nature

www.nature.com/nature

Vol 440 | Issue no. 7083 | 23 March 2006

Steering the future of computing

Computational power is surging thanks to insatiable consumers. Natural scientists should seize opportunities to stimulate computer science, to help everybody cope with huge volumes of data.

Sometime in the 2010s, if all goes well, the Large Synoptic Survey Telescope (LSST) will start to bring a vision of the heavens to Earth. Suspended between its vast mirrors will be a three-billion-pixel sensor array, which on a clear winter night will produce 30 terabytes of data. In less than a week this remarkable telescope will map the whole night sky with a greater speed and sensitivity than could have been imagined more than a decade or so ago, recapitulating with added detail the entire history of optical astronomy from Galileo to the Palomar Sky Survey.

And then the next week it will do the same again, looking for transient changes, adding new information and building up a database of billions of objects and millions of billions of bytes.

When looking at the future of scientific computing, as *Nature* does this week in a selection of News Features and Commentaries (starting on pages 398 and 409), it is easy to focus on the vast data architectures necessary for projects such as the LSST or the Large Hadron Collider at CERN, the European particle-physics laboratory near Geneva. The truly amazing story, though, is of the distributed power that ends up not in exceptional places such as the focal plane of a giant telescope, but spread out across the world; the power that allows data to be acquired from microfluidic chemistry sets and genome sequencers in labs around the world at astonishing rates, and allows the environment — or the human body — to be monitored in real time by vast arrays of sensors. The fact that everyday computing is getting exponentially cheaper promises to vastly increase data flows of all sorts and to revolutionize the practice of science.

It is this remarkable growth that has allowed projects such as the LSST to be imagined — and which will surpass them before they are very old. It is not driven by science, but it has been of immense use to scientists, and will continue to be, if they can change the way science is done to make use of the great potential.

Scientists will increasingly have to rely on automation to extract useful knowledge from these vast data resources. As with computer-aided proofs in mathematics, such automation challenges the processes by which scientists gain insight and generate theories. What's more, science will increasingly be done directly in the database, finding relationships among existing data while someone (or something) else performs the primary collecting role. And this means that scientists will have to understand computer science in much the same way as they previously had to understand mathematics, as a basic tool with which to do their jobs.

But scientists can be more than just passive responders to change. Although the great trends in computing are driven by economic and technical forces external to the scientific world, science can provide ideas and challenges that provoke the computer industry into moves it might not have made so quickly on its own. The World Wide Web, after all, grew out of the needs of scientific data users. It was years after Tim Berners-Lee had put his vision of hypertext onto the Internet that it revealed its capacity to revolutionize fields from bookselling to campaign financing.

> "Science can provide ideas and challenges that provoke the computer industry into moves it might not have made so quickly on its own."

The computer industry knows that scientists can come up with strange ideas and requirements that may well, in time, have broader commercial application elsewhere. This is one of the reasons why Microsoft is engaging the scientific community with its new 'Towards 2020 Science' report on computers in science (see http://research.microsoft.com/towards2020science). That report inspired this week's focus on computing in *Nature*. Microsoft is sponsoring free web access to our articles on the subject, although, as always, the content is exclusively *Nature*'s responsibility.

As computing gets ever cheaper, quicker and more powerful, scientists would do well to remember that, by being a demanding and stimulating 'user community' that engages the interest of companies such as Microsoft, Google and Intel, they can influence the development of the field, to everybody's benefit. ■

A scramble for Africa

Large dams benefit contractors and corrupt governments more than they aid the African people.

Towards the end of nineteenth century, Europe suddenly woke up to the riches that lay in the vast unexplored continent to its south, and the 'scramble for Africa' began. By the start of the First World War, almost all of the continent had been taken by European powers. The rights of Africa's own people, who lost land and many lives during this process, drew scant attention.

Why recall this episode today? Fleetingly, last summer, Africa was big news, when it became the central topic at a meeting in Scotland of the leaders of the G8 group of top industrialized nations, chaired by British prime minister Tony Blair. Yet the real action is being taken by a donor nation that isn't even a member of the G8: China.

The G8 nations — correctly, if belatedly — are considering how best to invest in Africa, so that the previous misappropriation and mismanagement can be avoided. China seems to have no such qualms. Across the continent, from Zimbabwe to Sudan, China is winning friends by lending money to Africa's most unsavoury regimes without asking awkward questions.

As a News story on page 393 of this issue illustrates, scientists and

engineers are sometimes complicit with this process. Sudan's Merowe dam on the Nile could be set to repeat the mistakes that have characterized previous large-scale hydropower projects in poor countries. Studies of how to resettle 50,000 people whose land will be flooded, and assessments of the project's environmental impact, were finished late in the day, and undertaken with insufficient rigour. They would have stopped the project going ahead if the World Bank, for example, was funding it.

Lahmeyer International, the German company that is coordinating the project, is disarmingly open about why this is so. It says that funders such as the World Bank make things too complicated. Thorough environmental and social impact assessments take years; Sudan wants power now. China is willing to invest, in part to cement closer ties with an important oil producer. And the Sudanese government lacks the political infrastructure — and probably the political will — to enforce proper safeguards. So, once again, thousands of poor people look set to suffer so that a big dam project can go ahead.

The project's backers have sought to portray Merowe as a necessary trade-off between the competing needs of development and the rights of local people. But there is no reason why both needs can't be met. Hydropower certainly has a role to play in Africa's development. Most of the continent's available hydropower resources are untapped and could, if properly harnessed, provide a valuable and renewable source of energy. But that doesn't mean that large dams need to be built. Successful projects in Asia and South America have shown that small hydropower projects can supply a few thousand local people without the need for big resettlement projects. Smaller projects can be run with more input from local people and are easier to combine with other renewable sources, such as solar power.

Unless the lessons of the past are thoroughly learned, large dam projects will sink over time in a morass of corruption, haphazard displacement of local people, lack of political accountability, and failure to plan properly for maintenance.

South Africa has, to its credit, tried to incorporate some of these lessons into a hydropower and water-supply project in Lesotho. The project is imperfect, but at least its administrators have sought to consult with local people and to run independent assessments of its environmental impact. But South Africa, with its wealth and its relatively sophisticated political system, is an exception.

"People's rights and needs are once again being sidelined in the stampede for wealth."

In many other African nations, there is little chance of proper safeguards being implemented. Chinese firms and government agencies will operate with few checks or balances. The same goes for the European companies involved in Merowe and elsewhere. They know that rigorous political consultation and environmental assessment are needed if big dam projects are to succeed. Yet they have been happy to engage in such projects in the absence of any such safeguards. The staff and shareholders of these firms are part of another scramble for Africa, in which local peoples' rights and needs are once again being sidelined in the stampede for wealth. ■

A colourful past

The production of dyes in the nineteenth century marked a turning point in the appliance of science.

The 150th anniversary of William Perkin's synthesis of aniline mauve dye (see page 429) is more than just an excuse to retell a favourite story from chemistry's past. To be sure, the tale contains much to delight in: Perkin's extraordinary youth and good fortune, the audacity of his gamble in setting up a business to mass-produce the dye, and the chromatic riches that so quickly flowed from an unpromising black residue of coal-gas production. But perhaps the most important aspect of the story is the relationship that it engendered between pure and applied science.

The demand for new, brighter and more colourfast synthetic dyes, along with new means of setting them on to fabrics ('mordanting'), stimulated manufacturing companies to set up their own research divisions, and thus cemented interactions between industry and academia that were just developing at the time.

Traditionally, dye-making was a craft, a combination of trial-and-error experimentation and the rote repetition of time-honoured recipes. The idea that chemical production required real scientific expertise did not arise until the eighteenth century, when the complexities of mordanting and multicolour fabric printing moved beyond the expertise of mere recipe-followers.

That was when the Scottish chemist William Cullen announced that if the mason wanted cement, the dyer a dye and the bleacher a bleach, "it is the chemical philosopher who must supply these". Making inorganic pigments preoccupied some of the greatest chemists of the early nineteenth century, notably Nicolas-Louis Vauquelin, Louis-Jacques Thénard and Humphry Davy. Perkin's mauve, however, was an organic compound and so, in the mid-nineteenth century, was rather more mysterious than metal salts. The drive to understand the molecular structure of carbon compounds during this period is often presented today as 'pure' chemistry, but in reality it owed much at the time to the profits that might ensue if the molecular secrets of organic colour could be unlocked.

"Dye companies, including Bayer, Ciba and Geigy, had seen the value of having highly skilled chemists on their payroll."

Both the need to understand molecular structure and the demand for synthetic methods were sharpened by chemists' attempts to synthesize indigo and alizarin (the natural colour obtained from the madder plant). When Carl Graebe and Carl Liebermann found a route to making alizarin in 1868, the Badische dye company, soon to become BASF, quickly acquired the rights. One of those who found a better route in 1869 was Ferdinand Riese, who was already working for Hoechst. Another was Perkin.

These and other dye companies, including Bayer, Ciba and Geigy, had seen the value of having highly skilled chemists on their payroll — something that was even more evident when they branched into pharmaceuticals early in the twentieth century. Thus, today's integration of scientific research into industry first began to take shape, as companies realized that good business needs good scientists. ■

CHAMPING AT THE BITS

Despite some remaining hurdles, the mind-bending and frankly weird world of quantum computers is surprisingly close. **Philip Ball** finds out how these unusual machines will earn their keep.

Five years ago, if you'd have asked anyone working in quantum computing how long it would take to make a genuinely useful machine, they'd probably have said it was too far off even to guess. But not any longer.

"A useful computer by 2020 is realistic," says Andrew Steane of the quantum-computing group at the University of Oxford, UK. David Deutsch, the Oxford physicist who more or less came up with the idea of quantum computation, agrees. Given recent theoretical advances, he is optimistic that a practical quantum computer "may well be achieved within the next decade".

This excitement is, however, tempered by the hurdles that have yet to be overcome. Building a quantum computer is still very, very hard to do. This is partly because it involves making quantum systems do things that don't come naturally to them. "There is progress, but it's still very slow," says physicist Chris Monroe of the University of Michigan in Ann Arbor.

