Neural Networks 23 (2010) 1135–1138

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

Editorial

Towards building a neural networks community

December 31, 2010 is my last day as the founding editor-in-chief of *Neural Networks* and also my 71st birthday! Such vast changes have occurred in our field since the first issue of *Neural Networks* was published in 1987 that it seemed appropriate to my editorial colleagues for me to reflect on some of the changes to which my own efforts contributed. Knowing where we as a community came from may, in addition to its intrinsic historical interest, be helpful towards clarifying where we hope to go and how to get there. My comments will necessarily be made from a personal perspective.

The founding of *Neural Networks* is intimately linked to the founding of the International Neural Network Society (INNS; http://www.inns.org) and of the International Joint Conference on Neural Networks (IJCNN; http://www.ijcnn2011.org). These events, in turn, helped to trigger the formation of other groups of scientists and engineers who are interested in our field, as well as other conferences to which they could contribute to share their ideas.

There was little in the way of infrastructure in our community thirty years ago. It was not obvious where to submit a modeling article for publication, or what conferences to attend where one could report work to interested colleagues. It was also not clear where one could go to school to learn how to model, or even to learn about was known in the field.

How did I come to be in a position to contribute to these developments? That happened in gradual stages over a period of over fifty years. My own beginnings in our field started accidentally in 1957, thirty years before *Neural Networks* was founded, when I was a 17 year-old Freshman at Dartmouth College. At that time, like thousands of other Freshmen, I took an introductory psychology course. The accident was my exposure to classical data about how humans and animals learn about the world. To me, these data were filled with philosophical paradoxes. And 17 year-olds of a certain disposition are particularly vulnerable to philosophical paradoxes!

For example, data about serial verbal learning in humans – that is, the learning of lists of language or action items through time – seemed to suggest that events can go "backwards in time". Indeed, the *non*-occurrence of future items in a list to be learned can dramatically change the distribution of performance errors at all previously occurring list positions. The so-called "bowed serial position curve" showed that the beginning and the end of a list can be learned more easily than the middle, with error gradients in performance that point forwards or backwards in time at the beginning or end of a list, and error gradients in the middle pointing in both directions. In what sort of space and time could events go backwards in time? What type of dynamics could represent the non-occurrence of an event? Exposure to such paradoxes triggered a passionate intellectual excitement that I had never before experienced.

Data about reinforcement learning in animals and about cognitive dissonance in humans kindled a similar passion. What sort of space and time representations could describe both a thought and a feeling? How could thoughts and feelings interact to guide our decisions and actions to realize valued goals? Just as in the case of serial learning, the non-occurrence of an event could have dramatic consequences. For example, the non-occurrence of an expected reward can trigger reset of cognitive working memory, emotional frustration, and exploratory motor activity to discover new ways to get the reward. What is an expectation? What is a reward? What type of dynamics could unify the description of cognition, emotion, and action? Within the reinforcement learning literature, there were also plenty of examples of irrational behaviors. If evolution selects successful behaviors, then why are so many behaviors irrational?

Such issues are still of current research interest. And all of them led me to the same theoretical method and modeling framework. The theoretical method, which I called the Method of Minimal Anatomies, showed how analysis of how an individual's behavior adapts autonomously in real time to a changing world can lead to models of how the brain works. In other words, brains embody a natural computational architecture for autonomous adaptation in real time to a changing world. By bridging the gap between behavioral experience and brain dynamics, the method opened a path towards solving the mind-body problem.

This method led me to introduce networks of neurons in which short-term memory (STM), medium-term memory (MTM), and long-term memory (LTM) traces exist. The STM and LTM traces are also known as cell activities, or potentials, and adaptive weights. In such a network, backward effects in time became easy to understand. These traces were described within Additive and Shunting Models that lie at the core of essentially all contemporary connectionist models of mind and brain. The Additive Model is now sometimes called the Hopfield model based on Hopfield's popular 1984 article on this topic. When I introduced the model in 1957–58, both the paradigm and the equations were new, and even revolutionary.

