Information and coding theory

Louis Wehenkel

Institut Montefiore, University of Liège, Belgium

ELEN060-2
Information and coding theory
February 2021

Outline

- Main objectives of this course
- Some practical contributions of information theory
- Outline of the way of thinking for this course

  The questions answered by Claude Shannon in 1948

  Where did these answers lead in the subsequent 60 years
- Course organization

  What we will do

  What you should do

  (Reciprocal) Evaluation

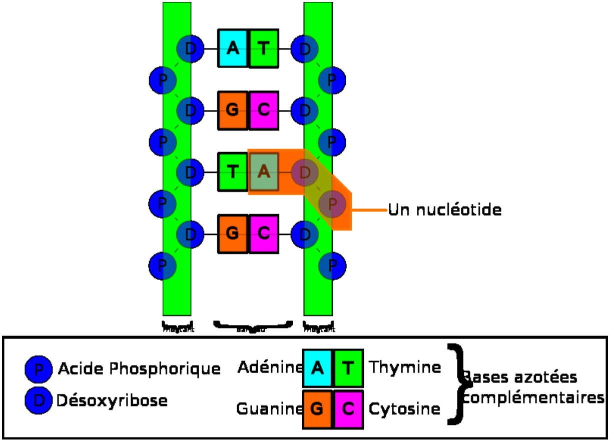
Main objectives of this course

- quantify the notion of information in a mathematically and intuitively sound way
- explain how this quantitative measure of information may be used in order to build efficient solutions to multitudinous engineering problems
- show its far-reaching interest in other fields (economics, biology, statistics, computer science, artificial intelligence...)

Telemedicine: multimedia information transmission

IT contribution 2

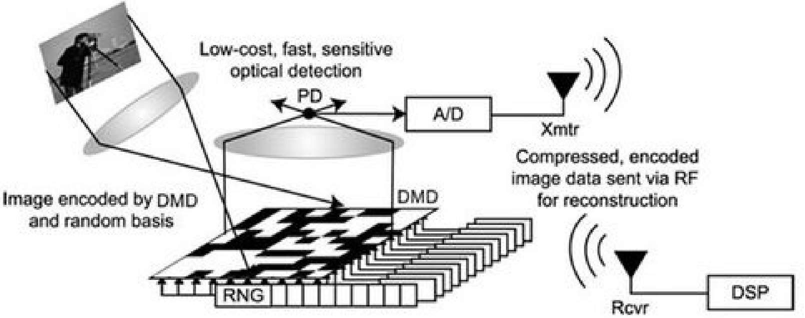
Virtual meeting: multimedia information transmission



Biology: better understanding biological information processing

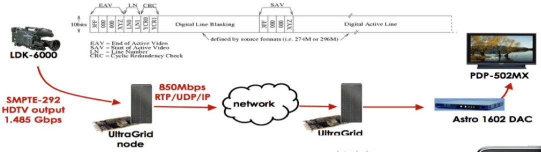
Data storage: increasing fidelity and lower cost of digital information storage

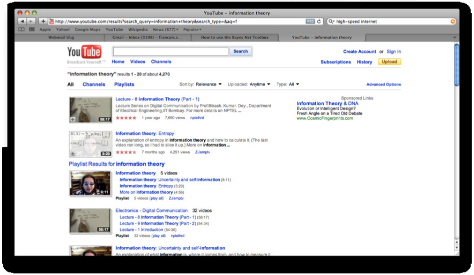
Security: encryption of private information



Compressed sensing: next generation cameras

Other IT contributions





and many more....

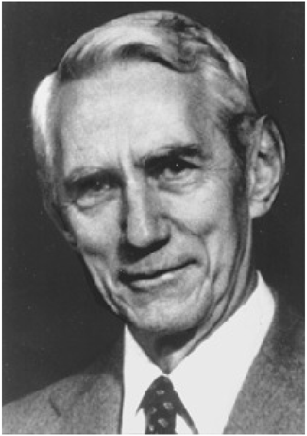
The main intuition behind information theory

- Information = reduction of uncertainty
- To measure information in a quantitative way, we need to be able to quantify uncertainty
- Uncertainty is dependent on the context and on the observer
- NB: the notion of uncertainty is different from the notion of randomness, but they are not unrelated

The original motivation behind information theory

- Claude Shannon (and others), in the first half of the last century, was struggling to formalize the problem of transmitting information over "noisy" channels.
- To understand the problem, he developed "A mathematical theory of information" to define well-posed problems in this context. Later, this theory appeared to be complete and became "The Information Theory".
- The seminal work of Claude Shannon (and others) has influenced in an irrevocable way our way of thinking in the fields of communications, statistics, reasoning under uncertainty, cybernetics, physics, etc.

Claude Shannon



From Wikipedia, the free encyclopedia

Claude Elwood Shannon (April 30, 1916 - February 24, 2001), an American electronic engineer and mathematician, is known as "the father of information theory".

Shannon is famous for having founded information theory with one landmark paper published in 1948. But he is also credited with founding both digital computer and digital circuit design theory in 1937, when, as a 21-year-old master's student at MIT, he wrote a thesis demonstrating that electrical application of Boolean algebra could construct and resolve any logical, numerical relationship.

It has been claimed that this was the most important master's thesis of all time.

Claude Shannon's seminal work in IT

Het internet Afbeeldingen Maps Nieuws Discussiegroepen Boeken Gmail meer ▼

Google wetenschap | C Shannon

Scholar **Alle artikelen** Recente artikelen — Resultaten **1** - **10** van circa

Tip: alleen in het **Nederlands** zoeken. U kunt uw zoektaal bepalen in Voorkeuren voor Scholar

A mathematical theory of communication - ▶ **kfupm.edu.sa** [PDF]
CE **Shannon** - ACM SIGMOBILE Mobile Computing and …, 2001 - portal.acm.org
T HE recent development of various methods of modulation such as PCM and PPM which exchange band- width for signal-to-noise ratio has intensified the interest in a general theory of communication. A basis for such a theory is **...**
Geciteerd door 32351 - Verwante artikelen - Alle 591 versies

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.

