ECE253/CSE208 Introduction to Information Theory

Lecture 1: Introduction & Overview

Dr. Yu Zhang

ECE Department
University of California, Santa Cruz

- Chap 1 of *Elements of Information Theory (2nd Edition) by Thomas Cover & Joy Thomas*.

Outline

- Course Information
- Overview of Information Theory

Course Information

Time

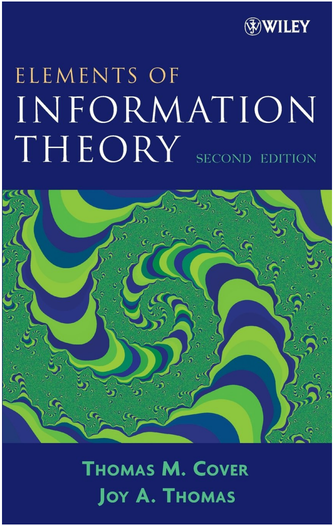
TuTh 5:20-6:55pm, 10/1/2020–12/11/2020, via Zoom meeting (recordings shared).
Course websites: Piazza & Gradescope.

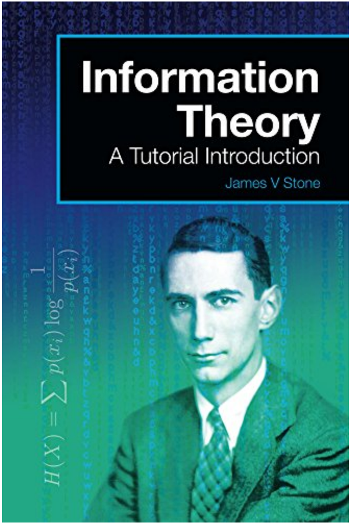
Instructor: Yu Zhang (zhangy@ucsc.edu). Office hour: Fri. 1–2pm via the lecture Zoom
TA: Kejun Chen (`kchen158@ucsc.edu`). Office hour: Mon. 3–4pm

`Supporting service for students with disabilities`: Please contact disability resource center (DRC, 831-459-2089, drc@ucsc.edu) for more information.

1. Students contact the DRC to request and receive their accommodation letter.
2. Students email the instructor with their accommodation letters, and/or any course related accommodation needs.
3. Students manage their own disclosure of disability status. Please maintain this confidential information according to UCSC data practices.

Textbooks


WILEY
ELEMENTS OF
INFORMATION
THEORY
SECOND EDITION
THOMAS M. COVER
JOY A. THOMAS


Information
Theory
A Tutorial Introduction
James V Stone
$H(X) = \sum p(x_i) \log \frac{1}{p(x_i)}$

Tentative Agenda

Week	Topic (chapters of R1)
Week-1	Course overview; probability theory revisited
Week-2,3	Entropy, mutual information, and inequalities (chap 2)
Week-4	AEP, entropy rate (chap 3-4)
Week-5,6	Data compression and gambling (chap 5-6)
Week-7	Channel capacity (chap 7);
Week-8	Differential entropy (chap 8); Gaussian channel I (chap 9); `midterm exam`
Week-9	Gaussian channel II (chap 9)
Week-10	Rate distortion theory (chap 10)
Week-11	Project presentations

Objectives and Prerequisites

Objectives:

- To understand the concepts of entropy, mutual information and channel capacity.
- To enhance the knowledge of probability and measures of information.
- To guide the student through the implications and consequences of fundamental theories and laws of information/coding theory with reference to the application in modern communication and computer systems.

Prerequisites:

1. Probability and stochastic processes, vector calculus, and theorem-proof exposition.
2. Optional: ECE261 *Error Control Coding* (by Prof. Hamid Sadjadpour).

Grading Policy

Workload

Homework	20%
Midterm exam	50%
Course project	30%

Grading scale

$[93, 100]$	A^+
$[86, 92]$	A
$[80, 85]$	A^-
$[75, 79]$	B^+
$[70, 74]$	B
$[60, 69]$	C
< 60	Fail

Academic integrity

- All assignments that are turned in must represent your own work (no plagiarism).
- Please ask for clarification when you are unsure if a particular behavior violates university and/or faculty policy.
- Complete information about academic integrity policy is available at `www.ue.ucsc.edu/academic_misconduct`.

Late Homework Policy

– Assignments and submissions are through Gradescope only.

– No late submissions are allowed.

– Instructor-granted extensions are only given in exceptional situations. The instructor and TA should be notified via email before the deadline.

Overview of Information Theory

What is Information?