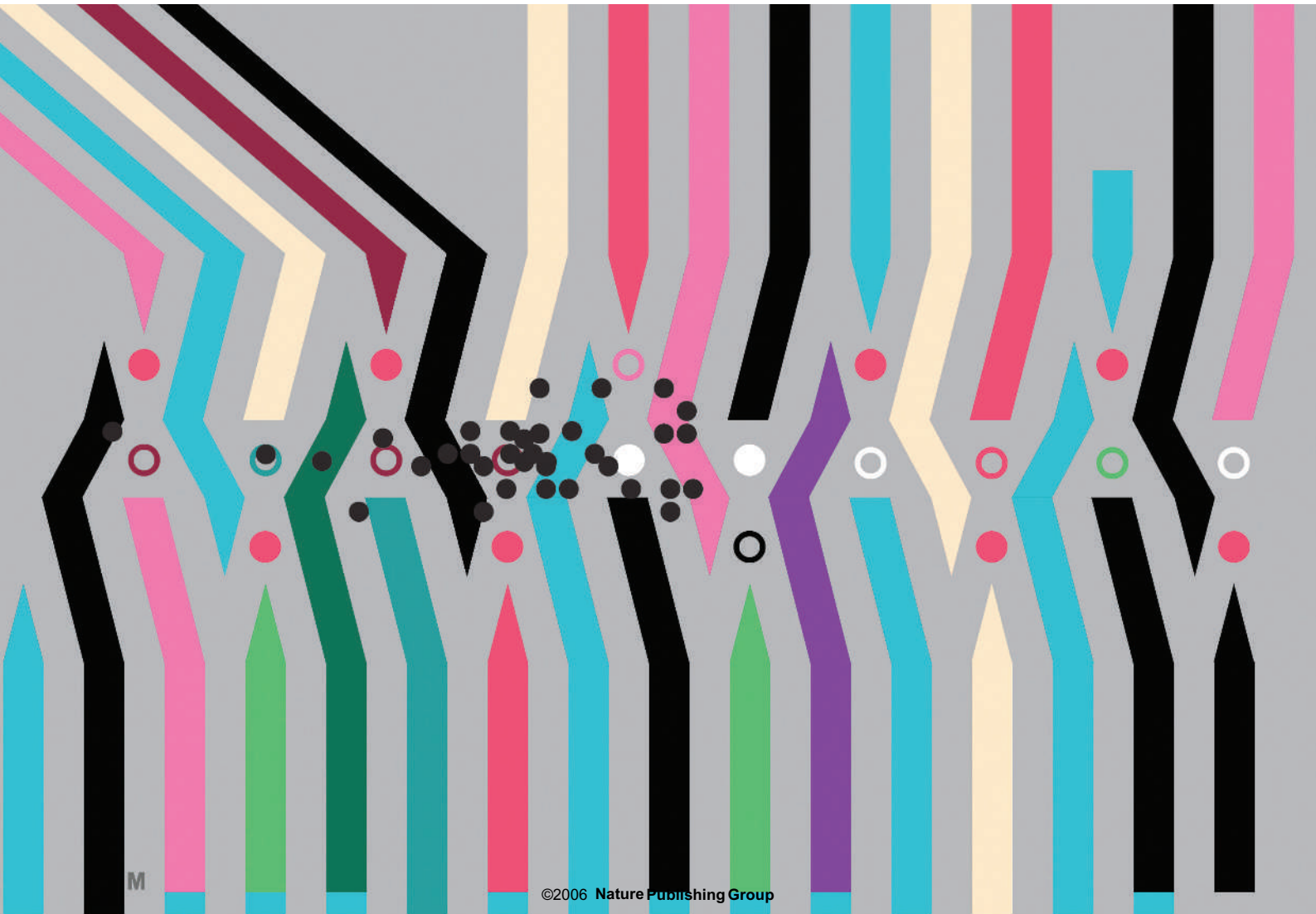
And even if we did have a working quantum computer today, there are hardly any programs that could run on it. In fact, it is likely that even once the machines are available, quantum computers are destined to remain niche products — excellent for certain tasks but not versatile devices like conventional personal computers. "Quantum computers will almost certainly never become general-purpose desktop machines," concedes Isaac Chuang, a quantum physicist at the Massachusetts Institute of Technology (MIT) in Cambridge.

Nevertheless, as a scientific research tool the quantum computer could be revolutionary because of its ability to simulate other quantum systems. In conventional, or classical, computers, information is stored as strings of bits: binary digits each of which can take the value of 0 or 1. The same is true for quantum computers, except that this time the binary digits — 'qubits' — are stored in the quantum states of microscopic systems, such as the electronic state of an atom or ion. So by its very nature, a quantum machine should be much better suited to simulating quantum systems than a classical computer.

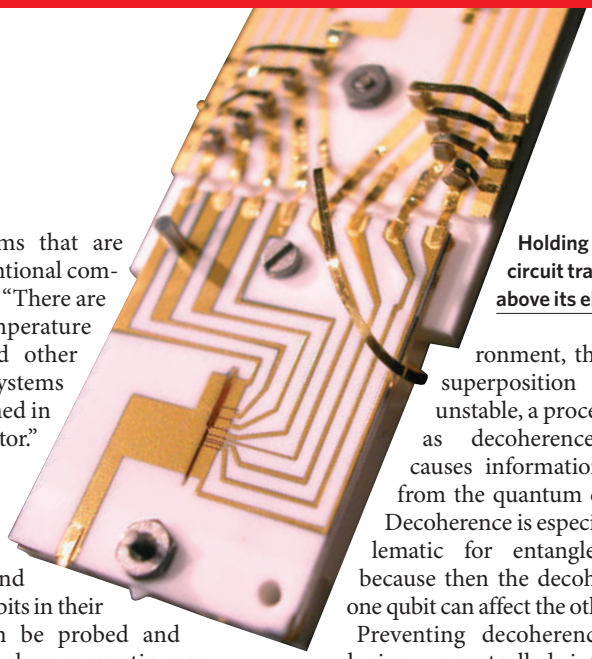
A quantum simulator would describe and predict the structure and reactivity of molecules and materials by accurately capturing their fundamental quantum nature. This is the sort of employment the early machines are likely to find: doing calculations of interest to chemists, materials scientists and possibly molecular biologists, says Steane.

"Just a few dozen qubits may shed light on

J. MAGEE



other physics problems that are intractable with conventional computers," notes Monroe. "There are models of high-temperature superconductivity and other condensed-matter systems that might be approached in such a quantum simulator."



M. BARRETT & J. JOST

Holding pen: this circuit traps ions above its electrodes.

In a spin

In fact, quantum simulations can already be done using atoms and molecules that store qubits in their nuclear spin and can be probed and manipulated using nuclear magnetic resonance (NMR) techniques. In their own terms, these 'computers' "run rings around any classical supercomputer", says Seth Lloyd, a theorist at MIT. He and his MIT colleague David Cory have been using this technique to simulate a variety of quantum systems in crystals of calcium fluoride and other materials. "As the crystal contains a billion billion spins, these simulations remain out of the reach of the most powerful classical computers," says Lloyd. The approach remains limited in terms of the different systems it can simulate, although Lloyd anticipates that fully functioning simulators will be readily available by 2020.

The key to the potential success of quantum computers is also the cause of the problems within the field: the quantum nature of data storage and manipulation. In classical computers, bits have clearly defined values of 1 or 0, but the laws of quantum mechanics allow qubits to exist in a 'superposition' of states — a mixture of both 1 and 0 that would be impossible in an everyday computer. This means that a quantum computer has much greater capacity for storing information.

A quantum processor can also compute with more than one qubit at once by exploiting another quantum property called entanglement, which makes qubits interdependent. The weird nature of the entangled state means that a measurement on one qubit instantly affects another, even though their previous individual states were undefined until that moment. Entangled states don't readily exist in nature: quantum engineers have to make them by allowing qubits to interact with one another.

By exploiting superpositions, a single quantum computer in effect mimics a whole suite of classical computers running at once, and by using entanglement these 'parallel computers' can be linked together. Unfortunately, this powerful parallel processor has an Achilles' heel. A quantum superposition has to remain stable for at least as long as it takes to do the computation. But as soon as qubits interact with their environment, the delicate superposition becomes unstable, a process known as decoherence, which causes information to leak from the quantum computer. Decoherence is especially problematic for entangled states, because then the decoherence of one qubit can affect the others too.

Preventing decoherence means reducing uncontrolled interactions with the environment. Cooling the quantum system to very low temperatures helps — but it may also be necessary to shield the qubits from stray electromagnetic fields. In practice, researchers have found it difficult to avoid decoherence of specific qubits for longer than a few seconds. But in principle it should be possible. "For qubits encoded in trapped ions, nobody really believes that we will ever be limited by coherence time," says Monroe.

Despite the fact that qubits need to be isolated from their environment to avoid decoherence, they must interact strongly with one another, to perform computations. And it must be possible for qubits in superposition to interact strongly with the environment when needed, so that the information can be read out. It is an extraordinarily delicate balancing act, which involves rules that defy intuition and aren't even completely understood.

An easy mistake to make

Decoherence also means that, as they process qubits using logic gates, quantum computers will inevitably incur errors at a much higher rate than classical computers. "The modern transistor has an error rate of less than 1 in 10^{14} or more switching events. In comparison, the best quantum gates we currently imagine will optimistically have an error rate of something like 1 in 10^{7}," says Chuang. Some researchers thought at first that this would make quantum computers too error-prone to be useful. But thanks to quantum error-correcting codes devised in the 1990s[1,2], it is now possible to correct error rates as high as 1 in 10^{5}.

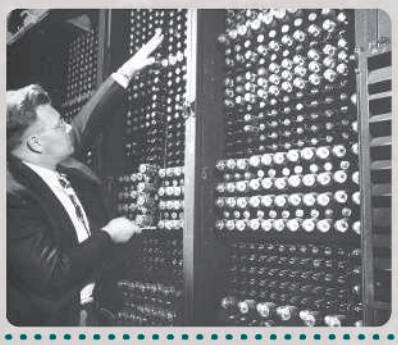
By 2002 the key principles behind a quantum computer had been sketched out by theorists (see 'How to build a quantum computer', overleaf), but how best to implement them in a real device remains a wide-open question. Much of the current effort is focused on making quantum computers using atoms or ions that are held in a trap. In an ion-trap computer, the qubits are encoded in the electronic states

MILESTONES IN SCIENTIFIC COMPUTING

PRE 1960s>>

>> PROTOTYPES...

1946 ENIAC, widely thought of as the first electronic digital computer, is formally unveiled. Designed to compute ballistics during the Second World War, it performs calculations in a variety of scientific fields including random-number studies, wind-tunnel design and weather prediction. Its first 24-hour forecast takes about 24 hours to do.



1951 Marvin Minsky, later of the Massachusetts Institute of Technology (MIT), builds SNARC, the first machine to mimic a network of neurons.

1954 John Backus and his team at IBM begin developing the scientific programming language Fortran.

1956 Building on earlier experiments at the University of Manchester, UK, and elsewhere, MANIAC at the Los Alamos National Laboratory in New Mexico becomes the first computer to play a full game of chess. In 1996, IBM's Deep Blue computer will defeat world chess champion Garry Kasparov.

1959 John Kendrew of the University of Cambridge, UK, uses computers to build an atomic model of myoglobin using crystallography data.

>>

of ions that are confined by an electromagnetic field. The ions interact with each other through electrostatic repulsion, and can be entangled by using laser beams to make them jiggle in unison. The quantum states of the ions can be read out by using other lasers to excite fluorescence, the wavelength of which depends on the ion's electronic state.

But the more qubits there are, the harder it is to read out their complex, collective vibrational states. One way to get round this is to hold most of the ions in a reservoir, and to perform each computational step using just a few of them, transferred from the reservoir to a processing chamber. This means that the ions have to be shuttled around without their quantum states being affected, so that they don't 'lose their memory' on the journey.

Charging ahead

Solving these problems is not easy, but recent progress has been encouraging. "Ion-trap chips look well placed to create useful computers before other methods," says Steane. Last December, for example, Monroe's team reported an ion trap built on a semiconductor chip using standard microfabrication techniques[3]. The trap held individual cadmium ions for more than an hour, while the researchers were able to move the ions smoothly between trapping sites.

David Wineland's group at the National Institute of Standards and Technology in Boulder, Colorado, is pursuing a similar idea in which the ions are trapped above electrodes etched into a chip's surface[4]. "Both methods have the advantage of using established fabrication techniques," says Wineland. "In the end, the winner might be determined simply by what is easier to fabricate."

Despite his success so far, Monroe is cautious about the long-term prospects. "Many groups are racing to build complex ion-trap chips," he says. "But it's less clear how trapped ions will ultimately compare with other quantum technologies."

Instead of ions, some researchers are encoding qubits using trapped neutral atoms. Atoms have the advantage that they interact more weakly with their environment than ions, but they also interact more weakly with each other. They can be trapped by laser beams — and by exploiting the interference pattern generated between crossed laser beams, hundreds of atoms can be held within an 'optical lattice', rather like an egg box. To make the atoms interact, the dimples in the egg box can be shifted closer together by adjusting the trapping beams.

