uc3m | Universidad Carlos III de Madrid

OpenCourseWare

Theory of Information and Communication

Sara Martínez Cardama

Lisandra Otero Borges

Information Theory: Origins and Development

Introduction

The theory of information emerged as a necessity to understand the informational and communicative phenomenon and how to transmit it more effectively with the new technological developments that marked the beginning of the past century. Several milestones marked the development of this understanding. Capurro (2014) points out that the modern concept of information emerged after World War II with the establishment of Information Science. This discipline arose from Library Science and is rooted in connections with Computer Science. The automation of bibliographic information management processes led to the development of Documentation Science and, in the Anglo-Saxon context, to the field known as Library and Information Science (LIS). Information is established as a key social value in the evolution of civilization, as we saw when studying the concept of the Information Society.

Information Theory, and more specifically, the work of Claude Shannon (1948), marks the beginning of modern Information Science. However, its emergence must be understood in a context where several factors and prior developments converged:

- Technological and armament development during World War II. Previously, technological innovations such as the telegraph and the telephone had already demonstrated the need to improve the effectiveness and efficiency of communications. Indeed, the arms race accelerated these processes for military purposes.
- Development of mathematics and logic. In 1842 and 1843, Ada Lovelace and Charles Babbage conceived the first idea of software and hardware by designing a calculator that used punch cards to perform analytical calculations. Although this project was never realized, it set an important precedent. However, this theory explicitly draws from the advances of Andrei A. Markov (1910) in the field of probability and Ralph V.L. Hartley (1927). The latter made a unique contribution to understanding the concept of information. He introduced what is known as Hartley's law, which states that the amount of information is proportional to the logarithm of the number of possible symbols.

As we can see, the information theory developed by Claude Shannon is built upon previous mathematical conceptual foundations. Within the framework of Computation Theory, Alan Turing's influence is particularly noteworthy as it intersects with Shannon's research in many aspects. Alan Mathison Turing (1912-1954) and Claude Shannon (1916-2001) developed their careers in sometimes convergent areas, such as the work conducted from Boolean formal logic in complementary lines on computability, in response to the Entscheidungsproblem (decision problem) posed by Hilbert in 1928, which questioned whether mathematics was computable, that is, if it was possible to design a mechanical procedure that, starting from a proposition, could determine in a finite number of steps whether that proposition was true or false (Rotger, 2013).

Turing replaced the formal language inspired by Gödel's arithmetic with the "computing machine" concept, developing the so-called "Universal Turing Machine." Almost simultaneously, in 1937, Shannon demonstrated in his doctoral thesis, developed at the Massachusetts Institute of Technology (MIT), that Boolean algebra statements could be expressed through the combination of relays and electrical circuits. This contribution was of great importance for the digital circuit design theory (Rotger, 2013).

During World War II, both Turing and Shannon focused their attention on Cryptography. Turing played a crucial role in accelerating the end of the conflict by deciphering messages from the Enigma machine, a Nazi encryption system. Meanwhile, Shannon laid the foundations of Information Theory and modern cryptology in his works Mathematical *Theory of Communication* (1948) and *Communication Theory of Secrecy Systems* (1949).

Both lines of research have produced significant contributions and practical developments, especially Turing's in Computation and Shannon's in communication and information transmission.

2. Information Theory: Development

Shannon's theory primarily addresses the technical challenges of transmitting messages and ensuring their quality upon reception rather than focusing on the messages' content. It is concerned with the mechanisms and efficiency of signal transmission.

Shannon's theory explains the process of communication by focusing on how messages, which are converted into signals, are transmitted through a channel. This channel, a key player in the process, can introduce noise or interference, posing significant challenges to the transmission. These encoded signals travel through this medium, and at the destination, a receiver interprets these electrical impulses, converting them back into the original message to ensure it reaches its intended target accurately. The theory's emphasis is on the efficiency and reliability of this transmission process rather than the meaning of the message itself.

This idea is shown in the following image:

