Artificial Neural Networks - A Science in Trouble

Asim Roy
Arizona State University
School of Information Management
Tempe, Arizona 85287-3606, USA
Phone: (480) 965-6324
Email: asim.roy@asu.edu

ABSTRACT

This article points out some very serious misconceptions about the brain in connectionism and artificial neural networks. Some of the connectionist ideas have been shown to have logical flaws, while others are inconsistent with some commonly observed human learning processes and behavior. For example, the connectionist ideas have absolutely no provision for learning from stored information, something that humans do all the time. The article also argues that there is definitely a need for some new ideas about the internal mechanisms of the brain. It points out that a very convincing argument can be made for a "control theoretic" approach to understanding the brain. A "control theoretic" approach is actually used in all connectionist and neural network algorithms and it can also be justified from recent neurobiological evidence. A control theoretic approach proposes that there are subsystems within the brain that control other subsystems. Hence a similar approach can be taken in constructing learning algorithms and other intelligent systems.

Keywords

Connectionism, artificial neural networks, brain-like learning, data mining, intelligent systems, automated learning.

1. INTRODUCTION

The fields of data warehousing, data mining, robotics and intelligent engineering systems are interested in tools that can ***automate the process of*** knowledge discovery and learning of rules from data. "***Automate the process***" in this context implies tools and algorithms that ***obviate the need for external human intervention of any kind***, beyond the specification of the problem at hand, in order to guide the detailed algorithmic processes. In the case of unsupervised learning or clustering, "*specification of the problem*" would mean providing a body of data that is to be grouped into classes. In the case of supervised learning for pattern classification, "*specification of the problem*" would mean providing a body of data that has the classification information for each data point and, perhaps, nothing else, unless classification errors of a certain type are to be avoided or reduced. This *model of problem specification* reflects the way humans learn; humans learn rules and discover knowledge based on this type of information and nothing more. *H**umans don't require any outside control of the learning processes in the brain.*** Because of these characteristics (of human learning), the fields of data mining, robotics and intelligent engineering systems look towards the science of artificial neural networks and connectionism to provide them with the tools for human-like automated, autonomous learning. Unfortunately, in spite of decades of research, the science of artificial neural networks and connectionism failed to deliver those tools. ***This article is about why they failed, pointing out the major flaws in their basic underlying concepts and assumptions***.

A glaring and fundamental *weakness* in the current theories of artificial neural networks and connectionism is the total ***absence of the concept of an autonomous system***. As a result, the field developed learning algorithms over the years that work well only when there is ***human intervention***. In other words, their learning algorithms need ***constant baby-sitting*** in order for them to work - learning rates need to be reset and readjusted, various network designs need to be tried so that they can generalize well, starting points need to be reset when they get stuck in a local minimum, and everything needs to be relearned from scratch when there is catastrophic forgetting in the network. And the list goes on. These are some of the common problems in both supervised and unsupervised neural network learning. *One cannot even imagine including these learning algorithms in future robots and other knowledge discovery and learning systems, whatever kind they maybe, that are supposed to be autonomous and learn on their own. One cannot have autonomous intelligent systems without autonomous learning algorithms.*

1.1 The Debate about How the Brain Learns

Can a whole body of science simply unravel when confronted by a few simple questions? Can a large body of scientists overlook some very simple facts for a long period of time? From the current debate on how the brain learns, the answer appears to be "yes" for at least one body of science – artificial neural networks and connectionism. A sampling of some recent comments might be a good indicator of this. The first open and public admission that much of the existing science is wrong came from Christoph von der Malsburg, a German neurobiologist and computer scientist affiliated with both the Ruhr University of Germany and the University of Southern California. In commenting on the challenge I posed, which claimed that neural networks do not embody brain-like learning, he remarked,"*I strongly believe that the current paradigm of neural network learning misses very essential aspects of the learning problem, and I totally concur with the assessment, expressed in your expose, that specific prior knowledge is required as a basis for learning from a given domain….I am glad the issue seems to start picking up momentum. Some Kuhnian revolution is required here, and as* *he (T. Kuhn) wrote, such scientific revolutions are always preceded by rising unrest in the community..*" And Malsburg is one of the founders of this field. Another founder of this field and a past president of the International Neural Network Society (INNS) confided to me that "*the neuro-boom is over*." But many other scholars have kept on fighting the arguments against the current science on brain-like learning. Another past president of INNS publicly disagreed with me at the recent debate in Alaska, saying: *"In brief, I disagree with everything he (Asim Roy) said."*