One way to perform quantum computing with atoms is to create discrete clusters of entangled atoms in a larger lattice. This was first suggested in the late 1990s by Hans Briegel, Ignacio Cirac, Peter Zoller and their colleagues at the University of Innsbruck in Austria. It is an approach to quantum computations that Deutsch describes as "far easier to implement physically" than other methods for handling qubits.

Unlike the standard approach, cluster computation does not involve manipulating individual particles. Instead, before the computation is run, several qubits are brought together in a many-particle entangled state. The answer is then read out at the same time as the computation is actually performed, by making a series of measurements on each individual qubit in the cluster. Its 'one-step' nature makes this an appealing approach, but Cirac, who is now based at the Max Planck Institute for Quantum Optics in Garching, Germany, admits that it requires many more qubits than other methods, and that the error-correction procedures are more elaborate.

Join the dots

There is no shortage of other ideas for building a quantum computer. Some are based on superconducting devices, exploiting the fact that superconductivity is itself a quantum phenomenon. Unlike the systems based on individual particles, the qubits here are superconducting circuits, which hold many-particle quantum states of electrical charge or magnetic flux and can interact through classical electromagnetic forces.

Others hope to create optical quantum computers, encoding qubits into the quantum states of photons, or to make qubits from tiny specks of semiconducting material called quantum dots. "I've been particularly impressed by the advances made in quantum-dot systems and by the superconductor-based approaches," says Chuang. Compared with ion-traps, he explains, they may scale up more easily and are perhaps easier to connect to traditional telecommunication systems for readout.

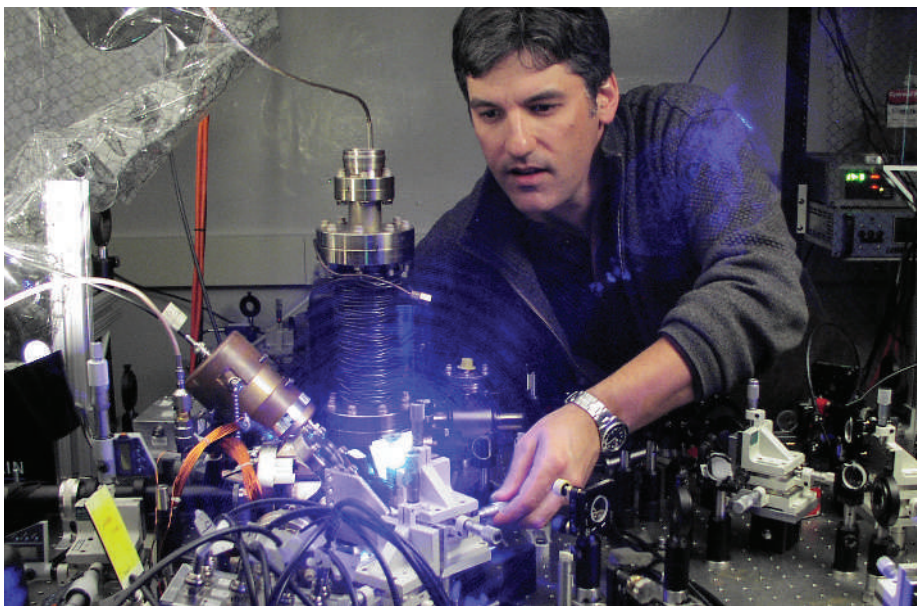
Quantum-dot systems may not produce the first useful computer, says Steane, but they have a naturally faster timescale — largely because the qubits are encoded in electrons, which are much lighter than ions — and so ultimately should outperform other systems such as ion-

How to build a quantum computer

The current US roadmap for the next decade of quantum computing (http://qist.lanl.gov) lists several requirements for a working machine[8]:

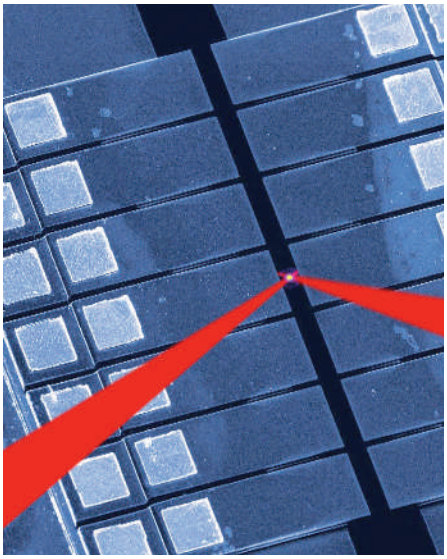
1. It must be scalable: it needs a set of qubits that can be added to indefinitely.
2. It must be possible to set all of the qubits to a simple initial state, such as all 0.
3. The interactions between qubits must be controllable enough to make quantum logic gates.
4. To perform operations using these gates, the decoherence times must be much longer than the gate-operation time (typically milliseconds to seconds).
5. There must be some readout capability.
6. To 'wire up' the computer's circuitry, it must be possible to convert memory qubits into processing qubits, and vice versa.
7. It must be possible to move processing qubits accurately between specified locations.

H. F. MONROE



Light touch: Chris Monroe aligns laser beams ready to trap ions for a quantum computation.

D. L. STICK



Caught in a trap: a single cadmium ion is held between two electrodes on a semiconductor chip.

trap chips. But, he adds, "the area is still sufficiently open that it is premature to slow efforts on any of the major contenders".

Because the hurdles in building the hardware are substantial, it is often suggested that the obstacles to making a quantum computer come down to engineering. But there is a bottleneck in theory too. So far, remarkably few specific computational problems have been translated into a form that a quantum machine could run and solve.

In fact, quantum computers are currently little more than two-trick wonders. In 1994, Peter Shor, now based at MIT, devised an algorithm that would allow quantum computers to factor numbers exponentially faster than conventional computers. Factorization is important in cryptography, where it is needed to make and break keys. And in 1996, Lov Grover, who is now at Lucent Technologies in Murray Hill, New Jersey, unveiled a quantum algorithm that can greatly speed up database searches.

Both of these algorithms have already been run using the NMR and optical techniques, but these methods are hard to scale up. At the end of last year, Monroe's group reported success with a Grover-type search using two cadmium ions in a trap[5]. Admittedly this meant looking through a database of just four entries — hardly a demanding task — but Monroe says the group plans to scale up to dozens of qubits over the next few years.

"It is striking that ten years have passed since Shor's invention, and very few new quantum algorithms have been developed," says Chuang. Among those that have appeared are methods for solving problems in number theory, drawn up by Sean Hallgren at NEC Laboratories in Princeton, New Jersey[6].

A major stumbling block for those trying to dream up new algorithms is that they first have to identify which problems will benefit most from quantum-computing methods. Theorist Michael Nielsen and his colleagues at the University of Queensland in Australia have recently made progress in this direction by showing that the general problem of finding quantum algorithms can be made easier by borrowing ideas from geometry[7].

In essence, the number of quantum operations, and thus the length of time, it takes to run an algorithm can be calculated by finding the shortest path between two points in a geometric space defined by all the possible sequences of quantum operations. "It really is a cool idea that has no classical analogue," says Lloyd. "It opens up a variety of methods for potentially creating new algorithms and for optimizing existing algorithms."

But not all quantum information processors will need complex algorithms. Many will be purpose-built tools that exploit quantum rules to improve on existing technologies such as atomic clocks and photonic technology. "We'll probably see rudimentary devices such as a 'quantum repeater' that converts photonic qubits to atomic qubits for error correction, and then back to photons to send them on their way down a long length of optical fibre," Monroe says.

> "Computers for specific applications are likely to come before general-purpose devices. But that doesn't rule out the possibility that we'll all be playing quantum *Grand Theft Auto* in the near future." — Seth Lloyd

If that seems a far cry from the quantum brains that are sometimes paraded as the next big thing, we may just have to get used to it. But Lloyd remains upbeat about the prospects. "I agree that quantum computers tailored for specific applications are likely to be built before general-purpose devices. But that doesn't rule out the possibility that we'll all be playing quantum *Grand Theft Auto* in the near future." ■

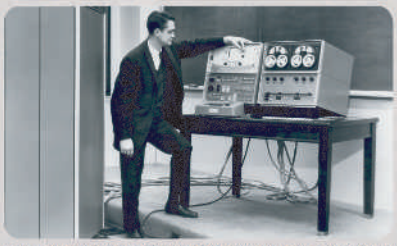
Philip Ball is a consultant editor for *Nature*.

1. Shor, P. W. *Phys. Rev. A* **52**, R2493–R2496 (1995).
2. Steane, A. M. *Phys. Rev. Lett.* **77**, 793–797 (1996).
3. Stick, D. *et al. Nature Phys.* **2**, 36–39 (2006).
4. Seidelin, S. *et al.* preprint at http://www.arxiv.org/quant-ph/0601173 (2006).
5. Brickman, K.-A. *et al. Phys. Rev. A* **72**, 050306(R) (2005).
6. Hallgren, S. in *Proc. 34th Annu. ACM Symp. Theor. Comput.* 653–658 (ACM, New York, 2002).
7. Nielsen, M. A., Dowling, M. R., Gu, M. & Doherty, A. C. *Science* **311**, 1133–1135 (2006).
8. DiVincenzo, D. P. *Fortschr. Phys.* **48**, 771–783 (2000).

>> MAINFRAMES . . .

1960s >>

1962 Charles Molnar and Wesley Clark at MIT's Lincoln Laboratory design the Laboratory Instrument Computer (LINC) for researchers at the National Institutes of Health. It is the first lab-based computer to process data in real time.



1963 In California, the Rancho Arm becomes the first artificial robot arm to be controlled by a computer.

1966 Cyrus Levinthal at MIT designs the first program to represent and interpret protein structures.

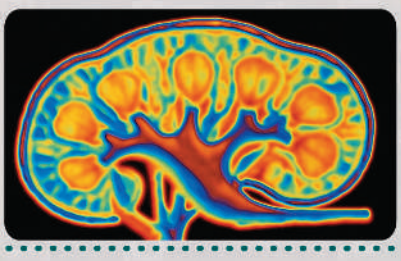
1967 ARPANET — the predecessor of the Internet — is proposed by the US Department of Defense for research networking.

1969 Results of the first coupled ocean–atmosphere general circulation model are published by Syukuro Manabe and Kirk Bryan, paving the way for later climate simulations that become a powerful tool in research on global warming.

>> WORKSTATIONS . . .

1970s >>

1971 Computing power shows its potential in medical imagery with a prototype of the first computerized tomography (CT) scanner.



1971 The Protein Data Bank is established at Brookhaven National Laboratory in Upton, New York.

1972 Hewlett Packard releases the HP-35, the first hand-held scientific calculator, rendering the engineer's slide rule obsolete.

>>

NATURE|Vol 440|23 March 2006

J. MAGEE

EVERYTHING, EVERYWHERE

Tiny computers that constantly monitor ecosystems, buildings and even human bodies could turn science on its head. **Declan Butler** investigates.

"What are you doing in the lab? Why aren't you out working in the field?" These are not the sorts of question you usually put to your computer. But they should be, according to the proponents of a new type of information technology sometimes known as 'smart dust'.

In their current, mostly desktop, incarnation, computers used for science usually come into their own quite late in the process of inquiry. Questions are asked, the data that might answer them identified, that data gathered — and only then does the computer start to play a role. In the future, this set up could be reversed. Computers could go from being back-office number-crunchers to field operatives. Twenty-four hours a day, year-in, year-out, they could measure every conceivable variable of an ecosystem or a human body, at whatever scale might be appropriate, from the nanometric to the continental.

