THE NEW ATLANTIS
A Journal of Technology & Society

Logic, DNA, and Poetry

Steve Talbott

In January, 1956, Herbert Simon, who would later win the Nobel prize in economics, walked into his classroom at the Carnegie Institute of Technology and announced, "Over Christmas, Allen Newell and I invented a thinking machine." His invention was the "Logic Theorist," a computer program designed to work through and prove logical theorems. Simon's casual announcement—which, had it been true, would surely have rivaled in importance the Promethean discovery of fire—galvanized researchers in the discipline that would soon become known as artificial intelligence (AI). The following year Simon spoke of the discipline's promise this way:

> It is not my aim to surprise or shock you.... But the simplest way I can summarize is to say that there are now in the world machines that think, that learn and that create. Moreover, their ability to do these things is going to increase rapidly until—in a visible future—the range of problems they can handle will be coextensive with the range to which the human mind has been applied.

There was good reason for the mention of surprise. Simon and his colleagues were, in dramatic fashion, surfing the shock waves produced by the realization that computers can be made to do much more than merely crunch numbers; they can also manipulate symbols—for example, words—according to rules of logic. The swiftness with which such programmed logical activity was equated, in the minds of researchers, to a human-like capacity for speech and thought was stunning. And, during an extended period of apparently rapid progress, their faith in this equation seemed justified. In 1965 Simon predicted that "machines will be capable, within twenty years, of doing any work that a man can do." M.I.T. computer scientist Marvin Minsky assured a *Life* magazine reporter in 1970 that "in from three to eight years we'll have a machine with the general intelligence of an average human being … a machine that will be able to read Shakespeare and grease a car."

***Steve Talbott** is a Senior Researcher at the Nature Institute and the editor of NetFuture, an online newsletter about technology and human responsibility (www.netfuture.org). He can be reached at stevet@oreilly.com.*

The story is well-told by now how the cocksure dreams of AI researchers crashed during the subsequent years—crashed above all against the solid rock of common sense. Computers could outstrip any philosopher or mathematician in marching mechanically through a programmed set of logical maneuvers, but this was only because philosophers and mathematicians—and the smallest child—were too smart for their intelligence to be invested in such maneuvers. The same goes for a dog. "It is much easier," observed AI pioneer Terry Winograd, "to write a program to carry out abstruse formal operations than to capture the common sense of a dog."

A dog knows, through its own sort of common sense, that it cannot leap over a house in order to reach its master. It presumably knows this as the directly given meaning of houses and leaps—a meaning it experiences all the way down into its muscles and bones. As for you and me, we know, perhaps without ever having thought about it, that a person cannot be in two places at once. We know (to extract a few examples from the literature of cognitive science) that there is no football stadium on the train to Seattle, that giraffes do not wear hats and underwear, and that a book can aid us in propping up a slide projector but a sirloin steak probably isn't appropriate.

We could, of course, record any of these facts in a computer. The impossibility arises when we consider how to record and make accessible the entire, unsurveyable, and ill-defined body of common sense. We know all these things, not because our "random access memory" contains separate, atomic propositions bearing witness to every commonsensical fact (their number would be infinite), and not because we have ever stopped to deduce the truth from a few more general propositions (an adequate collection of such propositions isn't possible even in principle). Our knowledge does not present itself in discrete, logically well-behaved chunks, nor is it contained within a neat deductive system.

It is no surprise, then, that the contextual coherence of things—how things hold together in fluid, immediately accessible, interpenetrating patterns of significance rather than in precisely framed logical relationships—remains to this day the defining problem for AI. It is the problem of meaning.

DNA's Ever-Receding Secrets

Artificial intelligence is not the only area of science that is susceptible to such misplaced confidence about solving the riddles of our humanity. On February 28, 1953, Francis Crick and James Watson burst into the Eagle pub in Cambridge, England, where (as Watson later recalled) Crick spilled the news that "we had found the secret of life." The secret, as the world now knows, lay in the double helical structure of DNA. Looking back on Crick and Watson's revelation fifty years later, the editors of *Time* would refer to "the Promethean power unleashed that day."