Many distinguished scholars have participated in the three open, public debates at the last two international conferences on neural networks. These debates centered on various aspects of brain-like learning as discussed later in this article. The first debate at the International Conference on Neural Networks (ICNN'97) in Houston, Texas in June, 1997, included four past presidents of INNS and five of the plenary speakers. The second debate at the World Congress on Computational Intelligence (WCCI'98) in Anchorage, Alaska in May, 1998, included five past presidents of INNS. The third debate at the International Joint Conference on Neural Networks (IJCNN'99) in Washington, D.C., in July, 1999 included several neuro and cognitive scientists, including three past presidents of INNS. A summary of the first debate has been published in the INNS Newsletter of May, 1998 [13], and on the Internet through various neural network-related mailing lists.

The debate about connectionism is nothing new. The argument between the symbol system hypothesis of artificial intelligence and the massively parallel system conjecture of artificial neural networks or connectionism has still not abated. Marvin Minsky of MIT characterized connectionism as "naïve" at the first international conference on neural networks in San Diego in 1988. And Minsky and Seymour Papert not only showed the limitations of the earlier simple neural networks, the perceptrons, but were also the first ones to raise the deeper question of computational complexity of learning algorithms ("Epilogue: The New Connectionism" in [8]). But the neural network field moved along heedlessly with its research agenda, ignoring all the deeper and more disturbing questions raised by thoughtful critics. However, a scientific field is destined to stumble sooner or later when it tries to skirt legitimate questions about its founding ideas. Now faced with fundamental challenges to the assumptions behind their brain-like learning algorithms, prominent researchers in the field are finally calling for a "*shake up of the field of neural networks*" and for its "*rebirth*."

2. SOME BACKGROUND INFORMATION ON ARTIFICIAL NEURAL NETWORKS

Connectionism or artificial neural networks (ANN) is the field of science that tries to replicate brain-like computing. The brain is understood to use massively parallel computations where each computing element (a neuron or brain cell in the terminology of this science) in the massively parallel system is envisioned to perform a very simple computation, such as $y_i = f(z_i)$, where z_i is assumed to be a real valued input, y_i is either a binary or a real valued output of the i^{th} neuron, and f a nonlinear function (see Figure 1). The nonlinear function f, also called a node function, takes different forms in different models of the neuron; a typical choice for the node function is a step function or a sigmoid function. The neurons get their input signals from other neurons or from external sources such as various organs of the body like the eyes, the ears and the nose. The output signal from a neuron may be sent to other neurons or to another organ of the body.

