limit of the parameters, the theory tends smoothly to general relativity in all its predictions. The 1960s vintage scalar–tensor theory of Brans and Dicke (which was based on earlier theories by Jordan and Fierz) is the classic example. There a parameter ω_{BD} is an inverse measure of the strength of a scalar field added to the usual spacetime metric tensor field. As ω_{BD} tends to infinity, Brans–Dicke theory tends to general relativity.

Non-simply connected theories are so qualitatively different from general relativity in their formulations, that they do not formally tend to it in any limit. Nevertheless, many theories in this class can be made to agree with general relativity in the weak-field, slow-motion regime of the Solar System.

Binary pulsar tests have proven to be deadly for non-simply connected theories, because their qualitative divergences from general relativity usually show up with a vengeance in strong-field and radiative situations. The best example of this was the 'bimetric' theory of Nathan Rosen, which agreed with Solar System tests but failed spectacularly in the binary pulsar[6].

The same cannot be said for simply connected theories. Not that they do not deviate from general relativity in binary pulsar systems, for they do, but they also generally deviate from it in the Solar System, and so are already constrained, usually more strongly than is currently achievable in binary pulsars. Current Solar System tests of light deflection and Shapiro time delay (both around 0.1 per cent) constrain the Brans–Dicke theory's ω_{BD} to be greater than 500 (ref. 4). This makes the theory so close to general relativity (roughly within factors of $1/\omega_{BD}$ in all its predictions) that it easily satisfies all the binary pulsar constraints[7]. Put differently, the results quoted for PSR1913+16 by Taylor *et al.*[1] provide the constraint $\omega_{BD} > 100$, which is not competitive with the Solar System bound.

This raises a theoretical question. Is it possible to find a simply connected theory of gravity that is not constrained by Solar System tests, but that can be constrained by binary pulsar tests? Damour and Esposito-Farese have shown that it is (T. Damour and G. Esposito-Farese *Class. Quant. Grav.*, manuscript submitted). They cook up a scalar–tensor theory *á la* Brans–Dicke in which two scalar fields appear, tuned so that their effects in the weak-field limit of Solar System tests cancel, making the theory indistinguishable from general relativity for that application. In the strong-field and radiative limits, however, the scalar fields combine to produce observable differences. Yet as the theory's two parameters β' and β'' tend to zero, the theory does tend smoothly to general relativity. As Taylor *et al.* report, the two binary pulsar systems together provide a strong constraint on these parameters, consistent with zero.

The Wolszczan–Frail planetary system[2] will not have much impact on general relativity, because the planets are too light to emit significant gravitational radiation, and their orbits are too circular to reveal a meaningful periastron. But it does present a challenge to astrophysicists to explain how and when such planets formed, and how they survived the various cataclysms that usually afflict the lives of millisecond pulsars. It also suggests that planetary systems in the Galaxy may be more common than has hitherto been evident from optical searches (which have found nothing). On the other hand, I wouldn't lay odds on the existence of life on *these* planets. □

Clifford M. Will is in the McDonnell Center for the Space Sciences, Department of Physics, Washington University, St Louis, Missouri 63130, USA.

1. Taylor, J. H., Wolszczan, A., Damour, T. & Weisberg, J. M. *Nature* **355,** 132–136 (1992).
2. Wolszczan, A. & Frail, D. A. *Nature* **355,** 145–147 (1992).
3. Bailes, M., Lyne, A. G. & Shemar, S. L. *Nature* **352,** 311–313 (1991).
4. Will, C. M. *Science* **250,** 770–776 (1990).
5. Wolszczan, A. *Nature* **350,** 688–690 (1991).
6. Will, C. M. & Eardley, D. M. *Astrophys. J. Lett.* **212,** L91–L94 (1977).
7. Will, C. M. & Zaglauer, H. W. *Astrophys. J.* **346,** 366–377 (1989).

NEURAL NETWORKS

Learning from your neighbour

Graeme Mitchison and Richard Durbin

WHAT can artificial neural networks tell us about the brain? One view is that they can be used to explore the consequences of different synaptic learning rules in a simplified formal setting. However, the most powerful learning algorithms, such as back propagation[1], need an external 'supervisor' to correct the mistakes made by the network, which is an unrealistic requirement especially for early stages of sensory processing. How, therefore, does one learn effectively without a supervisor? On page 161 of this issue[2], Becker and Hinton propose an answer to this question. Theirs is not the first unsupervised learning algorithm, but they take a new approach which has a paradoxical charm: in effect, different pieces of the inputs train each other.

