Computing

Solving problems in finite time

Philip W. Anderson

I remember with pleasure and some surprise the two weeks that my high-school maths teacher took out of teaching advanced algebra to introduce us to the kind of logical puzzle where, for instance, you may be given: three last names, say Smith, Jones and Doe; three houses, painted white, red and blue respectively; and three first names, Alice, Barbara and Carol. You are told a carefully rationed set of facts — such as Alice had a quarrel with Mrs Smith and Carol admired her friend's blue house — and asked to connect first and last names and house colours. It turns out (although it is not part of high-school maths even today) that these are all special cases of one of the classic problems of computer science, the 'K-SAT' problem. This is a good example of the type of problem where the computer time required to arrive at a solution may explode exponentially with the number of variables, or the number of propositions one is asked to analyse. On page 133 of this issue, Monasson et al.[1] present a new analysis of the K-SAT problem that throws light on when and why such computing tasks become prohibitive.

The K-SAT problem was the first to be assigned to the class of NP (nondeterministic polynomial time)-complete problems — those for which a potential solution can be verified in a reasonable amount of computer time; but the problem can cost exponential time to solve in the worst case. Thousands of computational problems from many different fields can be shown to be instances of the K-SAT problem. In fact, K-SAT problems are often used to benchmark the performance of search algorithms. These problems can always be put in a canonical form. There are usually N logical variables that may be given the values 'true' and 'false', and a set of M clauses, which are 'or' statements about the N variables. If the problem is pure K-SAT then these clauses connect K of the variables. For instance, in my little example above, the fact that there are three names translates into a three-'or' statement that the name can be Alice or Barbara or Carol. Then for the pure K-SAT problem one asks that the logical sum of all the clauses be true. Monasson et al.[1] also consider the problem of determining the minimum number of clauses that need to be falsified — crossed out — to make the sum of the clauses true.

This work is the latest of many attempts[2–4] to forge a connection between the classic problems of computer complexity theory and the statistical mechanics of random systems, such as the 'spin glasses' — a type of alloy containing a random distribution of magnetic atoms, the spins of which are 'frus-trated', which is to say they have conflicting interactions. In the case of K-SAT, the corresponding spin-glass energy function can simply be constructed by making the variables into spins — 'up' equals 'true', and so on — and by counting the number of unsatisfied clauses for any configuration of the spins. Because sophisticated mathematical methods have evolved for solving such statistical-mechanics problems, one might hope that this would have important implications for complexity theory. Unfortunately, these hopes have been mostly unsatisfied, because the goals of the two 'solutions' are so different. The computer scientist's solution is a particular configuration in a particular instance, and his proofs must encompass all difficult cases; the physicist is looking at a typical case among some universe of randomly chosen examples, and at the statistically average result for large N.

Nonetheless, some ideas from statistical mechanics have been helpful. Notably the 'simulated annealing' algorithm has been very effective in tackling the well-known 'travelling salesman problem', another NP-complete problem, which has many commercially important applications. This algorithm is based on the physics of phase transitions, specifically the way that the structure of a glassy solid relaxes as it is cooled. To reduce internal stresses the trick is to slowly cool the glass through the transition (the process of annealing), so that it is able to find its minimum energy state. If the glass is cooled too quickly it ends up in a higher energy state. Equivalently, if a search algorithm approaches a solution too rapidly, it may find a local, but not necessarily a global minimum.

The general idea of phase transitions turns out to apply to K-SAT as well, in that there is a critical value of α, defined by $\alpha = M/N$, for each value of K, below which almost all cases are satisfiable and above which they are almost all unsatisfiable. In this sense there is a discontinuous 'phase transition' across which drastic changes occur in the computational difficulty. The value of α for $K = 2$, as well as a number of other average properties, may be found analytically. It is an interesting and important fact that it is at values of α near the transition that the average cost of heuristic searches for the solution is maximum, making it extremely hard to find a solution[3]. But at values of α away from the boundary it becomes relatively easy to solve the formulae (by showing them to be satisfiable and unsatisfiable at low and high values respectively).

Monasson et al.[1] concentrate on '$2 + p -$ SAT' problems: where a fraction, $p < 1$, of the clauses have three variables, the rest two. The basic 2-SAT problem is not NP-complete, but 'polynomial' and therefore easy to solve, whereas 3-SAT is NP-complete. All problems with $p > 0$ are NP-complete according to mathematicians, but by empirically testing a large sample of randomly generated cases Monasson et al. find that, below $p = 0.41$, the computational cost scales linearly with N, and a critical value of α corresponding to a continuous transition from satisfiable to unsatisfiable solutions can be

Developmental biology

Bonsai flies

Miniature organisms are no longer the sole preserve of Japanese bonsai lovers. Ernst Hafen and colleagues (*Cell* 97, 865–875; 1999) have discovered that flies with mutations in a gene termed *chico* — meaning 'small boy' in Spanish — are less than half the normal size (compare the *chico* mutant fly with the wild-type fly below it). Not only are the cells in *chico⁻* flies smaller than usual, but there are fewer of them — defects that the authors attribute to decreased cell growth and proliferation rather than increased cell death.

