

Keyboards, Codes and the Search for Optimality

americanscientist.org (<https://www.americanscientist.org/article/keyboards-codes-and-the-search-for-optimality>) · by Robert Dorit



This Article From Issue

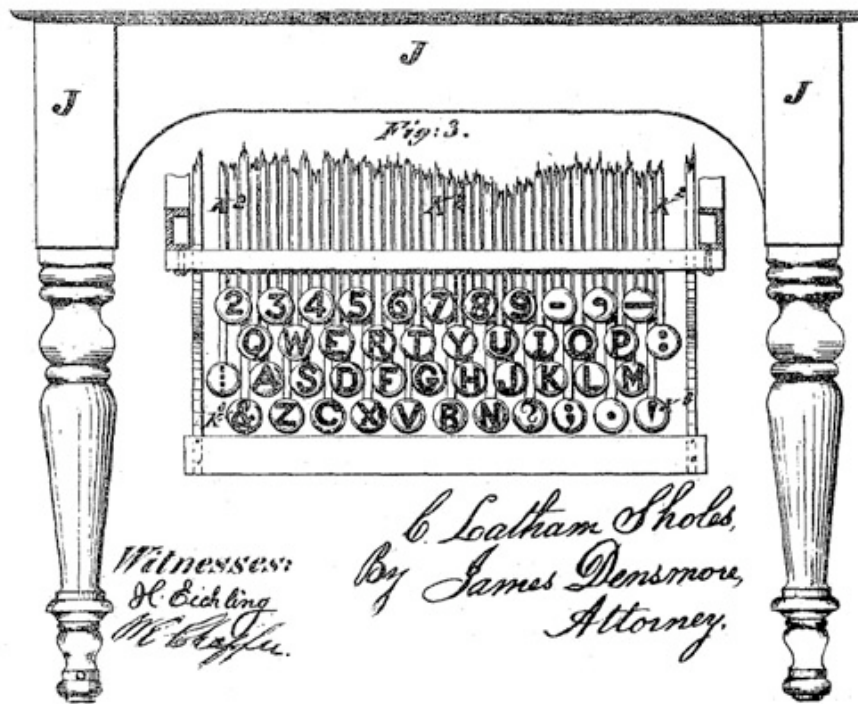
September-October 2009

Volume 97, Number 5

Page 376

- View Issue (<https://www.americanscientist.org/node/200>)

A simplified, and ultimately misleading, account of the evolutionary process argues that natural selection inexorably leads to optimal adaptation. According to this perspective, organisms face challenges presented by the environment, and ultimately, through the agency of natural selection, find the best solutions. From this point of view, the living world—from the three-dimensional structure of enzymes to the drag-minimizing shape of porpoises—could thus be described as a compendium of these supposedly ideal adaptations.