The LTM laws to which I was led included gated steepest descent learning laws that allow learned increases and decreases in weights, in order to learn spatially distributed weight patterns. From the outset, it was clear that the Hebb hypothesis that learned weights can only increase was incorrect. When in the 1970s von der Malsburg and I introduced self-organizing maps, such a gated steepest descent law was used in it, and Kohonen adapted this law in his applications of self-organizing maps in the 1980s.

The MTM law described how signals can be gated, or multiplied, by chemical transmitters or post-synaptic sites subject to activity-dependent habituation. This law started to become broadly used

0893-6080/$ – see front matter
doi:10.1016/j.neunet.2010.10.006

in the 1990s due to the work of Abbott, Markram, and their colleagues, often under the name of synaptic depression.

Thus, while I was in college, it became clear that, to link mind and brain, one needed nonlinear neural networks operating on multiple temporal and spatial scales. There was nothing like this going on in psychology departments then, which is why, with some trepidation, I decided to get my Ph.D. in mathematics. When I graduated from Dartmouth in 1961, the best research group in mathematical psychology, and one of the best departments in applied mathematics, were at Stanford University. I therefore went to Stanford to begin my graduate studies. While at Stanford, I tried to get my nonlinear neural networks simulated with the help of one of the best programmers at Stanford's computation center, without success.

Computers in those days were not up to the task of simulating nonlinear neural networks operating on multiple time scales. The learning in such networks made them nonstationary in time. The long-range interactions between neurons could be said to be nonlocal. In other words, computers in those days had problems simulating large networks of neurons undergoing nonlinear, nonstationary, and nonlocal interactions—what I called the Three N's of brain dynamics. My mathematical training helped me to overcome this roadblock by giving me enough tools to prove foundational theorems in the 1960s and 1970s about how STM, MTM, and LTM interact, including the first global theorems about content addressable neural memories, and theorems about how the bowed serial position curve arises.

The theoretical method also facilitated technology transfer to applications in neuromorphic engineering and technology. It did this by clarifying how brain mechanisms give rise to behavioral functions while an individual autonomously adapts to a changing world. Many outstanding problems in engineering and technology need to represent and autonomously control a changing world. And all applications need to link how a device works (its mechanism) with what it is for (its function). By helping to discover how brain mechanisms give rise to behavioral functions for autonomous adaptation to a changing world, the method discovered models that could be of use in technology.

When I arrived at Stanford in 1961, the statistical learning models of William Estes and his colleagues were being used to explain psychological data. Statistical approaches continue to the present day with the resurgence of Bayesian models. When I began my own theoretical struggles, I first tried to express my intuitions in term of statistical models. The failure of these models to explain how an individual autonomously adapts in real time to a changing world gradually let me to nonlinear neural networks with STM, MTM, and LTM traces. Indeed, one of my first discoveries about the Shunting Model was how distributed patterns of STM and LTM traces could self-normalize, much as the probabilities in a probability distribution. Neural dynamics could also carry out statistical prediction and decision-making. Thus, biological neural networks clarified why classical statistical methods were insufficient to understand the brain, but also why it was tempting to try to use them to do so.

These introductory remarks aim to illustrate how biological neural models represented a radical break with previous scientific methods, indeed a new paradigm for understanding biological intelligence. These models embodied both new intuitions and new mathematics to explain how autonomous intelligence adapts to a changing world. Paradigms in which new intuitions and new mathematics must simultaneously be developed are among the hardest ones to understand. In the case of mind and brain, they also required an interdisciplinary synthesis that crossed the disciplinary boundaries of psychology, neuroscience, mathematics, computer science, and even physics and philosophy.

