This bulletin began as a blog post at blog.linedandunlined.com/post/482893698/expansion-by-alphabet, and has been considerably, umm, expanded here.

Cover image: Emmett Williams in performance with "The Alphabet Symphony." Courtesy of the Emmett Williams Estate. With thanks to Ann Noël.

ABC

In 1941, following the British Royal Navy, the U.S. adopted the Joint Army/Navy Phonetic Alphabet in an attempt to standardize communications in each branch of its armed forces. Otherwise known as a "spelling alphabet," the system replaces key letters and numbers with words that can be recognized regardless of noise, radio static, pronunciation difference, etc. Future poet Emmett Williams enrolled in the U.S. Army in 1943, and would have learned the newly-implemented system, known informally as "Able Baker," after its first two letters, *A=able* and *B=baker*. He served until the end of WWII then finished his B.A. in English at Kenyon College. Upon graduation, he returned to Europe as soon as he could, visiting Paris on his honeymoon and then living in France, Switzerland, and Germany. Eventually, Williams settled in Darmstadt as features editor for the U.S. military newspaper *Stars and Stripes*. There, the so-called "Darmstadt Circle" was already active, including such figures as poets Donald Spoerri (whose *An Anecdoted Topography of Chance* Williams would translate into English), Claus Bremer, and future Fluxus honcho George Maciunas (who was working as a civilian graphic designer at a U.S. Air Force base in nearby Wiesbaden), along with composers attracted to the summer music courses including Pierre Boulez, Morton Feldman, Luigi Nono, and Karlheinz Stockhausen.

In 1956, in an attempt to broaden the phonetic alphabet to non-English speakers, the U.S. revised the spelling alphabet to its now-current form: the NATO Phonetic Alphabet, or "Alfa Bravo." The same year, Williams began his "ultimate poem." His first effort was "what":

what

art was out bones

out violin bones
art out image
devil let bones
the devil back on image

To make "what," Williams selected by chance 26 WORDS and substituted them for the 26 LETTERS of the alphabet. Here, *a=out*, *b=the* and so on. He selected a word that was not one of the 26 in his alphabet-of-words to be the title—"what." The first line spells out the title, according to this idiosyncratic key *w=art*, *h=was*, *a=out*, *t=bones*.

art was out bones	**w, h, a, t**

The second stanza then eats the first by converting each letter to its corresponding word in the cipher, and evolving every word of the first stanza to its own line in the second:

out violin bones	**a, r, t**
art out image	**w, a, s**
devil let bones	**o, u, t**
the devil back on image	**b, o, n, e, s**

In this way, Williams's initial selection of letters mapped to words grows ever outward, as the poem re-reads the previous word outputs as new letter inputs, resulting in more words and thus more letters. Williams called the technique "expansion by alphabet."

While it makes for an interesting study, "what" doesn't come with poetry's sense of music, and it leaves most of its 26 possible words untouched. Williams tried to address this shortcoming in five linked experiments grouped under the title, "rose is a rose."

In the first, he uses the phrase "rose is a rose is a rose is a rose is a rose," with two different alphabets-of-words. These attempts have no second stanza, but the words/lines repeat according to the words in the title.

In the first poem, *a=whereat* creates a vaguely legalistic run-on clause. In the second, *a=out* gives the sense of a Pentacostal sermon or possibly a bad acid trip. The next three poems then take this rhythmic sense and cycle through the other possible words in the lexicon, expressing fully the range of the alphabet.

rose is a rose is a rose is a rose is a rose

pickle dogs come down
man come
whereat
pickle dogs come down
man come
whereat
pickle dogs come down
man come
whereat
pickle dogs come down
man come
whereat
pickle dogs come down

So this first poem's first "rose is a" string,

pickle dogs come down	r, o, s, e
man come	i, s
whereat	a

then shifts backward one letter to form the second poem in the same cycle, with $c \rightarrow b$, $b \rightarrow a$, $a \rightarrow z$, etc., which now reads,

come supreme virgin swallow	r, o, s, e
experience virgin	i, s
hisses	a

In both, the position of the repeated letter *s* ("ros̲e is̲ a") provides a sense of structure against the changing lexicon. So, *s* = *come* in the first stanza becomes *r* = *come* (i.e. $s \rightarrow r$) in the second, while *s* = *virgin* in the second stanza becomes *r* = *virgin* in the third, and the system rolls on from there.

Williams called these 1956 experiments "by-products of a generative work called 'the ultimate poem.'" In other words, "the ultimate poem" is the program itself, the alphabetic cipher and the set of rules for how it is applied—the poem "what" is simply one output. Seen another way, though, "the ultimate poem" is an expression of the childhood adage "you are what you eat," ingesting the words like a kind of body whose health is dependent on its particular diet of options along with the initial conditions of its makeup. "What," along with the other poems in Williams's "ultimate poem," suggest a new possibility for what a poem can be. It is, in Williams's hands, more like a living thing.