JOIN MWU | GAMES | BROWSE THESAURUS | WORD OF THE DAY
FAVORITES

Merriam-Webster SINCE 1828

information

DICTIONARY THESAURUS

information noun

in·for·ma·tion | \ ˌin-fər-ˈmā-shən \

Definition of *information*

1 : the communication or reception of knowledge or intelligence

2 **a** **(1)** : knowledge obtained from investigation, study, or instruction

(2) : INTELLIGENCE, NEWS

(3) : FACTS, DATA

b : the attribute inherent in and communicated by one of two or more alternative sequences or arrangements of something (such as nucleotides in DNA or binary digits in a computer program) that produce specific effects

c **(1)** : a signal or character (as in a communication system or computer) representing data

(2) : something (such as a message, experimental data, or a picture) which justifies change in a construct (such as a plan or theory) that represents physical or mental experience or another construct

d : a quantitative measure of the content of information

specifically : a numerical quantity that measures the uncertainty in the outcome of an experiment to be performed

3 : the act of informing against a person

4 : a formal accusation of a crime made by a prosecuting officer as distinguished from an indictment presented by a grand jury

What is Information (Cont'd)?

- Information is a signal that has been processed within a context to give it meaning.
- Information is interpreted data.
- Information is meaningful.

Data + Meaning = Information

What is Information Theory?

Information theory

From Wikipedia, the free encyclopedia

Not to be confused with Information science.

Information theory studies the quantification, storage, and communication of information. It was originally proposed by Claude Shannon in 1948 to find fundamental limits on signal processing and communication operations such as data compression, in a landmark paper entitled "A Mathematical Theory of Communication". Applications of fundamental topics of information theory include lossless data compression (e.g. ZIP files), lossy data compression (e.g. MP3s and JPEGs), and channel coding (e.g. for DSL). Its impact has been crucial to the success of the Voyager missions to deep space, the invention of the compact disc, the feasibility of mobile phones, the development of the Internet, the study of linguistics and of human perception, the understanding of black holes, and numerous other fields.

Merriam-Webster SINCE 1828

JOIN MWU | GAMES | BROWSE THESAURUS | WORD OF THE DAY | FAVORITES

information theory

DICTIONARY | THESAURUS

information theory noun

Definition of *information theory*

: a theory that deals statistically with information, with the measurement of its content in terms of its distinguishing essential characteristics or by the number of alternatives from which it makes a choice possible, and with the efficiency of processes of communication between humans and machines

Communication of Information

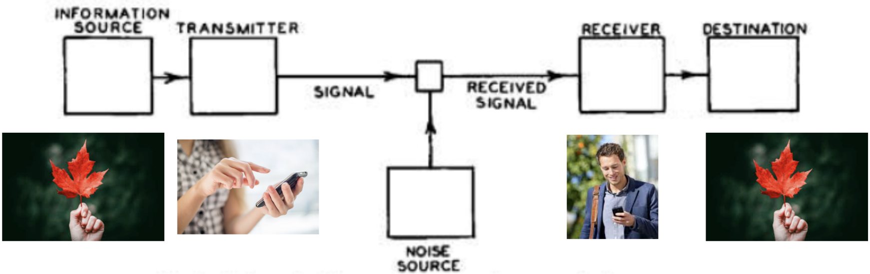


Figure: Credit to Prof. Tsachy Weissman at Stanford.

Key metrics for communication systems: Efficiency and Reliability.

Fundamental Questions

Information theory answers two fundamental questions in communication theory:

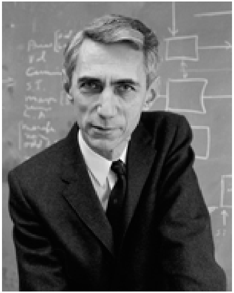
1. Q: What is the ultimate data compression?

   A: Entropy H.

2. Q: What is the ultimate transmission rate of communication?

   A: Channel capacity C.

Claude Shannon: The Father of IT



Reprinted with corrections from *The Bell System Technical Journal,*
Vol. 27, pp. 379–423, 623–656, July, October, 1948.

A Mathematical Theory of Communication

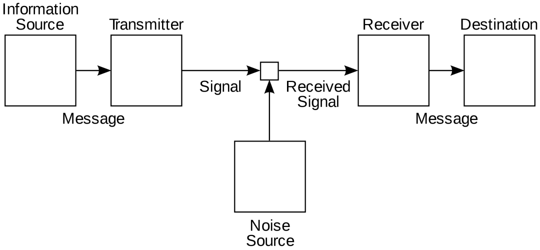
By C. E. SHANNON

INTRODUCTION

THE recent development of various methods of modulation such as PCM and PPM which exchange bandwidth for signal-to-noise ratio has intensified the interest in a general theory of communication. A basis for such a theory is contained in the important papers of Nyquist[1] and Hartley[2] on this subject. In the present paper we will extend the theory to include a number of new factors, in particular the effect of noise in the channel, and the savings possible due to the statistical structure of the original message and due to the nature of the final destination of the information.