The goal of an unsupervised learning algorithm is to extract meaningful features or variables from a set of input patterns. For example, we can try to find those features that allow the data to be reconstructed as faithfully as possible. This is the goal of principal component analysis, a standard tool of engineering and statistics. By identifying the combinations of inputs with maximum variance, it finds the variables that can be most effectively used to characterize the inputs. Remarkably enough it turns out that the first neurobiological learning rule to be formulated, Hebb's rule[3], is closely related to principal component analysis. Given a simple neural-network model consisting of a single unit, Hebb's rule results in that unit extracting the largest principal component, assuming some form of normalization of synaptic connection strengths[4,5]. With a small amount of modification, a set of units can be made to learn not just the largest component, but a set of components which together capture the greatest part of the variance[6,7].

Principal components appear in at least some cases to be involved in biological processes. In the local processing of visual images, for example, the principal components include edge segments, which are among the first features extracted in primary visual cortex[8]. However, other important variables, such as stereoscopic disparity, will not be explicitly extracted. Becker and Hinton show how one could set about extracting these more elusive variables. One way to describe their approach is that they assume that the interesting properties are more stable than the noise. For example, the depth of a surface, as measured by stereoscopic disparity, will tend to vary smoothly in scanning across an image, whereas the local pixel intensities may vary rapidly because of texture.

Consider a system looking at two neighbouring, non-overlapping patches, and suppose that, corresponding to each patch, there is a unit whose inputs come from that patch only. One could try to make the units extract a stable property by requiring that they both perform the same computation on their input and by minimizing the difference in their responses. But then they might end up both doing nothing (that is, give a zero response). To avoid this, one could try to mimic hebbian principal component learning, and ask the units to maximize the variance in their responses. Becker and Hinton combine these requirements by making the units maximize the variance of the sum of their outputs divided by the variance of their difference.

On the assumption that both the underlying variable and the noise have a gaussian distribution, this is equivalent to maximizing the mutual information of the two outputs. Here one can see particularly clearly how the algorithm works: the mutual information can be large only if, first, the units convey information (that is, they behave nontrivially) and if, second, they respond similarly, so they share this information. The Hebb rule essentially imposes the first constraint alone. By adding the second constraint the new rule allows information to be thrown away when it is not shared by other patches.

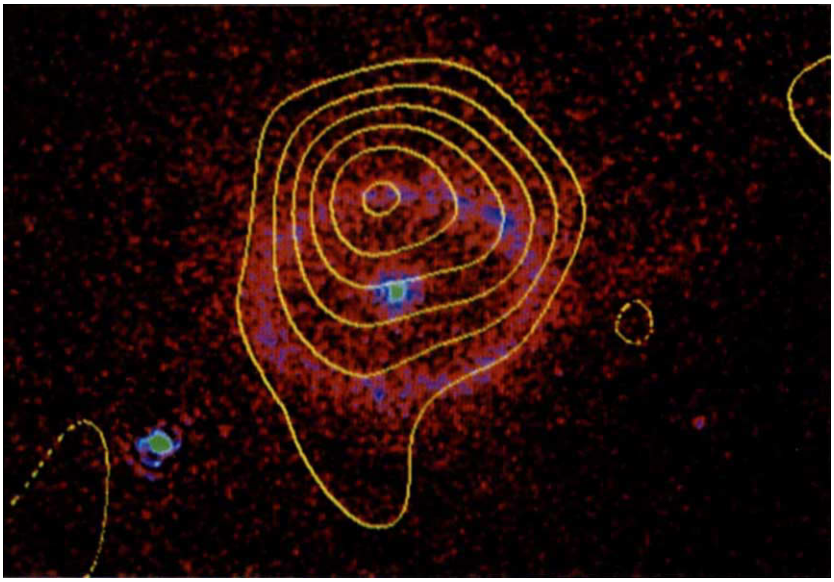
Maximizing mutual information can also be interpreted as prediction, because each unit can be used to predict the behaviour of neighbouring patches. The notion of prediction is more general than that of stability: we can look for properties that predict future inputs, or predict one set of sensory data through another sensory modality. Prediction can help to complete or interpret missing data, and where prediction fails something interesting is likely to be happening. For example, places where disparity changes sharply will usually correspond to the edges of objects.

