6

Keith Devlin

Executive Director and Senior Researcher

Center for the Study of Language and Information

Stanford University, USA

1. Why were you initially drawn to computational and/or informational issues?

In 1964, when a junior level high school student in Britain, I was fortunate to obtain a summer internship working at a nearby British Petroleum plant at the very time they purchased time on the first computer (an Elliott 803 mainframe) to be installed at the local university. (It was actually the first computer in the entire city.) Though my principal duty was initially data entry, I was totally intrigued by how the machine worked, and within a few weeks I had taught myself enough about *Algol* (the high level language the machine was equipped with) to be able to identify and fix a major flaw in the sales forecasting program the company was using. The following year BP hired me back for a second summer internship as a software developer, where I wrote a text editor for the on-site mainframe computer BP had just purchased, an Elliott Arch 9000, which came with a 9-instruction machine language and nothing else. (The instructions, called imaginatively using the digits 1 through 9, were all unary. It was pretty close to a Turing machine — and yes, input and output were by paper tape.)

Though the interest in computation this experience aroused in me never disappeared, it was overpowered by what I perceived to be (and I still think are) the far deeper intellectual challenges presented by mathematics, and I studied mathematics at university. It was only much later in my life, after spending many years as a university mathematician, that I began to look again at issues of computation. Having focused my research in mathematical logic, it was hardly surprising that my initial foray back into computation was in the area of artificial intelligence, but it was not long before I concluded that the original goal of "machines that think" was almost certainly not achievable (at least if those "machines" are digital computers), and my interest shifted to what I felt to be the more tractable, but still horrendously deep problem: what exactly *is* information?

I never did return to computing *per se*, either as practitioner or theorist, and accordingly this essay will focus almost entirely on information.

2. What example(s) from your work (or the work of others) best illustrates the fruitful use of a computational and/or informational approach for foundational researches and/or applications?

My first book about information, *Logic and Information*, published by Cambridge University Press in 1991, was based on Barwise and Perry's situation theory. (It actually began as a project to present their theory in a more mathematical fashion than they had in their earlier book *Situations and Attitudes*, published by MIT Press in 1983, but I ended up covering considerably more material, though in a less formal mathematical fashion than I had originally envisaged.) The focus of both books was on the use of situation theory to provide a framework for situation semantics, a theory of natural language semantics, that was particularly well-suited to capture, in particular, issues of indexicality.

Situation theory builds on our everyday, intuitive conception of information. This includes the kind of thing people seek when they approach an "information kiosk," perhaps by asking "Do you have any information about renting bicycles?" A number of ontological assumptions are required to get the theory off the ground. (In my answer to question 4, I'll explain some of the fundamental problems with the concept of information that motivated, in large part, the selection of the situation-theoretic ontology and the development of situation theory.)

First, we assume that information is in general distinct from its representation. Information is assumed to be a semantic concept (the representation being syntactic). We further assume that information is neither true nor false in of itself, but is made true by (or if you prefer, is true of) some part of the world — a *situation.* We write

$$s \models \sigma$$

to denote that the item of information σ is true in the context (i.e., situation) s.

Information is assumed to arise from (or be represented by) some configuration or event in the world by virtue of a *constraint*. Constraints can arise is various ways, from natural regularities (such as the constraint that dark skies are often followed by rain, and hence a dark sky provides the information that rain is likely), to conventions established by humans within a community (such as the convention that a bell ring in a school provides the information that class is over).

Formally, a constraint is defined to be a relation between two *situation types*. In turn, situation types are uniformities across situations. The "ringing bell indicates that the class is over" constraint would be a relation C between the type S of situation in which a bell rings and the type T of a class-ending (situation), written as

$$S \overset{C}{\Rightarrow} T$$

An agent who recognizes a situation s (say, audio input in the agent's current physical context, a linguistics class perhaps) to be of type S, written $s : S$, can infer that there is a situation t (perhaps the linguistics class) of type T, i.e., $t : T$.