The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have *meaning*; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is one *selected from a set* of possible messages. The system must be designed to operate for each possible selection, not just the one which will actually be chosen since this is unknown at the time of design.

If the number of messages in the set is finite then this number or any monotonic function of this number can be regarded as a measure of the information produced when one message is chosen from the set, all choices being equally likely. As was pointed out by Hartley the most natural choice is the logarithmic function. Although this definition must be generalized considerably when we consider the influence of the statistics of the message and when we have a continuous range of messages, we will in all cases use an essentially logarithmic measure.

The logarithmic measure is more convenient for various reasons:

The two (three) main questions posed by Shannon

1. How to measure (quantify) uncertainty in a logically sound way ?
2. How to exploit this measure of uncertainty, so as to quantify the best possible reachable performances in the context of information storage and transmission ?

(3.) Find out the main principles (in terms of abstract proofs of concept, working solutions) on how to build engineering systems reaching the best possible performances.

The measure of uncertainty proposed by Claude Shannon

- Entropy of a random variable.
- (Context: a random experiment seen from a particular observer)
- (Maths: exploit probability calculus)
- NB: other contributers: Kolmogorov, Chaitin, Picard , Von Neuman, etc.

The measure of uncertainty proposed by Claude Shannon

- Suppose you are only interested in the outcome of a certain binary variable: e.g. will you succeed or not in this course ?
- You are not sure about the outcome because of lack of information and because of some sources of (unavoidable) randomness....
- But, let us consider an extreme case: suppose that you (we) know in advance that everyone passes (pleasant situation for most people): it means no uncertainty.
- Let us consider another (hypothetical) extreme case (less pleasant, for most people in the room): assume that we know that everyone will fail: no uncertainty, either.
- For C. Shannon, and the theory that we are trying to learn in this course, these two situations are perfectly identical, because he (his theory) doesn't take care of "pleasure".

- let $\mathcal{X}$ be a random variable with two values $\{Fail, Pass\}$, with respective probabilities P_F, P_P.
- If $P_F = 0$ or $P_P = 0$ (NB: $P_F = 1 - P_P$), we have an uncertainty of 0.
- Logically, the maximum of the uncertainty is when $P_F = P_P = 0.5$. This will be considered as the unit of uncertainty measurements : 1 Shannon.
- Now, consider a second, independent, question of the same type: consider a student in this room and ask yourself whether he likes to eat fish or not.
- For either question, maximum uncertainty is 1 Shannon; for simultaneously considering both questions, we have in the most uncertain case, $P_{F,f} = 0.25; P_{P,nf} = 0.25 \dots$.

- Shannon proposed that the uncertainties of such two unrelated questions should be combined by summing them together.
- Thus, we have for the simultaneous case of two unrelated questions: Uncertainty = 2 Shannon.
- Shannon also asked that the uncertainty should change in a continuous way with the probabilities (and should depend only on these probabilities, because he didn't want to bother about pleasure...)
- Consequently: we have (and this is a theorem): $H(\mathcal{X}) = -\sum_{X_i \in \mathcal{X}} P(X_i) \log_2 P(X_i)$

This is the main formula of this course !

Information = the variation of the measure of uncertainty proposed by Claude Shannon

- Now consider that you are posing the same two questions, but that you are in a different context, where you already know which students in the class like to eat fish.
- What is your uncertainty now ?
- 1 Shannon! (only the uncertainty about fail or pass).
- Therefore, Shannon proposed that information quantity is measured as the change in uncertainty between two contexts: here the change of context is about 'whether you know about fish eating preferences'.

We will make this a bit more precise, but for the time being let us say that we have a measure of information which in the current case is of 1 Shannon (because, even if we know about fish preference, but we still have no idea about pass or fail).

- Question one: when we talk about students, in context 2, we don't need to say whether they like to eat fish: data compression (remove redundancies)
- Question two: when we transmit information over a noisy channel, we are limited in the number of information bits that we can transmit because the channel 'lies' in a random and non observable fashion (more subtle notion to be understood), but we can cope with that by repeating transmissions.
- When reasoning about uncertain outcomes (e.g. medical diagnostics, forensicks, economy, complex systems), information theory provides a principled way to adress a very broad set of questions.

- Our best effort to help you in learning about information theory.
- Provide you with material, during courses and over the web page.
- Be available for any kind of technical question related to this course:
  - Louis Wehenkel (L.Wehenkel@uliege.be) for questions related to the theory
  - Antonio Sutera (A.Sutera@uliege.be) for the practical work

- Use your time to learn as much as possible in this course!
- Participate in the lectures and do not hesitate to interrupt me and ask questions, if something is not clear or not well motivated.
- More general feedback about the course contents and the way we are trying to convey the knowledge to you are also very warmly solicited.

From your side: all comments are welcome.

From our side:

- Practical homeworks: in addition to classical exercise sessions we will ask you to do personal projects, with a short report to write and send back (more details will be provided soon).
- In June, you will have to pass a written exam. The written exam covers the material covered in the course and in the practicals via multiple choice questions and some open questions.

  Course grading:

  - Practical projects: $\sim 30\%$
  - Written exam: $\sim 70\%$

Project organization

Two projects mainly covering the following topics:

- information measures,
- source coding,
- data compression,
- channel coding.

For each project, you have to submit a written (short) report on the Submission Platform (`submit.montefiore.ulg.ac.be`).
You can already register to the course!