These new computers would take the form of networks of sensors with data-processing and transmission facilities built in. Millions or billions of tiny computers — called 'motes', 'nodes' or 'pods' — would be embedded into the fabric of the real world. They would act in concert, sharing the data that each of them gathers so as to process them into meaningful digital representations of the world. Researchers could tap into these 'sensor webs' to ask new questions or test hypotheses. Even when the scientists were busy elsewhere, the webs would go on analysing events autonomously, modifying their behaviour to suit their changing experience of the world.

> "We will be getting real-time data from the physical world for the first time on a large scale."
> — Gaetano Borriello

If this scenario sounds far fetched, imagine the owner of a mainframe in the 1970s asking why it wasn't sitting on millions of desks and laps worldwide. An absurd question — to which the answer was "it's just a matter of time". The world's stock of computing power, and the number of devices over which it is distributed, has increased exponentially since then, as has the capacity of networking technology. These trends show no sign of slowing down, and that makes pervasive sensor nets not so much possible as inevitable. One does not need to be a visionary to see that soon, tiny devices with the power of today's desktops will be cheap enough to put everywhere.

Gaetano Borriello, a computer scientist at the University of Washington in Seattle, argues that such widely distributed computing power will trigger a paradigm shift as great as that

brought about by the development of experimental science itself. "We will be getting real-time data from the physical world for the first time on a large scale."

Instead of painstakingly collecting their own data, researchers will be able to mine up-to-the-minute databases on every aspect of the environment — the understanding of diseases, and the efficacy of treatments will be dissected by ceaselessly monitoring huge clinical populations. "It will be a very different way of thinking, sifting through the data to find patterns," says Borriello, who works on integrating medical sensors — such as continuous monitors of heart rate and blood oxygen — with their surroundings. "There will be a much more rapid cycle of hypothesis generation and testing than we have now."

Mallikarjun Shankar, who works on sensor webs for military and homeland security at Oak Ridge National Laboratory in Tennessee, agrees. "If one looks at the landscape of computing, this is where it will link with the physical world — where computing science hits the tangible, the palpable world. It is the next frontier."

From virtual to actual

Things don't get much more palpable than the experience of having an offshoot of Europe's biggest ice sheet grinding you into the granite below. That is the lot in life of the sensor web that Kirk Martinez, of the University of Southampton, UK, has been running for the past few years. He is helping glaciologists to study the dynamics of the Briksdalsbreen glacier in northwest Norway, part of the Jostedalsbreen ice field, in the hope of better understanding the impact of climate change and weather patterns on the ice field[1].

Martinez's team uses a hot-water 'drill' to place pods — a dozen at any one time — at different locations deep under the ice. Each pod is equipped with sensors that measure variables such as temperature, pressure, and movement; the data collected are used to work out the rate of flow of the glacier and to model subglacial dynamics. The sensor web can be programmed in such a way that the individual pods cooperate. "You can get the pods talking to each other, and deciding that nothing much has happened recently as most of our readings have been the same, so lets the rest of us go to sleep and save batteries, with one waking us up if something starts happening," says Martinez.

Watchful eyes: tiny computers called motes may one day be in everything.

"This is where where computing science hits the tangible, the palpable world. It is the next frontier."
— Mallikarjun Shankar

Martinez himself is a computer scientist, not a glaciologist. He was drawn to the task of remotely monitoring a hostile environment around the clock because he wanted the challenge of trying to bring together the various different technologies such sensor webs need. "This is very, very technologically tricky stuff," he says.

For non-computer scientists, it is even trickier. Researchers can already buy pods the size of cigarette packets or credit cards off the shelf from a slew of new companies such as CrossBow and Dust Networks (both based in the Bay Area of California). But that's only the start. At present, creating a sensor web for a specific scientific application requires extensive customization, particularly in the programming.

Katalin Szlavecz, an ecologist at Johns Hopkins University, in Baltimore, Maryland, works on soil biodiversity and nutrient cycling. She was driven to sensor webs by frustration caused by trying to solve complex problems with limited data. Soil is the most complex layer in land ecosystems, but it is poorly understood, because sampling is limited by the fact that technicians must collect soil samples by hand, and analyse them back in the laboratory. "Wireless sensor networks could revolutionize soil ecology by providing measurements at temporal and spatial granularities previously impossible," she says.

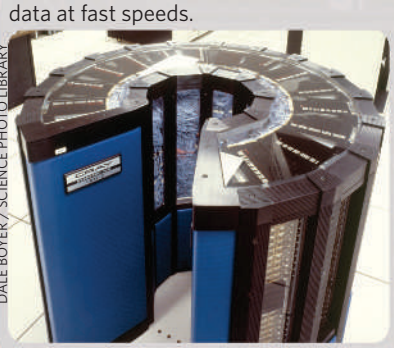
Data stream

Last September, Szlavecz deployed ten motes along the edge of a little stream just outside the university campus. Each mote measures soil moisture and temperature every minute, and the network transmits its data periodically to a computer in her office.

It sounds simple, but just to get the pods up-and-running she had to create a multidisplinary team, involving computer scientists, physicists and a small army of student programmers. Szlavecz says that she has "no doubt" that pod networks will take off widely, but that it won't happen until they are easier to deploy. And cheaper: "Each unit costs around US$300, but if you include all the hours of student time, each

>>

1976 At Los Alamos, Seymour Cray installs the first Cray supercomputer, which can process large amounts of data at fast speeds.



DALE BOYER / SCIENCE PHOTO LIBRARY

1980s>>

1983 Danny Hillis develops the Connection Machine, the first supercomputer to feature parallel processing. It is used for artificial intelligence and fluid-flow simulations.

1985 After receiving reports of a lack of high-end computing resources for academics, the US National Science Foundation establishes five national supercomputing centres.

1989 Tim Berners-Lee of the particle-physics laboratory CERN in Geneva develops the World Wide Web — to help physicists around the globe to collaborate on research.

MICHAEL QUAN/ZUMA PRESS

>> PERSONAL COMPUTERS...

1990s>>

1990 The widely used bioinformatics program Basic Local Alignment Search Tool (BLAST) is developed, enabling quick database searches for specific sequences of amino acids or base pairs.

1996 George Woltman combines disparate databases and launches the Great Internet Mersenne Prime Search. It has found nine of the largest known Mersenne prime numbers (of the form $2^n - 1$), including one that is 9,152,052 digits long.

>>

>> INTERNET...

KNUDSENS

Science of the future: researchers can keep a constant eye on the flow of a Norwegian glacier by tracking miniature sensors buried beneath the ice.

works out closer to $1,000." Despite this, Szlavecz says, the Microsoft-funded pilot was a success, revealing previously unobserved variations in soil microclimate, and showing how rain affects wetting and drying cycles[2]. (Szlavecz's colleague Alexander Szalay is an author of the commentary on page 413).

Handful of dust

Difficulties like Szlavecz's are all too common, says Kris Pister, founder and chief technology officer of Dust Networks. It was Pister who, at the University of California, Berkeley, coined the term smart dust to describe his vision of sensors smaller than the eye could see joined into networks larger than the mind could comprehend. Pister built prototype sensor webs with funding from the Defense Advanced Projects Research Agency, which is interested in the technology's military applications. But he says it was the desire to create more usable systems that led him to get the commercial backing to create Dust Networks. "It was very frustrating that while we could do spectacular demos, we couldn't give scientists something off-the-shelf, to put in a tree or a river. What people want is the ability to just put sensors out in the environment and get data back."

He likens the current stage of sensor-web development to the early days of computing. "There are a group of experts at the cutting edge of sensor webs, who have the time and expertise to go in and learn how to use the tools and do all the neat stuff," he says. "But for people who are not experts it has been difficult to get in and use it." He predicts, however, that just as the first web browser, Mosaic, and its successor, Netscape, sparked mass take-up of the World Wide Web, so future, more user-friendly sensor-web tools will generate interest.

Although Pister is interested in scientific applications, his key target is the lucrative industrial market for control systems. He has been contracted by the US Department of Energy to help build 'intelligent' self-monitoring lighting systems for factories, offices and homes; the 600 billion kilowatt-hours of lighting used for this purpose account for 30% of total building electricity use. "The next step is about really getting some standards and commercial adoption," he says. "That will drive cost down and performance up, and then scientific uses will take off."

Even with today's sensor-web technology, applications for research are proliferating. The Jet Propulsion Laboratory (JPL) in Pasadena, California, which is responsible for most of NASA's planetary science, has been running nine large sensor webs to study among other things, flooding, pollution and microclimates, in settings ranging from botanical gardens to the Sevilleta National Wildlife Refuge in central New Mexico. This month, Kevin Dellin, the head of the JPL sensor-web programme, spun it off into a company, Sensorware Systems.

> "We could do spectacular demonstrations but we couldn't give scientists something off-the-shelf, to put in a tree or a river."
> — Kris Pister

They mote be giants

But sensor webs currently have major limitations for people doing science in the field, says Deborah Estrin, director of the Center for Embedded Networked Sensing in Los Angeles, California, which operates a suite of land- and sea-based monitoring projects in collaboration with university groups. Estrin says that sensor webs alone are often not sufficient for all monitoring needs, and that the cost of sensors prohibits researchers from obtaining the pod densities often needed for detailed field experiments.

Estrin sees the sensor-web revolution as an important thread in a grander tapestry of global monitoring, which involves billions of dollars being poured into projects to monitor the continents and oceans. The US National Science Foundation's Ocean Observatories Initiative (OOI), for example, plans to spend $300 million over the next six years on gigabit 'backbones' — fibre-optic carriers of data — across the floor of the Pacific Ocean. On land, the planned National Ecological Observatory Network would enable research on terrestrial

ecosystems at regional to continental scales in real time. And underneath the surface, the EarthScope project would explore the four-dimensional structure of the North American continent from crust to core. Integrating local sensor webs and all these other networks is one of the biggest challenges facing the development of observational science, Estrin says.

Such mega-observatories may seem very different from Szlavecz's handful of little sensors alongside a stream in Baltimore. But the principles behind them are strikingly similar: to suck the best real-time data out of the environment as possible. "Instead of handling individual data files you will be handling continual streams of data", says Robert Detrick, chair of the OOI's executive steering committee. "You will be combining inputs from different sensors interactively to construct virtual observatories: sensors will be in everything."

The OOI's aim is to get around the fact that oceanographers tend to see only what their research vessels happen to be traversing at any given time. Checking in on the ocean fibre-optic backbone will be swarms of tiny autonomous submarines. These will carry sensors, go off on sampling missions and return to the backbone to recharge their batteries and upload data. "They can all communicate with one another acoustically," Detrick enthuses. "One can say, 'Hey, I've found an anomaly over here, so come on over'". Static sensors will monitor tectonic plates continuously along their entire length. Episodic events such as earthquakes, volcanic eruptions, and instabilities in ocean currents will be captured in real time, something that is impossible to do using ships.

Ecologist Katalin Szlavecz's mini computers tell her about minute-by-minute changes in soil moisture.

The existence of such large networks points to some major challenges down the line, says Estrin. Sensor webs will frequently be just single layers in a stack of data-collecting systems. These will extract information at different temporal and spatial scales, from satellite remote-sensing data down to *in situ* measurements.

Managing these stacks will require massive amounts of machine-to-machine communication, so a major challenge is to develop new standards and operating systems that will allow the various networks to understand each other. Sensors and networks of sensors will need to be able to communicate what their data are about, how they captured and calibrated them, who is allowed to see them, and how they should be presented differently to users with different needs. The lack of standards is not an insoluble problem for sensor webs, says Shankar "but it is slowing the field down by several years".

