Book Review

The Information: A History, A Theory, A Flood—James Gleick (New York: Pantheon Books, 2011, 526 pp.). *Reviewed by Sergio Verdú*

"*Sexily theoretical*" trumpets The New York Times[1] review of *The Information*—a book whose prologue regales us with some music to our ears:

> "*It may have far-reaching significance in electronics and electrical communications,*" *Bell Labs declared in a press release, and for once the reality surpassed the hype. The transistor sparked the revolution in electronics, setting the technology on its path of miniaturization and ubiquity, and soon won the Nobel Prize for its three chief inventors. For the laboratory it was the jewel in the crown. But it was only the second most significant development of that year. The transistor was only hardware. An invention even more profound and more fundamental came in a monograph spread across seventy-nine pages of The Bell System Technical Journal in July and October. No one bothered with a press release.*

Author of bestsellers such as *Chaos* and *Genius: The Life and Science of Richard Feynman*, James Gleick is, according to the book jacket, "our leading chronicler of science and modern technology." *The Information* is, by far, Gleick's most ambitious project to date, and the one where his gift for the craft of writing shines most vividly.

The reader is in for a sprawling historical ride with many heroes and a towering super-hero. If the bit is the main plot, we are treated to fascinating excursions far afield and recurring sub-plots, which stray, wisely, from a chronological account. This renders chapter boundaries fairly arbitrary, but makes *The Information* a very lively read, powered by Gleick's wide-eyed enthusiasm for the alluring notion of information and its many facets.

Thirty-five hundred years ago, the first alphabets containing about two dozen speech-coding letters were introduced by the Semitic tribes of the Eastern Mediterranean. Literature, law, and religion, up to then residing in volatile mortal memories, emerged as the killer applications of the new digital recording medium. Put a checkmark on communication across time; but how about communication across space beyond the fundamental limits imposed by the speed of bipeds or quadrupeds? Sub-Saharan Africans developed a curious drum language that conveys the tones of the vowels and then appends a sequence of sounds tailored to enable the recreation of the information lost in the missing consonants. Think of a systematic joint source/channel encoder for a deterministic channel. A few months after those immortal seventy-nine pages appeared in BSTJ, an English missionary published *The Talking Drums of Africa*, which is the subject of the first chapter.

The engineers enter the story in the chapter entitled *A Nervous System for the Earth*. A France-wide-web was developed during the time of Fourier and Laplace based on a mechanical semaphore whose inventor, Claude Chappe, called *le télégraphe*. Soon, hill-top telegraph towers were sprouting in places like Cairo, Calcutta, and San Francisco. But they suffered from two nasty problems: error accumulation and weather outages. Capitalizing on the recent invention of the relay by Princeton professor Joseph Henry, Cooke and

The reviewer is with the School of Engineering and Applied Science, Princeton University, Princeton, NJ 08544 USA (e-mail: verdu@princeton.edu).

[1]Janet Maslin, "Drumbeat to E-Mail: The Medium and the Message," *The New York Times*, March 6, 2011.

Wheatstone in England and Morse and Vail in New Jersey developed the first commercially successful electrical telegraphs in the 1830s. Speed (letters per minute) was the name of the game and Morse and Vail's motivation for their invention of a non-prefix variable-length source code. Earlier, a variable-length source code that did not take into account letter frequencies had been invented by none other than Gauss for his (and Weber's) short-distance relayless telegraph.

One of the heroes of *The Information* is the English mathematician Charles Babbage to whom Gleick devotes a whole chapter chronicling his valiant, ultimately unsuccessful, efforts to design a computing machine in the 19th century. By then, the logarithm, introduced two and a half centuries earlier by Napier, was already well-established as the indispensable tool for convenient multiplication. Manually computed logarithm tables were rife with errors, and Babbage set out to design an automated computing machine. He even foresaw programmable machines not driven by manual cranks but by steam. Babbage's young collaborator and programmer was not eligible for university admission on account of her gender. She was Lord Byron's daughter, Ada Lovelace, who had the brilliant insight that the automated computer should go beyond churning out numbers: It should be able to perform any sequence of logical operations. Before her untimely death, Ada came up with the software, but the machine tolerances of the day failed to deliver a working version of Babbage's hardware, funding dried out, and that was the end of that.