IBM

In 1966, Williams was invited to "do something" with a computer, and "the ultimate poem"'s source code was forked to become "IBM," combining the strategies from his earlier efforts, including the process of cyclical substitution for all possible words from "rose is a rose" and the title/first line/first stanza structure from "what." Williams described the new poem's title in a letter to his wife Ann Noël as "an understandable tribute to the muse's assistant."

Williams doesn't give further details on the offer or process of programming the computer itself, but as computers were then in very limited supply, he may have likely used the IBM 7090 at Princeton's Office of Information Technology. In 1963, Roald Buhler joined Princeton's OIT office from Educational Testing Services (ETS, also headquartered in Princeton). By 1965, Buhler was influential in brokering the partnership between IBM and the Ivy League. Up the road from Princeton in 1966, Bell Labs organized its historic "9 Evenings: Theater & Engineering," a series of commissions by contemporary artists, including Williams's friend John Cage, working with Bell's engineers.

Information technology, a domain whose study sprang from Claude Shannon's research at Bell Labs in 1948, was by 1966 looking to humanize itself. Williams writes in his autobiography *My Life in Flux—and Vice Versa* that he was "invited to work at Princeton with scientists from one of the great American corporations," possibly ETS. He continues,

They thought that a poet might be able to create programs to make home computers more interesting. Creative word games. There was lunch in the board room; the vice-president introduced me and held up several of my books. Then he shook my hand, and, with a smile from ear to ear, he said:

"I'm willing to be humanized, if you're willing to be mechanized." It was a delicious meal. But I never got mechanized.

Though the text of the "IBM" poem was finished in 1966, Williams did not complete the poem until after 1970, when he and Noël taught at CalArts, with Williams as a professor of art and Noël as director of the graphics workshop. Here in the workshop Williams found a Diatype, "that monster toy with its baffling range of type sizes. Now, every time a word is repeated, it gets larger, and creates real tension as it grows in stature."

The Diatype was a desktop phototypesetting machine from Berthold created in the 1960s. A single person could operate it by loading a sheet of film up to around A4/Letter size into the lightproof cassette. Fonts could easily be scaled between 4 and 36 points using knobs on the top of the machine. A selector at the front panel offered access to its 190 characters, while a lever on the left side triggered exposure of the cold type onto the film roll.

In a letter to Noël, Williams says that he prefers "this latest, handmade version of the poem," to the computer-generated version, and he promises "it will be the final version, too." In the "rose is a rose" experiments from 1956, the repeated word "a" served as a link, joining phrases together and as a watchword, an insistent warning. Now, as the reoccurring words grow larger, they fill with a kind of manic energy that is variously anxious, frightening, and hedonistic. The volume of these words increases with their size—black Black BLACK, white White WHITE, evil Evil EVIL, money Money MONEY, idiots Idiots IDIOTS—exposing the ecstatic yin-and-yang of this particular alphabet-of-words, words that came to Williams by chance (or as he suggests, "by design") during a year in which LSD was made illegal, the Black Panthers were founded, the Beatles began recording *Sgt. Pepper*, the Vietnam War was in full swing, and Dr. Richard Speck murdered eight student nurses in Chicago.

In the world of "IBM," the computer is the poem's conscious, unwaveringly logical mind; Williams is its unconscious, associative counterpart; and the Diatype machine is equal parts drug and microphone, pushing the whole thing into the rapid-fire output Williams published in 1973.

sex money idiots fear

fear black money sex

red going money black fear zulus money quivering naked

death fear action zulus

yes evil like money quivering naked

virgins sex action yes idiots

like evil quivering fear coming

perilous going sex going ticklish

quivering action idiots fear easy

like evil quivering fear coming

red going money black fear zulus money quivering naked

kool money yes idiots sex money ticklish white

ticklish fear hotdogs

like evil quivering fear coming

money easy money evil kool ticklish

death fear action zulus

By then he and Noël were at the Nova Scotia College of Art & Design in Halifax, but, before they left LA and while Williams was still completing "IBM" on the Diatype machine at CalArts, he asked Charles Bukowski to stop by and give a reading to his class. Recalling the encounter in *My Life in Flux*, Williams explains that Bukowski insisted on stopping at the liquor store before the reading to buy a bottle of scotch, which he swigged in front of Williams's students while Williams rushed to the finance office to cut him a check for his efforts. When he returned, Bukowski was hammered and proceeded to issue a "tirade against me and concrete poetry in general, which he regarded as so much crap, only in more concrete terms." Finally, he began to read—Williams sums up: "A great performance. A pretty good poet."

