Shannon, Orwell, and the Thinking Machine

Amites Sarkar

July 17, 2026

In 1948, Claude Shannon published his groundbreaking paper *A Mathematical Theory of Communication*. This work appeared, in two parts, in Volume 27 of the Bell System Technical Journal, and it's hard to overstate its importance. For example, it's instinctive for most of us to measure the size of a photo in megabytes, or the speed of an internet connection in megabits per second. But to do this is to silently acknowledge one of Shannon's key insights, namely, that communication can be understood and modeled as a *digital* process.

It certainly doesn't appear that way. Communication between people, mediated by vocal cords, tiny bones in the ear, and rods and cones in the retina, appears very *analogue*. And indeed, another theory of communication before Shannon was Norbert Wiener's *Cybernetics*, which was an analogue theory of feedback and control. Nonetheless, at the start of his paper, Shannon considers the *discrete channel*, by which he meant "a system whereby a sequence of choices from a finite set of elementary symbols $S_1 \cdots S_n$ can be transmitted from one point to another." For Shannon, the symbols were a dot, a dash, a letter space and a word space, and the system was telegraphy. For us, the symbols are 0 and 1, and discrete channels are everywhere.

All this is extremely well-known. Less well-known is Section 3 of Shannon's paper, entitled "The Series of Approximations To English." In this short section, which appears on pages 10 and 11 of his paper, Shannon includes the following text, as the result of a certain linguistic experiment he had conducted:

> THE HEAD AND IN FRONTAL ATTACK ON AN ENGLISH WRITER THAT THE CHARACTER OF THIS POINT IS THEREFORE ANOTHER METHOD FOR THE LETTERS THAT THE TIME OF WHO EVER TOLD THE PROBLEM FOR AN UNEXPECTED

What was this crazy experiment? Shannon is a little vague, but it seems that his method was more or less as follows. He chose a book (the title of which is lost to posterity), opened it to a random page, and selected a word at random from that page. Then he recorded that word, in capitals: THE. Next, he opened the book to another random page, and read, from the top, until he encountered the word "the". He then recorded the *next* word after "the", which must have been "head". After that, he repeated the process, opening the book to a third random page, finding the first occurrence of "head" on that page, and recording the next word, which

was apparently "and". (Shannon doesn't say what he did if the target word did not appear on the page at all, and the description I've given is modified from a procedure he does describe, but which involves letters instead of words.)

Section 3 is something of a digression in Shannon's paper, and it is not at first clear why he included it. He did not need it to motivate the key concept of entropy, which first appears 5 pages later, in Section 6. A clue is to be found still later, on page 20, when Shannon introduces the concepts of *relative entropy* and *redundnacy*. These are precise mathematical quantities, defined using equations, but, loosely speaking, the relative entropy measures the information density of a text, and redundancy is one minus the relative entropy. For instance, if a piece of text has relative entropy 30% (or 0.3), its redundancy is 70% (or 0.7). Shannon illustrates these concepts with the following engaging example:

> The redundancy of ordinary English, not considering statistical structure over greater distances than about eight letters is roughly 50%. This means that when we write English half of what we write is determined by the structure of the language and half is chosen freely. The figure 50% was found by several independent methods which all gave results in this neighborhood. One is by calculation of the entropy of the approximations to English. A second method is to delete a certain fraction of the letters from a sample of English text and then let someone attempt to restore them. If they can be restored when 50% are deleted the redundancy must be greater than 50%. A third method depends on certain known results in cryptography.

And it is this last sentence that provides the key. Shannon's work in communication was preceded and influenced by his wartime work in cryptography. This earlier work was still classified in 1948, but was published the year after, as *Communication Theory of Secrecy Systems*, again in the Bell System Techincal Journal. It turns out that entropy, the clarifying concept for discrete communication in the presence of noise, is also the fundamental factor for *discreet* communication in the presence of the enemy. Alan Turing, the distinguished British cryptographer, had visited Shannon at Bell Labs during the war, and, over lunch, they had discussed the possibility of constructing a thinking machine.

These conversations were not secret. They are recorded by Andrew Hodges in his masterful biography of Turing, *Alan Turing: The Enigma*.