Neural networks are inspired by real neurons, but is there is any reverse flow of inspiration? Might a rule such as this operate in the brain? It seems unlikely that neurons compute something as mathematically complex as the ratio of variances, let alone the determinants which occur in the more general expression for more than two units. Furthermore, some of the difficulties of back propagation apply to the multilayer version of this algorithm, which must somehow feed back a complex error signal to earlier stages in the neural pathway. But it is important not to be too intimidated by the mathematical formulation. After all, principal component analysis, which in its standard form requires matrix inversion, might seem an unlikely operation for neurons to accomplish. Yet it can be carried out by suitably organized hebbian machinery. It seems likely, in fact, that there are natural ways for neurons to carry out Becker and Hinton's kind of analysis, or something very close to it, and this may provide another clue to help us explore synaptic learning rules in the brain. □

Graeme Mitchison is in the Physiological Laboratory, University of Cambridge, Cambridge CB2 3EG, UK. Richard Durbin is in the MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

1. Rumelhart, D. E., Hinton, G. E. & Williams, R. J. *Nature* **323,** 533–536 (1986).
2. Becker, S. & Hinton, G. E. *Nature* **355,** 161–163 (1992).
3. Hebb, D. O. *The Organization of Behavior* (Wiley, New York, 1949).
4. Oja, E. *J. math. Biol.* **15,** 267–273 (1982).
5. Linsker, R. *Computer* 105–117 (March 1988).
6. Sanger, T. D. *Neural Networks* **2,** 459–473 (1989).
7. Földiák, P. in *Proc. Int. Joint. Conf. Neural Networks* Vol. 1, 401–405 (IEEE, New York, 1989).
8. Hubel, D. H. & Wiesel, T. N. *J. Physiol.* **160,** 106–154 (1962).

Radio days of a remnant supernova



IN the history of supernova 1987A, now almost five years old, radioastronomers have so far had a negligible role. Apart from a brief initial outburst of radio emission, lasting no more than a few days, the expanding nebula set into motion by the explosion has been quite invisible at radio frequencies, and the steady thinning and cooling of the ejected material, and its interaction with the circumstellar material that surrounded the progenitor, has been followed largely through ultraviolet, optical and infrared observations. But elsewhere in this issue (*Nature* **355**, 147–149; 1992). L. Staveley-Smith *et al.* describe their detection of radio emission from the remnant, illustrated here overlayed on an optical picture from the Hubble Space Telescope. Evolution of the radio remnant over the coming years will provide a new tool to dissect the progress of the expanding remnant.

The key to understanding the remnant of SN1987A lies in the nature of its unusual progenitor star, which was first a red-giant, then a blue giant, before it exploded. In its red-giant phase, the star threw off a dense, slow-moving wind, which was succeeded by a more tenuous but faster wind from the blue-giant. The circumstellar material of the progenitor at the moment of explosion therefore consisted of a hot thin gas cocooned inside a cooler, thicker shell, with a shock wave created at the boundary as the blue-giant wind ran into the red-giant wind.

The first brief flash of radio emission, reported by A. J. Turtle *et al.* (*Nature* **327**, 38–40; 1987), was a very minor part of the initial supernova outburst, and was probably attributable to the propagation of the shock wave from the explosion through the thin material immediately surrounding what had been the progenitor star. According to R. A. Chevalier (*Nature*, in the press), the emission now detected by Staveley-Smith and colleagues is due to the same expanding shock finally reaching the outer edges of the old blue-giant wind, just before it runs into the denser red-giant wind. Chevalier predicts that as the expanding ejecta passes through this boundary layer, the radio signal will rise and then diminish again, a signature which should be seen sometime during 1992.

After this transient appearance, SN1987A is unlikely to emerge as a fully formed radio supernova remnant for some time. The ages of radio remnants seen in other galaxies as well as our own are typically measured in hundreds of years at least, and there have been few opportunities for astronomers to observe a supernova at close enough hand to see the radio remnant arise from the expanding nebula. Before the advent of SN1987A, astronomers had to make do with studies of supernovae in other galaxies, and those that are detectable at radio frequencies have been either mature remnants or very new ones, which have faded within a few years.

Just before the radio recapture of SN1987A, however, J. Cowan, *et al.* (*Astrophys. J.* **379**, L49–L51; 1991) spotted the reappearance of a 20-year old supernova, SN1970G in the galaxy M101, that had been radio-bright for about three years after outburst but which had then sunk below detectability. It is thought that the progenitor of SN1970G was, like that of SN1987A, a fairly massive star, and the explanation for the reappearance of radio emission from the former after 20 years and from the latter after five may be essentially the same. **David Lindley**