Catching the moment

Despite the difficulties, the use of sensor webs continues to grow. For all the trouble her first ten pods caused her, Szlavecz is upgrading to a network of 200. And experts see no fundamental obstacles to their eventual ubiquity. "We are well on the road to getting there, and I would argue that on many points we are already there," says Borriello. By 2020, says Estrin, researchers using visualization tools like Google Earth will be able to zoom in, not just on an old satellite image, but on "multiple *in situ* observations of the Earth in real time".

Data networks will have gone from being the repositories of science to its starting point. When researchers look back on the days when computers were thought of only as desktops and portables, our world may look as strange to them as their envisaged one does to us. Although we might imagine a science based so much on computing as being distanced from life's nitty gritty, future researchers may look back on today's world as the one that is more abstracted. To them the science practised now may, ironically, look like a sort of virtual reality, constrained by the artificialities of data selection and lab analysis: a science not yet ready to capture the essence of the real world. ■

Declan Butler is a senior reporter for *Nature* based in Paris.

1. Martinez, K., Riddoch, A., Hart, J. & Ong, R. in *Intelligent Spaces* (eds Stevenson, A. & Wright, S.) Ch. 9, 125–138 (Springer, 2006).
2. Musualoiu-E., R. *et al. Life Under your Feet: A Wireless Soil Ecology Sensor Network* (submitted) The Third IEEE Workshop on Embedded Networked Sensors (2006).

>> IMPLICIT COMPUTING...

>>

1996 Craig Venter develops the shotgun technique, which uses computers to piece together large fragments of DNA code and hastens the sequencing of the entire human genome.

1998 The first working quantum computers based on nuclear magnetic resonance are developed.

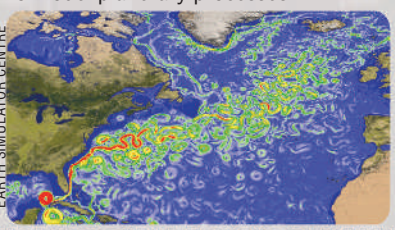
21st CENTURY>>

2001 The National Virtual Observatory project gets under way in the United States, developing methods for mining huge astronomical data sets.

SDSS

2001 The US National Institutes of Health launches the Biomedical Informatics Research Network (BIRN), a grid of supercomputers designed to let multiple institutions share data.

2002 The Earth Simulator supercomputer comes online in Japan, performing more than 35 trillion calculations each second in its quest to model planetary processes.



EARTH SIMULATOR CENTRE

2005 The IBM Blue Gene family of computers is expanded to include Blue Brain, an effort to model neural behaviour in the neocortex — the most complex part of the brain.

2007 CERN's Large Hadron Collider in Switzerland, the world's largest particle accelerator, is slated to come online. The flood of data it delivers will demand more processing power than ever before.

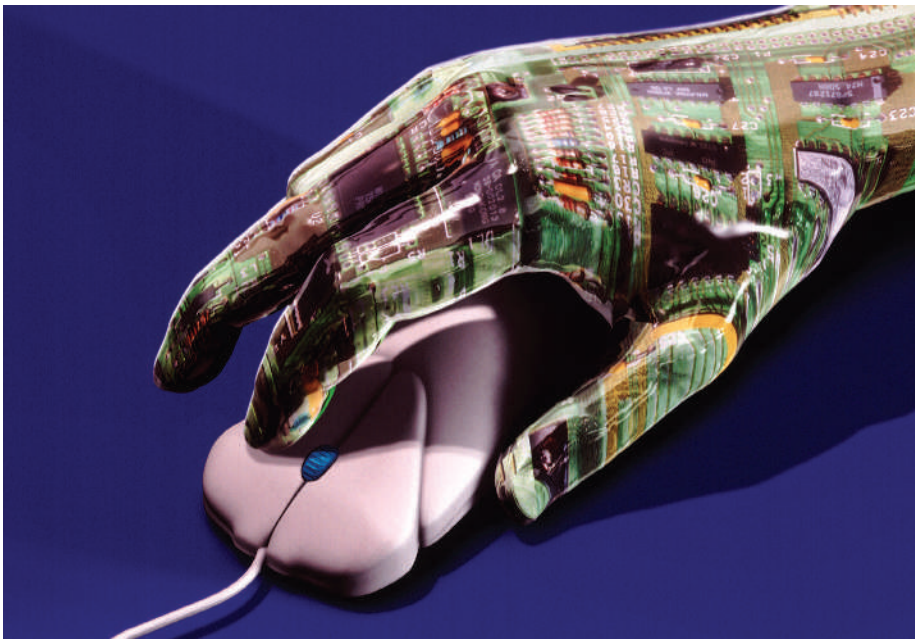
Jacqueline Ruttimann

COMMENTARY

Exceeding human limits

Scientists are turning to automated processes and technologies in a bid to cope with ever higher volumes of data. But automation offers so much more to the future of science than just data handling, says **Stephen H. Muggleton**.

The collection and curation of data throughout the sciences is becoming increasingly automated. For example, a single high-throughput experiment in biology can easily generate more than a gigabyte of data per day, and in astronomy automatic data collection leads to more than a terabyte of data per night. Throughout the sciences the volumes of archived data are increasing exponentially, supported not only by low-cost digital storage but also by the growing efficiency of automated instrumentation. It is clear that the future of science involves the expansion of automation in all its aspects: data collection, storage of information, hypothesis formation and experimentation (see table). Future advances will have the ability to yield powerful new forms of science that could blur the boundaries between theory and experiment. However, to reap the full benefits it is essential that developments in high-speed automation are not introduced at the expense of human understanding and insight.



FIREFLY PRODUCTIONS/CORBIS

Many aspects of science are already unsustainable without the aid of computers.

During the twenty-first century, it is clear that computers will continue to play an increasingly central role in supporting the testing, and even formulation, of scientific hypotheses. This traditionally human activity has already become unsustainable in many sciences without the aid of computers. This is not only because of the scale of the data involved but also because scientists are unable to conceptualize the breadth and depth of the relationships between relevant databases without computational support. The potential benefits to science of such computerization are high — knowledge derived from large-scale scientific data could well pave the way to new technologies, ranging from personalized medicines to methods for dealing with and avoiding climate change[1].

In the 1990s it took the international human genome project a decade to determine the sequence of a single human genome; but projected increases in the speed of gene sequencing imply that before 2050 it will be feasible to determine the complete genome of every individual human being on Earth. Owing to the scale and rate of data generation, computational models of scientific data now require automatic construction and modification. We are seeing a range of techniques from mathematics, statistics and computer science being used to create scientific models from empirical data in an increasingly automated way. For instance, in meteorology and epidemiology, large-scale empirical data are routinely used to check the predictions of differential-equation models concerning climate variation and the spread of diseases.

Meanwhile, machine-learning techniques from computer science (including neural nets and genetic algorithms) are being used to automate the generation of scientific hypotheses from data. Some of the more advanced forms of machine learning enable new hypotheses, in the form of logical rules and principles, to be extracted relative to predefined background knowledge. This background knowledge is formulated and revised by human scientists, who also judge the new hypotheses and may attempt to refute them experimentally. For example, within the past decade researchers in my group have used inductive logic programming (a subdiscipline of machine learning) to discover key molecular substructures within a class of potential cancer-producing agents[2]. Building on the same techniques, we have more recently been able to generate experimentally testable claims about the toxic properties of hydrazine from experimental data — in this instance, from analyses of metabolites in rat urine following low doses of the toxin[3].

CHANGES TO TRADITIONAL SCIENCE WITH AUTOMATION

Traditional science	Automated science
Hypotheses	Machine-encoded logical hypotheses
Chemical knowledge	Machine-encoded chemical algebra
Experiments	Chemical Turing machine programs
Experimental design	Decision theory

Mixing maths

In other sciences, the reliance on computational modelling has arguably moved to a new level. In systems biology, the need to account for complex interactions within cells — in gene transduction, signalling and metabolic pathways — is requiring new and richer systems-level modelling. Traditional reductionist approaches in this area concentrated on understanding the functions of individual genes in isolation. However, genome-wide instrumentation, including microarray technologies, are leading to a system-level

approach to biomolecules and pathways, and to the formulation and testing of models that describe the detailed behaviour of whole cells. This is new territory for the natural sciences and has resulted in multidisciplinary international projects such as the virtual E-Cell[4].

One obstacle to rapid progress in systems biology is the incompatibility of existing models. Often models that account for the shape and charge distribution of individual molecules need to be integrated with models describing the interdependency of chemical reactions. However, differences in the mathematical underpinnings of, say, differential equations, bayesian networks and logic programs make integrating these various models virtually impossible. Although hybrid models can be built by simply patching two models together, the underlying differences lead to unpredictable and error-prone behaviour when changes are made.

"Owing to the scale and rate of data generation, computational models of scientific data now require automatic construction and modification."

One encouraging development in this respect is the emergence within computer science of new formalisms[5] that integrate, in a sound fashion, two major branches of mathematics: mathematical logic and probability calculus. Mathematical logic provides a formal foundation for logic programming languages such as Prolog, whereas probability calculus provides the basic axioms of probability for statistical models, such as bayesian networks. The resulting 'probabilistic logic' is a formal language that supports statements of sound inference, such as 'The probability of A being true if B is true is 0.7'. Pure forms of existing probabilistic logic are unfortunately computationally intractable. However, an increasing number of research groups have developed machine-learning techniques that can handle tractable subsets of probabilistic logic[6]. Although it is early days, such research holds the promise of sound integration of scientific models from the statistical and computer-science communities

Miniature roboscientists

Statistical and machine-learning approaches to building and updating scientific models typically use 'open loop' systems with no direct link or feedback to the collection of data. A robot-scientist project in which I was involved offers an important exception[7]. Here, laboratory robots conducted experiments on yeast (*Saccharomyces cerevisiae*) using a process known as 'active learning'. The aim was to determine the function of several gene knockouts by varying the quantities of nutrient provided to the yeast. The robot used a form of inductive logic programming to select experiments that would discriminate between contending hypotheses. Feedback on each experiment was provided by data reporting yeast survival or death. The robot strategy that worked best (lowest cost for a given accuracy of prediction) not only outperformed two other automated strategies, based on cost and random-experiment selection, but also outperformed humans given the same task.

One exciting development that we might expect in the next ten years is the construction of the first microfluidic robot scientist, which would combine active learning and autonomous experimentation with microfluidic technology. Scientists can already build miniaturized laboratories on a chip using microfluidics[8] controlled and directed by a computer. Such chips contain miniature reaction chambers, ducts, gates, ionic pumps and reagent stores, and allow for chemical synthesis and testing at high speed. We can imagine miniaturizing our robot-scientist technology in this way, with the overall goal of reducing the experimental cycle time from hours to milliseconds. With microfluidic technology, each chemical reaction not only requires less time to complete, but also requires smaller quantities of input materials, with a higher expected yield. On such timescales it should become easier for scientists to reproduce new experiments and refute their hypotheses.

Today's generation of microfluidic machines is designed to carry out a specific series of chemical reactions, but further flexibility could be added to this tool kit by developing what one might call a 'chemical Turing machine'. The universal Turing machine, devised in 1936 by Alan Turing, was intended to mimic the pencil-and-paper operations of a mathematician. The chemical Turing machine would be a universal processor capable of performing a broad range of chemical operations on both the reagents available to it at the start and those chemicals it later generates. The machine would automatically prepare and test chemical compounds but it would also be programmable, thus allowing much the same flexibility as a real chemist has in the lab.

"There is a severe danger that increases in speed and volume of data generation could lead to decreases in comprehensibility."