It was, however, slightly strange for Crick and Watson to announce the revelation of a secret that came in the form of a code they did not understand and a text they did not possess. Yet the double helix, by all accounts, came in just that way. This is why we have been treated, during the intervening fifty years, to the celebration of one code-breaking and text-reading victory after another, culminating most recently in the Human Genome Project. Only now, we're told, has the full text of the deciphered Book of Life been laid out before the glittering eyes of genetic engineers.

The celebration—and also the expense—of this latest victory has been unparalleled in the history of science. So, too, has the orgy of self-congratulation and utopian prediction. The completion of the genome project, many scientists declared, would quickly enable us to slay the demons of genetically linked disease, after which we would employ designer genes to create an enhanced race of super-humans. The giddiness reached its silly zenith when Nobel laureate and molecular biologist Walter Gilbert observed that you and I will pocket a disk carrying the code for our personal genomes and say, "Here is a human being; it's me!"

But hold off on the celebration. Now, it appears, there's one, small, remaining obstacle on the path to unprecedented self-knowledge. Yes, we have discovered the alphabetic text of the Book of Life, but it turns out we still can't actually read it. For this, according to the current story, we need a new project—one that will dwarf even the human genome effort. We must unravel the functioning of the body's 100,000 or more proteins—molecules so deeply implicated in every aspect of the organism (including its genetic aspects) that the attempt to understand them looks suspiciously like the entire task we began with: to understand life.

The secret of life, it appears, is wrapped within layer after layer of mystery, each one requiring its own decoding, and each one extending further through the biochemistry of the whole organism. Where, then, is the single, controlling secret? If by their own admission they still cannot read the DNA text of the Book of Life, how can scientists pronounce so confidently on the nature and absolute importance of its meaning? And if they can achieve the reading only through recourse to everything else going on in the organism—that is, if they must in effect read the whole organism—then how can they know that the entire secret resides in one, small, still mostly undeciphered portion of the overall text?

Putting Genes in Context

Clearly there is a profound faith at work here. It is, in fact, the same faith that motivated Herbert Simon and his fellow AI researchers: once they laid hold of an apparently mechanizable logic, they just couldn't help themselves. The mechanism and logic *must* explain everything else. That is how they expected a mechanically conceived world to work, whether they were dealing with human speech and thought or the genetic text of the Book of Life.

What they thirsted after was a world of life and thought driven by a neatly controlling syntax that played itself out with something like cause-and-effect necessity. They imagined this causal necessity much as they imagined the external impact of particle upon particle, molecule upon molecule, where one thing "makes" another happen.

And if this is how things work, then why should they worry about what the Book of Life might turn out to say when they could actually read it? Their confidence that they had wrested the textual secret of life from the cell's nucleus without even a clue of the "translation" is proof that they were not really thinking in textual terms. It wasn't the unknown meaning of the text that excited them so much as their conviction that a cut-and-dried, mechanizable logic had been found for preserving certain "machine states" from one generation to the next. Surely, they thought, the discovery of such a mechanism—seductive and unqualified in its clarity and reassuringly necessary in the attractions and repulsions of its logical atoms—would explain everything.

It was in this spirit that Francis Crick articulated the Central Dogma of molecular biology in 1958. According to this deeply influential doctrine, genetic information flows in one direction only, from genes to proteins. As science historian Evelyn Fox Keller paraphrased the Central Dogma in her book *The Century of the Gene*: "DNA makes RNA, RNA makes protein, and proteins make us." The doctrine, with its genetic determinism and command-and-control view of DNA, paved the simplest, most direct highway to a mechanistic understanding of the organism.

