

SYMBOLS, SIGNALS
AND NOISE: *The Nature
and Process of Communication*

BY J. R. PIERCE

SYMBOLS, SIGNALS
AND NOISE: *The Nature
and Process of Communication*

BY J. R. PIERCE

HARPER MODERN SCIENCE SERIES
EDITED BY JAMES R. NEWMAN

HARPER & BROTHERS, NEW YORK

SYMBOLS, SIGNALS AND NOISE

COPYRIGHT © 1961 BY JOHN R. PIERCE

PRINTED IN THE UNITED STATES OF AMERICA

ALL RIGHTS RESERVED. NO PART OF THE BOOK MAY BE USED OR REPRODUCED IN ANY MANNER WHATSOEVER WITHOUT WRITTEN PERMISSION EXCEPT IN THE CASE OF BRIEF QUOTATIONS EMBODIED IN CRITICAL ARTICLES AND REVIEWS. FOR INFORMATION ADDRESS HARPER & BROTHERS, 49 EAST 33RD STREET, NEW YORK 16, N.Y.

FIRST EDITION

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 61-10215

Preface

WHEN JAMES R. NEWMAN suggested to me that I write a book about communication I was delighted. All my technical work has been inspired by one aspect or another of communication. Of course I would like to tell others what seems to me to be interesting and challenging in this important field.

It would have been difficult to do this and to give any sense of unity to the account before Claude E. Shannon published "A Mathematical Theory of Communication" in 1948. Shannon's communication theory, which is also called information theory, has brought into a reasonable relation the many problems that have been troubling communication engineers for years. It has created a broad but clearly defined and limited field where before there were many special problems and ideas whose interrelations were not well understood. No one can accuse me of being a Shannon worshiper and get away unrewarded.

Thus, I felt that my account of communication must be an account of communication theory as Shannon formulated it. The account would have to be broader than Shannon's in that it would discuss the relation or lack of relation of communication theory to the many fields to which people have applied it or tried to apply it. It would have to be narrower than Shannon's account in that it would have to be less mathematical.

Here came the rub. My account could be *less* mathematical than Shannon's, but it could not be *nonmathematical*. Communication theory is a mathematical theory. It starts from certain premises

that define the aspects of communication with which it will deal, and it proceeds from these premises to various logical conclusions. The glory of communication theory lies in certain mathematical theorems which are both surprising and important. To talk about communication theory without communicating its real mathematical content would be like endlessly telling a man about a wonderful composer yet never letting him hear an example of the composer's music.

How was I to proceed? It seemed to me that I had to make the book self-contained, so that any mathematics in it could be understood without referring to other books or without calling for the particular content of early mathematical training, such as high school algebra. Did this mean that I had to avoid mathematical notation? Not necessarily, but any mathematical notation would have to be explained in the most elementary terms. I have done this both in the text and in an appendix; by going back and forth between the two the mathematically untutored reader should be able to resolve any difficulties.

But just how difficult should the most difficult mathematical arguments be? Although it meant sliding over some very important points, I resolved to keep things easy compared with, say, the more difficult parts of Newman's *The World of Mathematics*. When the going is very difficult, I have merely indicated the general nature of the sort of mathematics used rather than trying to describe its content clearly.

Nonetheless, this book has sections which will be hard for the nonmathematical reader. I advise him merely to skim through these, gathering what he can. When he has gone through the book in this manner he will see why the difficult sections are there. Then he can turn back and understand them if he wishes. But, had I not put these difficult sections in, and had the reader wanted the sort of understanding that takes real thought, he would have been stuck. As far as I know, other available literature on communication theory is either too simple or too difficult to help the diligent but inexperienced reader beyond the easier parts of this book. I might note also that some of the literature is confused and some of it is just plain wrong.

By this sort of talk I may have raised wonder in the reader's

mind as to whether or not communication theory is really worth so much trouble, either on his part or on mine for that matter. I can only say that to the degree that the whole world of science and technology around us is important, communication theory is important, for it is an important part of that world. To the degree to which an intelligent reader wants to know something both about that world and about communication theory, it is worth his while trying to get a clear picture. Such a picture must show communication theory neither as something utterly alien and unintelligible nor as something that can be epitomized in a few easy words and appreciated without effort.