One can think of a chemical Turing machine as an automaton connected to a conveyor belt containing a series of flasks: the automaton can move the conveyor to obtain distant flasks, and can mix and make tests on local flasks. Just as Turing's original machine later formed the theoretical basis of modern computation, so the programmability of a chemical Turing machine would allow a degree of flexibility far beyond the present robot-scientist experiments, including complex iterative behaviour. In the same way that modern-day Turing machines (computers) are constructed from integrated circuitry, thereby combining the power of many components, a universal robot scientist would be constructed from a mixture of microfluidic machines and integrated circuitry controllers.

Human touch

This microfluidic Turing machine is not only a good candidate for the next-generation robot scientist, it may also make a good model for simulating cellular metabolism. One can imagine an artificial cell based on a chemical Turing machine being used as an alternative to *in vivo* drug testing. The program running this machine would need to contain algorithms both for controlling the experiment and for conducting the cell simulation. It would represent a fundamental advance in the integration of computation with its environment.

Some may argue that in the context of biological experimentation, the series of chemical reactions is the computation itself. However, one can imagine taking the integration between experiment and environment even further. In particular, by connecting the input and output ducts of the microfluidic Turing machine to the chemical environment of a living cell, one could conduct experiments on cell function. Such levels of close integration between computers, scientific models and experimental materials are, however, still a decade or more away from becoming standard scientific practice.

Despite the potential benefits, there is a severe danger that increases in speed and volume of data generation in science could lead to decreases in comprehension of the results. Academic studies on the development of effective human–computer interfaces[9] emphasize the importance of cognitive compatibility in the form and quantity of information presented to human beings. This is particularly critical for technologies associated with hypothesis formation and experimentation. After all, science is an essentially human activity that requires clarity both in the statement of hypotheses and their clear and undeniable refutation through experimentation. ■

Stephen H. Muggleton is in the Department of Computing and the Centre for Integrative Systems Biology at Imperial College London SW7 2BZ, UK.

1. *Towards 2020 Science* (Microsoft, 2006) http://research.microsoft.com/towards2020science
2. Sternberg, M. J. E. & Muggleton, S. H. *QSAR Combinatorial Sci.* **22**, 527–532 (2003).
3. Tamaddoni-Nezhad, A., Kakas, A., Muggleton, S. H. & Pazos, F. in *Proc. 14th Int. Conf. Inductive Logic Programming* 305–322 (Springer, 2004).
4. Takahashi, K. *et al. Bioinformatics* **13**, 1727–1729 (2003).
5. Halpern, J. Y. *Artif. Intell.* **46**, 311–350 (1990).
6. De Raedt, L. & Kersting, K. in *Proc. 15th Int. Conf. Algorithmic Learning Theory* 19–34 (Springer, 2004).
7. King, R. D. *et al. Nature* **427**, 247–252 (2004).
8. Fletcher, P., Haswell, S., Watts, P. & Zhang, X. Lab-on-a-Chip Micro Reactors for Chemical Synthesis. *Dekker Encyclopedia of Nanoscience and Nanotechnology* (2001).
9. Jacko, J. A. & Sears, A. *The Human–Computer Interaction Handbook* (Lawrence Eribaum Associates, 2003).

The creativity machine

What will emerge from using the Internet as a research tool? The answer, **Vernor Vinge** argues, will be limited only by our imaginations.

We humans have built a creativity machine. It's the sum of three things: a few hundred million computers, a communication system connecting those computers, and some millions of human beings using those computers and communications.

This creativity machine is the Internet. It has already changed the way we do science, most importantly by enhancing collaboration between researchers. The present-day Internet provides convenient connections between computerized labs, simulations and research databases. It also represents an enormous financial investment that is driven by the demands of hundreds of millions of consumers. As such, the total Internet software and infrastructure investment dwarfs the budgets of scientific research programmes and even of many government defence programmes. And more than any megaproject of the past, the essence of the Internet is to provide coordinated processing of information. For researchers seeking resources, these are facts worth considering.

For some disciplines, the Internet itself has become a research tool: grid computing has been used to exploit the power of millions of Internet-connected machines. Building on the popularity of SETI@home — an experiment that uses Internet-connected computers to search for extraterrestrial intelligence — and prime-number hunts, there are now physics, medical and proteomics projects enlisting the enthusiasm of people (and their computers) across the world. For linguists[1] and sociologists, new questions can be investigated simply by observing what occurs on the publicly available Internet. Even experimental sociology is possible: in their study of social influence on music preference, Salganik *et al.*[2] recruited more than 14,000 subjects through a popular website, ran online trials on these subjects, and then obtained results directly from their experiment website.

The possibilities do not end there. Even online games are attracting academic interest. Some games have millions of players. MMORPGs (massively multiplayer online role-playing games), such as World of Warcraft and EverQuest, feature vivid three-dimensional action involving both cooperation and combat. Another genre of MMORPGs lack a significant combat or quest element and are more often called 'virtual worlds'. For example, the virtual world Second Life has the visual realism of many MMORPGs, but it exists as a venue for the participants rather than as a predesigned adventure. Second Life provides a range of software tools, including a programming language, that gives participants the power to create artefacts according to their own designs. Thus the game depends on the skill and creativity of its participants to generate content. Such virtual worlds have already been used for educational projects, and are worthy of psychological and social research.

People power

The notion of enlisting users to create content is widespread on the contemporary Internet. Companies such as Google provide users with tools to integrate search and mapping services into their own websites. Interested users are numerous and have their own resources. In the 1990s, we had an early glimpse of the power of this new creativity machine: computers plus networks plus interested people delivered free and open-source software (FOSS) products of the highest quality, including the GNU/Linux operating system. FOSS products provide low-cost and flexible alternatives to proprietary software. For example, there is at least one open-source virtual-world platform, Croquet[3], which allows users to customize and extend its architecture at all levels. FOSS tools can be mixed and matched with proprietary software to deal with an enormous range of projects from quick, *ad hoc* combinations of data harvested from multiple locations[4] to large, long-duration experiments.

Participants in Second Life use software and creativity to build their environment.

All this points to ways that science might exploit the Internet in the near future. Beyond that, we know that hardware will continue to improve. In 15 years, we are likely to have processing power that is 1,000 times greater than today, and an even larger increase in the number of network-connected devices (such as tiny sensors and effectors). Among other things, these improvements will add a layer of networking beneath what we have today, to create a world come alive with trillions of tiny devices that know what they are, where they are and how to communicate with their near neighbours, and thus, with anything in the world. Much of the planetary sensing that is part of the scientific enterprise will be implicit in this new digital Gaia. The Internet will have leaked out, to become coincident with Earth.

How can we prepare for such a future? Perhaps that is the most important research project for our creativity machine. We need to exploit the growing sensor/effector layer to make the world itself a real-time database. In the social, human layers of the Internet, we need to devise and experiment with large-scale architectures for collaboration. We need linguists and artificial-intelligence researchers to extend the capabilities of search engines and social networks to produce services that can bridge barriers created by technical jargon and forge links between unrelated specialties, bringing research groups with complementary problems and solutions together — even when those groups have not noticed the possibility of collaboration. In the end, computers plus networks plus people add up to something significantly greater than the parts. The ensemble eventually grows beyond human creativity. To become what? We can't know until we get there. ■

Vernor Vinge is an emeritus professor of computer science at San Diego State University. His novel *Rainbows End* (2006) considers the Internet of 2025.

1. Kilgarriff, A. & Grefenstette, G. *Comput. Linguist.* **29,** 333–347 (2003).
2. Salganik, M. J., Dodds, P. S. & Watts, D. W. *Science* **311,** 854–856 (2006).
3. Croquet Project www.opencroquet.org
4. Butler, D. *Avian Flu Maps in Google Earth* http://declanbutler.info/blog/?p=16

Science in an exponential world

The amount of scientific data is doubling every year. **Alexander Szalay** and **Jim Gray** analyse how scientific methods are evolving from paper notebooks to huge online databases.

Scientists are trained early to keep careful records in their laboratory notebooks — recording both experimental procedures and observations, so that they can analyse their results and so that others can replicate what they have done. Galileo did it, Mendel did it, Darwin did it, and we are supposed to do it. This worked fine when small amounts of data were entered into notebooks and the analysis was computed alongside them. But data volumes are doubling every year in most areas of modern science and the analysis is becoming more and more complex, exceeding the capacity of the paper notebook. With data correlated over many dimensions and millions of points, none of the old steps — do experiment, record results, analyse and publish — is straightforward. Many predict dramatic changes to the way science is done, and suspect that few traditional processes will survive in their current form by 2020 (ref. 1).

Today, most scientists have replaced or enhanced their notebooks with desktop computers that record their results, provide a portal to the scientific literature, and link them to collaborators via e-mail. These computers also perform data analysis; Matlab, Mathematica and Excel are popular analysis tools. But none of these programs scale up to handle millions of data records — and they are primitive by most standards. As data volumes grow, it is increasingly arduous to extract knowledge. Scientists must labour to organize, sort and reduce the data, with each analysis step producing smaller data sets that eventually lead to the big picture. Analysing terabytes of data (one terabyte is 1,000 gigabytes) is a challenge; but petabyte data sets (of more than 1,000 terabytes) are on the horizon. One petabyte is equivalent to the text in one billion books, yet many scientific instruments, including the Large Synoptic Survey Telescope, will soon be generating several petabytes annually.

In response to this data deluge, the systematic use of databases has become an integral part of the scientific process. Databases provide tools to organize large data sets, find objects that match certain criteria, compute statistics about the data, and analyse them to find patterns. Many experiments today load their data into databases before attempting to analyse them. But there are few tools to properly visualize data across multiple scales and data sets. If we can no longer examine all the data on a single piece of paper, how can we 'see' a new pattern or find a data point that does not fit a hypothesis? Fortunately there are database tools, such as data cubes, that we believe can fulfil this role (see 'Data cubes' overleaf).

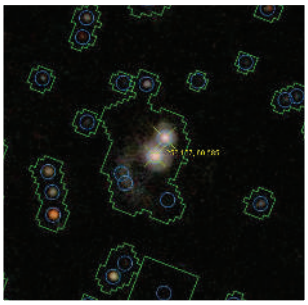
The same language

Experiments are themselves becoming electronic as computers become essential parts of scientific instruments; they are used not only to manage and analyse vast data sets, but also to acquire them in the first place. Procedures already involve instruments and software with myriad parameters. It is difficult to capture all the model numbers, software revisions, parameter settings and process steps in an enduring format. For example, imagine a measurement taken using a DNA-sequencing machine. The output is cross-correlated with a sequence archive (GenBank) and the results are analysed with Matlab. Fully documenting these steps would be arduous, and there is little chance that someone could repeat the exact procedure 20 years from now; both Matlab and GenBank will change enormously in that time. As experiments yield more data, and analysis becomes more complex, data become increasingly difficult to document and reproduce.

One might argue that complex biological experiments have always been difficult to reproduce, as there are so many variables. But we believe that with current trends it is nearly impossible to reproduce experiments. We do not have a solution for this problem, but it is important to recognize it as such, and to do what is possible to capture the workflows and to develop protocols for documenting instruments, procedures and measurements in ways that will be usable in several decades' time.

Increasingly, scientists are analysing complex systems that require data to be combined from several groups and even several disciplines. There are collaborations sharing data across departments and time zones, and important discoveries are made by scientists and teams who combine different skill sets — not just biologists, physicists and chemists, but also computer scientists, statisticians and data-visualization experts. It is important to realize that today's graduate students need formal training in areas beyond their central discipline: they need to know some data management, computational concepts and statistical techniques.

A collaboration involving hundreds of Inter-



A. SZALAY & J. GRAY

JOHNS HOPKINS UNIV.