Because of this break, suddenly, despite the excitement of continuous scientific discovery, I felt quite alone. I read scientific history to try to understand why it seemed so hard for people to understand discoveries which, at least to me, seemed natural and intuitively appealing. I gradually began to realize that I was contributing to a major scientific revolution whose groundwork was laid by great nineteenth physicists such as Hermann von Helmholtz, James Clerk Maxwell, and Ernst Mach. These interdisciplinary scientists were physicists as well as psychologists and physiologists. Their discoveries made clear that understanding the Three N's of the brain would require new intuitions and mathematics. As a result, the next generation of physicists exclusively pursued the major new paradigms of relativity theory and quantum theory, which could build on great new physical intuitions that were supported by known mathematics. Psychologists and physiologists were left with inadequate intuitions and mathematics with which to explain their data. Given inadequate tools with which to understand the Three N's, a century of controversy, along with anti-theoretical prejudices, unfolded in the mind-brain sciences as they collected huge, but typically unexplained, data bases.

Although these historical insights helped me to understand, and emotionally cope with, the frightening social forces to which I was exposed as a young man, it was still difficult to deal with them. Had I not been first in my class in school for many years, and were it not a time when the United States was investing large sums of money in student fellowships, I may not have survived the social pressures against the kind of research that I was doing.

This loneliness, combined with my passionate belief in the importance of this new paradigm shift, drove me to do everything that I could, when opportunities presented themselves, to build a community wherein future students and practicing scientists could readily learn about our field, as if it were the most natural thing in the world.

One of the most important things to provide in a paradigm shift is a place to publish new interdisciplinary work. I personally had a terrible time at first getting even my most important articles published. Editors of journals such as *Vision Research*, and *Psychological Review* routinely sent back my articles without review, saying that they did not publish "that sort of thing". I kept trying, and replied to these editors that it made little sense to publish data if articles that could quantitatively explain the data were not also published. One measure of the great change that has occurred is the large number of articles that I, and other modelers, have since published in these journals. Despite the gradual opening of disciplinary journals to interdisciplinary modeling research, these experiences made clear that new journals devoted entirely to behavioral and neural modeling articles were required. I founded *Neural Networks* to provide one such journal.

Other crucial needs of a new scientific community are interdisciplinary conferences at which to report new results, and interdisciplinary centers, programs, and departments at which students can learn about these results. What I want to emphasize here is that such developments often occur gradually, over a period of years or even decades, while the infrastructure needed to carry them out can get laid down. In addition, these various activities are often symbiotic and mutually energizing. Although this developmental process has already undergone several stages, it is not yet complete.

The earliest conferences that I helped to organize, starting in 1980, often piggy-backed on established societies, such as the American Mathematical Society, the Society for Industrial and Applied Mathematics, the Society for Mathematical Biology, the Society for Mathematical Psychology, and the Psychonomic Society, with funding from government agencies such as AFOSR and NSF. Gail Carpenter worked closely with me in organizing various of these conferences.

We were able to efficiently organize these conferences because I managed to establish the Center for Adaptive Systems (CAS) at

Boston University in 1982, with a permanent charter from the BU Trustees, in order to train postdoctoral fellows in the neural network field. My ability to do so depended upon the great good luck that John Silber was President of Boston University at that time, and had appointed wonderful administrators, notably Dennis Berkey, who served as Dean and Provost. These administrators allowed CAS to form and, even more remarkably, they supported my later efforts to found a Graduate Program in Cognitive and Neural Systems in 1988, and then a full-standing Department of Cognitive and Neural Systems (CNS) in 1991.

CNS obviously filled a real need, since it rapidly became one of the best departments of its kind in the world, attracting many gifted students who either wanted to understand how the brain works on its own terms, or wanted to study mind and brain as a source of new design insights for solving problems in neuromorphic engineering and technology. A key to the success of CNS was our development of a coherent interdisciplinary curriculum of 18 advanced undergraduate and graduate courses, each of which introduced psychological and neurobiological data, theoretical concepts, mathematical models, and computational methods in a key area of biological intelligence. The curriculum was, and remains, a jewel in our crown. It was aimed at trying to answer two key questions in a coordinated way: How does the brain control behavior? and How can technology emulate biological intelligence? This curriculum was designed to complement curricula in traditional departments, and our students took both CNS interdisciplinary courses and more traditional courses. They became highly prized by faculty in other departments, and went on to get excellent jobs in a wide range of fields related to biological and neuromorphic intelligence.