But the highway proved to be little more than a long, rutted detour. The straightforward, neatly determining logical structure envisioned by Crick—a structure that became a feverish obsession during the Human Genome Project—has progressively transformed itself into a seething cauldron of endlessly complex dynamic processes extending throughout the organism. The crucial problem for genetic determinism and the once-prevailing Central Dogma is that biochemical cause and effect within the cell, as in the organism as a whole, never proceeds in one direction alone. To put it coarsely: everything affects everything else.

The string of discoveries supporting this conclusion is not contested. We now know that one gene can produce many different proteins, depending on complex processes that are orchestrated not only by DNA, but also by proteins themselves. Moreover, one protein is not necessarily one protein. For example, depending on the presence of so-called chaperon proteins, a given chain of amino acids (the constituent elements of protein) may fold in different ways. These various foldings in turn shape the overall structure and functioning of cell and organism.

The supposedly linear structure of letters, words, and sentences into which DNA has been decoded simply does not articulate a clean, unambiguous, command-and-control authority sitting atop a hierarchical chain of command. Only a misguided preoccupation with an imagined set of well-defined syntactical relationships could have led researchers to dismiss the greater part of DNA—nearly all of it, actually—as "junk DNA." The junk didn't seem to participate in the neat controlling sequences researchers were focused on, and so it seemed irrelevant. But more recently, as cell biologist Lenny Moss describes, the erstwhile junk has been recognized as part of a "complex system of distributed regulation" in which "the spacing, the positioning, the separations and the proximities of different elements … appear to be of the essence."

But even more devastating for the centralized command-and-control view has been the discovery of "epigenetic" processes. These processes yield hereditary changes that are not associated with structural changes in DNA at all. Rather, they arise from alterations in how the rest of the organism marks and employs its DNA. And beyond this, researchers have been exploring effects upon DNA from the larger environment. In a dramatic reversal of traditional doctrine, investigations of bacteria show that gene mutations can arise from—can even be guided by—environmental conditions in a non-random way. In sum, genes are no more the self-determining cause of everything else in the organism than they are themselves the result of everything else.

Finally, we have seen a startling demotion of the human genome in size relative to other organisms. The most recent and near-final estimate by the Human Genome Project puts humans in possession of 20,000 to 25,000 genes—this compared to at least 25,000 for a tiny, primitive, semi-transparent worm, *Caenorhabditis elegans.* If genes constitute the one-way controlling logic or master program determining the potentials of the organism, then finding such unexpected gene counts is rather like discovering we could implement all the programs of the Microsoft Office software suite using only the minuscule amount of program logic required for a simple daily greeting program.

Reviewing the history of misdirection surrounding the gene, Moss writes:

> Once upon a time it was believed that something called "genes" were integral units, that each specified a piece of a phenotype [that is, a trait], that the phenotype as a whole was the result of the sum of these units, and that evolutionary change was the result of new genes created by random mutation and differential survival. Once upon a time it was believed that the chromosomal location of genes was irrelevant, that DNA was the citadel of stability, that DNA which didn't code for proteins was biological "junk," and that coding DNA included, as it were, its own instructions for use. Once upon a time it would have stood to reason that the complexity of an organism would be proportional to the number of its unique genetic units.

Today, as Evelyn Fox Keller tells us, the findings of the past few decades "have brought the concept of the gene to the verge of collapse." In fact, "it seems evident that the primacy of the gene as the core explanatory concept of biological structure and function is more a feature of the twentieth century than it will be of the twenty first."

Taking Our Words Seriously

To point out the failure of the Central Dogma will strike most geneticists today as anachronistic. "We long ago quit believing such a simplistic doctrine." And, in fact, you will find them regularly disclaiming the "gene-for" view—that is, the belief that for many or most traits of the organism there is a gene, or a few genes, that account for the trait. "We know it's much more complicated than that"—so the disclaimer runs. In the face of such protestations, recital of the history of misdirection begins to seem unfair. After all, scientists must be allowed to make mistakes, as long as they are willing to learn from them. What's important is the knowledge they eventually arrive at.