The process of writing this book was not easy. Of course it could never have been written at all but for the work of Claude Shannon, who, besides inspiring the book through his work, read the manuscript and suggested several valuable changes. David Slepian jolted me out of the rut of error and confusion in an even more vigorous way. E. N. Gilbert deflected me from error in several instances. Milton Babbitt reassured me concerning the major contents of the chapter on information theory and art and suggested a few changes. P. D. Bricker, H. M. Jenkins, and R. N. Shepard advised me in the field of psychology, but the views I finally expressed should not be attributed to them. M. V. Mathews provided the computer program in Chapter XI. Benoit Mandelbrot helped me with Chapter XII. J. P. Runyon read the manuscript with care, and Eric Wolman uncovered an appalling number of textual errors as well as making valuable suggestions. The reader is indebted to James R. Newman for the fact that I have provided a glossary, summaries at the ends of some chapters, and for my final attempts to make some difficult points a little clearer. To all of these I am indebted and not less to Miss F. M. Costello, who triumphed over the chaos of preparing and correcting the manuscript and figures.

J. R. PIERCE

About the Author

Dr. John R. Pierce was born in Des Moines, Iowa, in 1910 and spent his early life in the Midwest. He received his undergraduate education at the California Institute of Technology in Pasadena, and his M.A. and Ph.D. degrees in electrical engineering from the same institution.

Dr. Pierce has been with the Bell Telephone Laboratories since 1936, and is at present Director of Research in Communications Principles at their laboratory in Murry Hill, New Jersey. He lives with his wife and two children in Berkeley Heights, New Jersey.

Dr. Pierce's writings have appeared in *Scientific American*, *The Atlantic Monthly*, *Coronet*, and several science fiction magazines. His books include *Man's World of Sound* (with E. E. David); *Electrons' Waves and Messages*; *Waves and the Ear* (with W. A. van Bergeijk and E. E. David); and two highly technical books, *Traveling Wave Tubes* and *Theory and Design of Electron Beams* (the latter has appeared in a Russian translation).

Dr. Pierce is a member of several scientific societies, including the National Academy of Sciences, and is a fellow of the Institute of Radio Engineers. He has received the following awards: Eta Kappa Nu, Outstanding Young Electrical Engineer, 1942; Morris Liebmann Memorial Prize, 1947; Stuart Ballantine Medal, 1960.

Contents

	PREFACE	ix
I.	THE WORLD AND THEORIES	1
II.	THE ORIGINS OF INFORMATION THEORY	19
III.	A MATHEMATICAL MODEL	45
IV.	ENCODING AND BINARY DIGITS	64
V.	ENTROPY	78
VI.	LANGUAGE AND MEANING	107
VII.	EFFICIENT ENCODING	125
VIII.	THE NOISY CHANNEL	145
IX.	MANY DIMENSIONS	166
X.	INFORMATION THEORY AND PHYSICS	184
XI.	CYBERNETICS	208
XII.	INFORMATION THEORY AND PSYCHOLOGY	229
XIII.	INFORMATION THEORY AND ART	250
XIV.	BACK TO COMMUNICATION THEORY	268
	APPENDIX: ON MATHEMATICAL NOTATION	278
	GLOSSARY	287
	INDEX	295

CHAPTER I *The World and Theories*

IN 1948, CLAUDE E. SHANNON published a paper called "A Mathematical Theory of Communication"; it appeared in book form in 1949. Before that time, a few isolated workers had from time to time taken steps toward a general theory of communication. Now, twelve years later communication theory, or information theory as it is sometimes called, is an accepted field of research. Several books on communication theory have been published, and several international symposia and conferences have been held. The Institute of Radio Engineers has a professional group on information theory, whose learned *Transactions* appears quarterly, and the journal *Information and Control* is largely devoted to communication theory.

All of us use the words communication and information, and we are unlikely to underestimate their importance. A modern philosopher, A. J. Ayer, has commented on the wide meaning and importance of communication in our lives. We communicate, he observes, not only information, but also knowledge, error, opinions, ideas, experiences, wishes, orders, emotions, feelings, moods. Heat and motion can be communicated. So can strength and weakness and disease. He cites other examples and comments on the manifold manifestations and puzzling features of communication in man's world.

Surely, communication being so various and so important, a

theory of communication, a theory of generally accepted soundness and usefulness, must be of incomparable importance to all of us. When we add to *theory* the word *mathematical*, with all its implications of rigor and magic, the attraction becomes almost irresistible. Perhaps if we learn a few formulae our problems of communication will be solved, and we shall become the masters of information rather than the slaves of misinformation.

Unhappily, this is not the course of science. Some 2,300 years ago, another philosopher, Aristotle, discussed in his *Physics* a notion as universal as that of communication, that is, motion.