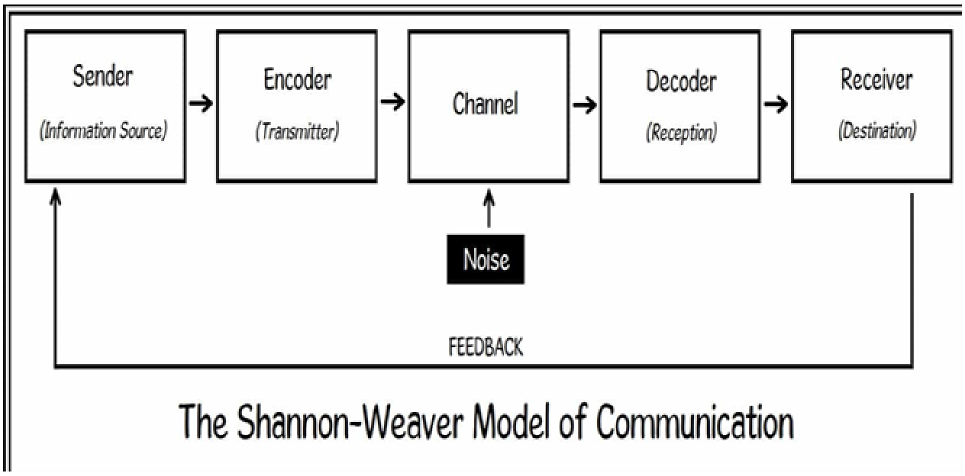

Image source: https://helpfulprofessor.com/shannon-weaver-model

Shannon mathematically modeled the communication process by applying the concept of probability to the transmission of information. In his theory, it did not matter whether the transmitted signal was a text, sound, or any other type of data; each signal was understood in probability. This approach allowed him to quantify the efficiency and reliability of communication, focusing on the likelihood of accurately transmitting messages through a channel, regardless of their specific content.Principio del formulario

In Shannon's information theory, the concept of entropy is borrowed from fields like thermodynamics, mechanics, and statistics, where it is generally understood as a measure of "disorder." In the context of information theory, entropy precisely measures the uncertainty or unpredictability of an information source. A higher entropy indicates more uncertainty or a more significant amount of possible information, while a lower entropy suggests more predictability.

Shannon also introduced the concept of redundancy, which represents the amount of predictable or repeated information within a message. Redundancy helps reduce errors during transmission by allowing the receiver to reconstruct the original message more accurately, even when noise or interference is present in the communication channel.

In Shannon's information theory, **entropy represents a message's** uncertainty or unpredictability, while redundancy is the portion of the message that is predictable or repetitive.

Let ‘s see this with two examples of different messages:

Concept	Message 1: UUUU	Message 2: ofbjfembsuy
Entropy	Low entropy: The message is highly predictable with no variation. The receiver can easily guess the next letter.	High entropy: The message has a random sequence, making it unpredictable. The receiver cannot easily guess the next letter
Redundancy	High redundancy: The repetition of "U" allows for easy error correction if some parts are lost or distorted.	Low redundancy: The lack of pattern or repetition makes correcting or guessing the information difficult if some parts are lost.

This table shows how **entropy** reflects a message's unpredictability, while **redundancy** represents its predictability and ability to withstand errors during transmission.

Here is another explanation of entropy and redundancy using the example of a die and a coin:

1. **Coin Flip:**
   - When flipping a coin, there are two possible outcomes: "Heads" (H) or "Tails" (T).
   - Each outcome has a probability of 1/2 (50%).
   - Since both outcomes are equally likely, the entropy is relatively low but not zero. The uncertainty of predicting whether it will be "Heads" or "Tails" is moderate because there are only two possible outcomes.
2. **Rolling a Die:**
   - When rolling a fair six-sided die, there are six possible outcomes: 1, 2, 3, 4, 5, and 6.
   - Each outcome has a probability of 1/6 (approximately 16.67%).
   - Because there are more outcomes, the entropy is higher. The uncertainty of predicting which number will appear is greater than the uncertainty of a coin flip.

Entropy measures a system's uncertainty or unpredictability; a die with more outcomes has higher entropy than a coin. Redundancy, which helps in error correction, is lower in the case of a die because there are more potential outcomes to consider if the message is distorted or incomplete.

In conclusion, He found a formula that allowed him to analyze the information in a message in bits. The more random a string of symbols is, the less predictable and less redundant it is, so it tends to contain a more significant amount of information per symbol.

Remember, Shannon's model was linear and closely linked to technology, particularly the telegraph. For Shannon, information was considered a statistical measure independent of its content, focusing purely on the transmission and reception of signals. Weaver extended this theory to encompass other forms of human communication, suggesting that the principles could apply beyond technological contexts.

In contrast, critics argued that Shannon's model overlooked semantic issues, such as meaning and context, which are essential in understanding human communication.

Shannon's theory remains highly relevant today, with significant applications across various fields. Telecommunications underpins the design and optimization of communication systems, including mobile phones and internet protocols. The theory's concepts of entropy are crucial for data compression techniques, such as ZIP files and JPEG images, which reduce file sizes while maintaining information quality. Additionally, Shannon's work forms the foundation for modern cryptography, network security, and advancements in machine learning and artificial intelligence.

In recent decades, integrating semiotic theories with information theory has led to a more nuanced understanding of communication. This includes examining how cultural, social, and contextual factors influence the interpretation of information, thus bridging the gap between the technical and semantic aspects of communication.

Umberto Eco (1972) significantly advanced this integration by exploring how semiotic theory applies to various forms of communication. Eco's work emphasized that meaning is not just encoded and decoded but is also influenced by cultural and contextual factors. His contributions highlighted the importance of interpreting signs within their broader semiotic contexts, bridging the gap between Shannon's technical model and human communication's richness with different interpretative dimensions.