

An Interview with Jack Cowan¹

Jack Cowan was born in Leeds, England in 1933. Educated at Edinburgh University, Imperial College and MIT, he is one of the pioneers of continuous approaches to neural networks and brain modelling. He has made many important contributions to machine learning, neural networks and computational neuroscience. In 1967 he took over from Nicolas Rashevsky as Chair of the Committee on Mathematical Biology at the University of Chicago where he has remained ever since; he is currently Professor in the Mathematics Department.

This is an edited transcript of an interview conducted on the 6th November 2006.

Phil Husbands: Can you start by saying a little about your family background, in particular any influences that might have steered you towards a career in science.

Jack Cowan: My grandparents emigrated from Poland and Lithuania at the turn of the last century; I think they left after the 1908 pogroms and they ended up in England, on my mother's side, and Scotland on my father's. My mother's parents had a clothing business in Leeds and my father's family sold fruit in Edinburgh. My father became a baker. My mother was clever and did get a scholarship to go to University but she had to decline because of the family finances. So I was the first member of my family to go to University.

PH: Did you get much influence from school?

JC: Yes, in that I went to a good school. My parents were very encouraging from an early age – my mother claims that I started reading when I was very young and that I was bossing the other kids in kindergarten! Anyway, we moved to Edinburgh from Leeds when I was six years old and I went to a local school there for about three years, but my parents could see that I had some aptitude so they got me into George Heriot's School, a very good private school. I got bursaries all the way through and ended up the top boy in the school – I was Dux of the school - and got a scholarship to Edinburgh University.

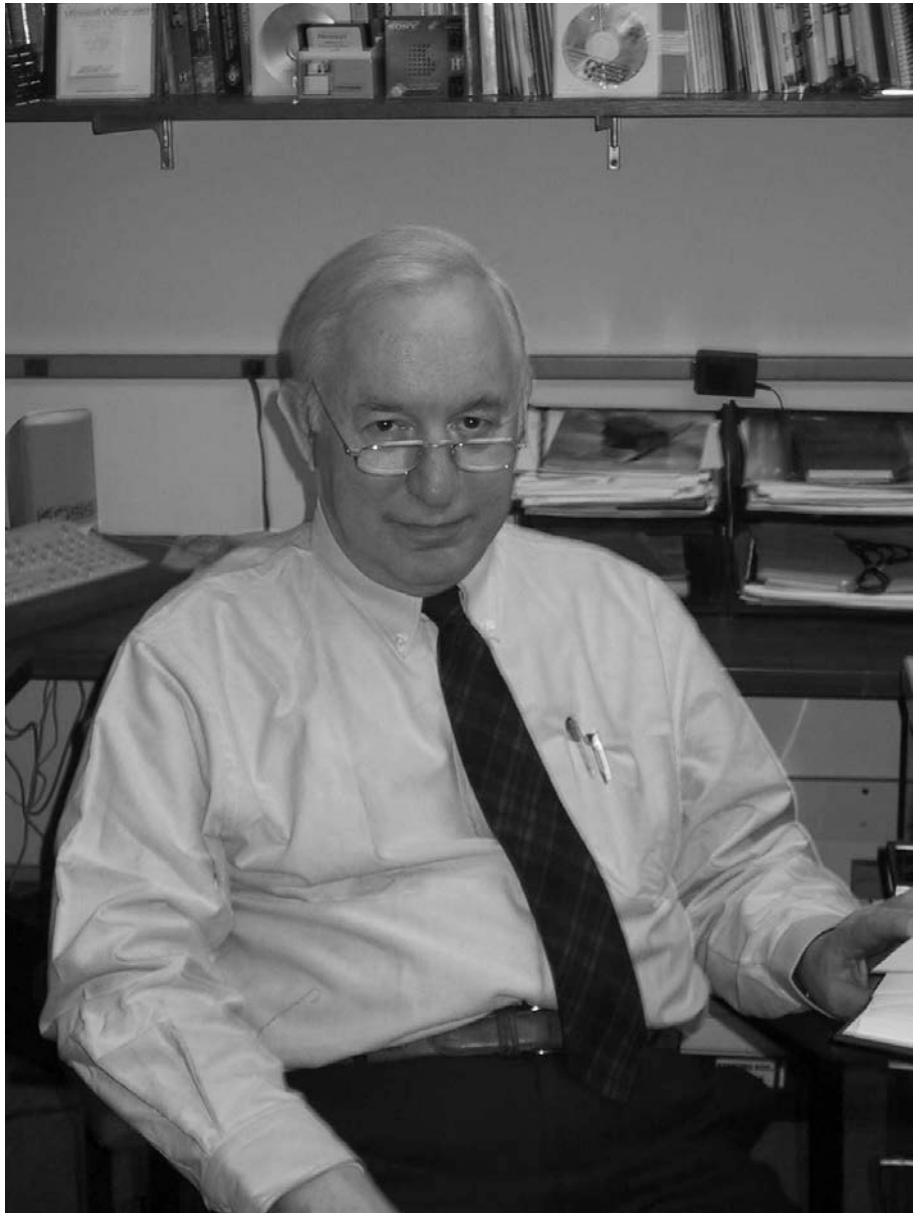
PH: What year did you go to university?