Aristotle defined motion as the fulfillment, insofar as it exists potentially, of that which exists potentially. He included in the concept of motion the increase and decrease of that which can be increased or decreased, coming to and passing away, and also being built. He spoke of three categories of motion, with respect to magnitude, affection, and place. He found, indeed, as he said, as many types of motion as there are meanings of the word *is*:

Here we see motion in all its manifest complexity. The complexity is perhaps a little bewildering to us, for the associations of words differ in different languages, and we would not necessarily associate motion with all the changes of which Aristotle speaks.

How puzzling this universal matter of motion must have been to the followers of Aristotle. It remained puzzling for over two millennia, until Newton enunciated the laws which engineers still use in designing machines and astronomers in studying the motions of stars, planets, and satellites. While later physicists have found that Newton's laws are only the special forms which more general laws assume when velocities are small compared with that of light and when the scale of the phenomena is large compared with the atom, they are a living part of our physics rather than a historical monument. Surely, when motion is so important a part of our world, we should study Newton's laws of motion. They say:

1. A body continues at rest or in motion with a constant velocity in a straight line unless acted upon by a force.
2. The change in velocity of a body is in the direction of the force acting on it, and the magnitude of the change is proportional to the force acting on the body times the time during which the force acts, and is inversely proportional to the mass of the body.

3. Whenever a first body exerts a force on a second body, the second body exerts an equal and oppositely directed force on the first body.

To these laws Newton added the universal law of gravitation:

4. Two particles of matter attract one another with a force acting along the line connecting them, a force which is proportional to the product of the masses of the particles and inversely proportional to the square of the distance separating them.

Newton's laws brought about a scientific and a philosophical revolution. Using them, Laplace reduced the solar system to an explicable machine. They have formed the basis of aviation and rocketry, as well as of astronomy. Yet, they do little to answer many of the questions about motion which Aristotle considered. Newton's laws solved the problem of motion as Newton defined it, not of motion in all the senses in which the word could be used in the Greek of the fourth century before our Lord or in the English of the twentieth century after.

Our speech is adapted to our daily needs or, perhaps, to the needs of our ancestors. We cannot have a separate word for every distinct object and for every distinct event; if we did we should be forever coining words, and communication would be impossible. In order to have language at all, many things or many events must be referred to by one word. It is natural to say that both men and horses run (though we may prefer to say that horses gallop) and convenient to say that a motor runs and to speak of a run in a stocking or a run on a bank.

The unity among these concepts lies far more in our human language than in any physical similarity with which we can expect science to deal easily and exactly. It would be foolish to seek some elegant, simple, and useful scientific theory of running which would embrace runs of salmon and runs in hose. It would be equally foolish to try to embrace in one theory all the motions discussed by Aristotle or all the sorts of communication and information which later philosophers have discovered.

In our everyday language, we use words in a way which is convenient in our everyday business. Except in the study of language itself, science does not seek understanding by studying words and their relations. Rather, science looks for things in nature, including

our human nature and activities, which, can be grouped together and understood. Such understanding is an ability to see what complicated or diverse events really do have in common (the planets in the heavens and the motions of a whirling skater on ice, for instance) and to describe the behavior accurately and simply.

The words used in such scientific descriptions are often drawn from our everyday vocabulary. Newton used force, mass, velocity, and attraction. When used in science, however, a particular meaning is given to such words, a meaning narrow and often new. We cannot discuss in Newton's terms force of circumstance, mass media, or the attraction of Brigitte Bardot. Neither should we expect that communication theory will have something sensible to say about every question we can phrase using the words communication or information.

A valid scientific theory seldom if ever offers the solution to the pressing problems which we repeatedly state. It seldom supplies a sensible answer to our multitudinous questions. Rather than rationalizing our ideas, it discards them entirely, or, rather, it leaves them as they were. It tells us in a fresh and new way what aspects of our experience can profitably be related and simply understood. In this book, it will be our endeavor to seek out the ideas concerning communication which can be so related and understood.

When the portions of our experience which can be related have been singled out, and when they have been related and understood, we have a *theory* concerning these matters. Newton's laws of motion form an important part of *theoretical physics*, a field called *mechanics*. The laws themselves are not the whole of the theory: they are merely the basis of it, as the axioms or postulates of geometry are the basis of geometry. The theory embraces both the assumptions themselves and the mathematical working out of the logical consequences which must necessarily follow from the assumptions. Of course, these consequences must be in accord with the complex phenomena of the world about us if the theory is to be a valid theory, and an invalid theory is useless.

The ideas and assumptions of a theory determine the *generality* of the theory, that is, to how wide a range of phenomena the theory applies. Thus, Newton's laws of motion and of gravitation

are very general; they explain the motion of the planets, the time-keeping properties of a pendulum, and the behavior of all sorts of machines and mechanisms. They do not, however, explain radio waves.

