MATHEMATICS

Proof at a roll of the dice

Bernard Chazelle

The PCP theorem encapsulates the idea that randomization allows the immediate verification of any mathematical proof. A simple route to this striking result was proposed earlier this year.

Many branches of mathematics have their signature numbers: geometry has the transcendental π; analysis Euler's exponential e; and algebra the imaginary unit, the square-root of –1, i. These numbers are the stars of exotic — but seemingly quite separate — worlds scattered across the vast expanse of the mathematical cosmos. Until, that is, one day a master stargazer spots a stellar alignment of breathtaking beauty, and nothing ever looks quite the same again. For these worlds, that alignment is Euler's identity, $e^{i\pi} = -1$, and it is proof, if proof were needed, that mathematics is about more than shape (π), change (e) and structure (i): it is also about magic.

Every science boasts its own brand of this wonder of unification: computer scientists, for their part, marvel at what they call PCP, shorthand for probabilistically checkable proof. Simply stated, this is the curious phenomenon that the mere ability to toss coins makes it possible to check the most complex of mathematical proofs at no more than a passing glance. This remarkable fact was discovered by Arora *et al.*[1,2] a decade ago, and proving it has relied on a vast arsenal of complex algorithmic and algebraic techniques. In the spring of 2006, however, Irit Dinur proposed an elementary proof[3]. In the short years of its existence, PCP has revolutionized the field of approximation algorithms — methods for finding nearly optimal solutions to problems that cannot be solved exactly within a reasonable time. Dinur's result is the latest chapter in one of the most engrossing chronicles of computer science.

To appreciate fully the significance of PCP, imagine you wake up one morning with your head full of a complete proof of the Riemann hypothesis. (This is arguably the greatest open problem in mathematics, and is a deep statement about the distribution of the prime numbers, the atoms of arithmetic.) Giddy with anticipation of certain fame, you leap out of bed and type up the proof, in its full 500-page glory. That done, you are seized by doubt. How wise is it to inflict the full brunt of your genius on your colleagues? Will anyone even listen, or be prepared to check your proof?

What you can do is re-format your write-up into a (somewhat longer) PCP proof. Anyone now wishing to verify your argument need only pick a handful of words at random, and follow a set list of instructions to conclude whether it is correct or not. An error might have slipped in on account of faulty working, buggy reformatting, or outright cheating. No matter, it will be caught with overwhelming probability. What the reformatting step does is smear any error all over the proof, making it easier to spot. In much the same way, a diligent sandwich maker will smear a smidgen of jam evenly over his bread, rather than leaving it concentrated in one corner, and so make the whole more savoury.

But before we look under the PCP hood, a word of reassurance to those allergic to maths: the theory extends far beyond mathematical proofs. In its full splendour, PCP asserts that any statement S whose validity can be ascertained by a proof P written over n bits also admits an alternative proof, Q. This proof Q has two appealing features: it can be derived from P in a number of steps proportional to n^c, where c is some constant; and P can be verified by examining only three bits of Q picked at random. If S is true, a correct P will satisfy the verifier with a probability of 99%. If it is not true, any alleged proof P will trigger a rejection from Q with a probability higher than 50% (ref. 4). Not impressed with this error rate? Then all you have to do is pick, instead of three bits of Q, as many bits as are contained in this line of text. The error probability will drop to one in a billion.

Let's take as an example the statement S that graph G is 'three-colourable'. (Thus, if it were a map, it could be coloured in with three colours — say red, green and blue — with no two neighbouring countries sharing the same colour.) The standard proof P of this statement specifies the colour of every vertex (country) in G, and it is verified by checking that no edge (border) is monochromatic. Dinur's formulation[3] of PCP, however, instructs the verifier how to build, relatively quickly, a new, error-smearing graph H. This second graph is constructed such that if G is three-colourable, so will H be; if G is not three-colourable, any colouring of H will leave at least a fixed fraction of its edges monochromatic (Fig. 1). In this way, equipped with H, the verifier will be able to probe any alleged colouring at random, with confidence that a monochromatic edge will be spotted if G is not three-colourable.