JC: I was an undergraduate from 1951 to 1955, studying physics. I remember when I was about fourteen we had the traditional argument between Jewish parents and their son – they wanted me to become a doctor or a dentist or lawyer or something like that and I kept telling them ‘No way, I’m going to be a scientist.’ So I decided early on that I wanted to do science and I can’t say there were any particular outside influences on this decision; it seemed to come from within.

PH: How were your undergraduate days?

¹ Preprint of P. Husbands (2008) An interview with Jack Cowan. In P. Husbands, O. Holland, M. Wheeler, (Eds), *The Mechanical Mind in History*, MIT Press, 431-446.

JC: Well from being top boy at Heriot's my undergraduate career was a disaster. I found the physics faculty and the lectures at that time really boring. I didn't do well at all.



But after that I was rescued by a man called J.B. Smith who was head of a section at Ferranti Labs in Edinburgh where I'd applied for a job. He has also been the school Dux at Heriot's - a decade or so before me - so I guess he took a chance and hired me. I was in the instrument and fire control section. I was there for three years from 1955, although in the middle of that I was sent to Imperial College for a year to work with Arthur Porter, one of the pioneers of computing in Britain. I also got to know Dennis Gabor, who I hit it off with. As well as being the inventor of holography, he had a lot of interest in cybernetics, machine learning and things like that. He worked on adaptive filters and introduced the idea of using gradient descent to solve for the coefficients of a filter that was learning by comparing the input with the output. I would say that Gabor was a huge influence on me.

PH: Was it going to Imperial that sparked the direction your work took, leading you into machine learning and neural networks?

JC: To a large extent. But before that what really got me started, and actually I think what impressed Smith, was that I had read Norbert Wiener's book on cybernetics. I picked it up in the library when I was an undergraduate and found it very very interesting. Also while I was still an undergraduate I heard a lecture by Gabor on machine learning which was very influential.

PH: What kind of work were you doing for Ferranti?

JC: The first project they gave me was to work out pursuit curves. Ferranti worked on the computer guidance systems for the British fighter planes of the time; there was a consortium of Ferranti, English Electric and Fairey Aviation involved in the computers that controlled air to air missiles. So they had me work with a couple of other people on the mathematical problems of prediction of missile trajectories and things like that. So I learned quite a bit of useful mathematical stuff doing that.

A year or two before I started at Ferranti, Smith and a colleague, Davidson, had built a machine that solved logic problems by trial and error. I got it working again and they arranged for me to take this machine down to London to the electrical engineering department at Imperial College to demonstrate it to Porter and Gabor. That developed my interest in automata theory and machine learning. Anyway, Ferranti arranged for me to spend a year at Imperial doing a postgraduate diploma in electrical engineering. Porter wanted me to stay to complete a PhD but I only had the year. I had started to play around with many-valued logics to try and solve logic problems in a better way than simple trial and error as embodied in Smith's machine. It was this that got Gabor interested in me and he became my mentor.

I met a lot of interesting people during that year: Wilfred Taylor, who was at University College and developed one of the very first learning machines, Raymond Beurle, from Nottingham University, who had written a very beautiful paper on the mathematics of large scale brain activity¹ and many others. So I met all these guys, which was very inspiring, and, in 1956, Ferranti also sent me to one of the earliest international meetings on cybernetics in Belgium² where I met Grey Walter, with his turtles, and Ross Ashby, and that is where I first met Albert Uttley, who was working on conditional probability approaches to learning, among other things. I also remember a very interesting lecture at Ferranti given by Donald Mackay. So by the time I was in my early twenties I'd already met most of the leading people in Britain working in the area that interested me. And there was a lot of very good work going on in Britain. As well as these interactions I came across a number of papers that would prove influential later – for instance, John Pringle's paper on the parallels between learning and evolution, which really set the foundation for competitive learning³ and Turing's work on the chemical basis for morphogenesis⁴ which would inspire my work a couple of decades on.

A little later, essentially through Porter and Gabor, I ended up with a fellowship from the British Tabulating Machine Company (BTM) to go to MIT. They ran a special

scheme to send graduate researchers from Britain to MIT. This attracted me, so I applied and got it.

PH: When did you start at MIT?