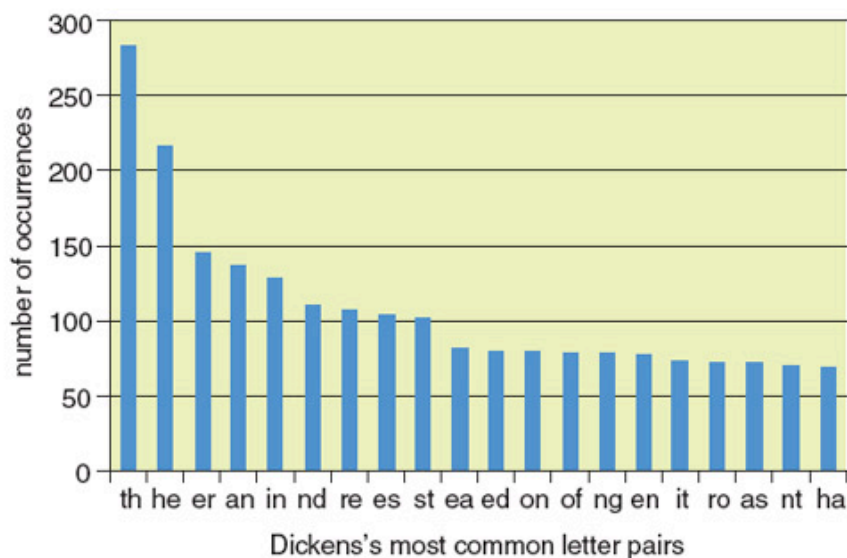
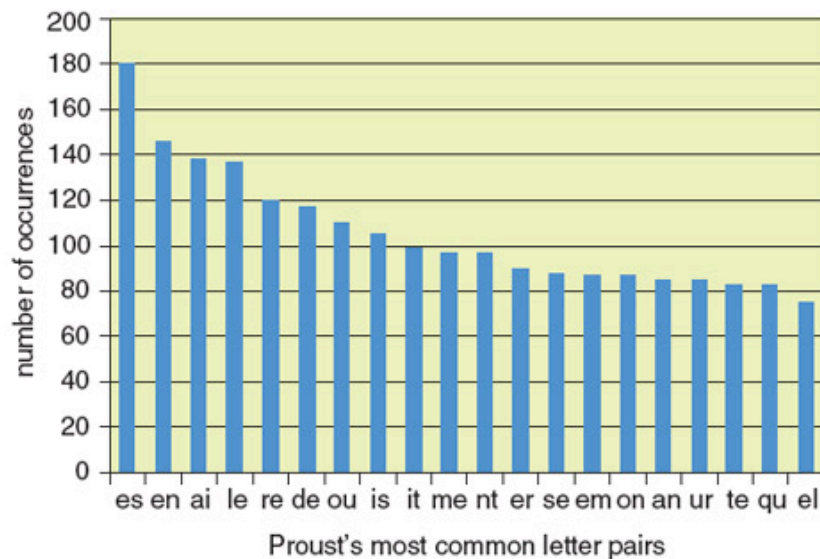
Christopher Latham Sholes of Milwaukee, Wisconsin received a patent for this keyboard layout in 1878. With few changes it became the layout that is still standard on computer keyboards today. The QWERTY keyboard, as it is called (after the first six letters in the top row of letters), may not optimal, but people have significant resistance to switching to something new.

This perspective beguiles in its simplicity, but in the end, it trivializes the complexities of the evolutionary process. Natural selection sorts among existing alternatives, but sometimes a good-enough solution may become inextricably locked in place. Evolution is not about what's best, but what works. Organisms do fit their environments exquisitely—and the task of contemporary evolutionary biology is to elucidate the interplay of history, chance and selection that shapes life on this planet. To be sure, we can ease our burden by downplaying the reach of history. We can even maintain that chance delays, but ultimately does not derail, the emergence of peak adaptation. And finally, we can dismiss what appears to be a suboptimal design by asserting that it simply reflects our lack of understanding of what is being optimized. But these are risky simplifications. In the end, life is more than a collection of adaptations, and evolution is more than the ascent to perfection. My job as an evolutionary biologist goes beyond simply imagining the plausible benefits that disembodied features might confer on individual organisms.

The 200th anniversary of Darwin's birth seems as good time as any to revisit the tension between history and optimality. I want to do so, however, by focusing on two seemingly disparate evolutionary narratives: The typewriter keyboard and the genetic code.

What's Your Type?

The typewriter keyboard is clearly the product of human ingenuity and is part of a technological revolution that continues to transform our lives. The genetic code, in contrast, resides at the very core of every living organism and has been evolving for four billion years. Yet both the keyboard and the code help illustrate the ways in which the persistence of features may not in itself be proof of their superiority. Instead, persistence may simply reflect the tenacious grip of a functional solution that emerges early in the process. In other words, a feature does not have to be flawless to endure—it just has to be good enough.



There are a number of stories that purport to give a rationale for the QWERTY layout, but perhaps the most compelling is that the arrangement separates commonly occurring letters pairs to minimize jamming in mechanical typewriters. If this were true, one would expect that keyboard layouts would vary depending on the language to be typed, and this is indeed the case, as one can see by comparing French and English. A comparison of the top 20 letter pairs in Marcel Proust's *À la recherche du temps perdu* (*top*) with those in Charles Dickens's *A Tale of Two Cities* (*bottom*) supports the argument, as it shows that only seven pairs appear with high frequency in both languages.

The typewriter is one of the great inventions of the late 19th century. The mechanizing of legible writing democratized communication and broke the stranglehold exercised by clerks on written business transactions. By allowing the influx of women into business and government offices, the invention of the typewriter also served as a bridge to the feminist revolution. At the heart of this

invention, since then also embedded in the modern computer, is a strange and idiosyncratic feature: the QWERTY keyboard. This particular arrangement of letters on a keyboard is the near-universal standard for the entire English-speaking world.