Although not in the way Babbage had envisioned, steam and logarithms would soon meet. Thermodynamics, "the theoretical study of the steam engine," is the subject of *Entropy and its Demons*. It is one of the least successful chapters in the book, perhaps not surprisingly in light of the pronouncement by the famous physicist Arnold Sommerfeld:

> *Thermodynamics is a funny subject. The first time you go through it, you don't understand it at all. The second time you go through it, you think you understand it, except for one or two small points. The third time you go through it, you know you don't understand it, but by that time you are so used to it, it doesn't bother you anymore.*

Entropy, Gleick writes, not quite accurately, "turned out to be a quantity as measurable as temperature, volume, or pressure." Sticking to the standard physics narrative, Gleick perpetuates the myth that Shannon's crowning achievement was the rediscovery of a well-known quantity:

> *Shannon reinvented the mathematics of entropy... To the physicist, entropy is a measure of uncertainty about the state of a physical system: one state among all the possible states it can be in. These microstates may not be equally likely, so the physicist writes*

$$S = - \sum p_i \log p_i. \tag{1}$$

> *To the information theorist, entropy is a measure of uncertainty about a message: one message among all the possible messages that a communications source can produce. The possible messages may not be equally likely, so Shannon wrote*

$$H = - \sum p_i \log p_i. \tag{2}$$

Actually, both Botzmann and Gibbs were in the business of writing integrals, not sums. However, in 1877 Boltzmann did come very close to writing (1) when, to add intuition, he imagined molecules that could take only a finite number of positions/velocities. The advent of quantum mechanics lent a firmer justification for such an approximation and, by

1948, Shannon was well aware of (1) as the expression for the statistical mechanical entropy in the quantum setting.

The description of the laws of physics has become increasingly probabilistic, and consequently more dependent on the role of information, or as Wheeler put it, with a hefty dose of poetic license: *it from bit*. Conversely, the imperialistic instincts latent in the physics community coupled with Shannon's "original sin" of borrowing a quantity from statistical mechanics made it inevitable that some would decree that *information is physical*. How our community has remained aloof from those slogans and debates, and, in fact, has developed a culture of steering clear of ventures beyond our original and natural habitat, would have made for a good side story.

Thankfully, Gleick avoids promoting Norbert Wiener to the exaggerated and distorted role that he occupies in all too many information theory histories, and, in particular, the balderdash that Wiener was to analog information what Shannon was to digital information. In *Cybernetics* (published contemporaneously with *A Mathematical Theory of Communication*), Wiener credits von Neumann with suggesting that (the negative of) the differential entropy of a density function f_1 is "a reasonable measure of the amount of information associated with the curve $f_1(x)$," maximizes (minimizes) it under a variance constraint, and considers the difference between unconditional and conditional differential entropies. In his autobiography *I am a Mathematician*, Wiener went on to write "the Shannon–Wiener definition of quantity of information (for it belongs to the two of us equally)." Unfortunately, both Shannon and Wiener referred to differential entropy as "entropy" so it is not surprising that Gleick fails to make the crucial point that differential entropy, unlike (2), is not a measure of information nor does it have an operational meaning (i.e., it is not the answer to an optimization problem of interest to the engineer). In fact, there is no evidence that Wiener ever grasped the notion, at the heart of information theory, of operational meaning lent by a coding theorem.

Among all the heroes in *The Information*, one actually had been certified as such (as Hero of Socialist Labor): Andrei Kolmogorov, who takes center stage in *The Sense of Randomness*, the excellent brief chapter devoted to algorithmic complexity. Gleick relates some of the major nuggets that tie the great Russian academician to information theory: his enthusiastic embrace (contrasting with his cold disdain for cybernetics) of Shannon's paper, initially viewed as suspect by hardcore mathematicians on both sides of the Atlantic; his foundation of the first mathematical school of information theory; his import of entropy into ergodic theory; and his definition of ϵ-entropy. A notable omission is Kolmogorov's pioneering introduction of the notion of universal data compression.