JC: I arrived at MIT in the fall of 1958 as a graduate student. I joined the Communications Biophysics group run by Walter Rosenblith. I was in that group for about 18 months then I moved to the McCulloch, Pitts and Lettvin group.

PH: How did that move come about?

JC: Well my interests were a bit more theoretical than what was going on in the Communications Biophysics group – they were mainly interested in auditory psychophysics which I didn't find as interesting as the more theoretical aspects of cybernetics. I had been working on many-valued logics at Imperial and through reading von Neumann's paper in Shannon and McCarthy's *Automata Studies* collection⁵ had got very interested in the problem of reliable computation using unreliable elements. So I started to work on applying many-valued logic to that problem. That kind of thing didn't really fit in the Rosenblith group.

In 1959, while I was still in his group, Rosenblith organized a very interesting meeting at MIT on sensory communication⁶. That was a great meeting for a graduate student like me to attend, there were all kinds of very interesting people there (I've got the proceedings here): Fred Attneave, Horace Barlow, Colin Cherry, Peter Elias, Licklider, Donald Mackay, Werner Reichardt, Willie Rushton, Pat Wall, to name a few! It was an amazing meeting. The stand-out talks for me were Horace Barlow's on *Possible Principles Underlying the Transformations of Sensory Messages*⁷, where he talked about the possible role of redundancy reduction in the nervous system, and Werner Reichardt's *Autocorrelation: a principle for the evaluation of sensory information by the CNS*⁸, in which he presented an early version of the famous Reichardt motion-detector model. That was also where I first heard about the Lettvin, Maturana, Pitts and McCulloch work on the frog's visual system^{9,10}, which was also extremely good, and that was what got me really interested in joining the McCulloch and Pitts group. McCulloch was also interested in the reliability problem, so I joined.

PH: What was MIT like at that period?

JC: In those days MIT was absolutely fantastic. I got to know a huge range of people; I consider myself to have been very lucky to have been there at that time. I remember the first day I got there I was taken to lunch by Peter Elias and David Huffman, Huffman of Huffman coding and Elias who was one of the big shots in information theory, and they said to me, 'You know, graduate school at MIT is not like in England. It's like a factory with an assembly line and you get on and it goes at a certain rate and if you fall off – too bad!' They warned me that it was very hard going and rigorous. They were right!

But it was an amazing place. As well as great names from cybernetics and information theory – Wiener, McCulloch, Pitts, Shannon – Noam Chomsky was down the hall, Schutzenberger was there working with him on formal linguistic theorems,

Roman Jakobson was around. Some of the classes were incredible – for instance, being taught by the great pioneers of information theory.

PH: Who were the major influences on you from that time?

JC: McCulloch, Pitts, Wiener and Shannon. I was very lucky that Shannon arrived at MIT from Bell Labs the year I got there. So I took courses on information theory with Bob Fano, Peter Elias and Claude Shannon, I had the benefit of a set of lectures from Norbert Wiener, and I interacted all the time with Warren McCulloch and also to quite an extent with Walter Pitts.

PH: So Pitts was still active in the lab?

JC: He was still sort of functional. In fact I was one of the last students to really talk to him at length about his interests and work. He and Wiener probably had the biggest influence on me because it was through talking with them – separately, because by then Wiener had fallen out with McCulloch and Pitts - that I decided to start working on trying to develop differential equations to describe neural network dynamics and to try to do statistical mechanics on neural networks. Pitts directly encouraged me to look at continuous approaches to neural networks.

PH: I seem to remember that you have an unfinished thesis by Pitts ...

JC: Well I don't have a thesis but what I have is a fragment of an unpublished manuscript which I copied. He gave it to me for a while and let me copy it. So I hand copied it, imitating his writing, and then gave it back to him. Jerry Wiesner, who was then head of RLE, the Research Lab of Electronics, to which we belonged, actually offered money to anyone who could get Pitts to write something up and publish it so that they could give him a degree. But unfortunately this thing was only a fragment, he never finished it.

PH: It was on the beginnings of a statistical mechanics treatment of neural networks wasn't it?

JC: Yes. It was the beginnings of Walter's attempt to do something but unfortunately it didn't go very far. But remember that when he did that, in the late '50s, this was long before any of the statistical mechanics techniques needed for solving the problem had been developed.

PH: Did you interact with Oliver Selfridge?

JC: I had some very nice talks with Oliver who was working on the Pandemonium research at that time¹¹. But he had also done some very nice earlier work with Wiener and Pitts on the origins of spirals in neural models with possible applications to cardiac problems¹². In fact some of the stuff I work on now is closely related to what they were doing. Marvin Minsky also got involved with that work. There is a very nice study by them on reverberators and spirals.

