News from New Zealand

by

C. S. Calude

Department of Computer Science, University of Auckland
Auckland, New Zealand
cristian@cs.auckland.ac.nz

1 Scientific and Community News

0. The latest CDMTCS research reports are (http://www.cs.auckland.ac.nz/staff-cgi-bin/mjd/secondcgi.pl):

432. Y. I. Manin. Zipf's Law and L. Levin's Probability Distributions. 02/2013

433. G. Altmann, I.-I. Popescu and D. Zotta. Stratification in Texts. 02/2013

434. C.S. Calude and K. Tadaki. Spectral Representation of Some Computably Enumerable Sets With an Application to Quantum Provability. 03/2013

435. K. Tadaki and N. Doi. Cryptography and Algorithmic Randomness. 04/2013

436. K. Tadaki. Phase Transition and Strong Predictability. 04/2013

437. A. Probert and M.J. Dinneen. Branchwidth, Branch Decompositions and b-parses. 04/2013

2 A Dialogue with Yuri I. Manin: My Life Is not a Conveyor Belt

Professor Yuri Manin, `http://www.mpim-bonn.mpg.de/node/99`*, is a member of three institutions based in three different countries: Max Planck Institut für Mathematik, Bonn, Germany, Steklov Mathematical Institute, Academy of Sciences, Moscow, Russia, and Northwestern University, Evanston, USA.*

Professor Manin was educated at Moscow University; his graduate studies have been supervised by I. R. Shafarevich. He obtained famous and important results in extremely diverse mathematical areas—algebraic geometry, number theory, mathematical logic, mathematical physics, informatics. "The field of quantum computing was first introduced by Yuri Manin in 1980 [2] and Richard Feynman in 1981 [3], [4]".[1]

Professor Manin is not "a monomaniac mathematician, but ... a deep scholar with wide interest, for whom penetration into the mystery of knowledge is much more important than professional success"[2]*. He has also published extensively in literature, linguistics, mythology, semiotics, physics, computer science, philosophy of science and history of culture. His very long list of honours and awards includes invited lectures at universities and congresses*[3]*, prizes, membership in learned academies, honorary degrees and visiting appointments from prestigious organisations around the world. The books* Mathematics and Physics *and* Mathematics as Metaphor *provide a deep insight in his philosophy of science. He supervised 49 PhD students, some outstanding mathematicians themselves.*

CC: What was your motivation to move from an area to a completely different one, not once, but several times? Did these transitions affect your "productivity" in the short term?

YIM: I love mathematics, this great vast realm of human spirit, I am interested in its various aspects, I would like to understand as much of it as I can, and there is no other way than studying and working in various areas in turn. Productivity?? My life is not a conveyor belt, this word is not in my vocabulary.

[1] Wikipedia, "Quantum Computer" `https://en.wikipedia.org/wiki/Quantum_computer`.

[2] `http://www-history.mcs.st-and.ac.uk/Biographies/Manin.html`.

[3] Among them, a plenary and five invited lectures at International Congresses of Mathematics, 1966–2006.

CC: The inscription (promoted by academic bureaucrats) "publish or perish" is on all Graduate Schools walls. There is no better rebuttal than your statement: "Productivity? This word is not in my vocabulary".

YIM: "Publish or perish" is a joke, of course... And a mild joke with a grain of sadness is the best way to cope with existential anxiety.

CC: Mathematical logic is a not a core subject for the working mathematician. It is not even taught in many mathematics departments: its home is nowadays in philosophy departments and mainly in computer science departments. How did get interested in mathematical logic?

YIM: I started thinking about mathematical logic when I have already published several dozens research papers in the domains I was trained (algebraic geometry, number theory) and felt the need to expand my overall view of mathematics. Moreover, it was time when Matiyasevich made the last decisive step in the proof of the theorem that all enumerable sets are Diophantine. I could easily understand what are Diophantine sets, but enumerable ones required some study. Turning to books and articles on Logic, I met again what was already a familiar problem: I could not achieve understanding by just reading, other people's texts did not tell me what I felt I needed to know.