More than a decade later, in 1985, Bukowski would consider switching from a typewriter to a computer. Weighing the pros and cons, he wrote "16-bit Intel 8088 chip," a poem that shares the same tension between programmatic and vital forces as Williams's "IBM." It begins,

with an Apple Macintosh
you can't run Radio Shack programs
in its disc drive.
nor can a Commodore 64
drive read a file
you have created on an
IBM Personal Computer.

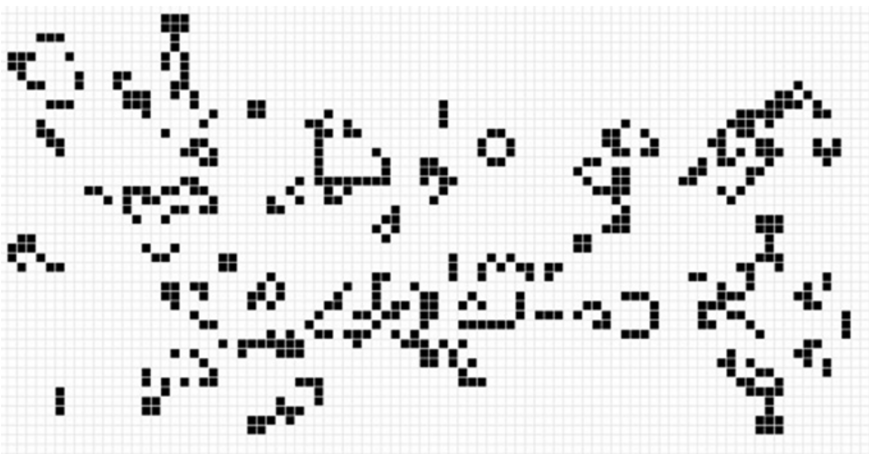
After pointing out a litany of other incompatibilities and conflicts, he concludes rather abruptly,

but the wind still blows over
Savannah
and in the Spring
the turkey buzzard struts and
flounces before his
hens.

Life, in other words, is pretty reliable software.

123

In 1970, as Williams prepared "IBM" on the Diatype machine at CalArts, math and science writer Martin Gardner was preparing a text of his own for *Scientific American*, part of his recurring "Mathematical Games" column, titled "The fantastic combinations of John Conway's new solitaire game 'life.'" Conway's Game of Life takes place on a 2D grid of square cells, either living (black) or dead (white). Each cell has eight surrounding neighbors, which influence its state over successive generations through the following rules: (1) Any live cell with fewer than two live neighbors dies, as if caused by under-population; (2) Any live cell with two or three live neighbors lives on to the next generation; (3) Any live cell with more than three live neighbors dies, as if by overcrowding; (4) Any dead cell with exactly three live neighbors becomes a live cell, as if by reproduction.



Conway's game resembles Williams's poem in two ways. First, it is the product of an initial condition acted on by a set of rules. In Williams's case, that initial condition is the alphabet-of-words used to start the poem. As Chris Funkhouser writes in his analysis of "IBM," the implications of this initial condition are significant:

You may discover immediately the benefits of using words with fewer letters. 65% of the words in Williams's original contain five letters or less and the outcome of such a restriction is clear ... [You might also want] to strategically manipulate the placement of words in the dictionary ... consciously select words which minimally repeat letters. If you want to use

all of the words in the vocabulary, you will have to incorporate every letter in the alphabet in the words you choose.

In Conway's game, Gardner observes,

You will find the population constantly undergoing unusual, sometimes beautiful and always unexpected change. In a few cases the society eventually dies out (all counters vanishing), although this may not happen until after a great many generations. Most starting patterns either reach stable figures—Conway calls them "still lifes"—that cannot change or patterns that oscillate forever. Patterns with no initial symmetry tend to become symmetrical. Once this happens the symmetry cannot be lost, although it may increase in richness ... A single organism or any pair of counters, wherever placed, will obviously vanish on the first move. A beginning pattern of three counters also dies immediately unless at least one counter has two neighbors.

Scanning the field of Life while it runs, certain patterns appear: stable shapes like boats, blocks, and beehives; oscillating shapes like blinkers, beacons, and pulsars; shapes moving ever-outward like gliders and spaceships; and shapes that, over many, many generations, stabilize or fade away completely.