PH: So when did your period at MIT end?

JC: 1962. So I was there for four years. During that period I recruited Shmuel Winograd, who went on to become a major figure at IBM, to the group. I was working on the reliability stuff with McCulloch, and Shmuel and I got interested in the capacity of computing devices. We developed a theory of how to design optimal reliable network configurations of computing elements. We came up with one of the earliest designs for a parallel distributed computing architecture. This work got us known and we wrote a monograph on it¹³.

PH: Would you say it was during this period that your interests started to move more towards biology?

JC: Yes. It was definitely at MIT, through the influence of McCulloch and others, that I moved from thinking about automata towards starting to think about the nervous system. So it was a defining period in that sense.

PH: At about that time approaches to machine intelligence began to diverge to some extent. Minsky and McCarthy and others were very active in exploring and promoting new directions in what they called artificial intelligence, and cybernetics was starting to wane. So things were at a cusp. What are your memories of the expectations people had?

JC: Well there was always this tremendous hype about artificial intelligence around Marvin and McCarthy and Newell and Simon and so on. I remember Herb Simon coming to give a talk and it was the same message we got from Marvin; if we had bigger and faster computers we would be able to solve the problems of machine translation and AI and all kinds of stuff. But they set up the AI Lab and were instrumental in the development of lots of useful technology.

Through McCulloch I got to know Marvin Minsky very well and in fact I recruited Seymour Papert to join our group, but by the time he arrived I'd gone back to England so he ended up working with Marvin.

PH: So what was the reaction in the McCulloch group to all the hype surrounding AI?

JC: Great scepticism.

PH: Do you remember what your own personal views were at the time on what was likely to be achieved and on what the important problems were?

JC: Well I was still in the middle of learning as much as I could and trying to think out what direction I should take. I had a strong bent towards applying the methods of theoretical physics and I was getting more and more interested in the nervous system and neural network models. As I mentioned earlier, Pitts and Wiener had influenced me to look in the direction of continuous approaches to neural networks and I started to think that the statistical mechanics of neural networks was a very important problem. I remember sitting in the office I shared with McCulloch and having the idea that there is an analogy between the Lotka-Volterra dynamics of predator-prey interactions in populations and excitatory and inhibitory neuron interactions in neural networks, and that set me going for the rest of my career.

PH: How was the transition back to England in 1962?

JC: Well in 1962 I was at a meeting in Chicago when I was approached by two gentlemen from the Office of Naval Research who asked me if I would like grant support. I said, ‘Well, yes!’ and so they gave me my own personal grant that I was able to take back to England with me. I had to go back to Britain for at least a year because that was part of the terms for the fellowship I had that funded me at MIT.

So I went back to Imperial as an academic visitor. Meanwhile I got a masters degree at MIT, but neither Shmuel Winograd nor I decided to brave the doctoral programme there on the advice of Claude Shannon. After Claude had written his first famous paper, on the application of Boolean algebra to switching networks, he took the doctoral qualifying exam in electrical engineering and failed; I think he failed the heavy current electrical engineering part. So he went to the Math department and did his PhD there. So we took his advice and Shmuel got his doctorate from NYU and I returned to Imperial without a PhD.

PH: How did your work develop at Imperial?

JC: So I went back to the electrical engineering department at Imperial and got involved in a number of things. I started doing a bit of teaching, labs on numerical methods and computing and things like that, and I started supervising students even though I was really technically still a student myself! I worked on the monograph on reliable computing from unreliable elements with Winograd which got published by MIT Press¹² after the Royal Society rejected it! We made the link between von Neumann’s work and Shannon’s work on the noisy channel coding theorem and introduced the parallel distributed architecture thirty years before its time. After we finished that I turned to the problem of neural network dynamics and by about 1964 I had the beginnings of a way to do the mathematics of neural networks using systems of non-linear differential equations. I did a version of it that led to a statistical mechanics¹⁴, but it wasn’t quite the right version; it was a special case, the anti-symmetric case. This came from the analogy with population dynamics where an excitor neuron is coupled to an inhibitor that is coupled back to it, so the weights are anti-symmetric. In this work I introduced the sigmoid non-linearity into neural models¹⁵. There was another special case that I didn’t follow up at the time but it was followed up fifteen or so years later by John Hopfield¹⁶; the symmetric case. Hopfield networks were the other special case of the network population that I introduced in about 1964. Anyway, when I was doing that work in the ‘60s I realised that there was clearly a relationship between what I had done and Raymond Beurle’s work on a field theory of large-scale brain activity – a kind of continuum model¹. So I spent quite a bit of time working on that and wrote it all up in a report for the Office of Naval Research.

PH: So who else in the UK were you interacting with at that time?

