

Introduction

1. Shannon's Information Theory
2. Source Coding Theorem
3. Channel Coding Theorem
4. Rate Distortion Theorem

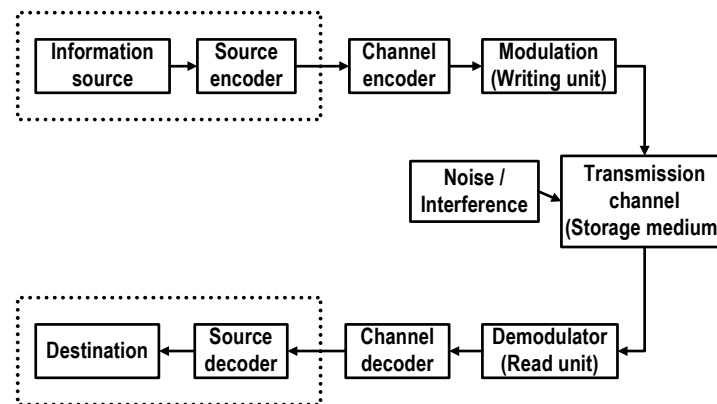
Appendix A: Historical Notes -- Shannon

Appendix B: Britannica Online Encyclopedia

⊙ Signals

<i>Amplitude</i>	<i>Time</i>	
	Discrete-time	Continuous-time
Discrete-amplitude	Digital	
Continuous-amplitude	Discrete-time (Sampled-data)	Analog (Waveform)

⊙ Block Diagram of a Communication System



- a. Source coding: reduce data redundancy
- b. Channel coding: reduce (channel) noise effect
- c. Modulation: select waveforms to cope with (analog) channel constraints (and noise)

⊙ Coding ↔ Mapping

1. Source coding:

Convert the original symbol stream to minimum binary representation

2. Channel Coding:

Convert the original symbol stream to (noise) robust symbol stream

3. Modulation:

Map the original symbol stream to waveforms of adequate form (meet channel constraints)

⊙ Nature of this Course

The theoretical principles here (digital communications) fall roughly into two categories – the first provides analytical tools for determining the performance of particular systems, and the second puts fundamental limits on the performance of any system. Much of the first category can be understood by engineering undergraduates, while the second category is distinctly graduate in nature. (Gallager, *Principles of Digital Comm*, Preface)

⊙ *Basic Questions in Information and Communication Theory* (Blahut, pp.2~3)

Information theory attempts to answer a number of very basic questions:

1. What is Information? That is, how do we measure it?
2. What are the fundamental limits on the transmission of information?
3. What are the fundamental limits on the extraction of Information from the environment?
4. What are the fundamental limits on the compression and refinement of information?
5. How should devices be designed to approach these limits?
6. How closely do existing devices approach these limits?

Information theory includes the quest for a theory of optimum communication waveforms. When restricted to this quest, the subject matter of information theory can be viewed as the existing answers to a sequence of successively weaker queries:

1. What is the best waveform family for transmitting k bits in T seconds through a specified channel? (For example, how should we send bits through an additive

gaussian-noise channel with an average power constraint and a bandwidth constraint on the transmitter?)

2. What is the performance of the best such waveform family for a continuous-time channel (even though the waveform itself may be unknown)?
3. What is the approximate performance of the best waveform family (or code) for transmitting information through a discrete memoryless channel such as a binary channel?

C.E. Shannon publishes a series of papers that attempt to answer the above questions. His papers stimulated a new research field, the so-called Shannon Theory or Information Theory. Essentially his theory has three main theorems: Source coding theorem, channel coding theorem and rate distortion theorem.

1. Source Coding Theorem (Shannon's first theorem)

A simplified version of this theorem can be stated as follows:

Given a discrete *memoryless* source (DMS) S of entropy $H(S)$, it is possible to encode it with an *instantaneous* code having the average code-word length L and $L \geq H(S) + \epsilon$, for any $\epsilon > 0$. Conversely, it is not possible to encode such a source with an arbitrary small error with an average length shorter than $H(S)$.

(Abramson, p.73; Blahut, p.75; Lafrance, p. 199; C&T, p. 113)

- The entropy is a measure of the *average information* content per source symbol. (Least number of bits to represent, on the average, an information source.)
- The source coding theorem is also known as the “noiseless coding theorem” in the sense that it establishes the condition for error-free encoding to be possible.

2. Channel Coding Theorem (Shannon's 2nd theorem)

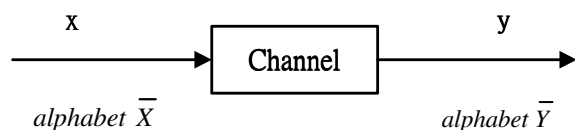
- A simplified version of this theorem is as follows:

Given a channel with capacity C , for $R < C$ and any $\varepsilon > 0$, there exists a block code C of length n and rate R whose block decoding error $P_E < \varepsilon$.

Conversely, if $R > C$, then P_E cannot be arbitrarily small.

(Abramson, p.166; Blahut, p.174,163; Lafrance, p. 208; C&T, p. 200; R&C, p.16, last paragraph)

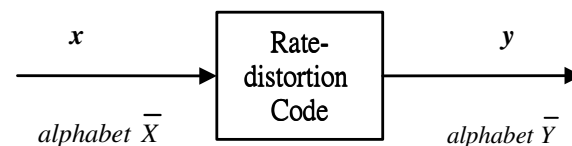
- The theorem specifies the channel capacity C as a fundamental limit on the rate at which the transmission of reliable error-free message can take place over a *well-defined* channel, say, DMC.



- The Channel Coding Theorem is also known as “noisy coding theorem”.