One of the things that we rapidly learned was that students who came with an interest in technology often became our best biological modeling students, and conversely. The trick was to create a community in which students and faculty could move freely between basic science and applications. This realization motivated a lot of our subsequent infrastructure development.

Gail Carpenter was founding Co-Director of CAS and Director of Graduate Studies of CNS. Without the many contributions by Gail and a few other individuals, I doubt that I could have done all that we managed to accomplish. Indeed, both CAS and CNS played a major role in the founding of INNS and IJCNN by providing the personnel who could help to support those activities. The CAS Assistant Director, Cindy Bradford, made major contributions by providing flawless leadership and support in scientific administration and conference planning. Cindy also became the founding Editorial Assistant of *Neural Networks* and still plays a major role in that capacity today, coordinating the three *Neural Networks* editorial offices of INNS, the European Neural Network Society (ENNS), and the Japanese Neural Network Society (JNNS). Cindy is also well known today to participants in two other activities that I later founded, again building on CAS/CNS infrastructure: the annual International Conference on Cognitive and Neural Systems (ICCNS: http://cns.bu.edu/cns-meeting/conference.html), which is having its fifteenth annual conference at CNS in May, 2011, and the Center of Excellence for Learning in Education, Science, and Technology (CELEST: http://celest.bu.edu/), of which she is the Administrative Director.

In parallel with these developments were activities led by other colleagues. I am proud of the fact that a number of these colleagues had been introduced to neural network research through my work. An important Tutorial Conference on Neural Modeling was organized in 1983 by Robert Hecht-Nielsen, David Hestenes, and Peter Killeen in Scottsdale, Arizona, with support by AFOSR, ONR, and Arizona State University. At this conference, I lectured all morning for a week on my research with the task of bridging all of the topics that other leading modelers would talk about in the afternoons. Around 50 modelers came to the meeting, which was considered a large meeting at that time! Attendees at that meeting included many future leaders of what came to be called the Connectionist Revolution.

Another parallel activity began in 1985, when I gave a year-long lecture series at MIT Lincoln Laboratory about the models that I and my colleagues had been developing at CAS. This invitation arose when several Lincoln Lab group leaders heard me speak about our models of how the brain sees at the annual meeting of the Optical Society of America. They thought that these models might solve some of the problems that they were having in processing data from artificial sensors, such as laser radar, synthetic aperture radar, infrared, and the like. This has turned out to be correct. My lectures also inspired Lincoln Laboratory to initiate the DARPA National Study on Neural Networks from 1987–1989, which had a large impact on organizing funding for neural network research.

With the experience of organizing a series of smaller interdisciplinary meeting behind us, we could start to plan meetings to reach a larger community. To this end, Bart Kosko and Robert Hecht-Nielsen initiated the planning of an international conference with IEEE. Since my work in neural networks strongly influenced their own research, I was asked by them to serve as General Chairman of what became the first International Conference on Neural Networks (ICNN), that was held in San Diego in 1987. I again worked closely with Gail Carpenter and our CAS colleagues to organize the conference program and publicity. Our conference work led to an experience in conference planning that made clear where the buck stops when you are in charge. The budget for ICNN'87 program development was basically nil, and various people at IEEE expressed their doubts that ICNN could be brought off at all. In order to get the word out, Gail and I put together the conference brochure and program, stuffed thousands of brochures into envelopes, and went to the post office near our home in Newton Highlands, MA, on the day before Christmas in 1986 to have the post office put stamps on the envelopes. Our request was denied, so we spent that afternoon licking stamps "for the cause". As it turned out, ICNN'87 was a brilliant success with almost 2000 people in attendance.