The remedy was also by this time well-rehearsed by me: I taught a course on mathematical logic, this time not even in Moscow University where I served as Professor (although my principal job was at the research Steklov Institute for Mathematics), but at the Moscow Institute of Electronic Engineering. My notes of the course were published in 1974, and they became the first draft of my book on Mathematical Logic published by Springer in 1977.

CC: You have developed your own formulation of Kolmogorov complexity in the idiosyncratic book on mathematical logic. In spite of being incomputable, you have successfully applied Kolmogorov complexity to mathematics, physics and linguistics. Your ideas, presented a few years ago at the CiE conference, the largest and arguably most important meeting dedicated to computability theory, have attracted a lot of interest.

YIM: But the mathematical theory of computations is interesting *exactly* because it delineates precise boundaries of the realm of computable, and most interesting things happen when we cross these boundaries! Kolmogorov complexity turned out the great bridge from the land of computable to the vaster realm of mathematics, unconstrained by computability, and, I hope, to physics as well.

I am very happy that during the last several years this vague feeling found justification in three papers of mine where Kolmogorov complexity plays the role of "energy" in three very different contexts: renormalisation in computation, asymptotic boundaries for error-correcting codes as phase transition curves (joint work with Matilde Marcolli), and quite recently, a mathematical explanation of Zipf's law[4].

CC: Please explain one result in which Kolmogorov complexity works as "energy".

YIM: Consider a finite alphabet A consisting of q letters. An error-correcting block code is a subset $C \subset A^n$, $n \geq 1$ (I am speaking about unstructured codes, but the results I will explain hold, with appropriate modifications, also for linear codes etc.). Each such code C defines a point in the unit square of the plane (*R:= transmission rate, δ := relative minimal distance*). Now, what will you see if you look at the cloud of *all* code points (in a fixed alphabet)?

In 1981 I have proved that this cloud is everywhere dense below the graph of a certain continuous function, $R \leq \alpha_q(\delta)$ whereas above it each graph point is isolated. This curve $R = \alpha_q(\delta)$ ("silver lining" of the cloud) is called *the asymptotic bound.* In dozens of papers various upper/lower estimates of this function were obtained, but up to now, its exact values are unknown (outside the trivial range), and even whether this function is theoretically computable is open.

In our paper with Matilde we constructed a partition function on the set of all codes, in the sense of statistical physics, in which the energy of the code is exactly its (logarithmic) Kolmogorov complexity. It turned out that after a simple renaming of coordinates, the asymptotic bound becomes the phase transition curve in the (*temperature, density*) plane.

CC: What led you to think about quantum computing?

YIM: First, contemporary computers were electronic devices. They were supposed to produce *exact* results at the level of single bits, hence they had to be designed to suppress quantum effects inherent to any electronic device. Ongoing micro–minituarization was making this task more and more difficult, so it was natural to think about *using* quantum effects rather than suppressing them.

Second, as I have written then, *"the quantum configuration space is much more spacious that the relevant classical one: where in classical physics we have N discrete states, in the quantum theory allowing their superposition there is about c^N states. The union of two classical systems produces N_1N_2 states, whereas the quantum version has $c^{N_1N_2}$ ones."*

CC: The ancient Greek poet Archilocus observed that "the fox knows many things, but the hedgehog knows one big thing". For Freeman Dyson, foxes are mathematical birds (who "fly high in the air and survey broad vistas of mathematics out to the far horizon") and hedgehogs are mathematical frogs (who "live in the mud below and ... solve problems one at a time"). Mathematics needs both birds and frogs.

[4] Zipf's law states that in a natural language corpus, the frequency of a word is inversely proportional to its rank in the frequency table.