> Here at least was something they could speak of freely. Once Alan said at lunch, 'Shannon wants to feed not just *data* to a Brain, but *cultural* things! He wants to play *music* to it!' And there was another occasion in the executive mess, when Alan was holding forth on the possibilities of a 'thinking machine'. His high-pitched voice already stood out above the general murmur of well-behaved junior executives grooming themselves for promotion within the Bell corporation. Then he was suddenly heard to say: 'No, I'm not interested in developing a *powerful* brain. All I'm after is just a *mediocre* brain, something like the President of the American Telephone and Telegraph Company.' The room was paralysed, while Alan nonchalantly continued to explain how he imagined feeding in facts on prices of commodities and stock, and asking the machine the question 'Do I buy or sell?' All afternoon the phone was ringing in his laboratory, with people asking who on earth it was.

So let's return to Shannon's literary experiments. In modern language, Shannon used word pair frequencies (which measure how often two words appear next to each other), from a single book, to predict the next word in a sentence. Perhaps Shannon in 1948 had access to a computer. But he didn't need one. Nor do you need a computer if you wish to repeat his experiment, and, if you do, you will discover what Shannon himself noted, namely that

> ...these samples have reasonably good structure out to about twice the range that is taken into account in their construction. [...] The particular sequence of ten words "attack on an English writer that the character of this" is not at all unreasonable.

Shannon had noticed something strange about his small language model: its performance appears much better than it actually is. By looking *just one word back*, one can generate a sequence of *ten* words which might have come from a real piece of writing. Shannon concludes Section 3 as follows:

> It would be interesting if further approximations could be constructed, but the labor involved becomes enormous at the next stage.

That was in 1948, the same year that John Bardeen, Walter Brattain and William Shockley invented the transistor.

Around the same time that Claude Shannon was writing about communication, George Orwell was writing about the same thing. To write about poverty, he crawled through a coal mine. To write about war, he fought in one, and was shot in the neck. By 1945, he had turned his attention to writing itself, of which he had, by that time, done plenty. But it was reading other writers, such as Harold Laski, that inspired him to write his famous essay "Politics and the English Language" (PEL), published in 1946.

Harold Laski must have been an interesting character. Once the most prominent intellectual on the British Left, he had, by 1944, lost much of his influence, due to his support for Soviet Russia. Still, in 1945, he was chairman of the British Labour Party, and his books were widely read. In one of these books, *Faith, Reason and Civilization* (1944), he compares the Russian Revolution to early Christianity, and writes, in conclusion:

> But no one who compares the old Russian civilisation, stagnant, pessimistic, convinced of its own futility, with the new, so energetic, so ardent, so confident in its aspirations, can easily deny that there is in the inner ethos of the Russian Revolution the clue to the secret which each race of men has ever pursued and from which it derives the dynamic of a revitalised freedom.

After the war, Clement Attlee defeated Churchill in a landslide, and Laski was quietly sidelined. Later, Orwell depicted Laski's new, energetic, ardent and confident society in his dystopian novel *Nineteen Eighty-Four*. But first, he addressed Laski's writing. The following famous passage appears towards the end of PEL:

> In our time, political speech and writing are largely the defense of the indefensible. [...] Defenseless villages are bombarded from the air, the inhabitants driven out into the countryside, the cattle machine-gunned, the huts set on fire with incendiary bullets: this is called *pacification*. Millions of peasants are robbed of their farms and sent trudging along the roads with no more than they can carry: this is called *transfer of population* or *rectification of frontiers*. [...] The great enemy of clear language is insincerity.

But PEL is not just about insincerity. Like Norbert Wiener, Orwell identifies a *feedback mechanism*:

> [The English language] becomes ugly and inaccurate because our thoughts are foolish, but the slovenliness of our language makes it easier for us to have foolish thoughts. The point is that the process is reversible. Modern English, especially written English, is full of bad habits which spread by imitation and which can be avoided if one is willing to take the necessary trouble. If one gets rid of these habits one can think more clearly, and to think clearly is a necessary first step toward political regeneration: so that the fight against bad English is not frivolous and is not the exclusive concern of professional writers.

What are these bad habits, and how do they start? Orwell has answers:

> ...modern writing at its worst does not consist in picking out words for the sake of their meaning and inventing images in order to make the meaning clearer. It consists in gumming together long strips of words which have already been set in order by someone else, and making the results presentable by sheer humbug. The attraction of this way of writing is that it is easy.