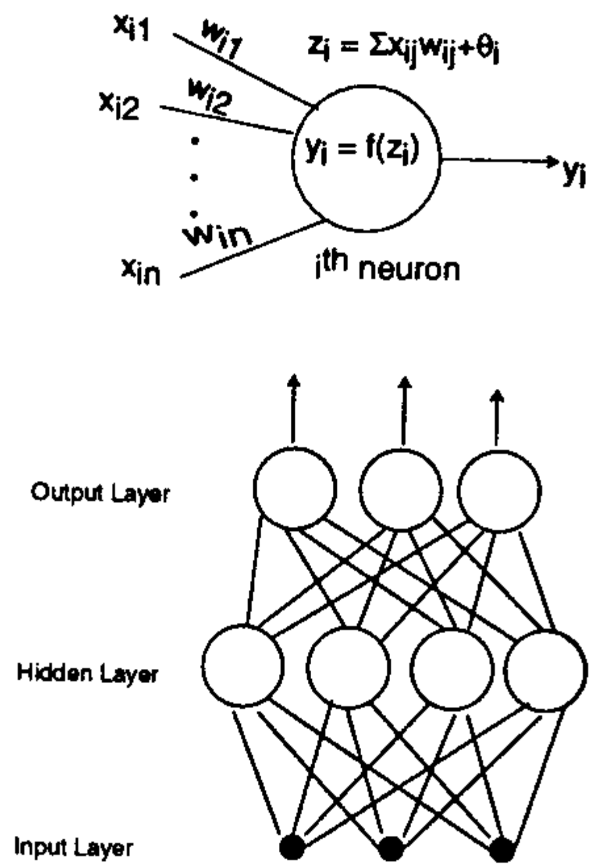

Figure 1

Studies in neuroscience and neurobiology show that different parts of the brain perform different tasks such as storage of short or long term memory, language comprehension, object recognition and so on. A particular task is performed by a particular network of cells (hence the term neural networks) designed and trained for that task through the process of learning or memorization. These networks, when invoked to perform a particular task, then send their outputs to other parts of the brain or to an organ of the body.

A network can have many layers of neurons, where the outputs of one layer of neurons become the inputs to the next layer of neurons. And a network can have more than one output signal; thus the output layer can have more than one neuron. Different neural network models assume different modes of operation for the network, depending somewhat on the function to be performed. A neural network model for pattern classification is often conceived to be a feedforward type network where the input signals are propagated through different layers of the network to produce outputs at the output layer. On the other hand, a neural network model for memory is often conceived to be of the feedback type (also called a recurrent network or a nonlinear dynamical system) where the outputs of the network are fed back to the network as inputs. This process of feedback continues until the network converges to a stable set of output values or continuously cycles among a fixed set of output values.

Let $\mathbf{x_i} = (x_{i1}, x_{i2}, \ldots, x_{in})$ be the vector of input signals to the i^{th} neuron, the inputs signals being from other neurons in the network or from external sources. Neural network models assume that each input signal x_{ij} to i^{th} neuron is "weighted" by the strength of the i^{th} neuron's connection to the j^{th} source, w_{ij}. The weighted inputs, $w_{ij}\ x_{ij}$, are then summed to form the actual input z_i to the node function f at the i^{th} neuron: $z_i = \Sigma\ w_{ij}\ x_{ij} + \theta_i$, where θ_i is a constant, called the threshold value. As mentioned before, some typical node functions are (1) the step function, where $f(z_i) = 1$ if $z_i \geq 0$, and $f(z_i) = 0$ otherwise, and (2) the sigmoid function, where $f(z_i) = 1/(1 + e^{-z_i})$.

A network of neurons is made to perform a certain task (memory, classification and so on) by designing and training an appropriate network through the process of learning or memorization. The **design** of a network involves determining (a) the number of layers to use, (b) the number of neurons to use in each layer, (c) the connectivity pattern between the layers and neurons, (d) the node function to use at each neuron, and (e) the mode of operation of the network (e.g. feedback vs. feedforward). The **training** of a network involves determining the connection weights $[w_{ij}]$ and the threshold values $[\theta_i]$ from a set of training examples. For some learning algorithms like back-propagation [14,15], the design of the network is provided by the user or some other external source. For other algorithms like Adaptive Resonance Theory (ART) [5], reduced coulomb energy (RCE) [10], and radial basis function (RBF) networks [9], the design of the network is accomplished by the algorithm itself, although other parameter values have to be supplied to the algorithm on a trial and error basis to perform the design task.