Dyson called you a bird. While "globally" this seems to be true, I think that "locally" you have alternated between a bird and a frog. Are you a Francis Bacon's mathematical bee (who "extracts material from the flowers of the gardens and meadows, and digests and transforms it by its own powers")?

YIM: The Swiss writer Max Frisch published in 1953 the ironic comedy "Don Juan oder die Liebe zur Geometrie"[5]. I find its title a wonderfully concise description of my mathematical personality.

Yes, I am a mathematical Don Juan. I still love all my loves, and when I meet my old flame, she can seduce me once again.

Yes, perhaps, all these loves are just incarnations of my deep love for "Geometry" (the latter including algebraic geometry, homotopy topology, quantum field theory . . . and what not).

CC: In a memorable interview published in 1998, you said that "the mathematics of the 20th century is best presented around programmes". Is this trend visible in the 21th century? Is the same true for computer science?

YIM: Probably, it is too early to speak about the trends of the 21th century: imagine an interview on the trends of the 20th century in 1913 . . .

But anyway, I see the full of energy development of programmes I most cherished in the 20th century mathematics: Grothendieck's vast enterprise expanding in many directions thanks to efforts of many strong minds; Langlands' programme; Turing and von Neumann's programmes (each new computer virus is a poisonous descendant of von Neumann's imagination).

CC: The "computer-assisted proofs, as well as computer-unassisted ones, can be good or bad. A good proof is a proof that makes you wiser" is nowadays less controversial than fifteen years ago when you made it. How can a computer-assisted proof make you wiser? What about a quantum computer proof?

YIM: A good proof starts with a project connecting your expected theorem with other results, opening vistas to interesting variations and generalisations. When you develop a detailed plan of it, it might happen that on the road you will have to make a complete list of some exceptional cases, or marginal situations, in which something is not quite as it is "generally", and that making such a list requires a search in a finite but vast set of a priori possibilities.

Then a good computer program that makes this work for you will not spoil the quality of your proof. If the computer is quantum one, then of course you must additionally convince yourself and others that the answer given with "probability close to one" is in fact a correct one.

CC: A paraphrase by C. Anderson (Wired Magazine) falsely attributed to Google's research director Peter Norvig, claims that "all models are wrong, and

[5] Don Juan, or the Love of Geometry.

increasingly you can succeed without them". An example is Google capability to match ads to content without any knowledge or assumptions about the ads or the content, and to translate languages without actually "knowing" them. Why? Google doesn't know why the webpage A is better than the webpage B. However, "if the statistics of incoming links say it is, that's good enough. No semantic or causal analysis is required". From here to the aggressive proclamation of the death of science was just one step: data deluge makes the scientific method obsolete.

YIM: I'll start with stressing that we are speaking about the science rather than market managing.

Now, what Chris Anderson calls "the new availability of huge amounts of data" by itself is not very new: after spreading of printing, astronomic observatories, scientific laboratories, and statistical studies, the amount of data available to any visitor of a big public library was always huge, and studies of correlations proliferated for at least the last two centuries.

Charles Darwin himself collected the database of his observations, and the result of his pondering over it was the theory of evolution.

Even if the sheer volume of data has by now grown by several orders of magnitude, this is not the gist of Anderson's rhetoric.

What Anderson actually wants to say is that human beings are now – happily! – free from thinking over these data. Allegedly, computers will take this burden upon themselves, and will provide us with correlations – replacing the old–fashioned "causations" (that I prefer to call scientific laws) – and expert guidance.