But does the painfully repetitive history of genetics and AI suggest that they have in fact learned from their mistakes? The best way I know how to answer this question is to elucidate the central misdirection in the history under discussion.

The real significance of the overheated rhetoric of the Human Genome Project lies in the seemingly unstoppable appeal by geneticists to language and thought—that is, to book, word, letter, code, translation, transcription, message, signal, and all the rest. Or, to employ the most universal term today: information. This resort to a terminology so brazenly *mental* in origin appears to be a stunning reversal. Just a few decades ago we still lived within the long historical era during which it was unpardonable for the natural scientist to draw his explanatory terms from intelligent activity. What changed?

Crucially, the age of cybernetics and computation arrived. This brought with it, for many researchers, the promise of the mechanization of language and thought. Suddenly it became respectable to invoke human mentality in scientific explanation because everyone knew you were not really talking about mentality at all—certainly not about anything remotely resembling our actual mental experience. You were invoking computational mechanisms. So the change was less a matter of assigning human intelligence to the mechanically conceived world than of reconceiving human intelligence itself as mechanical performance.

Of course, we have seen that the equation of mechanical computation with mentality was based on the extraordinarily naïve assumption that machine logic is the essence of thinking and language. But if we can look past this reductionism, what we find is that geneticists have glimpsed more truth than they realize, and the reason for their confusion is that, due to their mechanistic compulsions, they cannot bring themselves to accept their own inchoate insight. If they have been driven to textual metaphors with such compelling, seemingly inescapable force, it is because these metaphors capture a truth of the matter. The creative processes within the organism are *word-like* processes. Something does speak through every part of the organism—and certainly through DNA along with all the rest. Geneticists are at least vaguely aware of this speaking—and of the unity of being it implies—and therefore they naturally resort to explanations that seem to invoke a *being who speaks.*

The problem is that their insistence upon textual *mechanisms* blinds them even to the most obvious aspects of language—aspects that prove crucial for understanding the organism. If I am speaking to you in a logically or grammatically proper fashion, then you can safely predict that my next sentence will respect the rules of logic and grammar. But this does not even come close to telling you what I will say. Really, it's not a hard truth to see: neither grammatical nor logical rules determine the speech in which they are found. Rather, they only tell us something about *how* we speak, not *what* we say or *who* we are as speaking beings.

If geneticists would reckon fully with this one central truth, it would transform their discipline. They would no longer imagine they could read the significance of the genetic text merely by laying bare the rules of a molecular syntax. And they would quickly realize other characteristics of the textual language they incessantly appeal to—for example, that meaning flows from the larger context into the specific words, altering the significance of the words. This is something you experience every time you find yourself able, while hearing a sentence, to select between words that sound alike but have different meanings. The context tells you which one makes sense.

The role of context is pervasive. As poets know very well, even the word "prophet" in the two phrases, "old prophets" and "prophets old," carries different ranges of meaning (a point made by Owen Barfield in his seminal study of "poetic diction"). If DNA is like a text, then plasticity of the gene must be one of the rock-bottom, fundamental principles of heredity.

Conversation and Poetry

There is no need for geneticists to endure lectures from philologists, however. As we have seen, this is exactly what their own discoveries of the past fifty years have been shouting at them. In the ongoing conversation between word and text, part and whole—and contrary to the Central Dogma—we find the context of the organism *informing* the genetic text at least as much as the genes can be said to inform the organism. This is the underlying truth that science historian Lily Kay elaborates when she writes: "once the genetic, cellular, organismic, and environmental complexities of DNA's context-dependence are taken into account," we might find that genes "read less like an instruction manual and more like poetry, in all their exquisite polysemy [multiplicity of meaning], ambiguity, and biological nuances."