Why this arrangement, out of the 4×10^{26} possible ways in which the English alphabet might have been laid out on a typewriter? Surely the inventors of the typewriter, and the generations that have followed, had some explicit criteria that make the QWERTY layout the unrivaled keyboard design. As with many important inventions, the legends surrounding QWERTY far outnumber the facts. One legend has it that the QWERTY keyboard was designed to *reduce* typing speed (and perhaps increase accuracy) by dispersing the most common letters. Another is that this keyboard design enabled early salesmen to type certain words (including the word “typewriter”) without leaving the top row. These are plausible just-so stories, but they are hard to verify.

Perhaps the most compelling account of the origin of the QWERTY design begins with a mechanical constraint: jamming keys. Readers below the age of 35 may need to be reminded that depressing a key on a mechanical typewriter causes a metal type hammer with a letter at the end of it to rise up in an arc, strike the ink ribbon and the paper and return to its original position. The most favorable keyboard would therefore be laid out in a way that separates the most commonly occurring pairs of letters, thus minimizing the likelihood of jamming.

This explanation for the layout of the keyboard has undeniable appeal: It identifies a specific constraint, proposes a mechanism and defines optimality. If true, we would expect to see a different layout of letters in, say, a French keyboard—where common letter pairings are different than those in English—and we do. The French AZERTY keyboard positions the keys quite differently than the QWERTY.

But we cannot confuse a plausible explanation with a correct one. In order to test this scenario for the evolution of the keyboard, let’s examine the first 1,500 or so words of two well-known 19th century novels: Charles Dickens’s *A Tale of Two Cities* and Marcel Proust’s *À la recherche du temps perdu*. The charts in the second figure show digraph frequencies, reflecting how often certain letter combinations appear within words. The figure shows that the top 20 letter pairs occurring in

each of these novels (and in the two languages) are indeed different (*th/he/er* in English, as opposed to *es/en/ai* in French), although some pairs are present in both charts (such as *re*, *es* and *it*).

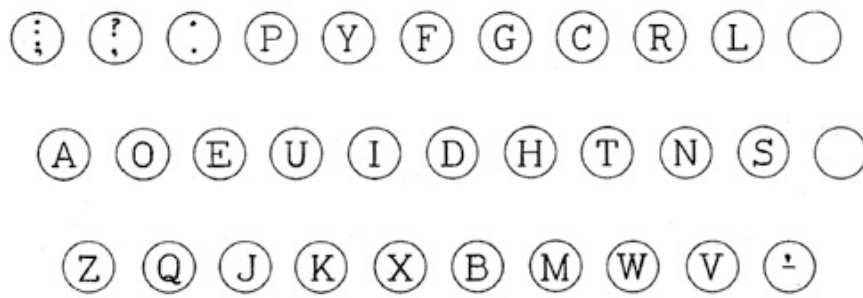
The QWERTY and AZERTY keyboards certainly appear to do a reasonable job of keeping co-occurring letters apart on the keyboard, far better than a simple alphabetical arrangement (or a random arrangement) of letters would. Even on these keyboards, however, certain very common pairs occur in close proximity (*th*, *er* and *ed*, for instance). Even assuming that minimizing jamming is the only criterion for optimality—in itself a questionable assumption—the QWERTY/AZERTY arrangements are not the best possible solutions. They are good-enough solutions.

The Early Bird

As evolutionary biologists, we know how to test optimality, even in keyboards. If an optimal keyboard embodies the ideal response to a specific set of challenges imposed both by the machine (avoid the jamming of hammers) and by the users (enhance typing speed and accuracy), we would predict that a significant change in these selective forces would surely change the resulting optimum. Two such tests of keyboard optimality are already underway.

The first test involves the advent of the computer, which has done away with the problem of jamming hammers. Freed from this constraint, the keyboard should, in principle, now be free to evolve to a new optimal design. The second, even more radical experiment is tied to the phenomenon of texting—using handheld devices to type and send written messages. Texting represents a radical shift in the interaction between humans and keyboards: Digits that were previously deemed only suited for pressing the space bar (thumbs) have become central players in communication.

Yet despite these radically changed circumstances, one look at a computer or a handheld device reminds us that the keyboard layout has not changed much since Sholes and Densmore were granted a patent for it in 1878.