- This landmark paper has almost 130,000 citations by Oct 1, 2020!!
- Shannon defined in mathematical terms what information is and how it can be transmitted in the face of noise.

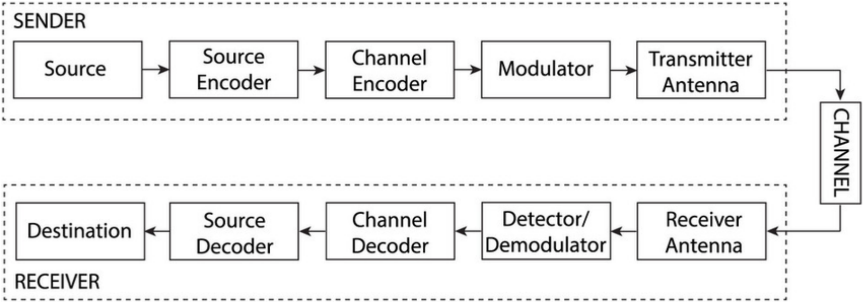
Shannon's Diagram of a General Communication System



Shannon's article laid out the basic elements of communication:

- An information source that produces a message.
- A transmitter that operates on the message to create a signal sent via a channel.
- A channel, which is the medium over which the signal is sent.
- A receiver, which transforms the signal back into the message intended for delivery.
- A destination (a person or a machine), for whom or which the message is intended.

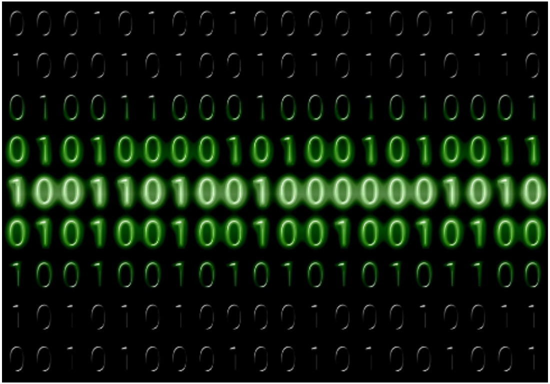
Communication System Diagram



Scratched CD

Music and data CDs are coded using error correcting codes and thus can still be read even if they have minor scratches using error detection and correction.

Information Bits

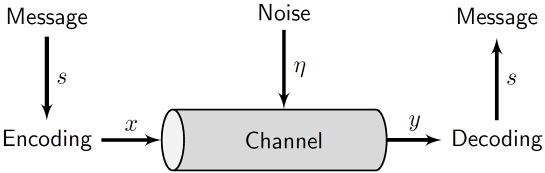


- succinct representation in bits (compression)
- effective and reliable communication of bits (across unreliable media)

Shannon discovered the two, and the mechanism that combines them is optimal.

Information: Physical Quantity

A basic idea in information theory is that information can be treated very much like a physical quantity, such as mass or energy.
Caude Shannon, 1985.



Shannon Information and Entropy

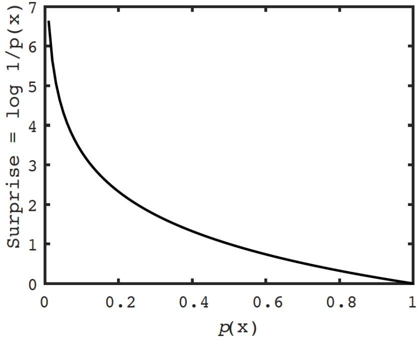
Shannon information: Consider a coin which lands heads up 90% of the time (i.e. $\Pr(x_h) = 0.9$). The less likely a particular outcome is, the more surprised we are. The Shannon information (surprisal of each outcome) of x_h is:

$$\boxed{\text{Shannon information} := \log_2 \frac{1}{\Pr(x_h)} = -\log_2 \Pr(x_h) \text{ bits}}$$