What this means practically is that, as Craig Holdrege put it in his book *Genetics and the Manipulation of Life*:

> We gain a knowledge of genes … only through knowledge of the organism as a whole. The more knowledge we have of the organism as a whole, the more information we have. This information is not in the genes; it is the conceptual thread that weaves together the various details into a meaningful whole.

The weaving together is a *conversation*, not a merely mechanical unrolling of a logically compelling sequence. When we speak of such things as *messenger* RNA, the conversational context should be obvious. It makes no sense—or, at least, no sense that biologists have yet explained—to speak of a message without a recipient capable of a certain understanding, and without a context for determining how the message is to be construed. If we eliminate these things from the picture, we have a message without meaning, which is no message at all. The question, then, is whether geneticists really believe their own terminology.

They ought to. Everything we have been learning about the genome points to the significance of its conversational context. As Lenny Moss puts it: "If the sum total of coding sequences in the genome be a script, then it is a script that has become wizened and perhaps banal. It wouldn't be the script that continued to make life interesting but rather the ongoing and widespread *conversations about it*" within the biochemistry of the organism.

Actually, it is not so much the script that is banal as the reduced, syntactic reading of it. As Moss himself reminds us, the script is a dynamic one, subject to continual and rapid changes with profound significance—"transpositions, amplifications, recombinations, and the like, as well as modulation by direct chemical modification." There is a lively conversation going on here, but it is one in which our genes are caught up, not one they are single-handedly dictating.

Words of Explanation

Language is the very soul and substance of explanation itself. The reason for this can only be that the world we are explaining has something language-like about it. When we offer a scientific explanation for some aspect of the world, we necessarily assume that the meaning of our words is at the same time the meaning of the chosen aspect of the world. If this so-called "intentionality" of language—its being about something else and not just about itself—were not born of the world's word-like character, then our scientific explanations could tell us nothing about reality. The world must in some sense be a text waiting to be deciphered. This is why the scientist can, in fact, decipher it into the text of a scientific description.

So in reality *all* scientific explanation is founded upon an appeal to the word. The irony lies in the fact that precisely where the computer scientist and geneticist resort explicitly to "word" and "text," what we actually see is a concerted attempt to substitute wordless logic and computational mechanisms for language.

Most fundamentally, this stance takes the form of an attempt to explain words themselves as if they were objects. No longer standing consciously within the transparent meaning of the words we speak, allowing the world to become visible and meaningful through their transparency, we instead take these words as additional things in the world to be explained. That is, we want to understand our explanatory words as if they themselves were nothing more than causal results of processes going on in the world they explain. There is something gravely misconceived in this effort to explain explanation itself—and all the more when the effort involves an appeal to mechanisms stripped as far as possible of their word-like (and therefore of their explaining) nature. It is rather like trying to prove the validity of logic—or, in other words, trying to *prove* the validity of the instruments of proof—and to do so by invoking physical laws. A fool's task.

We can recognize something like the fruitless struggle to explain explanation in the difficulties that beset twentieth-century physics when the attempt was made to understand light—that is, to illuminate illumination. But light is that by which the world becomes manifest, so that the attempt to understand it in terms of manifest entities—for example, in terms of materially conceived particles or waves—led only to universally acknowledged confusion.

The discipline of artificial intelligence went down an analogous path when computer scientists came to believe they could explain speech (and thought) as manifestations of computational devices. Their aim was to explain our powers of explanation by appealing to something not having the essential character of explanation. The result could only be nonsense, which is why the researchers quickly began arbitrarily projecting language back into their wordless explanatory devices.

At its worst, the projection of language in this way became extraordinarily crude. All one needed to do was to label programs and data structures with terms like UNDERSTAND and GOAL, and then mindlessly assume that the programs actually had something to do with understanding or goal-seeking. Such nonsense eventually became downright embarrassing. In 1981 computer scientist Drew McDermott published an essay entitled "Artificial Intelligence Meets Natural Stupidity" in which he ridiculed the use of "wishful mnemonics." He wondered aloud whether, if programmers used labels such as G0034 instead of UNDERSTAND, they would be equally impressed with their clever creations.