JC: Mainly Gabor, Uttley and Mackay at that stage and a little bit with Christopher Longuet-Higgins and David Willshaw who were doing interesting neural network research in Edinburgh – associative memory work. I also used to interact a bit with Richard Gregory who I got on very well with.

PH: You went and worked with Uttley's group didn't you?

JC: Yes. I spent four years at Imperial, '62-'66, and then in '66-'67 I split my time – about a day a week at Imperial and the rest at the National Physical Laboratory at Teddington. Albert Uttley had invited me to go out there to work in his Autonomics Division. I mainly worked with Anthony Robertson who was a neurophysiologist working in that group.

PH: What did you think of Uttley's ideas at that time?

JC: Well I always liked Uttley's ideas; I think he was under valued. He had some very good ideas which were precursors to more modern work on machine learning. He had the right ideas – for instance, using a conditional probability approach¹⁷ – he just didn't have a clean enough formulation. Of course this was long before people discovered the relationship between statistics and neural networks.

PH: What about the wider field of theoretical biology that was gaining strength in Britain at about this time?

JC: Yes, that was another group of very interesting people I was involved in. My link to that started back in Edinburgh when I was growing up. One of my friends was Pearl Goldberg who got married to Brian Goodwin, the theoretical biologist. We met up in Boston when I was at MIT and through me they ended up staying with McCulloch for a while. Anyway, Brian had developed a statistical mechanics approach to cell metabolism. I liked that a lot and I realised my Lotka-Volterra thoughts on neural networks could be done the same way. So Brian's work was a trigger to my first statistical mechanics approach to neural networks. When I got back to London in '62 I'd meet up with Brian, who was at Edinburgh University, and through him I got to know Lewis Wolpert, the developmental biologist. And so we had a discussion group on theoretical biology, which Michael Fisher used to come to occasionally, and so that's when I really started to get into the wider field. Then Waddington, who was also in Edinburgh, organized the *Towards a Theoretical Biology* meetings¹⁸, and through Brian I got to go to those. That was quite an interesting collection of people. The mathematicians René Thom and Christopher Zeeman were there, and so were Ernst Mayr, John Maynard Smith and Dick Lewontin, the evolutionary biologists, and Lewis Wolpert, Donald Michie, who at that time was still working in genetics, Christopher Longuet-Higgins, Brian and me.

Now Lewontin was on the look out for someone to take over from Rashevsky at the University of Chicago. Rashevsky had set up the Committee on Mathematical Biology¹⁹ in the late 1930s but by 1965 he had resigned and they were looking for a replacement. They settled on either Brian Goodwin or me and Brian wasn't interested as he had not long before moved to Sussex University. So I went for a long walk with Lewontin and Ernst Mayr in the woods outside the Villa Serbelloni, where we were having the meeting, which overlooked Lake Como. I remember Ernst was amazing, pointing out every animal and insect and plant in the woods. Anyway they talked me into thinking seriously about taking the job. At that time I wanted to go to Sussex to work with Brian. I had applied to the UK Science Research Council for a grant to work on the statistical mechanics of large scale brain activity and told them that if I didn't get the funding I'd have to go to the US. And they didn't give me the funding.

The referees, who included Donald Mackay, claimed it was too speculative. So I ended up taking the job and moving to Chicago.

I'd been appointed a professor and chairman of the Committee on Mathematical Biology at Chicago and I still didn't have a PhD. So I decided it really was time and I took a week out to write up some of my work into a thesis, and I had a viva exam with Gabor as my internal examiner and Raymond Beurle as the external. The viva lasted two minutes and then we drank some champagne! So I got my PhD on the statistical mechanics of neural networks, the first ever PhD in that area.

I arrived in Chicago with my wife, who was seven months pregnant, the day after a monster snow storm in the winter of 1967.

PH: Had the intellectual climate changed much in the time you'd been away? I'm wondering if the AI bandwagon had had a negative impact on funding in the areas you were interested in?

JC: Yes and no. It didn't do anything to mathematical biology but it did damage the field of neural networks. When Minsky and Papert published their attack on the perceptron in 1969²⁰, and they'd been giving talks on that stuff for a while before, they made the claim that you couldn't solve the perceptron training problem. In retrospect I had invented the machinery necessary to solve the problem, and show that they were wrong, in the mid '60s – the sigmoid model I used in my Lotka-Volterra like network dynamics model. But I didn't work on that aspect; I put it aside. There were two major things that I should have done but didn't at the time. One, as I've already mentioned, was to do the other case of the Lotka-Volterra network, which is essentially what Hopfield did, and the other was to use the sigmoid model to do perceptron training, which is what Rumelhart, Hinton and Williams did in 1986²¹. So I kick myself for not doing either.