August Dvorak, an educational psychologist at the University of Washington, and his brother-in-law William L. Dealey received a patent for this alternate keyboard layout in 1936. The keyboard was designed to increase typing speed while reducing finger movements. Although it came to the typing scene relatively early, it was still too late to displace the QWERTY keyboard as the industry standard, although the Dvorak keyboard continues to have some stalwart fans.

The persistence of the QWERTY keyboard, and of its related but subtly different counterparts in other languages, testifies to the power of history. The design persists not because it is the best of all possible keyboards, but because it is good enough to do the job. Over time, QWERTY has prevailed because that keyboard has become deeply embedded in the technology of manufacture, in the training of users, and in our muscular and technical consciousness. QWERTY links us, inextricably, to a vast network of mechanical, muscular and symbolic interactions that would make replacing the arrangement of the keyboard—even by a more optimized design—an enormously complex task. Even so, the experiment has been tried. In 1936 a patent was issued for the Dvorak keyboard, a keyboard design that the patent stated to be:

a scientific plan of arranging the keys which will decrease the possibility of typing errors ... [It will] increase operating speed by eliminating awkward sequences ... lessening the fatigue of the typist, because of fewer interruptions due to errors, because of better arrangement of the keys for typing the sequences most frequently used, and the rhythmical flow of typing induced thereby, and because of more evenly distributed labor for the individual fingers and the two hands.

To be sure, its inventors assumed that the factors that defined the ultimate keyboard could be identified, isolated and optimized. (Nowhere, for instance, do the inventors consider the possibility that errors or operator fatigue might emerge

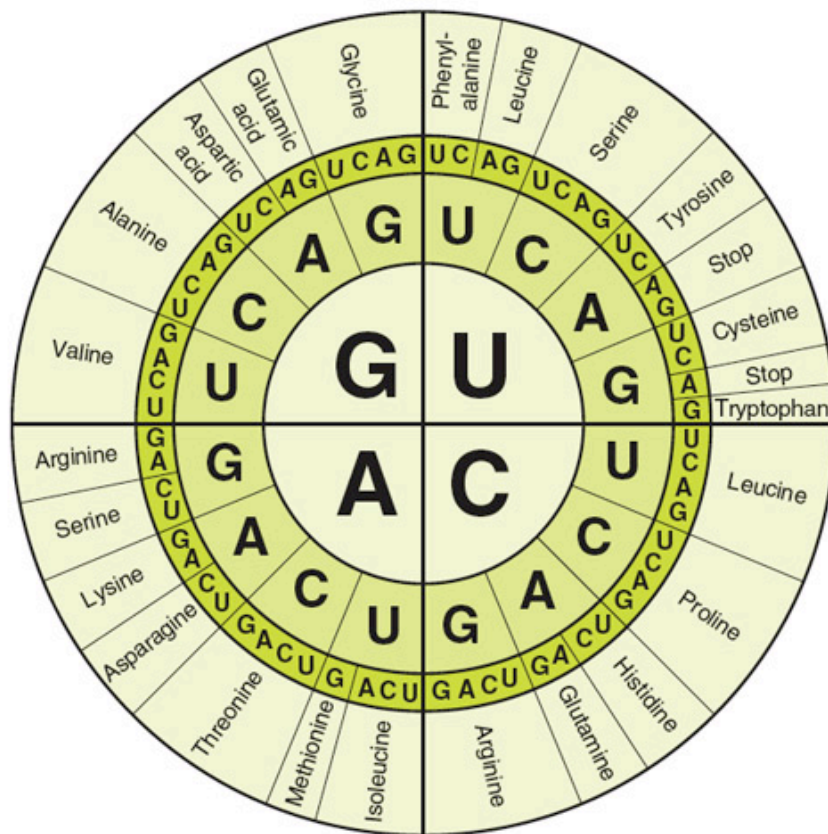
from the alienating character of the typist's job in the then-emerging machine economy).

The Dvorak keyboard sought to incorporate what was known then about the physiology of muscular action, the ergonomics of the hand, and the frequency of individual letters and of letter pairs in English. Although there is still some question about just how much of an improvement the Dvorak keyboard represented, its layout certainly allowed users to type most common letters without leaving the home (central) row and to alternate hands when typing the most common letter pairs. It was a beautiful attempt to design an optimal keyboard from first principles. But, perhaps most important, it was a layout that arose too late in the game, because 73 years later, the Dvorak keyboard remains the province of the few and the fiercely committed. In engineering, as in biology, optimality seldom trumps history.