Maxwell's equations¹ explain all (non-quantum) electrical phenomena; they are very general. A branch of electrical theory called *network theory* deals with the electrical properties of electrical circuits, or networks, made by interconnecting three sorts of idealized electrical structures: resistors (devices such as coils of thin, poorly conducting wire or films of metal or carbon, which impede the flow of current), inductors (coils of copper wire, sometimes wound on magnetic cores), and capacitors (thin sheets of metal separated by an insulator or dielectric such as mica or plastic; the Leyden jar was an early form of capacitor). Because network theory deals only with the electrical behavior of certain specialized and idealized physical structures, while Maxwell's equations describe the electrical behavior of any physical structure, a physicist would say that network theory is *less* general than are Maxwell's equations, for Maxwell's equations cover the behavior not only of idealized electrical networks but of all physical structures and include the behavior of radio waves, which lies outside of the scope of network theory.

Certainly, the most general theory, which explains the greatest range of phenomena, is the most powerful and the best; it can always be specialized to deal with simple cases. That is why physicists have sought a unified field theory to embrace mechanical laws and gravitation and all electrical phenomena. It might, indeed, seem that all theories could be ranked in order of generality, and, if this is possible, we should certainly like to know the place of communication theory in such a hierarchy.

Unfortunately, life isn't as simple as this. In one sense, network theory is less general than Maxwell's equations. In another sense,

¹ In 1873, in his treatise *Electricity and Magnetism*, James Clerk Maxwell presented and fully explained for the first time the natural laws relating electric and magnetic fields and electric currents. He showed that there should be *electromagnetic waves* (radio waves) which travel with the speed of light. Hertz later demonstrated these experimentally, and we now know that light is electromagnetic waves. Maxwell's equations are the mathematical statement of Maxwell's theory of electricity and magnetism. They are the foundation of all electric art.

however, it is more general, for all the mathematical results of network theory hold for vibrating mechanical systems made up of idealized mechanical components as well as for the behavior of interconnections of idealized electrical components. In mechanical applications, a spring corresponds to a capacitor, a mass to an inductor, and a dashpot or damper, such as that used in a door closer to keep the door from slamming, corresponds to a resistor. In fact, network theory might have been developed to explain the behavior of mechanical systems, and it is so used in the field of acoustics. The fact that network theory evolved from the study of idealized electrical systems rather than from the study of idealized mechanical systems is a matter of history, not of necessity.

Because all of the mathematical results of network theory apply to certain specialized and idealized mechanical systems, as well as to certain specialized and idealized electrical systems, we can say that in a sense network theory is *more* general than Maxwell's equations, which do not apply to mechanical systems at all. In another sense, of course, Maxwell's equations are more general than network theory, for Maxwell's equations apply to all electrical systems, not merely to a specialized and idealized class of electrical circuits.

To some degree we must simply admit that this is so, without being able to explain the fact fully. Yet, we can say this much. Some theories are very strongly *physical* theories. Newton's laws and Maxwell's equations are such theories. Newton's laws deal with mechanical phenomena; Maxwell's equations deal with electrical phenomena. Network theory is essentially a *mathematical* theory. The terms used in it can be given various physical meanings. The theory has interesting things to say about different physical phenomena, about mechanical as well as electrical vibrations.

Often a mathematical theory is the offshoot of a physical theory or of physical theories. It can be an elegant mathematical formulation and treatment of certain aspects of a general physical theory. Network theory is such a treatment of certain physical behavior common to electrical and mechanical devices. A branch of mathematics called *potential theory* treats problems common to electric, magnetic, and gravitational fields and, indeed, in a degree to aerodynamics. Some theories seem, however, to be more mathematical than physical in their very inception.

We use many such mathematical theories in dealing with the physical world. Arithmetic is one of these. If we label one of a group of apples, dogs, or men 1, another 2, and so on, and if we have used up just the first 16 numbers when we have labeled all members of the group, we feel confident that the group of objects can be divided into two equal groups each containing 8 objects ($16 \div 2 = 8$) or that the objects can be arranged in a square array of four parallel rows of four objects each (because 16 is a perfect square; $16 = 4 \times 4$). Further, if we line the apples, dogs, or men up in a row, there are 2,092,278,988,800 possible sequences in which they can be arranged, corresponding to the 2,092,278,988,800 different sequences of the integers 1 through 16. If we used up 13 rather than 16 numbers in labeling the complete collection of objects, we feel equally certain that the collection could not be divided into any number of equal heaps, because 13 is a prime number and cannot be expressed as a product of factors.