The Chico protein is, in fact, very similar to the vertebrate IRS1–4 proteins, which are involved in cellular metabolism. Consistent with this, Hafen's group find that *chico⁻* flies also have metabolic defects — despite their diminutive stature, they contain twice as much lipid as normal. It also turns out that *chico⁻* flies are similar in size to normal flies reared without proper nutrition. So, the authors propose that Chico may belong to a nutritional sensing system which, by regulating cell proliferation, growth and metabolism, controls growth in response to nutritional conditions. **Alison Mitchell**
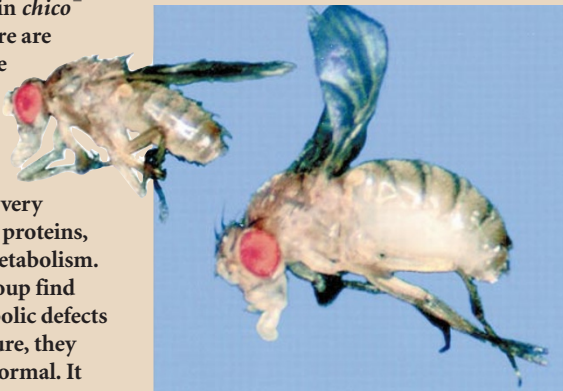
found analytically. But at $p = 0.41$ the phase transition changes character and simultaneously the computational cost of heuristic searches becomes exponentially great. The authors have defined an order parameter for the statistical phase transition, and above $p = 0.41$ they show that this phase transition jumps discontinuously at the transition. But its variation near the transition suggests that there must be critical fluctuations. (The incorrect statement is made[1] that this is new, but in fact it resembles markedly and unexpectedly the Anderson–Yuval–Kosterlitz–Thouless type of defect-mediated transition[5], used to describe melting in two dimensions.)

In addition to the fascinating implications of this unexpected structure for computational complexity theory, where it marks a long-suspected distinction between

problems that are in some sense 'usually' simple to solve rather than 'usually' difficult, this is the first direct correlation of a transition in computational complexity with a critical point of a statistical-mechanics transition. For statistical physicists, this 'new' type of transition (but see my parenthetical comment above) may hold unexpected implications. □

Philip W. Anderson is in the Joseph Henry Laboratories of Physics, Princeton University, Princeton, New Jersey 08544-0708, USA.
e-mail: pwa@pupgg.princeton.edu

1. Monasson, R., Zecchina, R., Kirkpatrick, S., Selman, B. & Troyansky, L. *Nature* **400,** 133–137 (1999).
2. Mezard, M., Parisi, G. & Virasoro, M. A. *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987).
3. Kirkpatrick, S. & Selman, B. *Science* **264,** 1297–1301 (1994).
4. Monasson, R. & Zecchina, R. *Phys. Rev. Lett.* **76,** 3881–3885 (1996).
5. Thouless, D. J. in *Topological Quantum Numbers in Non-Relativistic Physics* 109–115 (World Scientific, Singapore, 1998).

Alzheimer's disease
Antibody clears senile plaques

Peter H. St George-Hyslop and David A. Westaway

People suffering from Alzheimer's disease develop a progressive dementia in adulthood, accompanied by three main structural changes in the brain: diffuse loss of neurons in the hippocampus and neocortex; accumulation of intracellular protein deposits termed neurofibrillary tangles; and accumulation of extracellular protein deposits termed amyloid or senile plaques, surrounded by misshapen nerve terminals (dystrophic neurites). A main constituent of these amyloid plaques is the amyloid-β peptide (Aβ), a 40–42-amino-acid protein that is produced through cleavage of the β-amyloid precursor protein (APP). Although Alzheimer's disease can be treated, we can currently neither prevent nor cure it. On page 173 of this issue, however, Schenk *et al.*[1] show that, in a mouse model of Alzheimer's disease, immunization with Aβ inhibits the formation of amyloid plaques and the associated dystrophic neurites.