But how do we efficiently face-lift G into the desired H? First of all, in a pre-op phase, we inject extra edges into G to make it look random. The surgery itself alternates between two procedures. First, we add edges to make neighbourhoods of G more tightly connected. The intended effect is to amplify the

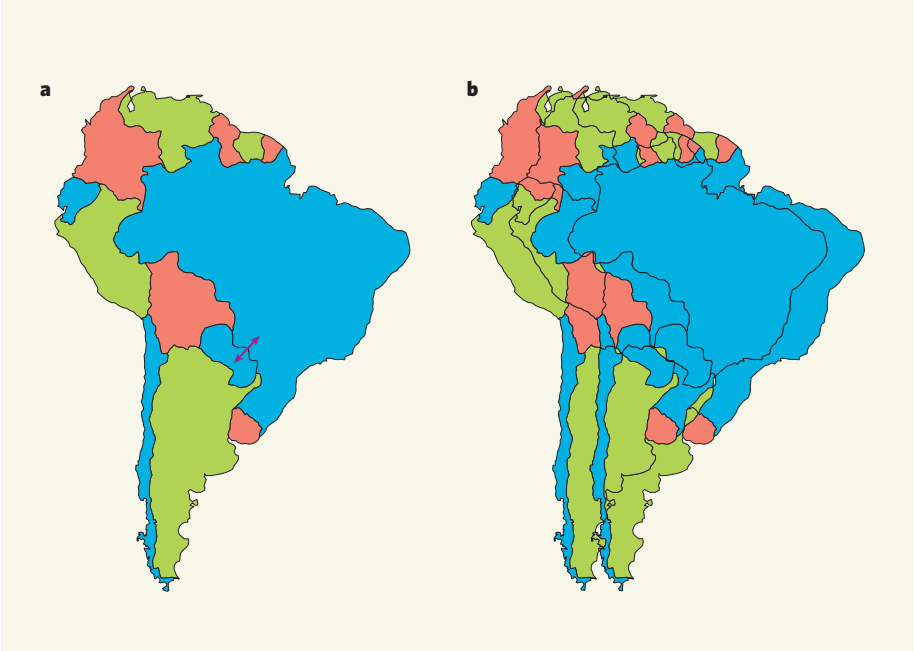


Figure 1 | Probabilistically checkable colouring. The theorem that, on a two-dimensional map, no more than four colours are needed so that no two adjacent countries have the same shading is one of the most notorious in mathematics, and the first to be proved by a computer[5]. Using only three colours is more ticklish, and (**a**) runs into difficulties even in a finite region of limited boundaries. PCP can be considered as offering an alternative way — rather than laboriously checking every boundary — of ascertaining whether a map claimed to be three-colourable really is. The process is analogous to (**b**) 'smearing' the map and its error out: in the original, only one error (adjacent countries of the same colour; arrow) occurs, whereas in the smeared map, it is replicated in many areas. Establishing the validity — or not — of the original map with high statistical certainty thus requires the checking of only a small, randomly chosen subregion of the smeared map.

monochromaticism of any faulty colouring. Cosmetic enhancement comes at a price: here, the addition of new colours. So, in a second step, to restore the graph to its original hues of red, green and blue, we call upon error-correcting codes, the redundancy-adding devices invented by coding theorists to make signals immune to noise. This is no coincidence: just as an encoded string with not too many erroneously flipped bits can, through error correction, be restored to its original state, so a PCP proof with not too many errors can be seen to encode a correct proof. Indeed, any error that is not smeared widely enough to be easily spotted is *de facto* inconsequential.

If you think that no one besides children and cartographers has any interest in colouring graphs, think again: the Riemann hypothesis, protein folding, cryptography and most questions in artificial intelligence can be reduced to the three-colourability of some graphs. That universality is another wonder in the computing cosmos.

PCP, and its elementary proof by Dinur[3], is the culmination of 40 years of research in the field of computational reduction. Keep in mind that this is all about verifying proofs, not about understanding them — with only three bits! — let alone discovering them. That must still be done the hard way. In the end, PCP boils down to one revolutionary insight: in the art of persuading others of the truth, the ultimate weapon is a set of dice. ■

Bernard Chazelle is in the Department of Computer Science, Princeton University, 35 Olden Street, Princeton, New Jersey 08540-5233, USA.
e-mail: chazelle@cs.princeton.edu

1. Arora, S. & Safra, S. *J. Assoc. Comput. Machin.* **45,** 70–122 (1998).
2. Arora, S., Lund, C., Motwani, R., Sudan, M. & Szegedy, M. *J. Assoc. Comput. Machin.* **45,** 501–555 (1998).
3. Dinur, I. *Proc. 38th Annu. ACM Symp. Theor. Comput.* 241–250 (2006).
4. Håstad, J. *J. Assoc. Comput. Machin.* **48,** 798–859 (2001).
5. Appel, K., Haken, W. & Koch, J. *J. Math.* **21,** 439–567 (1977).