3. Rate Distortion Theorem (Shannon's 3rd theorem)

- What is the minimum bit rate representing a source when a certain amount of distortion is allowed? This may be viewed as an extension of the noiseless source coding theorem. A simplified version of the so-called *rate distortion theorem* is as follows.
- A *rate distortion code* is a mapping from the original message to its approximation. Because the reconstruction is not perfect, a distortion is thus introduced.



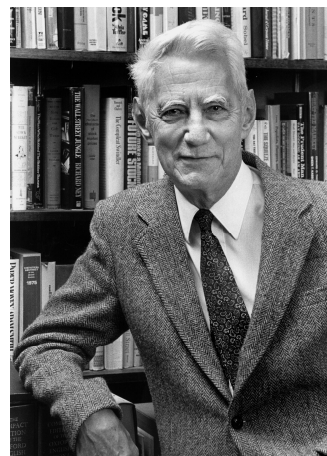
- Theorem:** Given a DMS X and single letter distortion measure $d(x, y)$, and let $R(D)$ be the rate distortion function of X with respect to $d(\cdot, \cdot)$. Then, for any $\varepsilon > 0$ and D_0 , there exist a code of block length n with the coding rate bounded by $R < R(D_0) + \varepsilon$ and with the dis-

tortion bounded by $D < D_0 + \varepsilon$. Conversely, if the rate is less than $R(D_0)$, then the distortion must exceed D_0 .

(Blahut, p.225; Lafrance, p. 385; C&T, p. 307)

- $R(D)$ is characterized by the joint-prob of X and Y (and $d(x,y)$).
- In transmitting continuous-amplitude signals using digital (binary) representations, it is unavoidable to allow distortion. The *entropy* (distortionless representation) of a real (floating) number may be infinity, but its rate-distortion function (with a reasonable small amount of distortion) is well-defined (finite value).

▪ Appendix A. Historical Notes: Shannon



Claude Elwood Shannon
(1916-2001)

- The seeds of Shannon's information theory appeared first in a Bell Labs classified memorandum dated 1 September 1945, "A Mathematical Theory of Cryptography" (revised form was published in BSTJ, 1949.)
- Shannon's work on cryptography problem at Bell Labs brought him into contact with two scholars, R.V.L. Hartley and H. Nyquist, who were major consultants to the cryptography research.
- Nyquist (1924) has shown that a certain bandwidth was necessary in order to send telegraph signals at a definite rate.
- Hartley (1928) had attempted to quantify of information as the logarithm of the number of possible message built from a pool of symbols.
- Shannon was aware of Norbert Wiener's work on cybernetics (He cited Wiener's book in his two 1948 articles on information theory). Wiener had recognized

that the communication of information was a problem in statistics.

- Shannon took a mathematics course from Norbert Wiener when he was studying at MIT, and Shannon had accessed to the "Yellow Peril" report (Wiener's book before publication) while he worked on cryptographic research.

Major Publications of Shannon

- A Mathematical Theory of Communication, I, II, published in Bell Syst. Tech. J, 1948, pp.379-423, 623-656.
- Communication Theory of Secrecy Systems, published in BSTJ, 1949, pp.656-715.
- Coding Theorems for a Discrete Source with a Fidelity Criterion, IRE Nat. Cov. Rec. Part 4, 1959, pp.142-163.
- Paperback edition of "The Mathematics Theory of Communication" was published by University of Illinois Press, 1963. 32,000 copies have been sold up to 1964.
- In a review of the accomplishments of Bell Labs. in communication science, Millman in his edited book (1984), "A history of engineering and science in the Bell system: Communication Science (1925-1980)" says: "Probably the most spectacular development in communication mathematics to take place at Bell Laboratories was the formulation in the 1940's of infor-

mation theory by C. E. Shannon."

- In his 1948 articles on information theory, Shannon credits J.W. Tukey (a professor at Princeton Univ.) with suggesting the word "bit" as a measure of information.
- In 1984, J.L. Massey wrote a paper "Information Theory: the Copernican System of Communications." Published in *IEEE Communications Magazine*, vol. 22, No. 12, pp. 26-28.
- The Shannon's theory of information is the scientific basis of communications in the same sense that the Copernicus' heliocentric theory is the scientific basis of astronomy.
- Commemorate Issue, "Fifty Years of Shannon Theory," *IEEE Transaction on Information Theory*, Oct. 1998. Vol.44, No.6, (1948-1998)

▪ **Appendix B.** Britannica Online Encyclopedia
Information Theory

Main

A mathematical representation of the conditions and parameters affecting the transmission and processing of information. Most closely associated with the work of the American electrical engineer in the mid-20th century, information theory is chiefly of interest to communication engineers, though some of the concepts have been adopted and used in such fields as psychology and linguistics. Information theory overlaps heavily with communication theory, but it is more oriented toward the fundamental limitations on the processing and communication of information and less oriented toward the detailed operation of particular devices.

- Historical background
- Classical information theory
- Applications of information theory
- Additional Reading
- Related Articles
- External Web sites
- Citations

(<http://www.britannica.com/>)

References

1. N. Abramson, *Information Theory and Coding*, McGraw-Hill, 1963.
2. P. Lafrance, *Fundamental Concepts in Communications*, Prentice-Hall, 1990.
3. R.E. Blahut, *Principles and Practice of Information Theory*, Addison-Wesley, 1987, 1990
4. T.M. Cover and J.A. Thomas, *Elements of Information Theory*, John Wiley and Sons, 2nd ed., 2006.
5. R.G. Gallager, *Principles of Digital Communications*, Cambridge, 2008.
6. R. W. Yueng, *Information Theory and Network Coding*, Springer, 2008.
7. S. Verdu, "Fifty Years of Shannon Theory," *IEEE Transaction on Information Theory*, pp.2057-2078, Oct. 1998.