

The bit juggler



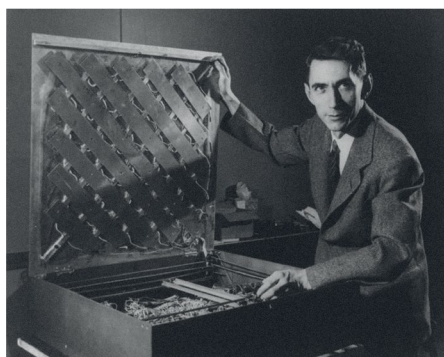
75 years ago, Claude E. Shannon's 'A mathematical theory of communication' appeared in Bell Labs's research journal, marking the birth of the discipline of information theory. This month, we celebrate the influence of Shannon's work, as well as his eccentric attitude.

It's 1985, Brighton, United Kingdom. At the evening banquet of the International Symposium on Information Theory, the meeting's chairman persuades a shy, tall and white-haired man to jump onstage and address a packed and cheerful audience. Unable to make himself heard over the jolly applause, the man shouts "This is ridiculous!", proceeds to extract three balls from his pockets, and starts juggling. That man was Claude E. Shannon, the father of information theory, and he had not attended a conference on the subject in decades.

Information is arguably the defining concept of our age. The functionality of modern society heavily relies on large-scale digital communications. An email, a text, a streamed movie or song, are not merely containers of words, images or sounds, but also carry a certain amount of a more general object – information. Although the idea of quantifying the content of any message in terms of fundamental units of information, or bits, may sound obvious today, it required a huge conceptual leap to be established.

Shannon lived in a world already equipped with intercontinental communications networks. Technologies like the telegraph, telephone, radio and television were concrete realities at various stages of development. However, disentangling the concept of information from the specific meaning of a transmitted message and the technological intricacies of communication devices took decades of refinement from a generation of scientists and engineers. Shannon's landmark paper (C. E. Shannon, *Bell System Tech. J.* **27**, 379–423; 1948) represented an overarching synthesis of this collective work.

Drawing on the work of earlier pioneers, Shannon constructed a minimal model of communication. It consisted of a transmitter encoding information into a signal, a source of



random noise deteriorating the signal across the communication channel, and a receiver. The elegance of the model lies in its universality – one can always break down a communication process into these simple conceptual blocks: source, transmitter, channel, receiver.

The properties of his model depended crucially on Shannon's insights on the probabilistic nature of information. The generation of a specific message is, at its core, a selection between a given set of symbols – letters forming specific words, or a sequence of dots and dashes in a telegraphic message. But not all choices are equally likely. In probabilistic terms, not all coins are fairly weighted. In fact, most symbols do not behave like fair coins at all, as to give a message any meaning one has to resort to some form of limitation. For instance, grammar is an ensemble of constraints that skews the distribution of possible letter choices when composing text.

From this probabilistic perspective, Shannon redefined information as the reduction of uncertainty about something. When we measure the length of an object, we are effectively decreasing the uncertainty about that object. The more a message diminishes uncertainty about its subject, the more information it carries. At the same time, the informational value of a message depends on how surprising it is. If we already know the answer to a question, or if it's highly predictable, getting that answer does not provide much new information.

A well-known feature of Shannon's pioneering paper is the definition of the fundamental unit of information (or uncertainty) – the bit – as a portmanteau of 'binary digit'. Shannon relied on the bit to model the uncertainty of an information source through the concept

of information entropy, which, in the simplest case, corresponds to the number of bits required to encode a message. A central result of his work was a formula for computing the maximum rate of reliably transmitted information across a noisy channel, which established a fundamental limit in communication theory.

Another key contribution was recognizing that information could be compressed efficiently by taking advantage of a message's redundancy. As symbols are not fair coins, and can even carry no information if perfectly predictable, some can be omitted. Some letters, for instance, can easily be removed from English words without compromising the intelligibility of the sentence.

But redundancy can also be a valuable resource. The transmission of bits through a channel can be prone to errors that appear independently and with a given probability. Noise may flip a bit in a two-bit string, resulting in a loss of the intended message's meaning. Yet, for longer and more redundant words, things are easier. It doesn't take much to recognize a mistake if we have additional bits of information.

These ideas, which appeared in Shannon's paper as rigorous mathematical proofs, laid the foundations of a wide range of research branches. Data compression techniques, error correction codes, cryptography, and even artificial intelligence, are just a few areas that rely on his fundamental insights to this day. And despite that, Shannon did not receive the accolades and praise that we normally attribute to the greatest thinkers of our time. He was undoubtedly an academic celebrity, but shied away from the spotlight.

He ceased publishing papers on information theory by the end of the 1950s, and refrained from attending conferences on the topic for decades until his startling appearance at the International Symposium on Information Theory in 1985. But it was not a matter of lack of interest. He was just too interested in other things, like robotics, artificial intelligence and juggling machines. Shannon's legacy goes beyond the enormous impact he had in shaping the digital world we live in. It is a true testament to intellectual freedom, and to doing science for the sheer pleasure of it.

Published online: 13 July 2023