A formal calculus of informational entities (*infons*), situation types, and constraints makes it possible to develop these ideas to a point where they can be applied to analyze linguistic utterances and information flow, and situation *semantics* met with some initial success. But the real descriptive/analytic power of situation *theory* was not realized until Duska Rosenberg and I applied it to analyze linguistic data gathered in the course of an ethnographic study of workplace communication.

In our monograph *Language at Work – Analyzing Communication Breakdown in the Workplace to Inform Systems Design*, published by CSLI Publications in 1996, Rosenberg and I used situation theory to organize the ethnographic data Rosenberg had collected in a lengthy workplace study. Our approach was inspired by an analytic approach to language developed by the ethnomethodologist Harvey Sacks. Based on Sacks' work, we created an analytic methodology called Layered Formalism and Zooming (LFZ analysis), and that was what we used to analyze the data collected from the ethnographic study. Our analysis led to specific, implementable recommendations to the company for increasing efficiency.

The key to our success was that, although situation-theoretic descriptions and analyses of linguistic communication are, in all but extremely simple, "toy" examples, nothing like as precise as the mathematical descriptions and analyses employed in, say, physics, they nevertheless prove to be very effective in bringing an unprecedented degree of mathematical precision to the analysis of everyday human–human communication from a social science perspective. In short, situation theory turned out to be ideally suited for bringing greater formality to analyses of complex sociolinguistic data.

While I believe that our monograph did make fundamental scientific contributions to our understanding of information, its greatest initial impact was through the practical application that was the primary focus of our study. Several industrialists expressed interest in the work, but a common complaint was that our account was too heavily mathematical, which for many potential readers made it inaccessible. Accordingly, a short while later I brought out another book on the topic aimed squarely at the business community, called *Infosense*. In that account, I reduced the mathematical formalisms to an absolute minimum, and provided many examples from everyday life.

3. What is the proper role of computer science and/or information theory in relation to other disciplines, including other philosophical areas?

I don't agree with the unstated assumption behind this question. I do not think there is such a thing as a "proper role." Terms I would agree with are "useful role" or "appropriate role." Science is about understanding and engineering is about building, and whatever helps either is justifiable. For example, Stephen Wolfram has been promoting the idea of natural science based on computation (his "a new kind of science"). As it happens, I don't think that the approach he outlines in his book by that name succeeds as a viable alternative to current (property-/relation- based) science, but that is because of his particular approach, not because I think there is a philosophically preferred framework for doing science.

Earlier in my career I would have answered this question differently. Academics are in the business of acquiring and disseminating knowledge. To do this in a systematic way requires an epistemic framework, and that framework governs practically everything a professional in a particular academic discipline does. In the case of the discipline (or disciplines, if you carve things more finely) generally referred to as "Foundations of Mathematics," there is an innate linear ordering — or perhaps a tree — that attempts to impose (or reflect) an order in which one concept is built on another. Thus, the notions of a formal logic and an abstract set are often taken as basic, axioms for both are introduced, and structures such as the natural numbers and the higher number systems are built up within set theory. This viewpoint carries with it, for those who pursue foundational studies, a sense of being "the way things are," or even more strongly, "the way things have to be." Moreover, because the foundations of mathematics is almost universally viewed and practiced as a *synthetic* discipline, practitioners frequently develop a sense that the theory is *prior* to any applications, particularly applications in the real world. Indeed, the very use of the term "applications" reflects this viewpoint.

An alternative approach is to take the world as we encounter it, both the physical world and the social world, and regard the various academic frameworks as simplifying filters that aid our understanding and facilitate analysis. For much of my career, I had the former view; of late I have aggresively adopted the latter. From this "materially pragmatic" perspective, there are no "proper" frameworks; the only metric is efficacy.

