broadly speaking, a communication system is a system which sends information from one place to another. Examples include, but are not limited to, the telephone network, radio, television, cellular telephones, local-area computer networks, and so on. Storage systems are systems for the storage and later retrieval of information. In a sense, such systems may be regarded as communication systems which transmit information from *now* (the present) to *then* (the future). Examples include magnetic and optical disk drives, magnetic tape recorders, video tape players, and so on.

Both of these types of systems may be represented abstractly by the block diagram given in Figure 1.1.1. In all cases, there is a *source* from which the information originates. Information from the source is processed by a system for encoding and modulating the information. This encoder/modulator processes the information into some form of signal, which is designed to facilitate the transmission (or storage) of the information in physical form. In communication systems, this function is often referred to as a transmitter while in storage systems it is often called a recorder or a writer.

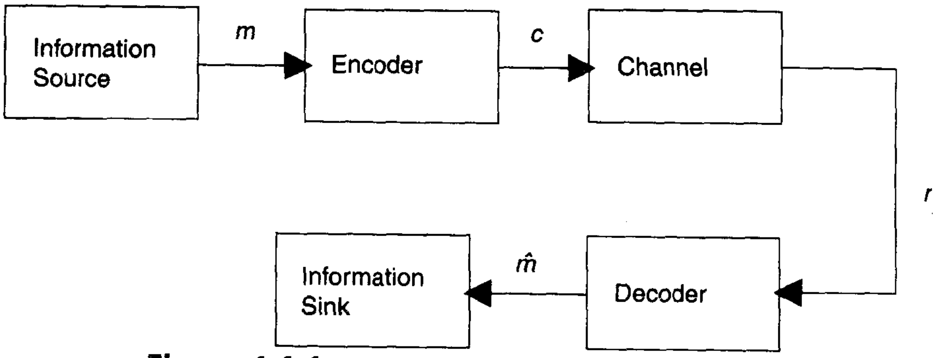


Figure 1.1.1: Basic Information Processing System

The output of the encoding system is then transmitted through some physical communication channel (in the case of a communication system) or stored in some physical storage medium (in the case of a storage system). Examples of the former include wireless transmission using electromagnetic waves and wire transmission using copper telephone cables or fiber optic cables. Examples of the latter case include magnetic disks, such as those used by a floppy disk drive, magnetic tape, and optical disks, such as those used by a CD-ROM or a compact-disk player. Regardless of the explicit form of the medium, we shall refer to it as the "channel."

Information conveyed through (or stored in) the channel must be recovered at the destination and processed to restore the original representation of the information. This is the task of the decoder/demodulator. In the case of a communication system, this device is often referred to as the receiver. In the case of a storage system, this block is often called the playback system or the reader. The signal processing performed by the decoder can be viewed as the inverse of the function performed by the encoder. The output of the decoder is then presented to the final user or destination, which we call the information sink.

The physical channel typically produces a received signal r which differs from the original input signal c. This is because of signal distortion and noise introduced by the channel. Consequently, the decoder can only produce an estimate $\hat{m}$ of the original information message m. The goal of all well-designed system is to attempt to reproduce m as reliably as possible while, at the same time, sending as much information as possible per unit time (communication system) or per unit storage (storage system).

In the typical case, the source message m consists of a time sequence of *symbols* emitted by the information source. The source is said to be a continuous-time source if this sequence is continuous in time. Otherwise, the source is said to be discrete-time. An example of a continuous-time source would be a speech waveform. Examples of discrete-time sources include data sequences from a computer or the printed text on this page (which is an example of a data *storage* system).

The symbols emitted from the source can also be characterized as continuous-amplitude or as discrete-amplitude. Speech is an example of a continuous-amplitude source, since a model of a speech waveform consists of real-valued signal amplitudes. This text is again an example of a discrete-amplitude source since it draws its characters from a finite symbol alphabet.

In this introductory text, we will be primarily concerned with discrete-time/discrete-amplitude sources since these sources have the simplest mathematical treatment and since practically all new communication or storage systems currently fall into this category. Extension of the theory presented here to the continuous case is more appropriately dealt with in an advanced course.

1.2 DISCRETE INFORMATION SOURCES AND ENTROPY

1.2.1 Source Alphabets and Entropy

Information theory is heavily based on the concepts and mathematics of probability theory. This is because the term *information* carries with it a connotation of unpredictability in the transmitted messages. The information content of a message is directly related to the amount of "surprise" conveyed by the message. For example, suppose someone were to say to you, "The capitol of the United States of America is Washington, D.C." Once, early in your life, you did not know this, and so the first time you heard of it, it was *informative.* Seeing it just now, in the previous sentence, this message was completely uninformative (at least for any resident of the United States!). From the point of view of information theory, the statement above held zero information for you after the word "is."

There is a distinction to be made between *information* and *knowledge.* They are not the same thing although one gives rise to the other. Knowledge has the following definitions according to Webster's dictionary: (1) The fact or condition of knowing something with familiarity gained by experience or association; (2) the fact or condition of being aware of something; (3) the sum of what is known. That which adds to our knowledge may be said to be *informative.* Information is therefore distinguished by the property that it adds to our knowledge. That which is not informative delivers no information. Consequently, that which is informative carries with it this element of surprise or uncertainty. It also follows from this character of information that data and information are not the same thing. Were I to recite the English alphabet to you, I would be supplying you with data but, assuming you can read this text, I would be providing no information to you.

An information source is defined by the set of output symbols it is capable of producing and the probability rules which govern the emission of these symbols. A finite discrete source is one for which there is a finite number of unique symbols. The symbol set is frequently called the *source alphabet.* For a source alphabet with M possible symbols, we represent the symbol alphabet as a set

$$A = \{a_0, a_1, \ldots, a_{M-1}\}. \qquad 1.2.1$$

The number of elements in a set is called its *cardinality* and is written

$$M = |A|.$$

The source outputs symbols in a time sequence represented by the notation

$$\bar{a} = (s_0 s_1 \ldots s_t \ldots), \qquad 1.2.2$$

where $s_t \in A$ is the symbol emitted by the source at time t. In this text, we shall take t to be an integer time index unless stated otherwise. At any given time index, the probability that the source emits symbol a_m is written as $p_m = \Pr(a_m)$. If the set of probabilities

$$P_A = \{p_0, p_1, \cdots, p_{M-1}\} \qquad 1.2.3$$

is not a function of time, the source is said to be *stationary.* Since it is certain that the source emits only members of its alphabet A, we have

$$\sum_{m=0}^{M-1} p_m = 1. \qquad 1.2.4$$

Mathematically, the simplest sources to treat are *synchronous* sources which emit a new symbol at a fixed time interval T_s. An *asynchronous* source is one in which the time interval between emitted symbols is not fixed. Such a source can be modeled in an approximate fashion by defining one of its symbols, say a_0, to be a *null character.* If the

此为试读，需要完整PDF请访问：www.ertongbook.com