These results raise the possibility of vaccination with Aβ against human Alzheimer's disease. But before this can be seriously entertained, several questions must be answered. Schenk and colleagues[1] found that high levels of anti-human Aβ antibody were necessary for the effect to be seen in mice. So, can injection with human Aβ induce enough of the antibody? Will immune tolerance (whereby the immune system does not react against the body's own proteins) frustrate this, as it has often done with attempts to target cancers with antibodies? Conversely, is it safe to immunize people with high levels of a protein that is widely expressed outside the protective confines of the blood–brain barrier? When Schenk *et al.* immunized the mutant mice (known as PDAPP mice, because they over-

express a human APP transgene bearing the pathogenic valine-to-phenylalanine mutation at position 717) with human Aβ peptide, they did not see any autoimmune responses. But then, the human Aβ antigen induced much lower levels of antibodies to the endogenous mouse Aβ or APP.

The most critical question is whether depletion of the amyloid plaques is accompanied by an improvement in the behavioural/neurophysiological impairments,

and a reduction in the nerve cell death of Alzheimer's disease? In other words, does immunization with Aβ simply clear a neuropathological by-product or can it cure the disease? This question may be difficult to answer because all of the current animal models (based on overexpression of human APP and/or presenilin-1 transgenes bearing missense mutations associated with Alzheimer's disease) provide only a partial model of the human condition. So, although these animals accumulate increased levels of Aβ in the brain and have many amyloid-plaque deposits, they have only subtle behavioural and electrophysiological deficits. More problematically, these animals do not develop neurofibrillary tangles or show significant neurodegeneration (refs 2–4 and D. A. Westaway *et al.*, unpublished observations).

A second set of questions concerns the mechanism by which immunization with Aβ blocks the formation of amyloid plaques. Antibodies against Aβ might act as an artificial chaperone for extracellular Aβ, possibly by binding to Aβ and preventing it from aggregating or from changing into β-pleated-sheet conformation. Alternatively, these antibodies could accelerate clearance of Aβ from the central nervous system through one of several peripheral mechanisms (targeting Aβ for destruction by the peripheral reticuloendothelial system, for example, or reducing the production of Aβ in the periphery). Finally, Aβ might affect immune modulation of inflammatory mechanisms that are thought to be activated in Alzheimer's disease.
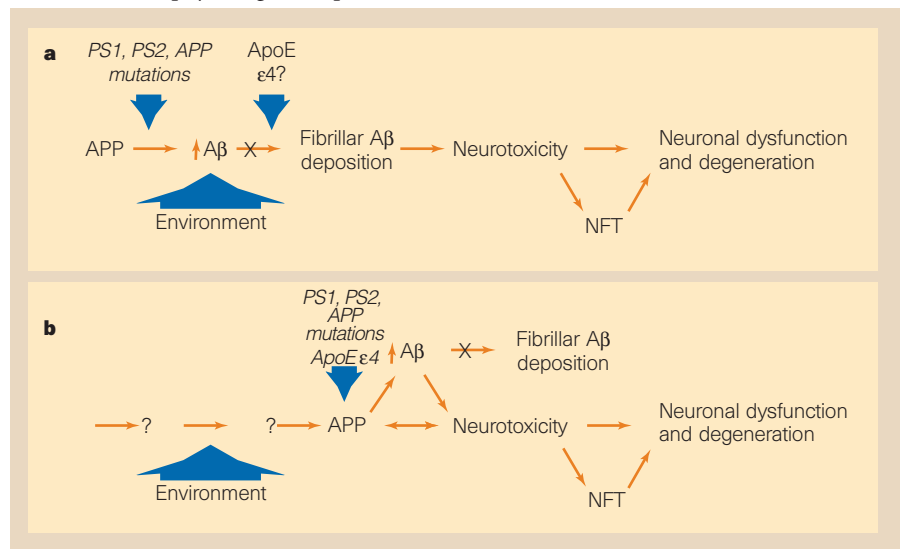

Figure 1 Controversies about the neuropathological alterations in Alzheimer's disease. a, In the 'amyloid cascade' hypothesis, generation of extracellular amyloid-β peptide (Aβ) from the β-amyloid precursor protein (APP) is thought to be central to the pathogenesis of Alzheimer's disease. b, An alternative point of view is that the extracellular plaques containing Aβ are an invariant but peripheral event. Large arrows depict putative entry points of environmental and genetic causes of Alzheimer's disease in these pathways. Genetic causes include mutation or polymorphisms in presenilin-1 (PS1), presenilin-2 (PS2), APP and Apolipoprotein E (ApoE). Schenk *et al.*[1] have found that, in a mouse model of Alzheimer's disease, immunization with Aβ presumably blocks (marked with an X) the extracellular deposition of this protein. (NFT, neurofibrillary tangles.)