Orwell's theory of communication was decidedly *non-mathematical*, so it's interesting to compare it with Shannon's. And indeed there are connections. For both cryptography and word prediction, *redundancy*, in the form of unnecessary words and phrases, is good. But for Orwell, it was bad: no less than 4 of his 6 famous rules for writing aim to reduce the redundancy (i.e., to increase the relative entropy) of writing.

Orwell objected to the mindless recycling of stale words, phrases, analogies and images to create an artificial consensus with no factual basis. By contrast, LLMs thrive on any type of consensus: they need it to work, and, by working, they reinforce it. Indeed, any statistical biases in an LLM's training text will be reflected in its internal connections, and thus reproduced in its output. This bias is unconscious for the LLM, because it is not conscious. But it's also unconscious for the *user*, who mistakes the confident pronouncements of an LLM for the truth.

Thanks to LLMs, gumming together long strips of words has never been easier. That's why people use them. But Orwell's analysis points to a new problem: exactly how reversible is this process? We'll only know once a generation of children has grown up with LLMs. And by then, it may be too late.

Starting around 8000 BCE, in what is now Turkey, Syria, Iraq and Iran, people recorded property (cattle, barley etc) using small clay tokens: spheres, cylinders, cones, discs and tetrahedra, about 1-3 cm in size. Many new tokens were invented when the first cities appeared. Soon afterwards, around 3300 BCE, clay envelopes, in the shape of hollow clay balls, were invented to store tokens, and, after 3100 BCE, clay tablets with impressed or incised signs started to appear. People had realized that the tokens could be impressed onto the outsides of the soft clay balls before they were placed inside - for example, a cone could make a triangle, a cylinder could make a rectangle, and a disc could make a line. After that, it was no longer necessary to use the balls at all, since a simple tablet would suffice. This is one theory of the origin of writing.

The theory is due to Denise Schmandt-Besserat. She introduces her article *How Writing Came About* as follows:

> Writing makes it possible for us to store information, thereby creating a "superintelligence" — a fund of knowledge which is well beyond the ability of any single human to remember yet, at the same time, is available to all of us. Ancient civilisations considered writing a gift of the gods. The Egyptians attributed it to Thoth, the god of science and magic; the Babylonians to Nebo, son of Marduk; and the Greeks to Hermes.

For some people, AI is also a superintelligence. In his beautiful New Yorker article, *Will the Humanities Survive Artificial Intelligence?*, D. Graham Burnett describes the results of an experiment he conducted in his lecture course at Princeton:

> The assignment was simple: have a conversation with a chatbot about the history of attention, edit the text down to four pages, and turn it in.
>
> Reading the results, on my living-room couch, turned out to be the most profound experience of my teaching career. I'm not sure how to describe it. In a basic way, I felt I was watching a new kind of creature being born, and also watching a generation come face to face with that birth: an encounter with something part sibling, part rival, part careless child-god, part mechanomorphic shadow—an alien familiar.

Burnett goes on to describe the assignment of Paolo, a music-composition major.

> Paolo asked it to write a song that would make him cry.
>
> It tried. Paolo sent me a note: "The system failed the test."
>
> But I was crying, there on the couch, reading.

If my own colleagues are crying, it is likely for different reasons. "I want to fire my entire class," one of them reported, "they are all cheating with ChatGPT." "What will you actually do?" I asked. "Retire" came the answer. Another colleague reported overhearing one student tell another that they had several assignments due the next day, but it was not a problem, since

"Chat" could write them all. In my own classes, many students are turning in work that, if not AI-generated, is certainly AI-inspired.

Another troubling trend I've noticed is that students seem less inclined not just to write, but also to *read*. I teach a course on the history of mathematics, in which I'm required to grade student writing. To make this more interesting, we read some famous essays in class and try to figure out what makes them so good. Since I am not George Saunders, this is a trying exercise even at the best of times. But this is the worst of times. "What is he even talking about?" asks a student, in response to an essay that begins "Most people who bother with the matter at all would admit that the English language is in a bad way, but it is generally assumed that we cannot by conscious action do anything about it." And why should they read anything at all, since, with the rise of AI, everything sounds the same, coming at us in the confident monotone of HAL, the computer from *2001: A Space Odyssey*?