PH: How did things pan out in the Committee on Mathematical Biology?

JC: Well I was chairman for six years and I built it into a department of theoretical biology. I recruited people like Stuart Kaufmann and Art Winfree, who both went on to become very prominent, and various other people. It actually had quite a decent influence on theoretical biology in the US and elsewhere. But then we merged with the biophysics department, because it was thought that small departments were not so viable, but that proved to be a mistake. The merged department then got further merged to become part of something that also accommodated genetics and molecular biology and other branches of biology. So in 1980, or thereabouts, I moved to the mathematics department and I've been there ever since.

PH: What was the main focus of your work from the late 1960s?

JC: So my idea of correcting and extending Beurle's work paid off and I was very fortunate to recruit a very good post-doc, Hugh Wilson, to work with me on that. So Wilson and I published a couple of papers, in '72 and '73, which triggered a great deal of activity^{22,23}. We basically gave the first non-trivial and useful field theory, what we would now call a mean field theory, for looking at large scale brain dynamics. But even then I knew that that work wasn't really the answer to the

problem I'd set myself of doing statistical mechanics of neural networks. Even when, in the late '70s, Bart Ermentrout and I showed that you could apply modern mathematical techniques to calculate the various patterns that could form in networks of that kind²⁴, which turned out to be useful for various kinds of applications, it still wasn't really getting to grips with what might be going on in the nervous system. So I made a start on trying to do that in 1979 and discovered the key to doing it in 1985 while working at Los Alamos with two physicists, Alan Lapedes and David Sharp, and got a first version going in about 1990. I worked on it a bit more with a student, Toru Ohira, but it is only in the last two or three years, with a really bright graduate student named Michael Buice, have we actually solved the problem. So now we are in possession of a field theory for large scale brain activity which is exactly the kind of object that Norbert Wiener and Walter Pitts were clearly pointing at nearly fifty years ago. We've solved the problem that was put to me by Pitts and Wiener all those years ago. We finished the first paper on this only last week, so it will see the light of day in due course. It uses Wiener path integrals as well as all the machinery of modern statistical mechanics and field theory, and it's made exactly the right contact with physics that I was hoping for and it's relevant to data at every level of analysis. It's a great boon at my age to be in middle of all this new stuff. It might be the Rosetta Stone that unlocks a lot of how large scale brain activity works.

PH: That sounds very exciting, I look forward to reading more about it. Can we just go back a little in that trajectory and talk about your work in pattern formation in neural networks and how it links to Turing?

JC: Well the bulk of that research goes back to 1979 when I was working with another extremely bright graduate student, Bart Ermentrout. I went to a conference that Hermann Haken organized in Germany in 1977 on what he called synergetics – a modern version of cybernetics, but stressing the role of excitation. While at that meeting I realised that Turing's 1952 work on the chemical basis of morphogenesis⁴ could be applied to neural networks. I realised that the stuff I'd done with Hugh Wilson was an analogue of the reaction-diffusion networks that Turing had worked on. There was a very good talk at that meeting by an applied mathematician from the US called David Sattinger showing how to apply the techniques of non-linear analysis, bifurcation theory as it's called, in the presence of symmetry groups, to things like fluid convection. And I realized there was an analogue of that in the nervous system. When I got back I mentioned this to Bart and he immediately saw what I saw.

We realized that we could apply it to the problem of what is going on in the cortex when people see geometric patterns when they are hallucinating. This happens after taking hallucinogens, or through meditation or sometimes in other conditions. The Chicago neuropsychologist, Heinrich Klüver, did a lot of field work in the '60s to classify these types of geometric hallucinations. He mainly experimented on himself, using peyote²⁵. Anyway he discovered that there were only four classes of patterns; they were the same for everyone experiencing these kinds of hallucinations. So we produced a first treatment of why people see these patterns – tunnels, funnels, spirals and honeycombs – in the visual field. Applying the Turing mechanism we showed what kind of neural architecture would spontaneously give rise to these patterns²³ and showed that is was consistent with the neuroanatomy that had been discovered by Hubel and Weisel and others going back to Sholl²⁶. In recent years we've followed

that up, working with Paul Bressloff, Martin Golubitsky and some of my students, and we now have more detailed explanations²⁷. We have a series of papers that will come out in due course that extend the model to cover hallucinations involving colour, depth and motion. We've extended the analysis to look at why people see themselves falling down tunnels with light at the end and so forth. We believe this work tells us quite a lot about what the architecture of the relevant parts of the brain must be like to generate these things. I was at a computational neuroscience and vision conference recently and I discovered that some of the techniques we have introduced in this work may be very relevant to computational vision, and that there may be some deep links between the field equations Wilson and I introduced and problems in vision such as colour matching. So this is a new direction I am going to collaborate in.