This seems not to depend at all on the nature of the objects. Insofar as we can assign numbers to the members of any collection of objects, the results we get by adding, subtracting, multiplying, and dividing numbers or by arranging the numbers in sequence hold true. The connection between numbers and collections of objects seems so natural to us that we may overlook the fact that arithmetic is itself a mathematical theory which can be applied to nature only to the degree that the properties of numbers correspond to properties of the physical world.

Physicists tell us that we can talk sense about the total number of a group of elementary particles, such as electrons, but we can't assign particular numbers to particular particles because the particles are in a very real sense indistinguishable. Thus, we can't talk about arranging such particles in different orders, as numbers can be arranged in different sequences. This has important consequences in a part of physics called *statistical mechanics*. We may also note that while Euclidean geometry is a mathematical theory which serves surveyors and navigators admirably in their practical concerns, there is reason to believe that Euclidean geometry is not quite accurate in describing astronomical phenomena.

How can we describe or classify theories? We can say that a theory is very narrow or very general in its scope. We can also distinguish theories as to whether they are strongly physical or

strongly mathematical. Theories are strongly physical when they describe very completely some range of physical phenomena, which in practice is always limited. Theories become more mathematical or abstract when they deal with an idealized class of phenomena or with only certain aspects of phenomena. Newton's laws are strongly physical in that they afford a complete description of mechanical phenomena such as the motions of the planets or the behavior of a pendulum. Network theory is more toward the mathematical or abstract side in that it is useful in dealing with a variety of idealized physical phenomena. Arithmetic is very mathematical and abstract; it is equally at home with one particular property of many sorts of physical entities, with numbers of dogs, numbers of men, and (if we remember that electrons are indistinguishable) with numbers of electrons. It is even useful in reckoning numbers of days.

In these terms, communication theory is both very strongly mathematical and quite general. Although communication theory grew out of the study of electrical communication, it attacks problems in a very abstract and general way. It provides, in the *bit*, a universal measure of amount of information in terms of choice or uncertainty. Specifying or learning the choice between two equally probable alternatives, which might be messages or numbers to be transmitted, involves one bit of information. Communication theory tells us how many bits of information can be sent per second over perfect and imperfect communication channels in terms of rather abstract descriptions of the properties of these channels. Communication theory tells us how to measure the rate at which a message source, such as a speaker or a writer, generates information. Communication theory tells us how to represent, or *encode*, messages from a particular message source efficiently for transmission over a particular sort of channel, such as an electrical circuit, and it tells us when we can avoid errors in transmission.

Because communication theory discusses such matters in very general and abstract terms, it is sometimes difficult to use the understanding it gives us in connection with particular, practical problems. However, because communication theory has such an abstract and general mathematical form, it has a very broad field of application. Communication theory is useful in connection with

written and spoken language, the electrical and mechanical transmission of messages, the behavior of machines, and, perhaps, the behavior of people. Some feel that it has great relevance and importance to physics in a way that we shall discuss much later in this book.

Primarily, however, communication theory is, as Shannon described it, a *mathematical* theory of communication. The concepts are formulated in mathematical terms, of which widely different physical examples can be given. Engineers, psychologists, and physicists may use communication theory, but it remains a mathematical theory rather than a physical or psychological theory or an engineering art.

It is not easy to present a mathematical theory to a general audience, yet communication theory is a mathematical theory, and to pretend that one can discuss it while avoiding mathematics entirely would be ridiculous. Indeed, the reader may be startled to find equations and formulae in these pages; these state accurately ideas which are also described in words, and I have included an appendix on mathematical notation to help the nonmathematical reader who wants to read the equations aright.

I am aware, however, that mathematics calls up chiefly unpleasant pictures of multiplication, division, and perhaps square roots, as well as the possibly traumatic experiences of high-school classrooms. This view of mathematics is very misleading, for it places emphasis on special notation and on tricks of manipulation, rather than on the aspect of mathematics that is most important to mathematicians. Perhaps the reader has encountered theorems and proofs in geometry; perhaps he has not encountered them at all, yet theorems and proofs are of primary importance in all mathematics, pure and applied. The important results of information theory are stated in the form of mathematical theorems, and these *are* theorems only because it is possible to prove that they are true statements.

Mathematicians start out with certain assumptions and definitions, and then by means of mathematical arguments or proofs they are able to show that certain statements or theorems are true. This is what Shannon accomplished in his "Mathematical Theory of Communication." The truth of a theorem depends on the validity