From this perspective, what we call "foundational studies" are a finer-grained *analysis* of the epistemically-prior mathematical framework we use to study and understand the world, and the "foundations of mathematics" are not foundations in the sense of the foundations of a building, they are more like the roof. Interestingly, I believe this is precisely the view that the vast majority of mainstream mathematicians always had of foundational studies. For instance, most mathematicians were not the least bit phased by Gödel's Incompleteness Theorems or by Cohen's undecidability results in set theory, which they viewed as dealing with the icing on the top of the cake rather than something truly foundational on which their own work depended. My present philosophy simply extends this view from foundational studies to all of mathematics.

4. What do you consider the most neglected topics and/or contributions in late 20th century studies of computation and/or information?

I'll concentrate entirely on information, since this is what my work has mostly focused on.

I think that it took far too long a time before theorists began to recognize the fundamental problems associated with the notion of information, and to view information as a social construct. (Many still do not.) In particular, Shannon and Weaver's use of the term "information theory" to describe their quantitative, entropy-based approach to communication delayed the development of genuine "*information* theories" considerably. For one thing, the real focus of their theory was a (highly useful) notion of channel capacity, measured in bits, not on information as most people commonly use that word, which is to refer to the "information" that those bits carry. Put simply, counting bits in a signal tells you almost nothing about the information (i.e., what a person would typically call "information") that the signal may carry, which depends entirely on contextual factors. To take a simple example, two people can establish a convention whereby a signal of a single bit can carry an enormous amount of information. Admittedly, the ShannonÐWeaver theory can handle this, by considering the convention as part of the system, and this can be philosophically justified, but to my mind the result is unsatisfyingly contrived, and not at all as mathematically crisp as their framework suggests.

A more appropriate approach, I believe, is to start with the everyday notion of information, make that as precise as possible, and then proceed to analyze the way that information arises, is stored, and is transmitted. This is precisely the approach adopted by situation theory, as outlined above. (It is also, of course, consistent with my overall philosophy of mathematics as outlined in the previous section.)

It is when you approach information in an analytic fashion that you soon find yourself mired in complexity. For instance, at first blush, if I come up to you and say "I just won a major award for my paper on information," and someone were then to ask you what information my statement conveyed to you, almost certainly you would say "Devlin just won an award for a paper he wrote on information." You are less likely to give any of the following replies:

- "Devlin speaks English with a Yorkshire accent."
- "The man who just spoke to me is alive."
- "The man who just spoke to me was nervous."
- "The man who just spoke to me was lying."
- "The man who just spoke to me was intoxicated."

Yet my utterance could equally have conveyed each of those other pieces of information, depending on the circumstances. (It would always convey the information about my accent and the information about my being alive.) Indeed, under the appropriate circumstances, any one of those alternative pieces of information, and an endless sequence of further possibilities, could be said to be the *primary* item of information you, as listener, acquired from my utterance. (Consider immigration officials at airports, who often ask a returning passenger "What was the purpose of your trip?" The official couldn't care less what you were doing, he or she simply wants to see if you display any signs of unusual nervousness that might indicate a problem. In this situation, the key information being sought, and often obtained, is not encoded in what you say, but how you say it.)

In fact, by taking advantage of, or establishing, the appropriate circumstances, practically any signal can be used to store and convey any piece of information, the famous knotted handkerchief being a familiar everyday example of a one-bit representation that can mean one thing one day, another thing the next.

Once you recognize that information depends fundamentally on the circumstances — that words, objects, actions, etc. can convey pretty well any information we want them to — you have to admit that a study of information has to be carried out in a framework that captures enough of the social context in which it arises, is transmitted, and is consumed. How much is "enough" in that last sentence? That was one of the questions the early developers of situation theory had to address. I always felt they did a pretty good job for the first pass — which is why I decided to throw my lot in with the situation theory camp not long after they got their theory off the ground, in the early 1980s.

5. What are the most important open problems concerning computation and/or information and what are the prospects for progress?