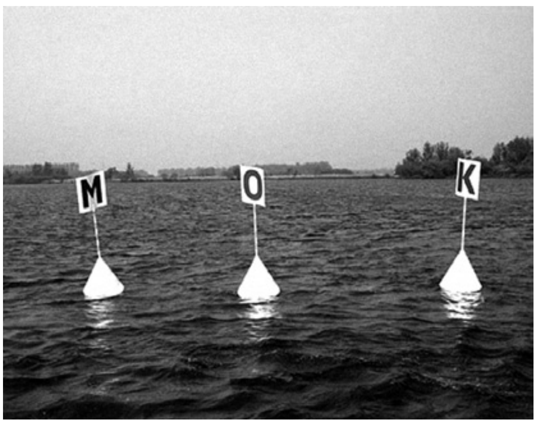
The poem and the game are also linked through their early reliance on the computer. In Funkhouser's study *Prehistoric Digital Poetry*, "IBM" appears as one of his earliest and most significant examples, one of the first dozen or so digital poetic attempts, which spans multiple categories within the genre including "procedural," or rule-based poems; "multi-media," or visual/sonic/kinetic poems; and "hypertextual," or poems that involve linked texts (like Williams's pre-loaded lexicon). Likewise, Gardner notes in his article that some of Conway's early cataloguing of forms and behaviors within his game were done using "a PDP-7 computer with a screen on which [Conway] can observe the changes ... Without its help some discoveries about the game would have been difficult to make." Soon after the game was announced, it became a popular pastime with hobbyists buying their first microcomputers, who would let it run for hours at night while their machines were unused to see if new patterns might emerge. Eventually, mathematicians proved The Game of Life is actually a computer itself—the rules allow the game to function as a Universal Turing Machine, a machine that can emulate any other machine. It is also capable of self-replication. A pattern called Gemini, which constructs a copy of itself after destroying its original, was announced in 2010.

DNA

Soon after Gemini's announcement, the poet Christian Bök made reference to it in a lecture at the University of Pennsylvania. Conway's Game of Life was of great evolutionary interest, he explained, as he showed a series of poetic texts transformed into QR codes—those gridded, black-and-white slices of binary information that evoke a momentary phase of Life. Treating these QR codes as Life's initial condition, Bök supposed, we might translate "into this regime a variety of poems by different writers, from different schools, in order to compare how they might persist under the rules of this algorithm, in order to see which one of them constitutes the fittest for survival within this universe."

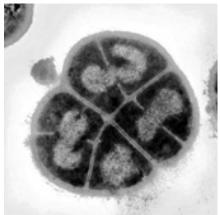
As he spoke, he ran one of his own poems through its paces, with clusters of cells vanishing, freezing, and blinking indefinitely. In doing so, Bök follows Williams's expansion of poetry's potential with one of his own, radically redefining its purpose: POETRY IS LANGUAGE THAT SURVIVES. Rhyme, meter, stanzas, all of the structural components of early poetic forms and verses, these add not solely to the pleasure of poetry but also to its dissemination, durability, and reproduction. In fact, in a directly generative way, the experiential pleasure intrinsic to the act of reading or seeing a poem is what triggers our desire to spread it. And to do that, we must have its words close at hand: their sound must be remembered, their form must reproducible, or they must be capable of being retrieved with a simple cell phone camera snap. Whatever the storage medium, though, a poem is a text that *expresses its enclosure with pleasure,* and that pleasure in turn enables ubiquity, which aids discoverability, which is then reinforced once more by pleasure, and so on.

"The discursive structures of epidemiology, as seen in such notions as viral marketing, and viral computing," Bök observed in his talk, "might apply to the transmission of ideas throughout our culture." "In fact," he continued, "We might be able to transmit messages across stellar distances, even epochal intervals, so that, unlike any other cultural artefact so far produced (except perhaps for the Pioneer or Voyager probes) such a poem … might conceivably outlast terrestrial civilization itself, persisting like a secret message in a bottle flung at random into a giant ocean. Even though poets may pay due homage of the immortality of their heritage, few of us have ever imagined creating a literary artifact capable of outliving the existence of our species. A literary artifact that could remain upon the planet until the very hour when, at last, the sun itself explodes." Bök's image of language unbound, encapsulated, floating in a vast cosmic sea—continuing to signal and even self-assemble though no one is left to utter or inscribe it—calls to mind a scene from 2007, in a northern Dutch canal, where three striped buoys floated, topped with props oscillating from the gusts offshore.



The two-sided props, installed by designer Karel Martens and writer Kees't Hart, used a total of six letters (T/M, O/A, L/K) to power a "word machine" capable of spelling 16 different Dutch words, from LOT ("fate," as in "lottery") to TAL ("number," as in "tally"). Locked in this shrunken lexicon, bobbing in the chilly current, the buoys might have been mistaken for a physical version of the blinking cells in Conway's Life endgame—a kind of ultimate poem.