PH: I wonder what your views are on the correct level of abstraction for brain modelling. There is an awful lot more known today about some of the low-level biochemical details, but still the higher level overall picture is rather obscure.

JC: It's a very interesting question. We now have at Chicago Stephen Smale, who is a great mathematician – Fields Medallist for his work on the Poincaré conjecture many years ago and many other honours – who has got interested in machine learning and vision recently. He's starting to work with a number of people in these areas and he has a very abstract way of thinking, but a very powerful way. There is a group of mathematicians who work on differential geometry and topology who are getting very interested in what goes on in the nervous system. There are many different levels of mathematical abstraction that can be applied to brain modelling. I think there are going to be rich developments over the coming decades in this area and we may see some rather different styles of modelling emerge than have been used to date.

PH: Historically, one of problems faced by theoretical work in neuroscience is indifference, or sometimes hostility, from the majority of those working in neuroscience. Do you see that changing?

JC: Well this is something I've had to struggle with for nearly fifty years but I think it is changing. Most of the new young people coming through have a different attitude. Many are much better educated than their equivalents were even twenty five years ago. I think more and more biologists will become at least open to mathematics whilst remaining very good empirical scientists. But the fact that experimental tools and methods have become more precise means that there is a lot more data that cries out for mathematical approaches. So I think attitudes are changing.

PH: If you put yourself back at the start of your career, right back to Ferranti, and try and remember your general expectations then, are you surprised at how far machine intelligence has come, or hasn't come?

JC: Well something that has always surprised me is how many times ideas in this field are rediscovered by the next generation. For example I recently heard a very nice lecture from Tommy Poggio, who has been in the game a good while himself, on early vision. He used a mathematical device that actually had been invented in the 1950s by Wilfred Taylor at University College. Tommy wasn't aware of that. A lot of the ideas and machinery that is current now has actually been sitting in the field for a very long time. It's just that we haven't always seen the implications or how to use

them properly. But am I surprised at how difficult it has turned out to do real machine intelligence? No, not at all. I always thought it would be much harder than the people in strong AI claimed. Now back in about 1966, Frank Schmitt, who ran the neuroscience research programme at MIT, organized one of the first meetings on sensory coding and Shannon was at that meeting. I remember Shannon said something very interesting during the meeting. He said that he thought that while initially strong AI might make some interesting progress, in the long run bottom-up work on neural networks would prove to be much more powerful. He was one of the few people at MIT in 1958 who responded positively to a lecture Frank Rosenblatt gave on the perceptron. Most were extremely negative in their response to it, and I have to say it was a pretty bad lecture with too many wild claims, but not Shannon. He said, ‘There could be something in this.’ I consider him to be amazingly perceptive, much more so than most others in the field. Him and McCulloch, Wiener and Pitts.

PH: What do you think are the most interesting developments in machine learning at the moment?

JC: Well there is some very interesting work on data mining that the mathematician Raphy Coifman at Yale and others have been involved in. If you have a very large database from which you want to extract information, you can get a big advantage if you can map the data space onto some lower dimensional manifold in a systematic way. What they found was that simple things like averaging operators, smoothing Laplacian operators, and things connected to diffusion, are immensely powerful for doing that. In a strange way it’s connected to what underlies the solution to the Poincaré conjecture because that involves the smoothing of manifolds and smoothing plays a key role in this work on data mining. I’ve recently proposed that the resting state of the brain is Brownian motion, which is also closely related to that kind of operator. So I think there is something going on in the nervous system and something going on to enable machine learning that may be related and which will prove to be very interesting.

PH: Finally, is there a particular piece of your work that you are most proud of?

JC: Well I think that the work I’m doing now with Michael Buice, which we discussed earlier, and which is the culmination of many years’ work, is what I’m going to end up being most proud of. Even though I’m in my anecdote, as they say, I like to look forward and what I’m doing now I find is most interesting to me.

Notes and references

¹ Beurle, R. L. (1956). *Properties of a Mass of Cells Capable of Regenerating Pulses*. *Phil. Trans. Royal Soc. London B*, **240**:55-94.

² *The First International Congress on Cybernetics*, Namur, Belgium, June 26-29, 1956. Proceedings published by Gauthier-Villars, Paris, 1958.

³ Pringle, J. (1951). On the parallel between learning and evolution. *Behaviour* **3**: 174-215.

⁴ Turing, A.M. (1952). The chemical basis of morphogenesis, *Phil. Trans. R. Soc. London B.* **237**:37-72.

⁵ von Neumann, J. (1956). Probabilistic Logics and the Synthesis of Reliable Organisms from Unreliable Components. In C.E. Shannon & J. McCarthy (Eds.), *Automata Studies*, Princeton: Princeton University Press, 43-98.