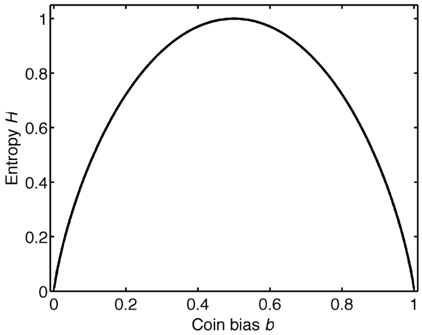
Entropy is Average Shannon information. We are often interested in how much surprise, on average, is associated with the entire set of possible values. The average surprise of a random variable X is defined as the entropy $H(X)$:

$$\boxed{H(X) := \Pr(x_h) \times \log \frac{1}{\Pr(x_h)} + \Pr(x_t) \times \log \frac{1}{\Pr(x_t)} \text{ bits}}$$

Shannon information and entropy (cont'd)

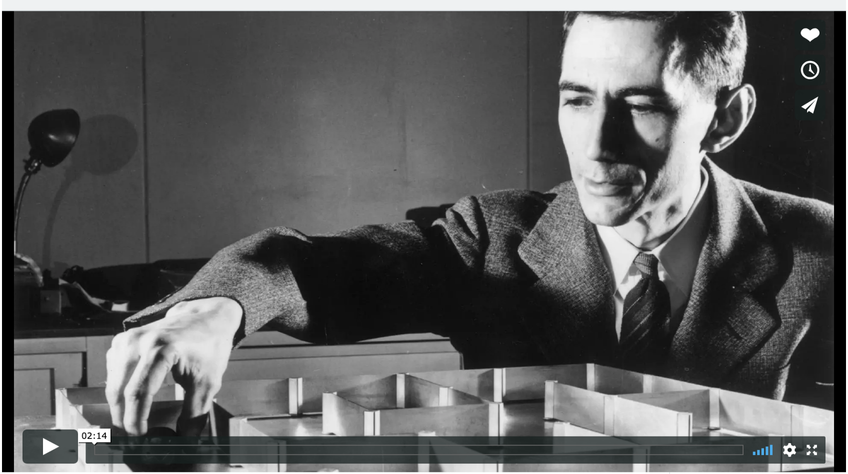


(a)



(b)

The Bit Player Trailer



https://vimeo.com/323615460

Interaction with Other Fields

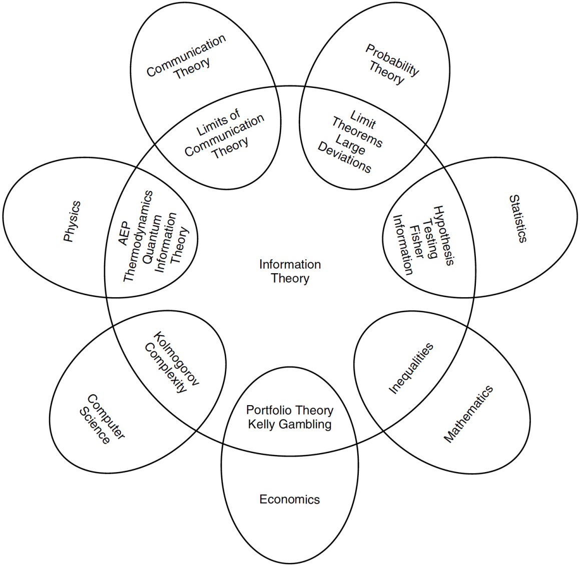


FIGURE 1.1. Relationship of information theory to other fields.

Interaction with Other Fields (cont'd)

- EE (Communications). It was thought to be impossible to send information at a positive rate with negligible probability of error. Shannon proved that the probability of error could be made nearly zero for all rates below channel capacity.
- CS (Kolmogorov Complexity). The complexity of a string of data can be defined by the length of the shortest binary computer program for computing the string (minimal description length). This definition is universal; i.e., computer independent.
- Physics (Thermodynamics). Statistical mechanics is the birthplace of entropy and the second law of thermodynamics. Entropy always increases. The second law allows one to dismiss any claims to perpetual motion machines.
- Economics (Investment). Repeated investment in a stationary stock market results in an exponential growth of wealth. The growth rate of the wealth is a dual of the entropy rate of the stock market.
- Philosophy of Science (Occam's Razor). William of Occam said "simpler solutions are more likely to be correct than complex ones."

Broad Topics

- **Information measures (entropy, divergence, mutual information):** Convexity, monotonicity and continuity properties. Extremization, saddle point, capacity as information radius.
- **Lossless data compression:** Variable length and fixed length. Linear compression. Slepian-Wolf problem. Ergodic sources: Shannon-McMillan and Birkhoff-Khintchine theorems. Basics of universal data compression. Optimality of Lempel-Ziv.
- **Binary hypothesis testing:** Bounds for finite number of samples. Asymptotics: Stein and Chernoff exponents. Large deviations: Sanov, I-projection, tilting.
- **Channel coding:** Achievability and converse bounds. Asymptotics: Capacity, strong converse, error-exponents, channel dispersion. Gaussian channels (parallel, with intersymbol interference, minimal energy-per-bit, continuous time). Coding with feedback: Zero and non-zero error capacities.
- **Lossy data compression:** Scalar quantization and Panter-Dite approximation. Vector quantization and rate-distortion theorem. Separation principle.