In 1956, as Williams drafted the rules for his "ultimate poem," a scientist named Arthur Anderson discovered a remarkable bacterium after repeated attempts to sterilize rations of tinned meat. It was unusually durable, surviving extremes of cold, absence of water and air, and exposure to acid and gamma radiation. Almost half a century later, Pak Wong led a team of scientists at the Pacific Northwest National Laboratory to successfully encode the lyrics of the song *It's a Small World* on the DNA of Anderson's bacteria, *Deinococcus radiodurans*.



Christian Bök was intrigued when he learned about this discovery—with *Deinococcus radiodurans*, he imagined a poetic vessel truly built to last, a bottle built tough enough to send any message, and he soon set about composing his own original poem for insertion into the bacterium. By 2007, Bök was fully engrossed in his own effort to weave an alphabet of triple-based DNA codons (made of A, C, G, or T) into a coherent poem he dubbed the *Xenotext*.

A project created with artist Micah Lexier for a window display at Printed Matter in New York earlier that year produced two short texts, each an anagram of the other. "Two Equal Texts" makes explicit reference to the authors themselves, as well as to the efforts each took to present this composite work in which one text recodes the other. This referencing and recoding became an essential aspect of the *Xenotext*. It was not enough for Bök to simply input the text into the DNA of *Deinococcus radiodurans*—this had already been done by Dr. Wong and others. Any organism processes DNA codons into corresponding mRNA strings, which in turn produce a protein sequence of amino acids. For Bök's poem to truly express its medium, he reasoned, his DNA codon must not *only* render an mRNA string. Instead, the protein produced by that DNA codon input string should produce both an equally readable mRNA string output, and, in turn, be *expressed* by the organism itself.

BoTSL#4 2012 NOV 11 2:53 PM

Finding such a text, however, was no simple task. Bök has shown that mutually assigning each letter of the alphabet produces 8 million possible combinations. But *usable* combinations, where one meaningful English word can be mutually encoded back and forth to another meaningful English word, are dramatically more limited. In an interview with *The Believer*, Bök illustrates how restricting this approach can be by using the word LANGUAGE as an example:

You would think that an eight-letter word like LANGUAGE would probably have many counter-parts ... But as it turns out, there are only three: foxtrots, toxicoid, and copyboys — each of which has its own set of unique ciphers.

Just as in "IBM," as words get longer the poem increases in complexity, and implementing the rules becomes more difficult; no words in the *Xenotext* are longer than five letters. To help find these words, Bök wrote a computer program which takes a cipher plus "some heuristic methods" as an input and outputs a list of possible words from which to assemble the poem.

Finally, and perhaps most complexly, Bök's final mRNA string correlates to a protein string that emits a red glow, like the "faery" in the poem, here with the DNA text on the left and the mRNA "xenotext" on the right:

any style of life	**the faery is rosy**
is prim	**of glow**

Bök's effort with the *Xenotext* was exhaustive. However, the text itself, either in its DNA or mRNA output, is potentially eclipsed (at least at the present time, though perhaps not in Bök's very long-term, post-solar future) by the language Bök has produced to explain the effort itself. Here is a selection of claims from his initial abstract announcing the project:

The complex, but orderly, structure of every living system arises spontaneously out of underlying principles of self-organization — principles no less important than the laws of selective evolution ... I hope that my poem might urge readers to reconsider the aesthetic potential of science, causing them to recognize that, buried within the building blocks of life, there really

does exist an innate beauty, if not a hidden poetry—a literal message that we might read, if only we deign to look for it.

Far from unusual, Bök argues, the instinct to marry language and meaning, code and life, is utterly natural. Talking about Sunday newspaper puzzles and cryptograms, he wonders why they always tend to be non-sensical gibberish, "Why don't the puzzle designers create a meaningful sentence that could in turn be deciphered into yet another meaningful sentence?" And this is exactly his project with the *Xenotext*, "Two Equal Texts," and other works. To begin not with "gibberish" but with meaningful language as an initial condition, from which the poetic constraints, rules, or processes generate an equally meaningful secondary condition. For Bök, "meaning" is opposed to "gibberish," and meaningful texts, in the end, produce identifiable, often physical changes as a result of their utterance, visibility, or insertion. Asked to respond to the work of artist Eduardo Kac, who coded an excerpt from the book of Genesis into the DNA of an *E. coli* bacterium, Bök uses a colorful metaphor to host his critique:

[It] does not seem radically different from the act of inserting a copy of the bible into the saddlebag of a donkey, and then letting the donkey wander on its own through a minefield. I think that, if possible, the inserted text must change the behavior of the donkey in some profound way, perhaps converting it to Christianity, if you like.