Since Open AI released ChatGPT, I've made my students complete 5 in-class writing assignments, the results of which I can then compare with their longer, typed, take-home essays. Moreover, in the syllabus, I ask them not to use AI. Last year, for the first time, a student submitted a term paper that didn't match his in-class writing at all. I asked him to come to my office hours. After a couple of last-minute cancellations, there he was. We talked about his other classes, and I brought up the issue at hand.

He explained that English wasn't his first language, and so he used a program, Quillbot, to help with grammar and spelling. Fine, I said. For your next assignment, I want you to submit not just the program's output, but also the input: what you type into it. And, for the next assignment, I received two almost identical essays, one with extremely bad grammar, but still obviously not the student's own work, and a corrected version of the same.

This time, I was more blunt. My student quickly confessed that "a friend" had helped him write even the first version. Who was this friend, I asked. No response. I then asked him about a particularly complicated paragraph, to which his response was "I don't remember writing that."

I told him: "I'm not going to judge you, since I don't know your story. But I'll tell you mine. English is my only language. But it was my mother's fourth language out of seven. She taught English as a second language to other immigrants, so they could fill out forms, so they could have rights." I don't think grammar and spelling are that important, I added.

I continued: "Quillbot is not my responsibility. Your friend is not my responsibility. You are my responsibility. Your writing is my responsibility. It is my responsibility to help you improve just one aspect of your writing, any aspect. I can't do that if I don't get to see your writing. So I will give you a passing grade. In return, you will take just half of what you've already given me, and tell me, in your own words, paragraph by paragraph, what you really think."

Recently, Eric Schmidt, the former CEO of Google, has been arguing that AI is not over-hyped, but under-hyped. To understand why, I watched a short interview he gave at last year's RAISE Summit at the Louvre in Paris. The interviwer was Henri Delahaye, and the interview can be found on YouTube under the title "Eric Schmidt: AI and the Genesis of a New Epoch."

Watching this interview, with the intention of satirizing it, I was instead charmed. Schmidt talks about mathematics *a lot*. When asked to explain why AI might not be a bubble, his thoughts turn to — mathematics. He's interested in proof assistants and scale-free systems. He makes bold analogies. A man after my own heart.

I was particularly interested in something called "The San Francisco Consensus," which emerges about 3 minutes into the video:

> There's a group of people that I work with, they are all in San Francisco, and they've all basically convinced themselves, that in the next two to four years - the average is three years - the entire world will change.

I wonder: *how* will the world change? Schmidt instead explains *why*.

> And here's roughly how they make their argument. [...] First you have language to language, LLMs, ChatGPT is the best one, they did a great job, the other folks get started, and they're catching up, in various forms, in that department. Meanwhile, you have the development of reasoning, the adoption of memory inside these systems, and so, the agentic revolution, which everybody understands here what that means, because many of you are in the middle of the agentic revolution, can be understood as: language in, memory and language out.

Now I wonder: this can be understood *by whom*? *Whose* language, *whose* memory, in and out of *what*? An example might help, and Schmidt provides one.

> So a typical example is, I'll use an example, I have a house in California, I want to build another one, I have an agent that finds the lot, I have another agent that works on what the rules are, another agent that works on designing the building, another selects the contractor, and, at least in America, you have an agent that then sues the contractor when the house doesn't work. [...] Now why would I give you such a stupid example? I just gave you a workflow example that's true of every business, every government, and every group human activity.

At this point, my thoughts turn to a specific group of humans, 2 million in number, with no home, no agency, and no voice.

In the agentic revolution, who has agency? Who gets a job, a home, a voice, an education, a choice, a life? So much has been promised, to so many, by so few. And if this is the inevitable future that only those few people actually want, then it is no longer the exclusive concern of professional writers.

References

[1] Burnett, D. G., Will the humanities survive artificial intelligence?, *The New Yorker*, April 26, 2025.

[2] Hodges, A. P., Alan Turing: The Enigma, Hutchinson, 1983.

[3] Laski, H. J., Faith, Reason and Civilisation, Victor Gollancz, 1944.

[4] Orwell, G., Politics and the English language, *Horizon*, April 1946.

[5] Orwell, G., Essays, with an introduction by John Carey, Everyman's Library, 2002.

[6] Schmandt-Besserat, D., How writing came about, *Zeitschrift für Papyrologie und Epigraphik* **47** (1982), 1–5.

[7] Shannon, C. E., A mathematical theory of communication, *Bell System Technical Journal* **27** (1948), 379–423, 623–656.