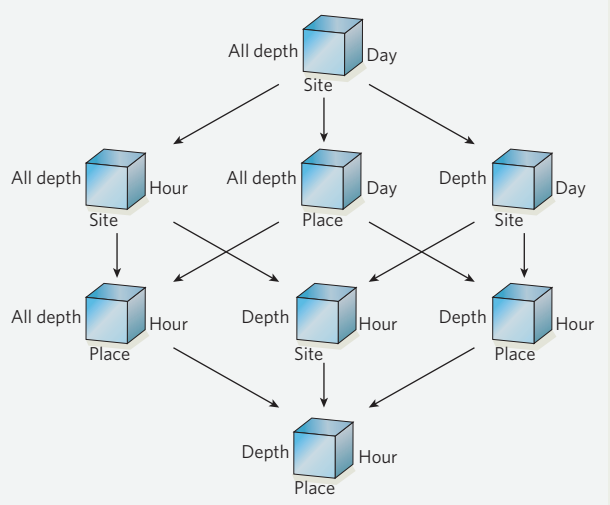
Automated systems will transform data collections, from astronomy (left) to sampling soil properties under our feet.

DATA CUBES

Traditional notebook and analysis tools are being challenged not just by data volumes, but also by data complexity. For example, the three-dimensional structural representation of a complex protein is not easily transcribed into a notebook.

Complex scientific data are often organized as a collection of independent variables and their dependent measurements. For example, meteorological data (temperature, pressure, humidity, wind velocity and direction) are collected at various times and locations (latitude, longitude and altitude). These data can be thought of as a multidimensional cube in which time and place exist in four dimensions and the measurements are shown by vectors at each point in the cube (pictured right).

A meteorologist may ask: show me the minima, maxima and average winds for Australia aggregating over times (hour, day, week, year) and volumes (per square kilometre of the atmosphere). The data cube makes it easy to express such user queries and to compute the answers, even to seemingly complex questions.



C. STOLTE & P. HANRAHAN, STANFORD UNIV./TABLEAU

net-connected scientists raises questions about standards for data sharing. Too much effort is wasted on converting from one proprietary data format to another. Standards are essential at several levels: in formatting, so that data written by one group can be easily read and understood by others; in semantics, so that a term used by one group can be translated (often automatically) by another without its meaning being distorted; and in workflows, so that analysis steps can be executed across the Internet and reproduced by others at a later date.

Standards for sharing data are crucial, for example, in understanding soil ecosystems. We are helping to build a system for measuring long-term environmental trends that affect soil biodiversity (www.lifeunderyourfeet.org; see also News Feature, page 402). This system integrates local environmental data from a sensor network with regional data on hydrology, climate, biodiversity and biogeochemistry. For these data to be useful to others, they must be published using a controlled vocabulary and in standard forms, and the instruments and measurements must be well specified. Fully documenting the sensors and data-collection process is arduous, and there are few standards for us to draw on.

Data gold-mine

Multidisciplinary databases also provide a rich environment for performing science; that is, a scientist may collect new data, combine them with data from other archives, and ultimately deposit the summary data back into a common archive. Many scientists no longer 'do' experiments the old-fashioned way. Instead they 'mine' available databases, looking for new patterns and discoveries, without ever picking up a pipette.

But this data-rich approach to science faces challenges. The speed of the Internet has not kept pace with the growth of scientific data sets. And so large data archives are becoming increasingly 'isolated' in the network sense — one can copy gigabytes across the Internet today, but not petabytes. In the future, working with large data sets will typically mean sending computations to the data, rather than copying the data to your workstation. But the management of distributed computations raises new questions of security, free access to public data and cost. Few data archives address these issues today.

Are we reaching the limits of what one scientist, or one lab, can expect to achieve in data handling and analysis? If so, this will have implications for how we review and publish our work. For example, a data-mining paper needs to include the explicit description (database query) of how the data that were analysed in the paper were collected and filtered, but not the data themselves. In this way, a reviewer with access to public data could reproduce the data sets and analysis procedures. For the analysis to be repeatable in 20 years' time requires archiving both data and tools.

The publication process itself is increasingly electronic, with new ways to disseminate scientific information (such as the preprint repository arXiv.org). But there is, as yet, no standard for publishing large volumes of data. Paper appendices cannot hold all the data needed to reproduce the results. Some disciplines have created their own data archives, such as GenBank; others just let data show up, and then disappear, on individual scientists' websites. Astronomers created the International Virtual Observatory Alliance (www.IVOA.net), integrating most of the world's medium and large astronomy archives. This required new standards for data exchange, and a semantic dictionary that offers a controlled vocabulary of astronomy terms.

To encourage data sharing, it should be rewarded. Public data creators and publishers should be given credit, and archives must be able to automatically provide provenance details. Current databases have a long way to go to achieve this ideal.

For how long will this exponential growth in scientific data continue? Desktop computers today are as powerful as the super-computers of 10 years ago. Similar progress is happening with scientific instruments — they quickly become obsolete and are replaced by better and often cheaper ones. Likely computer-performance improvements by 2011 include tenfold more processing, storage and network bandwidth per dollar. So we can expect ten times more data.

Smaller is faster

However, not all experiments will experience exponential growth. There is reason to believe that it will be the smaller experiments, not the big multibillion-dollar facilities, that will grow the fastest. Exponential growth occurs when a new generation of instruments leapfrogs the previous generation, which become obsolete. There are two trends in science today, scaling up and scaling out. Some scientists are building billion-dollar facilities, such as astronomy's Large Synoptic Survey Telescope or the Large Hadron Collider, which are only affordable as international collaborations. Such facilities are not easily leapfrogged. And once these peta-scale experiments are switched on they will produce roughly the same amount of data each year — merely linear growth. But in the scaling-out model, experiments that deploy an array of small instruments can exploit the coming explosion in cheaper commodity technology. The wireless sensors that were US$300 a year ago are $100 today, and will be $30 next year. A similar phenomenon occurred with DNA chips and gene sequencers. It is important to recognize this pattern; it is universal. And so although some sub-disciplines may reach a plateau in data generation, other technological innovations will take their place. Scientists in 2020 will continue to work in an exponential world. ■

Alexander Szalay is in the Department of Physics and Astronomy, Johns Hopkins University, Baltimore, Maryland 21218, USA.
Jim Gray is at Microsoft Research, San Francisco, California 94105, USA.

1. *Towards 2020 Science* (Microsoft, 2006); http://research.microsoft.com/towards2020science

Can computers help to explain biology?

The road leading from computer formalisms to explaining biological function will be difficult, but **Roger Brent** and **Jehoshua Bruck** suggest three hopeful paths that could take us closer to this goal.

It doesn't require vast prophetic vision to identify developments in computers and information technology that will greatly affect the practice of biology. By 2020 we expect that biologists will use computers, numerous 'omic' data types (ref. 1) and a greatly expanded biological literature to design experiments, generate and analyse new data, and think about their own work. But we will leave forecasting about PubMed and Google, metadata and the semantic web to others. Instead, we wish to consider some of the formalisms offered by computer science[2] that developed alongside computing machines. The search for biologically relevant formalisms has a chance to greatly affect the understanding of biological function, in ways we are just starting to imagine.

Today, by contrast with descriptions of the physical world, the understanding of biological systems is most often represented by natural-language stories codified in natural-language papers and textbooks. This level of understanding is adequate for many purposes (including medicine and agriculture) and is being extended by contemporary biologists with great panache. But insofar as biologists wish to attain deeper understanding (for example, to predict the quantitative behaviour of biological systems), they will need to produce biological knowledge and operate on it in ways that natural language does not allow.

Living computers

We begin with what we know in 2006: the trajectory of living systems through developmental time and space is highly determined by the actions and interactions of functional molecules encoded by their genomes. These encoded molecules are further influenced by external perturbations. Because the dynamic behaviour of biological systems is highly determined by a central stored program, living systems differ profoundly from all other naturally occurring, time-evolving systems. The weather has no genome.

Some aspects of biological systems, such as the sequence of encoded proteins (which determines their structure), arise directly from the genome. But others, including most biological functions, arise from the genome by considerably more complex routes, with the consequence that function typically occurs simultaneously at multiple levels[3,4]. These levels include the biochemical activity of an individual protein, the function of that protein in cellular processes involving other proteins, and the developmental trajectory of those processes within a multicellular organism[4,5]. None of these levels is more true or fundamental than the other.

If biology and information science continue with business as usual, then, by 2020, most of the natural-language stories of 2006 about biological function will be subsumed into more sophisticated narratives, which will be better organized and accessed by computers. But the outlines of most of these stories will probably remain unchanged. Here, however, we imagine ways that formalisms from computer science might contribute to a deeper understanding of biological function. These approaches will not bear fruit without deliberate and difficult work.

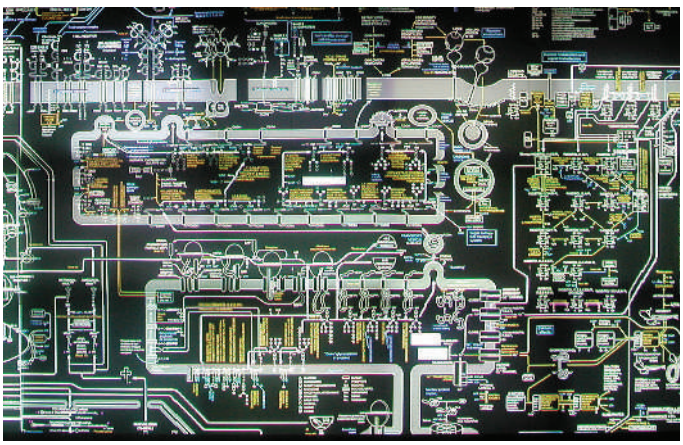
The development of computer science required both new formalisms to capture reasoning in natural languages and ways to implement those formalisms in physical devices[6]. In 2006, it seems reasonable to compare living systems to 'von Neumann', or stored-program, computers, with processing systems (here encoded by the stored program), various external and internal inputs, and outputs in the form of execution. In this view, the biological system is not primarily a factory or a chemical plant but an assemblage that takes information, processes it, decides and executes.

As yet there is no theory that can specify the meaning or purpose of a string of computer code, but Sussman has suggested how elements of such theory might arise[7]. He points out that mathematics had its roots in a workaday human activity, that of Egyptian surveyors redefining the boundaries of fields after the Nile floods receded. Rigorous thinking about this activity led eventually to mathematics: geometry, trigonometry, algebra and beyond. In the same way, a workaday activity — the design and use of procedural imperative languages to write code ('do this, now, do that, if such a thing happens, then do this!') to program computers — may lead to new formalisms describing information processing and eventually to new mathematics.

Blurred boundaries

In biological systems it seems reasonable to view the DNA script in the genome as executable code, code that could have been specified by a set of commands in a procedural imperative language. And in the same spirit, we can view any signal-transduction pathway as a collection of protein machines that takes inputs from inside and outside the cell, performs processing operations on those inputs to arrive at decisions, and communicates those decisions to an apparatus that executes it.

However, to make the analogy between biological systems and von Neumann computers is to reveal important differences between them (Fig. 1). At the level of cells and organisms, biological systems differ from computers in many ways, including (but not limited to): lack of modularity and boundaries in code; lack of fixed order of execution in code; self-



ARGONNE NATL UNIV.

Biology needs to move beyond natural-language descriptions of biomolecules and pathways. Adopting new formalisms from computing may lead to greater insight than even graphical displays (such as this wall chart) can offer.