ETA

Of course, donkeys aren't the only animals germane to genetics, gibberish, or instructions. The history of the familiar theory that an infinite army of monkeys equipped with typewriters might, if given ample time, almost certainly produce the text of Shakespeare's *Hamlet* is examined by Jorge Luis Borges in his 1939 essay "The Total Library." Borges traces the history of the idea back to Aristotle, who included components of the theory—of a contingent, equivalent set of elements and a sufficient degree of chance—in *Metaphysics* and *On generation and corruption*. Three hundred years later in *On the nature of the gods*, Cicero synthesized these ideas in order to counter Aristotle's "atomist" worldview: if a set of golden letters were tossed into the air an infinite number of times, he supposes, "fortune could not make a single verse" of Ennuis' *Annales*. Borges's "total library" is an effort to imagine such an imaginary place, a place that includes all possible texts, which is an interest he shares with predecessors Jonathan Swift, Blaise Pascal, and others. Within the library's infinitude, of course, is a specific subset of texts:

The detailed history of the future, Aeschylus' *The Egyptians*, the exact number of times that the waters of the Ganges have reflected the flight of a falcon, the secret and true nature of Rome, the encyclopedia Novalis would have constructed, my dreams and half-dreams at dawn on August 14, 1934, the proof of Pierre Fermat's theorem, the unwritten chapters of *Edwin Drood*, those same chapters translated into the language spoken by the Garamantes ...

This specific subset—*Hamlet*, or anything else—as typed by monkeys was an image first conjured a quarter century before Borges's essay by French mathematician Émile Borel in his book *Chance* (the output of a million monkeys typing ten hours a day could not equal the books of the richest libraries in the world). The image was then popularized fifteen years later by British physicist and amateur poet Arthur Eddington in his book *The Nature of the Physical World* (an "army of monkeys" typing constantly "might write all the books in the British Museum").

But as Borges and others keenly understand, the signal-to-noise ratio of this library would be very Very VERY low, approaching zero:

for every sensible line or accurate fact there would be millions of meaningless cacophonies, verbal farragoes, and babblings. Everything: but all the generations of mankind could pass before the dizzying shelves—shelves that obliterate the day and on which chaos lies—ever reward them with a tolerable page.

This "chaos" Borges describes is the entropic state of the total library: with information so dispersed, it is buried in a chasm of "meaningless cacophonies"—the same "gibberish" that Bök describes in his Sunday puzzles.

BoTSL#4 2012 NOV 11 2:53 PM

Those puzzles are worked out, in part, with statistics—the well-known phrase ETAOIN SHRDLU is the approximate order of the alphabet's 12 most common letters in the English language. Bök's own *Xenotext* cipher maps one alphabetic frequency (f) close to the other: *a* (f = 8.12 per 40,000 words) maps to *t* (f = 9.10), *b* (f = 1.49) maps to *v* (f = 1.11), *c* (f = 2.71) maps to *u* (f = 2.88), and so on. Bök describes using "heuristic methods" to arrive at his chosen cipher from his 8 million possible options; likely these letter frequencies were in his mind.

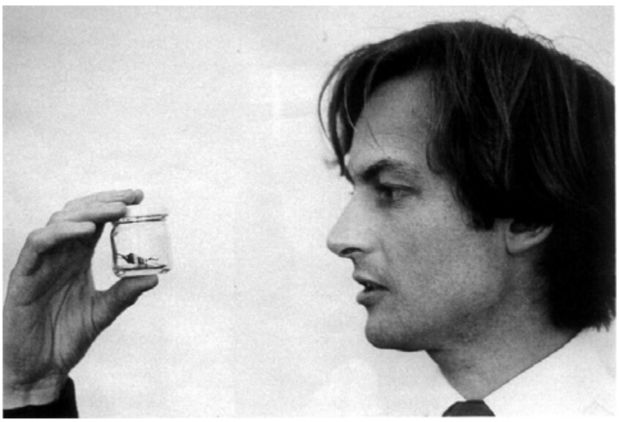
What this illustrates, then, is Bök's *prior knowledge of English* in composing his poem, and that with prior knowledge, you can do a lot. In a 1950 paper on "Prediction and Entropy of Printed English," mathematician and information theorist Claude Shannon shared the results of an experiment in which participants were asked to guess the letters of an unknown phrase. After each guess, the participant is either told their guess was correct or is told the letter that would have been correct then instructed to guess the next letter. The sample phrase had a 26-letter alphabet plus a space, with a total of 129 letters in the text: THE ROOM WAS NOT VERY LIGHT A SMALL OBLONG READING LAMP ON THE DESK SHED GLOW ON POLISHED WOOD BUT LESS ON THE SHABBY CARPET. English-speakers guessed 89 of these letters correctly, or 69%.