If the tensions between the constraints of history and notions of optimality are so apparent in the evolution of a human invention, how do these tensions play out in the evolution of biological features? Technological evolution is, after all, quite different from biological evolution. There are designers and a purpose behind the arrangement of letters on the keyboard; neither drives the evolution of biological features. Nonetheless, the hand of history is clearly visible both in technological and in biological evolution.

Going with What Works

The vagaries of history inform the most fundamental act of living systems: the interpretation of genetic information. The genetic code, the Rosetta stone of all living systems, is hardly the best of all possible codes. Instead, like the QWERTY keyboard, the correspondences between the information in DNA bases and their amino-acid meanings are not optimal, but in the final analysis, are good enough.



This circular diagram of RNA bases depicts the near-universal genetic code. It is read like a tree diagram, starting at the center, giving the three bases that code for each of the resulting amino acids. Starting with U, for instance, and choosing U again for the second of the three bases, either a U or C in the third position results in the amino acid phenylalanine. This redundancy helps to temper the effects of mutations. Much like the QWERTY keyboard, this genetic code, which evolved early in the history of life (and perhaps even prebiotically) may not be optimal, but it is good enough.

The ribosome, the cellular machinery that synthesizes proteins, reads the information in RNA by interpreting the meaning of non-overlapping but adjacent triplets of bases, the alphabet of this nucleic acid (the well-known A, U, C and G). In one of the most compelling arguments for the common descent of all living organisms, the genetic code—the translation table that matches particular amino acids to particular triplets—is virtually universal. GGG codes for the amino acid glycine in you, in the bacteria in your gut, in the yeast in your bread and in every other living thing. But although the universality of the genetic code speaks to the common ancestry of all life on this planet, it is still not obvious why GGG should code for glycine.

Why does GGG, and not, for example, ACA, code for glycine? Is the universal genetic code simply the result of a random assignment of amino-acid meanings to the 64 possible triplets, or has natural selection operated on the genetic code? A quick inspection of the code answers that question: Natural selection is at work. Thus, for instance, triplets that differ in their third position (such as GGG, GGC or GGA) all code for glycine: The assignment is not random. This pattern, known as the redundancy of the genetic code, acts to temper the effects of the inexorable mutations that occur in DNA. Because of redundancy, between a quarter and a third of mutations in a gene, on average, do not affect the sequence of the resultant protein. Redundancy alone suggests that the genetic code is the result of evolution and not simply the consequence of a random assignment of amino-acid meanings to triplet coding sequences.

Dismissing the idea of random assignments, however, does not mean embracing optimality. The genetic code does indeed contain many features that dampen the impact of mutation on the sequence of proteins and hence, presumably, on the fitness of organisms. But a computer program can come up with hypothetical genetic codes that do a better job of dampening the impact of mutation: The actual genetic code is not ideal; it is good enough. Moreover, the presumption that the genetic code has been shaped solely by a single set of selective forces is just that—presumption. Likewise, when we assert that the genetic code must be optimal, but in ways we do not yet understand, we are skating past the intellectual core of evolutionary biology. A richer narrative suggests that once a functioning genetic code evolved, a complex and interlinked set of biological features, including the entire machinery for protein synthesis, evolved around it and cemented it in place. In effect, the genetic code is nature's QWERTY keyboard. But unlike the QWERTY keyboard—which could, in theory, be replaced by a better alternative—the genetic code is so embedded in the molecular machinery that evolution will not tolerate its disruption.

Neither our technology nor our biology can evade the hand of history. History underlies the configurations of the QWERTY and the AZERTY keyboards. Within the keyboards, vestiges of an even deeper history remain: The central row preserves traces of the original alphabetical arrangement that existed in the very first typesetting machines.

History, too, accounts for the universality of the genetic code. And here again, even deeper traces of prebiotic evolution can be found: Certain base triplets show particular physical attraction for the amino acids they would eventually come to encode. Everything has a past, though it may sometimes be concealed. The power of the evolutionary perspective resides in its acknowledgment of the importance of that past. Perhaps more subtly, evolutionary logic makes a profound distinction between history and destiny. We may find great comfort in the idea of inexorable progress, but the products of the evolutionary process, like the products of human ingenuity, are not about perfection.

[americanscientist.org \(https://www.americanscientist.org/article/keyboards-codes-and-the-search-for-optimality\)](https://www.americanscientist.org/article/keyboards-codes-and-the-search-for-optimality) · by Robert Dorit