Likewise, McDermott commented on Herbert Simon's "GPS" program, written as a much more ambitious successor to the Logic Theorist: "By now, 'GPS' is a colorless term denoting a particularly stupid program to solve puzzles. But it originally meant 'General Problem Solver,' which caused everybody a lot of needless excitement and distraction." He went on to say:

> As AI progresses (at least in terms of money spent), this malady gets worse. We have lived so long with the conviction that robots are possible, even just around the corner, that we can't help hastening their arrival with magic incantations. Winograd … explored some of the complexity of language in sophisticated detail; and now everyone takes 'natural-language interfaces' for granted, though none has been written. Charniak … pointed out some approaches to understanding stories, and now the OWL interpreter includes a 'story-understanding module.' (And, God help us, a top-level 'ego loop.')

The geneticist's strategy was much like the computer scientist's—unsurprisingly, given that DNA is often imagined as a genetic *program.* All that was needed was to put a label on the gene saying that it was the *gene for* such-and-such a trait and—presto!—the gene now spoke a meaningful language. The only problem is that these neatly labeled genes are forever disappearing as rapidly as they are discovered—or, rather, they lose their neat, causal identity against a background of extraordinary complexity. What stands on the biochemical and supposedly causal side of the relation never clearly relates to the trait, and certainly fails to explain it in any adequate sense. This is because the trait—whether it is dark skin, green eyes, cancer, or an aggressive tendency—is quite properly understood in qualitative and meaningful (word-like) terms, whereas the "causal" gene remains at the level of mechanism, not language. Causes and mechanisms, as we prefer to have them, do not *mean.*

The Destruction of Real Science

After centuries of doing its best to ignore the self-contained, illuminating reality of the language with which it describes phenomena, science is now broadly engaged in the task of trying to seize the word by killing it. This is to destroy the means of all explanation, and the result can only be a continuing loss of coherence within scientific discourse. This helps us to understand how we could possibly encounter statements like this one from geneticist Maxim Frank-Kamenetskii:

> Under [DNA's] inexorable laws, the ill fate that befell the father may also threaten the son.

Here, within a single sentence, "inexorable" mutates unaccountably into "may threaten." What does he mean to say? An inexorable law is one thing, and a possible threat quite another. To juxtapose the two as if one were equivalent to the other is sloppiness one should hardly expect to find in a prominent scientist.

It's the same degeneration of the word we've been noting in the language of geneticists generally, and also in the language of artificial intelligence researchers. The unhappy truth is that many scientists, secure amid their precise, mechanically conceived formulas, can scarcely bring themselves to worry very much about the meaning of their words. Indeed, they are doing their best to demonstrate that words are not vessels of meaning at all, but rather genetic or silicon mechanisms for delivering inexorable, cause-and-effect results.

Here, then, is a formula for the destruction of science as a discipline of understanding rather than merely of effective technique. There could hardly be a surer indication of the insecure and disturbed foundations of science than we find in all the confusion over word and text. How much confidence can we place in the understandings conveyed through an enterprise whose verbal and conceptual instruments of understanding are so badly damaged? If there is to be a scientific Prometheus for our day, he must bring the fire of meaning into our various theoretical languages—languages that, in their current desiccated state, are like dry tinder eager for the blaze. And it is almost as if geneticists, with their ceaseless invocation of word and text, have been unconsciously calling down the tongues of flame.

Such a conflagration will doubtless consume a great deal. But it may also purify and transform. If, as we heard Evelyn Fox Keller suggest, the concept of the gene has been brought to the verge of collapse, we can hope that in our revitalized understanding, the gene will truly speak with all the creative and clarifying power of the word—because the entire organism speaks through it. Then its language of wholeness will belong as much to the poet as to the scientist, and we will hear within its rhythms and cadences a song of destiny in which we ourselves are singers.