The paper extends Shannon's earlier essay, "A Mathematical Theory of Communication" from 1948, which details the monkeys-on-typewriters scenario in which all 27 symbols are equally probable (XFOML RXKHRJ FFJUJ ZLPWCFWKCYJ FFJEYVKCQSGHYDQPAAMKBZAACIBZ LHJQD) and includes prior knowledge of English frequencies (OCRO HLI RGWR NMIELWIS EU LL NBNESEBYA TH EEI ALHENHTTPA OOBTTVA NAH BRL), double-letter pairings (ON IE ANTSOUTINYS ARE T INCTORE ST BE S DEAMY ACHIN D ILONASIVE TUCOOWE AT TEASONARE FUSO TIZIN ANDY TOBE SEACE CTISBE), triple-letter pairings (IN NO IST LAT WHEY CRATICT FROURE BIRS GROCID PONDENOME OF DEMONSTURES OF THE REPTAGIN IS REGOACTIONA OF CRE). Since many words are three letters or longer, Shannon then jumps to words. He arrives at a phrase that's remarkably close to English in only two additional iterations.

A bit is the fundamental binary unit of information. The arrival of a bit acts on prior knowledge by reducing uncertainty by half—on/off, black/white, yes/no. Suppose you are guessing a card drawn randomly from a deck, the king of diamonds. If limited to yes-or-no questions, you might attempt to reduce your uncertainty through the following sequence: Is it black? No. Is it a heart? No. Is its value 7 or higher? Yes. Is it a face card? Yes. Is it the king of diamonds? Yes. You have used five bits to determine the answer. (Mathematically speaking, $\log_2 52 = 5.7$, and indeed the sequence may take up to 6 guesses to determine the card with total certainty.)

GOD

The card example was used by evolutionary biologist Richard Dawkins to discuss Shannon's ideas in a 1998 article "The Information Challenge."



Dawkins applies the complex uses of Shannon's theories to genetics to refute a creationist challenge that in order for evolution to function as Darwin posited, it should increase the information in the genome through mutation in each successive generation. Dawkins counters this claim through several close readings of Shannon's theories, including a careful unpacking of his ideas about the redundancy of English and its entropy. Redundancy in English is around 50%—if alternate letters are deleted in a text, 80–100% of subjects can restore the text to its original state. Entropy in English is anywhere from 0.6 to 1.5 bits per letter, which averages to about 1 bit per letter, and roughly corresponds to the 50% redundancy rate.

What this means is that each individual letter of English may or may not add to the overall information value of the message it's part of, just as each individual DNA codon may or may not add to the overall information value of the organism in which it resides. While it's true that more complex organisms generally have more information in their DNA (Dawkins uses the example of a copy-and-paste millipede versus a highly articulated lobster) it is not true in all cases (Dawkins points out that humans have less information in their genome than the crested newt). He continues with a computational metaphor:

The crested newt has a bigger "hard disc" than we have, but since the great bulk of both our hard discs is unused, we needn't feel insulted. Related species of newt have much smaller genomes. Why the Creator should have played fast and loose with the genome sizes of newts in such a capricious way is a problem that creationists might like to ponder.

Dawkins continues,

The "information challenge" turns out to be none other than our old friend: "How could something as complex as an eye evolve?" It is just dressed up in fancy mathematical language—perhaps in an attempt to bamboozle. Or perhaps those who ask it have already bamboozled themselves, and don't realise that it is the same old—and thoroughly answered—question.

Dawkins has already attempted to answer this question (about the human eye's evolution) himself with yet another animal and yet another bit of genetics, gibberish, and instruction. In *The Blind Watchmaker*, he begins by invoking the typewriting monkeys, directing them to produce not the entire text of *Hamlet* but a simple fragment. In the play, Hamlet asks Polonius if he sees a cloud shaped like a camel. Polonius agrees. Then Hamlet, feigning madness, changes his mind: "METHINKS IT IS LIKE A WEASEL." How quickly, Dawkins wonders, could the monkeys produce this phrase? He calculates 10^{40} randomly-occurring variations of this 28-letter phrase (including spaces), and even if they typed millions of combinations per second without stopping, the monkeys would still be typing until the end of the universe.