⁶ Rosenblith, W.A. (1961). *Sensory Communication*, MIT Press.

⁷ Barlow, H.B. (1961) Possible principles underlying the transformations of sensory messages. In W. A. Rosenblith (Ed), *Sensory Communication*, MIT Press, 217-234.

⁸ Reichardt, W. (1961). Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. In W.A. Rosenblith (Ed.), *Sensory Communication*, MIT Press, 303-317.

⁹ Lettvin, J. Y., Maturana, H. R., McCulloch, W. S., & Pitts, W. H. (1961) Two Remarks on the Visual System of the Frog. In W.A. Rosenblith (Ed.), *Sensory Communication*, MIT Press, 757–776.

¹⁰ Lettvin, J. Y., Maturana, H. R., McCulloch, W. S., & Pitts, W. H. (1959) What the frog's eye tells the frog's brain, *Proceedings of the I.R.E.* **47**: 1940-1959.

¹¹ Selfridge, O.G. (1959), Pandemonium: A paradigm for learning. In Blake, D., Uttley, A., (Eds), *The Mechanisation of Thought Processes. Volume 10 of National Physical Laboratory Symposia*. Her Majesty's Stationery Office, London, 511-529.

¹² Selfridge, O.G. (1948) Some Notes on the Theory of Flutter, *Arch. Inst. Cardiol. Mexico*, **18**, 177.

¹³ S. Winograd and J.D. Cowan (1963). *Reliable Computation in the Presence of Noise*, MIT Press.

¹⁴ J.D. Cowan (1968). Statistical Mechanics of Nervous Nets. In: *Proceedings of 1967 NATO Conference on Neural Networks*, E.R. Caianiello (Ed.), Springer-Verlag, 181-188.

¹⁵ The sigmoid function is a differentiable non-linear ‘squashing’ function widely used as the transfer function in nodes in artificial neural networks (to compute node output from input). It turns out that this kind of function is necessary for various multi-layered learning methods to work, including the back-propagation method (see note 21); the original perceptron used linear transfer functions.

¹⁶ Hopfield, J. (1982) Neural Networks and Physical Systems with Emergent Collective Computational Abilities, *Proc. Nat. Academy of Sciences*, **79**:2554-2558.

¹⁷ Uttley, A. (1956) Conditional Probability Machines and Conditioned Reflexes. In C.E. Shannon & J. McCarthy (Eds.), *Automata Studies*, Princeton: Princeton University Press, 253-275.

¹⁸ C.H. Waddington (Ed). *Towards a Theoretical Biology, vol. I: Prolegomena*, Edinburgh University Press, 1968. [Several volumes in this series were produced]

¹⁹ Nicolas Rashevsky, a Russian physicist who arrived in America after various scrapes and near escapes during the civil war in his home country, set up the Committee on Mathematical Biology at the University of Chicago in the 1930s. An influential pioneer in mathematical biology, he mentored many important theoretical biologists. He set up and edited *The Bulletin of Mathematical Biophysics* which, among other notable works, published the pioneering papers by McCulloch and Pitts on neural networks.

²⁰ Minsky, M. L. and Papert, S. A. (1969) *Perceptrons*, MIT Press.

²¹ Rumelhart, D.E. and Hinton, G.E. and Williams, R.J. (1986) Learning representations by back-propagating errors, *Nature*, **323**:533-536. [This method for learning in multi-layer networks, which overcame the limitations of perceptrons, had been independently described previously by Paul Werbos in his 1974 Harvard PhD Thesis, and a similar method had been proposed by Shun-ichi Amari in 1967]

²² H.R. Wilson and J.D. Cowan (1972). Excitatory and Inhibitory Interactions in Localized Populations of Model Neurons. *Biophysical J.*, **12**: 1-24.

²³ H.R. Wilson and J.D. Cowan (1973). A Mathematical Theory of the Functional Dynamics of Cortical and Thalamic Nervous Tissue. *Kybernetik*, **13**: 55-80.

²⁴ G.B. Ermentrout and J.D. Cowan (1979). A Mathematical Theory of Visual Hallucination Patterns. *Biological Cybernetics*, **34**: 137-150.

²⁵ Klüver, H. (1966) *Mescal and the Mechanisms of Hallucination*, Chicago: University of Chicago Press.

²⁶ Sholl, D.A. (1956) *The Organization of the Nervous System*, McGraw-Hill.

²⁷ Bressloff, P. and Cowan, J. and Golubitsky, M. and Thomas, P. and Wiener, M. (2001) Geometric visual hallucinations, Euclidean symmetry and the functional architecture of striate cortex, *Phil. Trans. Royal Soc. London B*, **356**:299-330.