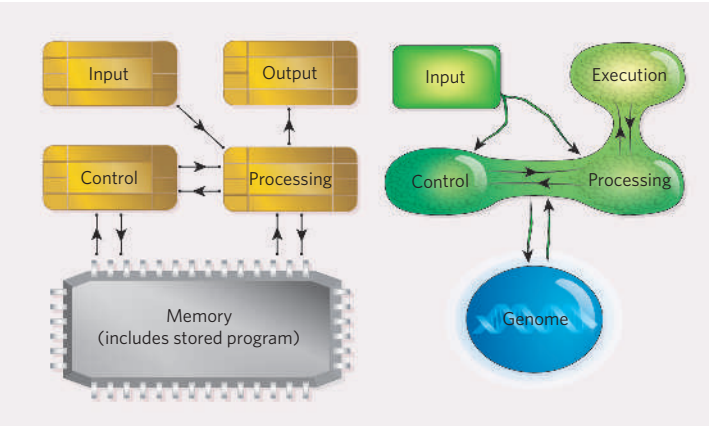


Figure 1 | Biological systems (right) have similarities and differences to von Neumann computers (left). In biological systems there is no distinction between processor and output, as function, phenotype and selection act at many levels.

assembly of encoded components; lack of intelligible sentient design; and lack of crisp boundaries between memory, processor, input and output components.

Most importantly, biological systems usually lack a clear boundary between processing apparatus and output. This distinction arises because function in biology is a consequence of selection, and selection usually acts at many different levels. Thus, the simplest human question 'what does the system do?' (which translates into 'what was the system selected for?') usually has simultaneous multiple correct answers. This fact will continue to frustrate analyses of biological systems in terms of the 'objective functions' they are 'optimized' to 'execute'. Because of this difference, it is unlikely that even a mature theory of stored-program machines will be adequate to explain biological systems. However, even in the absence of grand theory, one can work on intermediate steps. Here we describe three avenues worth exploring.

One fruitful approach formalizes cause-and-effect relationships between named proteins and regulatory sites by translating these into defined chemical reactions undergone by defined molecular species. These reactions can be modelled as differential equations constrained by the rules of chemical kinetics, more formally codified as the 'chemical master equation'[8]. In biology, differential-equation models have a mixed history; they were vital for understanding transmission of the nerve impulse[9] and for helping to identify reaction types before the channel molecules were discovered, but were less successful in circadian-rhythm research until biologists identified molecular entities and relevant reactions.

Measure of meaning

For most biological narratives, the resulting sets of differential equations are too complex to be analytically tractable. But their dynamic behaviour can be approached down a second path — by simulating approximate numerical and stochastic methods[10]. These simulations already constitute 'theory', in the narrow sense that they can generate hypotheses that can then be tested by direct experiment. Equally important, they have inspired mathematicians and computer scientists to apply existing means to reduce complexity and seek new ones. For example, biological reaction networks do not have an order of execution, but probabilistic methods can be used to explore the most likely chains of reactions executed by a given network (M. Riedel and J. Bruck, personal communication).

A third path to better understanding of function begins with deeper analysis of the natural language now used to describe it. The cause-and-effect stories of function of proteins and regulatory sites use an impoverished vocabulary: many proper nouns, few verbs and some prepositional phrases denoting location. Like information in Geographical Information Systems, which also have a limited vocabulary, biological narratives of cause and effect are readily systematizable by computers. There are at least four commercial companies working to provide such systematizations, which are already providing some insight[11].

"We imagine ways that formalisms from computer science might contribute to a deeper understanding of biological function."

But for biological function, just beyond cause-and-effect narratives and before the ultimate truths of fitness and selection, there lies a muddy patch of ground known as 'teleology'. Teleology is hard to avoid: it is difficult to explain why the lens of the eye is transparent without at some point mentioning that the eye is 'for' seeing. But in that mud there may be hope.

The twentieth-century architects of information theory[6] deliberately restricted their concept of information because they were limited by their ability to define and measure it. Information theorists wanted to build a theory that involved the meaning (the semantic content) of messages, but could not measure meaning in sender, recipient or at any point in between. Instead, they chose a meaning for information that was restricted to the carrying capacity of communications channels such as telephone lines — the information technology of their era.

Similarly, biologists would like to cast their descriptions in terms of meaning and purpose, but are limited in their ability to measure those things. As we have said, biology does offer a clear definition of meaning ('it was selected'), but the multiple levels at which selection acts means that meaning is always difficult to determine.

Deeper understanding

Happily, there is considerable interest in wanting to build one element of biological semantics — the passage of time — into information theory. Formalizations of information processing that embodied this and other semantic concepts relevant to biology might help biologists to go beyond quantifying reaction rates and molecular species of biological systems to understand their dynamic behaviour. They might also help to suggest new experiments — perhaps on synthetic biological systems engineered to have a crisper division between process and output, which could then be evolved by artificial selection. This approach might bring a deeper understanding of function at its most fundamental level of fitness and selection.

However marvellous developments in computation are by 2020, if their impact is limited to information generation, handling, visualization and integration, it will mean that their potential contribution to a more predictive understanding of biological function will have failed. By laying out three paths from current computer science that might lead to deeper insights, we at least hope to stir things up. But we also observe growing frustration with business as usual. If we knew better how biological systems worked, we could better perturb existing ones (such as ours, for human medicine) and we could design and build better ones. The fact that both possibilities and frustrations are now starkly evident should make the next 16 years interesting indeed. ■

Roger Brent is at The Molecular Sciences Institute, Berkeley, California 94704, USA; Jehoshua Bruck is at the California Institute of Technology, Pasadena, California 91125, USA.

1. Weinstein, J. N. *Science* **282,** 687 (1998).
2. Simon, H. R. *The Sciences of the Artificial* 1st edn (MIT Press, Cambridge, Massachusetts, 1969).
3. Fields, S. *Nature Genet.* **15,** 325–326 (1997).
4. Kirschner, M. W. *Cell* **121,** 503–504 (2005).
5. Brent, R. *Cell* **100,** 169–183 (2000).
6. Shannon, C. E. & Weaver, W. *The Mathematical Theory of Communication* (Univ. Illinois Press, 1948).
7. Sussman, G. J. *The Legacy of Computer Science* NCR/CSTB Symposium on the Fundamentals of Computer Science (2001).
8. Gillespie, D. T. *Physica A* **188,** 404–425 (1992).
9. Hodgkin, A. L. & Huxley, A. F. *J. Physiol. Lond.* **117,** 500–544 (1952).
10. Gillespie, D. T. *J. Phys. Chem.* **81,** 2340–2361 (1977).
11. Calvano, S. E. *et al. Nature* **437,** 1032–1037 (2005).

Acknowledgements We are grateful to L. Lok, K. Tahashashi, D. Endy, O. Resnekov, P. Rabinow, D. Gillespie, M. Cook and M. Riedel for useful discussions and comments on the manuscript.

A two-way street to science's future

To view the relationship between computing and science as a one-way street is mostly untrue today, argues **Ian Foster**, and will be even less true by 2020.

A growing number of sciences, from atmospheric modelling to genomics, would not exist in their current form if it were not for computers. A simplistic analysis of this relationship focuses on hardware, and sees science as largely a passive beneficiary of the computing industry's relentless innovation, acquiring and applying to its own ends the fastest computers, largest disks and most capable sequencing machines. In this view, science and computing (as an intellectual discipline) have little to say to each other: it is the computer industry that drives the advances that have an impact on science.

A more sophisticated narrative says that science is increasingly about information: its collection, organization and transformation. And if we view computer science as "the systematic study of algorithmic processes that describe and transform information"[1], then computing underpins science in a far more fundamental way. One can argue, as has George Djorgovski, that "applied computer science is now playing the role which mathematics did from the seventeenth through the twentieth centuries: providing an orderly, formal framework and exploratory apparatus for other sciences."[2]

Information overload

Why this shift? Of course, there is more information, as technology allows us to collect, store and share vastly more data than before. Equally significant is that science is becoming less reductionist and more integrative, as researchers attempt to study the collective behaviour of larger systems. To quote Richard Dawkins: "If you want to understand life, don't think about vibrant, throbbing gels and oozes, think about information technology."[3]

Such system-level approaches are emerging in fields as diverse as biology, climate and seismology. A frequent goal is to develop high-fidelity computer simulations as tools for studying system-level behaviour. Computer science, as the 'science of complexity', has much to say about how such simulations — which can be considered a new class of experimental apparatus — should be constructed, and how their output should be analysed and compared with experiment[4]. Similarly, information theory provides formidable insight into how biological systems encode, transform and transmit information.

The sciences rely on computers, but the benefits are two-way; each is driving the other forwards.

Both the data deluge and system-level science demand computing technology in all its forms — hardware, software, algorithms and theory. The growing importance of computing has several implications for the science of 2020, of which I explore three here.

First, the scientist of 2020 will be adept in computing: not only will they know how to program, but they will have a solid grounding in, for example, the principles and techniques by which information is managed; the possibilities and limitations of numerical simulation; and the concepts and tools by which large software systems are constructed, tested and evolved. This knowledge has been picked up on the job by many pioneering scientists and will hopefully be instilled in the next generation by more formal training. The idea that you can be a competent scientist without such training will soon seem as odd as the notion that you need not have a solid grounding in seventeenth-century mathematics (such as algebra).

Fruitful partnerships

Second, successful science collaborations of 2020 will include computer scientists as key members. All scientists will be adept at applying existing computational techniques, but they will also understand that progress in their fields will require innovation in computing technology. So they will work with computer scientists to identify computational problems, much as today's experimentalists and theorists understand the strengths and weaknesses of their favoured methods and know to partner with others when new techniques are needed. Indeed, this fruitful interplay is already occurring: for example, the communication challenges inherent in far-flung physics collaborations inspired the development of the World Wide Web, and the need for efficient indexing of terabytes of digital astronomy data has spurred new approaches to organizing spatial data in relational databases[5]. Elsewhere, the scientific opportunities that arise from using wireless sensor networks for, say, continuous habitat monitoring are driving innovations in network protocols and algorithms.

Third, the scientific disciplines and institutions of 2020 will need to train, attract and reward researchers whose focus is on producing the computing innovations required for science to advance: what we might term 'applied computing'. Thus, we see new organizational structures, such as the Computation Institute at the University of Chicago and Argonne National Laboratory (for which I work) and Harvard's Institute for Innovative Computing, that aim to bridge the distinct concerns of computer science and other sciences. Academic departments are hiring faculty with strong computational inclinations. National laboratories have established computational directorates. It will be interesting to see which of these interdisciplinary structures work best.

The growing importance of applied computing also has implications for computer science. Indeed, just as during the early days of the sciences, scientific concerns drove mathematics forward (think of the origins of the calculus), so the many challenging problems posed by modern science can help to focus and motivate research in computing. In my view, it is no accident that some of the most vibrant areas in computing today are those tightly coupled to scientific problems. These dynamics occur in fields as diverse as sensor networks, data integration and grid computing. It's a two-way street, and always has been. ■

Ian Foster is director of the Computation Institute at the University of Chicago and Argonne National Laboratory, Illinois, USA.

Acknowledgements To Anne Rogers for her comments.

1. Denning, P. *et al. Commun. ACM* **32,** 9–23 (1989).
2. Djorgovski, G. S. *Virtual Astronomy, Information Technology, and the New Scientific Methodology.* 7th Int. Workshop Comput. Arch. for Machine Percept. 125–132 (2005).
3. Dawkins, R. *The Blind Watchmaker* (Norton, 1986).
4. Cohen, J. *PLoS Biol.* **12,** 2017–2023 (2005).
5. Szalay, A. S. *et al. SIGMOD Rec.* **29,** 451–462 (2000).