Leaving aside such questions as how "correlations" might possibly help us understand the structure of Universe or predict the Higgs boson, I would like to quote the precautionary tale from J. Groopman. The Body and the Human Progress, *New York Review of Books*, Oct. 27, 2011:

"[...] *in 2000 Peter C. Austin, a medical statistician at the University of Toronto, and his colleagues conducted a study of all 10,674,945 residents of Ontario aged between eighteen and one hundred. Residents were randomly assigned to different groups, in which they were classified according to their astrological signs. The research team then searched through more than two hundred of the most common diagnoses of hospitalization until they identified two where patients under one astrological sign had a significantly higher probability of hospitalization compared to those born under the remaining signs combined: Leos had a higher probability of gastrointestinal haemorrhage while Sagittarians had a higher probability of fracture of the upper arm compared to all other signs combined.*

It is thus relatively easy to generate statistically significant but spurious cor-

relations when examining a very large data set and a similarly large number of potential variables. Of course, there is no biological mechanism whereby Leos might be predisposed to intestinal bleeding or Sagittarians to bone fracture, but Austin notes, 'It is tempting to construct biologically plausible reasons for observed subgroup effects after having observed them.' Such an exercise is termed 'data mining', and Austin warns, 'Our study therefore serves as a cautionary note regarding the interpretation of findings generated by data mining' [. . .]"

Hence my answer to Anderson's question: "What can science learn from Google" is very straightforward: "Think! Otherwise no Google will help you."

CC: Ramsey theory has shown that complete disorder (true randomness) is an impossibility. Every large database (of numbers, points or objects) necessarily contains a highly regular pattern. Most patterns are not computable. Can content–based correlations be distinguished from Ramsey-type correlations?

YIM: I am not an expert. I did not wander far away from the notorious motto[6] about "lies, damned lies, and statistics".

CC: Richard Hamming famously said that "the purpose of computing is insight, not numbers". Do you agree? Do you think that mathematics will continue to be relevant to computer science?

YIM: Yes, and yes.

CC: Will mathematics die? But linguistics?

YIM: Archilocus' fable on the Fox and Hedgehog was re–introduced in our contemporary cultural household by Isaiah Berlin. Berlin had a keen ear for compressed wisdom, and called another of his book of essays after Immanuel Kant: "The Crooked Timber of Humanity". Berlin's message was that all global social projects were doomed: one cannot build a house from crooked timber.

However, we want to be optimists and to believe that human civilisation as we know it for the last two thousand years survives. Then mathematics will survive as well. It is incredibly resilient! My favourite example recently was Pappus' hexagon theorem (Alexandria, about 330 AD), a jump through millennia from Euclid to modernity.

CC: How interesting. Can you explain some details?

YIM: Trying to explain its statement without a picture, I would first suggest to imagine six points in plane, numbered cyclically ("vertices of a hexagon"). Two consecutive points define a line, passing through them, "one side" of this hexagon, there are all in all six sides. Two opposite points define another line, "a diagonal" of this hexagon, there are all in all three diagonals. For each diagonal, there are exactly two sides intersecting this diagonal *not* in vertices. I will say that this diagonal has *Pappus property* if this diagonal and the respective two sides

[6] Benjamin Disraeli.

have just one common point. *The Pappus Theorem* now says that if two of three diagonals have Pappus property, than the third one has it as well.

What immediately strikes anyone looking at the Pappus theorem, is its totally “non–Euclidean” character: neither its statement, nor its proof depends on angles and distances. In fact, it took more than a millennium to understand that Pappus theorem refers to the (real) projective plane, uses only the relation of incidence between lines and points, and, in a hidden form, basic properties of addition and multiplication of real numbers.

A couple centuries later, it became clear that Pappus plane’s combinatorics is *completely equivalent* to the axiomatics of abstract fields and abstract projective geometry over them: essentially, his statement taken as an *axiom* is equivalent to the fact that the combinatorics of the incidence relation is an instance of (linear) projective geometry.

Then the whole non–linear algebraic geometry over algebraically closed commutative fields was rewritten in the incidence terms, vastly generalising Pappus, using the theory of models, a chapter in mathematical logic.

And during the last twenty years the abstract Pappus theorem/axiom was used in order to achieve an essential progress in the Alexander Grothendieck’s *anabelian programme*.

CC: Many thanks.