A subject whose founding paper states in its second paragraph:

> [The] semantic aspects of communication are irrelevant to the engineering problem

was bound to meet with less than a universal welcome. With admirable scholarship and keen sense of humor, Gleick captures the dismal reception accorded to information theory by the social scientists: no semantics, meaningless beep beeps, the horror! The social science take on information/communication receives plenty of attention, with a liberal sprinkling of quotes by Marshall McLuhan, and whole chapters devoted to the data deluge and to memes. Genes are to living organisms what memes are to thoughts, and therefore, subject to the evolutionary dictum of the survival of the fittest—the *Journal of Memetics* was an early victim.

There are many other fascinating side histories to be found in *The Information*, such as those of writing, the dictionary, Wikipedia and its paper forefathers, the DNA code, cryptography, and quantum information/computing. But let us finally get on with the super-hero of the story. *The Information* makes a courageous effort to elevate Claude Shannon from his underrated status in the mainstream popular perception of the history of science. Superbly researched, Gleick chronicles Shannon's early years in Michigan; his graduate studies at MIT including the most important Master's thesis of the century, his work on the differential analyzer, and his unpublished doctoral thesis on an algebra for genetics; his sojourn at the Institute for Advanced Study at Princeton; his work on cryptography, and the gestation of *A Mathematical Theory of Communication* at Bell Labs.[2] Claude Shannon's marriages to Norma Levor and Betty Moore, as well as his passion for gadgets, unicycling, and juggling, get their dues. Also mentioned are his contributions to the design of computers with unreliable components, to the max-flow min-cut theorem, and to programming a machine to play chess. But Shannon disappears from the story abruptly. The reader hopes, in vain, to find out what ever happened to the genius who, in his early thirties, revolutionized the world.

If anything, Gleick reinforces the lay misconception that Shannon's salient contribution is his introduction, or rediscovery, of entropy. Nothing is said on the key role of measuring with bits not just likelihood but dependence. Moreover, Gleick fails to make the point that information theory is about theorems, rather than definitions; that it is about posing, and answering, questions on the best performance of algorithms for information compression and transmission and beyond. The reader is told about redundancy and the Shannon limits, but why is channel capacity indeed more relevant to Qualcomm than the speed of light is to Ferrari?

In view of his obvious gifts, I wish Gleick had attempted a lay narrative of what we have learned since 1948 about the Shannon limits themselves (or, in the language of this Transactions, Shannon theory), and of those problems which remain stubbornly elusive. Along the way, he may have uncovered the cult of the pretty formula that we, the People of the Article, profess, and how those formulas have a way to inspire engineers. Most unfortunately, particularly in view of Gleick's engaging account of the various 19th century information technologies and his repute as a leading popularizer of modern science and engineering, *The Information* makes no effort to explain how technology has evolved in the last 63 years in pursuit of the Shannon limits. No attempt is made to elucidate the algorithms that convert images and sounds into bits. Or the captivating race to convert bits into sounds at faster and faster speeds to send data through the telephone wire. Nothing on the fascinating story that starts with Hamming and culminates with sparse-graph codes. Ending with Huffman, the lossless data compression narrative misses one of the crown jewels of information theory: the magical Lempel–Ziv algorithm, implemented everywhere and conqueror of the Shannon source coding limit. Nothing on the renaissance of information theory brought about by the wireless revolution. And how about some elementary iTheory for the iPhone? The reader would have loved to find out that Shannon is the answer to the question: Why do more bars result in faster downloads?

Although we are told that

> Shannon's theory [...] led to compact discs and fax machines, computers and cyberspace, Moore's law and all the world's Silicon Alleys

the reader leaves with the impression that information theory is very much a thing of 1948. We are going to need a bigger book.

---

[2]Likely based on his interview of Betty Shannon, Gleick reveals that Claude Shannon was among the last researchers to move from the Bell Labs New York city building in the West Village to the gigantic new facility in Murray Hill, NJ. So, in fact, information theory did not come into existence in the state that gave us the lightbulb, the transistor, the Morse code, FM, and color TV, but in the city that gave us abstract expressionism and hip hop.