BoTSL#4 2012 NOV 11 2:53 PM

How, then, do genes become more information-rich? The same way Shannon's sentences do—through prior knowledge. Knowing its target phrase, with randomly-occurring mutations, "WDLTMNLT DTJBKWIR ZREZLMQCO P" evolves into "METHINKS IT IS LIKE A WEASEL" in around 40 generations. Dawkins wrote the computer model in BASIC; it ran while he was out to lunch. When he rewrote it in Pascal, it took 11 seconds. The Weasel Program is now written for nearly every computer programming language imaginable.

```
 0: WSC UDZHGQQFBSJCVKOBBXP AVND -- score: 3
 1: WSC UDZHGQQFBSJCVKOBAXP AVND -- score: 4
 2: WSC UDZHGQQFBSJCVKOBAXP AVED -- score: 5
 3: WST UDZHGSQFBSJCVKO AXP AVED -- score: 7
 4: MST UDZHGSQFBSHCVKO AXB AVED -- score: 8
 5: MST UDZHGSQFBSHCVKO AXBEAVED -- score: 9
 6: MST UDZHGIQFBSHCVKO AXBEAVED -- score: 10
 7: MST UDZHGIQFBSHCIKO AXBEAVED -- score: 11
 8: MDT IDSHGIQFBSHCIKO AXWEIVED -- score: 12
 9: MDT IDQH IQFBSNCIKN AXWEIVED -- score: 13
10: MDTHIDQH IQFBSNCIKN AXWEIVED -- score: 14
11: MDTHIDQH IQFBSQCIKN AXWEIVED -- score: 14
12: MDTHINQH IQFBSQCIKN AXWEIVED -- score: 15
13: MDTHINQH ITFBSQCIKN AXWEIVED -- score: 16
14: MDTHINQH ITFBSQLIKN AXWEIVED -- score: 17
15: MDTHINDH ITFBSQLIKN AXWEIVED -- score: 17
16: MDTHINDS ITFBSQLIKN AXWEIVED -- score: 18
17: MDTHINDS ITFBSQLIKN AXWEMVED -- score: 18
18: MDTHINLS ITFBSQLIKN AXWEMVEL -- score: 19
19: MYTHINLS ITFBSQLIKE AXWEMVEL -- score: 20
20: METHINLS ITFBSQLIKE AXWEMVEL -- score: 21
21: METHINLS ITFBSQLIKE AXWEACEL -- score: 22
22: METHINNS ITFBSQLIKE AXWEASEL -- score: 23
23: METHINNS ITFBSQLIKE AXWEASEL -- score: 23
24: METHINNS ITFBSQLIKE AXWEASEL -- score: 23
25: METHINNS ITFBSQLIKE AXWEASEL -- score: 23
26: METHINNS ITFBSQLIKE AXWEASEL -- score: 23
27: METHINNS ITFBSQLIKE AXWEASEL -- score: 23
28: METHINNS ITFBSQLIKE AXWEASEL -- score: 23
29: METHINNS ITFBSQLIKE AXWEASEL -- score: 23
30: METHINKS ITFBSQLIKE AXWEASEL -- score: 24
31: METHINKS ITFBSQLIKE AXWEASEL -- score: 24
32: METHINKS ITFBSQLIKE AXWEASEL -- score: 24
33: METHINKS ITFBSQLIKE ANWEASEL -- score: 24
34: METHINKS ITFBSQLIKE ANWEASEL -- score: 24
35: METHINKS ITFBSQLIKE ANWEASEL -- score: 24
36: METHINKS ITFBSQLIKE ANWEASEL -- score: 24
37: METHINKS ITFBSQLIKE ANWEASEL -- score: 24
38: METHINKS ITMBSQLIKE ANWEASEL -- score: 24
39: METHINKS ITMBSQLIKE A WEASEL -- score: 25
40: METHINKS ITMBSQLIKE A WEASEL -- score: 25
41: METHINKS ITJBSQLIKE A WEASEL -- score: 25
42: METHINKS ITJBSQLIKE A WEASEL -- score: 25
43: METHINKS ITJBS LIKE A WEASEL -- score: 26
44: METHINKS ITJIS LIKE A WEASEL -- score: 27
45: METHINKS ITJIS LIKE A WEASEL -- score: 27
46: METHINKS ITJIS LIKE A WEASEL -- score: 27
47: METHINKS IT IS LIKE A WEASEL -- score: 28
```

But even Dawkins cautions placing too much emphasis on the ultimate poetry of the Weasel Program:

Although the monkey/Shakespeare model is useful for explaining the distinction between single-step selection and cumulative selection, it is misleading in important ways. One of these is that, in each generation of selective "breeding," the mutant "progeny" phrases were judged according to the criterion of resemblance to a distant ideal target, the phrase METHINKS IT IS LIKE A WEASEL. Life isn't like that. Evolution has no long-term goal. There is no long-distance target, no final perfection to serve as a criterion for selection, although human vanity cherishes the absurd notion that our species is the final goal of evolution. In real life, the criterion for selection is always short-term, either simple survival or, more generally, reproductive success.

*

BoTSL#4 2012 NOV 11 2:53 PM