The end of my response to question 4 provides my answer to this question. Although there have been a number of attempts to develop formal theories of information (in the everyday sense of the word "information"), to my mind, none have truly succeeded. I believe there is considerable scope for advancement in this area, but it has to be accepted from the start that, since information (as understood by most people, and as I approach it) is a social construct, there is no possibility of developing formal, mathematical theories that resemble, say, classical logic. I believe formality and mathematical precision can be best (and possibly only) achieved through a zooming methodology of the kind Rosenberg and I describe in our monograph.

Such an approach is very different from the currently-accepted conception of mathematical formality, which seeks to provide a (ground-level) formal theory. Instead, the formality lies in the *process* used in an endless *activity* of analysis.

Such a development would be an instance of what I believe will be a general shift in the way we bring mathematics to bear in analyzing social issues. Seduced — with very good reason, I may add — by over two millennia of incredible success in developing and using mathematics to understand the physical world, we came to accept the way mathematics works its magic, with depth, precision, formality, and finality. It was, and remains, the case that when we try to apply mathematics in the social realm, however, things do not work out anything like as well, and attempts to emulate physics in the study of social concepts such as information are doomed to fail.

Economists recognized this long ago, and now make sophisticated use of the latest mathematical models *along with other forms of reasoning that cannot be captured in an equation.* A similar appreciation of the limits of mathematics has yet to be realized by — to pick on just one particular group by way of example — many in the artificial intelligence community, who still purport to believe in a future, mathematically-specified, (digital-) machine intelligence that will equal or surpass the human mind. (You will gather that I do not share that view. I explained my reasons at length in my book *Goodbye Descartes*, so will not repeat here what I wrote there. Many others have articulated similar objections to GOFAI — "Good Old-Fashioned Artificial Intelligence". I should perhaps add that I have always been impressed by some of the real advances in software systems made under the banner of "AI"; it is the original bold goal of GOFAI that I object to.)

Part of the lesson that must be learned in order for new mathematics to be developed that will enable us to truly gain better understanding of social issues such as information, is that rigor does not require mathematical formalism (axioms, formal proofs, and the like). Moreover, when it comes to understanding many social phenomena, the goal is not "perfect understanding" but "better (i.e., deeper, more precise, more illuminating, more useful) understanding."

To take just one example of many possible, Chomsky's mathematical theory of *syntactic structure*, first outlined in his famous 1957 book by that title, provided a mathematical formal description of certain important features of sentence structure. We learned a great deal about language by virtue of Chomsky's mathematics. Not because his theory captured language the way the atomic theory of matter captured (or modeled, if you prefer) the material world around us. It most obviously did not do that. Rather, we learned more about language by seeing the *extent* to which real language both *conforms to* and *differs from* Chomsky's mathematical descriptions.

A second example that comes to mind is Paul Grice's "maxims of everyday language usage,"[1] where he adopts a decidedly Euclid-like axiomatic approach to explaining how people use language to communicate. In the case of Chomsky, the mathematics actually *looks like* (symbolic) math; Grice's work, in contrast, does not easily lend itself to symbolic presentation, yet his approach is clearly "mathematical".

Both Chomsky's work on syntax and Grice's observations on communication are examples of what I am convinced will be a growing trend of using a mathematical or mathematically-inspired approach to increase our understanding of social phenomena. My work with Rosenberg cited earlier falls into the same category.

I believe that the 21st century will see considerable progress in understanding information in this vein, but I suspect that the majority of scholars currently active in "foundational studies" would not, were they to live long enough, recognize or endorse such work as "of their own". Rather, what will transpire, I foresee, is yet another instance of the oft-cited observation of Max Planck that a new scientific paradigm comes to ascendency not because the new turks convince the old guard, rather that the old guard simply die off.

Whether the new guard will refer to such work as "mathematics" I would not hazard a guess, and I don't think it really matters. Whatever it is called, it will give us a greater understanding of information, how it arises, how it is transmitted, and how to process it.

---

[1] See his paper "Logic and conversation", available in Cole, P. and Morgan, J. (eds.) *Syntax and Semantics, Vol 3* New York: Academic Press, and also downloadable from several websites.