As ICNN was being planned, I came increasingly to feel that our field needed its own society to build a neural modeling community with democratic procedures for electing its leaders. Such a society would already have the new *Neural Networks* journal to serve its members and the world community. I therefore founded the International Neural Network Society in 1987, became its first President, and invited leading neural network researchers of multiple persuasions to form its Board of Governors. The initial INNS Board of Governors read like a Who's Who of leading researchers. Harold Szu, a long-time friend from my graduate days at The Rockefeller University, where I finished my Ph.D. after Stanford, was a great help in planning the initial INNS infrastructure. With some chutzpa, I announced the formation of INNS at my plenary lecture at ICNN'87. INNS went on to stimulate the formation of ENNS and JNNS. All three societies today share *Neural Networks* as their archival society journal.

The final step in forming INNS was to plan an annual INNS meeting that would serve as an international forum where people from multiple disciplines, ranging from the biological to the technological, could regularly come together to exchange ideas and results. Gail and I cooperated again to plan the first annual INNS meeting, which was held in Boston in 1988, and drew a large and enthusiastic audience. Fourteen other societies agreed to cooperate with INNS in this venture, thereby supporting the interdisciplinary community that INNS hoped to form. My introductory remarks at the INNS'88 meeting noted that, during the 14 months of my INNS presidency, 3071 individuals joined INNS at a steady rate of 200 members a month. At that time, there was no sign of saturation in the INNS growth rate. There were members from 38 countries around the world, and 49 states of the United States, with 20% of the members in the life sciences, 19% from the computer and information sciences, 27% from the different branches of engineering sciences, 2% from business, and the remaining 7% from a variety of other fields. This is the sort of distribution of practitioners that INNS seeks to maintain today.

IJCNN arose as a fusion of the ICNN and INNS annual meetings in 1989. IJCNN is currently the largest meeting in the world that is devoted to neural network research, in all of its manifold manifestations.

A critical problem for our field in the 1980s was to provide interdisciplinary education whereby to improve modeling literacy. CAS contributed towards solving this problem in a small way by training a series of gifted postdoctoral fellows. However, it became clear that, no matter how smart these individuals were, it was not possible to give them a broad enough training so late in their careers. A new interdisciplinary curriculum was needed to more fully train scientists in the biology and technology of neural modeling, and to do so at a sufficiently early stage of their careers. After several years of administrative negotiations and approvals, I managed, as noted above, to found the Graduate Program in Cognitive and Neural Systems (CNS) in 1988, which became the Department of Cognitive and Neural Systems in 1991. Several other interdisciplinary departments also began to get founded during the ensuing years, but most did not then have, and still do not have, a systematic curriculum aimed at training students in advanced biological and technological neural modeling.

Thus, even today, although our field now has a lot of valuable infrastructure of which it can be proud, one of its critical problems is how to improve interdisciplinary education and literacy. The revolution of the Three N's is not over! I believe that multiple approaches should continue to be tried, and that *Neural Networks*, INNS, IJCNN, and other programs, departments, and conferences in our field should facilitate these activities through their coordinated action. We need more development of undergraduate and graduate interdisciplinary programs with in-depth modeling curricula; summer schools where the most important modeling breakthroughs of the past year are presented in detail for evaluation and discussion by other modelers, independent of their clique affiliations; multi-hour tutorials about the most important modeling concepts at international conferences, including the principal conferences attended by experimental psychologists, neurobiologists, or technologists; workshops where practitioners of different modeling approaches come together for extended and frank comparative discussions of the strengths, weaknesses, and opportunities offered by their respective approaches; and creation of web-based curricula that begin with exciting data to capture the imagination of students of all ages, and then progress in small steps from intuitive to increasingly mathematical explanations of these data using behavioral and neural models.

I am personally looking forward to continuing my research on understanding mind and brain, to which I remain as passionately committed today as I was when I began it 53 years ago. I also will help in whatever ways I can to facilitate the healthy maturation of our exciting and transformative field. I thank my colleagues across the world for many years of collegial and productive cooperation towards reaching this worthy goal.

Founding Editor-in-Chief
Stephen Grossberg
Boston University
E-mail address: steve@bu.edu.
URL: http://cns.bu.edu/~steve.