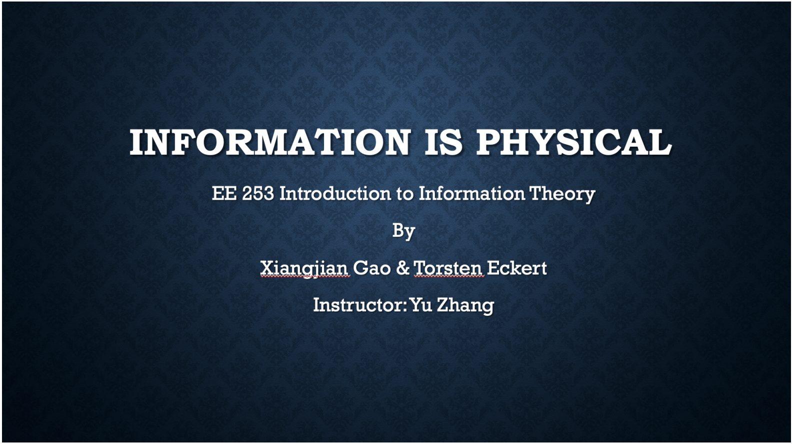
Potential Topics of Course Project

- Universal source coding
- Arithmetic coding
- Zero-error channel capacity
- MIMO channel capacity
- Capacity of channels with memory
- Algorithms for computing channel capacity and the rate-distortion function
- Bounds on the performance of block channel codes (error exponents)
- Low-density parity check (LDPC) codes
- Quantum information theory
- Information theory and biology

Should not be limited to these topics[1]: Any topics on information theory and data science with sufficient scope are good.

[1] Credit to Dr. Nihar Jindal

Sample Projects



Information Theory and Evolutionary Biology

By Chang Kim, Skyler Stewart and Kavin Subramanyam

Salient Region Detection

Jiahao Luo, Sanjay Damani

EE253 - June 5, 2019

Quantum Computing

G5: Kevin Bowden and Bryan Tor

Deep Learning from an Information-Theoretic Point of View

Shih-Ming Wang, Wei-Lin Wu

06-07-2019

Sample Based Estimation of Information Theoretic Measures

Applied to Bayesian Inference

Rene Gutierrez[1] Milad Gholami[1]

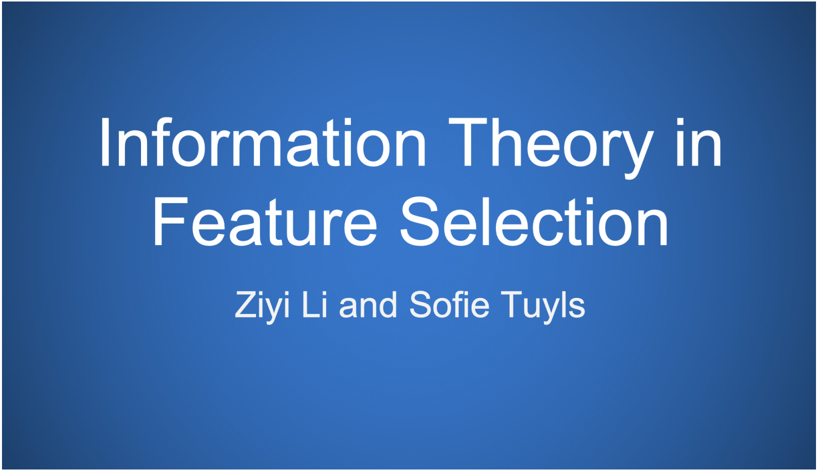
[1]Statistics, UCSC
[2]Electrical and Computer Engineering Department, UCSC

Spring 2019

A survey on classic Entropy Coding Algorithms and Asymmetric Numeral Systems

Howard Guo
Yifan Luo

Sample Projects



UNIVERSITY OF CALIFORNIA
SANTA CRUZ

Demystifying the Connection between Deep Neural Networks and Information Compression

Hadi Zarkoob and Jing Xiong

University of California, Santa Cruz

Dec 5, 2019

Thank You!

Email: <zhangy@ucsc.edu>

Homepage: https://people.ucsc.edu/~yzhan419/