STRUCTURAL BIOLOGY

Dangerous liaisons on neurons

Giampietro Schiavo

Crystal structures show that botulinum toxins bind simultaneously to two sites on neurons. This dual interaction allows them to use a Trojan-horse strategy to enter nerve terminals, with deadly effect.

Botulinum neurotoxins (BoNTs) are some of the most deadly substances known to mankind. By blocking nerve function, they cause botulism, a severe condition that may ultimately lead to muscular and respiratory paralysis. These sophisticated bacterial proteins owe their toxicity to their extraordinary specificity for neurons and to their enzymatic activity. In this issue, papers by Jin *et al.*[1] and Chai *et al.*[2] describe the mechanisms by which BoNT/B — a toxin that causes human botulism — recognizes the surface of neuron junctions (synapses)*. This work provides insight into how other BoNTs may exert their lethal action, and describes a mode of binding that might be used by other biological compounds.

Once inside a neuron, a single molecule of BoNT is, in principle, capable of deactivating the whole synapse. BoNTs consist of two protein segments, known as the heavy and light chains. It is the light chain that deactivates neuromuscular junctions — the synapses that connect muscles to their controlling neurons — by specifically inhibiting members of the SNARE protein family[3]. SNARE proteins are distributed over the membranes of all animal and plant cells and are also found on the membranes of synaptic vesicles, the bubble-shaped organelles that store and release neurotransmitter chemicals at neuron terminals. SNARE proteins are essential for membrane fusion, during which vesicles merge with the cell membrane and release their load. Once the synaptic vesicles have done this, they are recycled by the neuron for further use.

So how do BoNTs enter neurons? The heavy chain is most likely to be responsible. One half of the heavy chain mediates binding to neurons by interacting with lipid molecules (polysialogangliosides, PSGs) in the cell membrane, and with either one of two integral membrane proteins — synaptotagmin I or synaptotagmin II — found in synaptic vesicles. A dual-receptor model for these toxins was proposed long ago[4], but experimental validation of this theory has required a worldwide effort. The model predicts that the interaction of BoNTs with both PSGs and protein receptors is necessary to explain their awesome potency[3], with a different protein receptor being recognized by each BoNT.

Evidence for protein involvement in BoNT binding was scarce until it was discovered[5] that BoNT/B binds to both PSGs and the part of synaptotagmins that lies inside synaptic vesicles, in the area known as the lumen. More recently, the specific regions of synaptotagmins that bind BoNT/B have been identified[6],

*This article and the papers concerned[1,2] were published online on 13 December 2006.

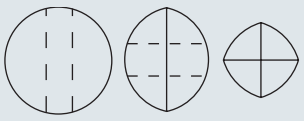
50 YEARS AGO

The recent passion for the progressive splitting up of genera has spread through most branches of zoology and latterly to botany as well. In fact, the present phase in taxonomy is to lump species and split genera... At present, names are very often proposed on flimsy grounds and the work of proving their validity or otherwise is handed down to future workers. New genera are established using merely the specific characters of the unique species included without any additional evidence. The result is that two species such as *Cypraea tigris* L. and *C. pantherina* Sol. are placed in different genera, whereas they are only just distinct species which actually hybridize. Such absurdities should have no standing in nomenclature.

From *Nature* 22 December 1956.

100 YEARS AGO

"Cutting a Round Cake on Scientific Principles" — Christmas suggests cakes, and with these the wish on my part to describe a method of cutting them that I have recently devised to my own amusement and satisfaction. The problem to be solved was, "given a round tea-cake of some 5 inches across, and two persons of moderate appetite to eat it, in what way should it be cut so as to leave a



minimum of exposed surface to become dry?" The ordinary method of cutting out a wedge is very faulty in this respect... The cuts shown on the figures represent those made with the intention of letting the cake last for three days, each successive operation having removed about one-third of the area of the original disc. A common India-rubber band embraces the whole and keeps its segments together.

From *Nature* 20 December 1906.

50 & 100 YEARS AGO