The **training** of a network is accomplished by adjusting the connection weights $[w_{ij}]$ by means of a **local learning law**. A local learning law is a means of gradually changing the connection weights by an amount Δw_{ij} after observing each training example. A learning law is based on the general idea that a network is supposed to perform a certain task and that the weights have to be set such that the error in the performance of that task is minimized. A learning law is **local** because it is conceived that the individual neurons in the network are the ones making the changes to their connection weights or connection strengths, based on the error in their performance. Local learning laws are a direct descendent of the idea that the cells or neurons in the brain are autonomous learners. The idea of "autonomous learners" is derived, in turn, from the notion that there is no homunculus or "a little man" inside the brain that "guides and controls" the behavior of different cells in the brain. The "no homunculus" argument says that there couldn't exist a distinct and separate physical entity in the brain that governs the behavior of other cells in the brain. In other words, as the argument goes, there are no "ghosts" in the brain. So any notion of "extracellular control" of synaptic modification (connection weight changes) is not acceptable to this framework. Many scientists support this notion (of cells being autonomous learners) with examples of physical processes that occur without any external "control" of the processes, such as a hurricane.

So, under the connectionist theory of learning, the connection weight $w_{ij}(t)$, after observing the t^{th} training example, is given by: $w_{ij}(t) = w_{ij}(t-1) + \Delta w_{ij}(t)$, where $\Delta w_{ij}(t)$ is the weight adjustment after the t^{th} example and is determined by a local learning law. Donald Hebb [6] was the first to propose a learning law for this field of science and much of the current research on neural networks is on developing new learning laws. There are now hundreds of local learning laws, but the most well-known among them are back-propagation [14,15], ART [5] and RBF networks [9]. To give an example, the back propagation learning law is as follows: $\Delta w_{ij}(t) = -\ \eta(\partial E/\partial\ w_{ij}(t)) + \alpha\Delta w_{ij}(t-1)$. Here η is the learning rate (step size) for the weight update at step t and α is a momentum gain term. E is the mean-square error of the whole network based on some desired outputs, in a supervised mode of learning, where a teacher is present to indicate what the correct output should be for any given input. Back-propagation is a steepest descent algorithm and $-\partial E/\partial\ w_{ij}(t)$ is the steepest descent direction (negative of the gradient).

2.1 The Distinction between Memory and Learning

Two of the main functions of the brain are memory and learning. There are of course many categories of memory (short term, medium term, long term, working memory, episodic memory and so on) and of learning (supervised, unsupervised, inductive, reinforcement and so on). In order to characterize the learning behavior of the brain, it is necessary to distinguish between these two functions. Learning generally implies learning of rules from examples. Memory, on the other hand, implies simple storing of facts and information for later recall (e.g. an image, a scene, a song, an instruction). Sometimes these terms are used interchangeably in the literature, and in everyday life: memory is often confused with learning. But the processes of memorization are different from that of learning. So memory and learning are not the same.

2.2 Learning or Generalization from Examples

Learning of rules from examples involves generalization. Generalization implies the ability to derive a succinct description of a phenomenon, using a simple set of rules or statements, from a set of observations of the phenomenon. So, in this sense, the simpler the derived description of the phenomenon, the better is the generalization. For example, Einstein's $E = MC^2$ is a superbly succinct generalization of a natural phenomenon. And this is the essence of learning from examples. So any brain-like learning algorithm must exhibit this property of the brain - the ability to generalize. That is, it must demonstrate that it makes an explicit attempt to generalize and learn. In order to generalize, the learning system must have the ability to design the appropriate network.

3. THE PROBLEMS OF CONNECTIONISM - SOME MAJOR MISCONCEPTIONS ABOUT THE BRAIN

3.1 A Misconception - No Synaptic Change Signals are allowed to the Cells from Other Sources within the Brain

The notion that each neuron or cell in the brain is an “autonomous/independent learner” is one of the fundamental notions of this field. Under this notion, it is construed that individual cells modify their synaptic strengths (connection weights) solely on the basis of their “input and output behavior.” The input and output information of a cell may include information about the error in the performance of a given task by the network and an individual cell’s contribution to that error; see for example the back-propagation learning law in the last section. This notion implies that no other physical entity external to the cell is allowed to “signal” it to adjust its connection strengths in a certain way. All of the well-known local learning laws developed to date most faithfully adhere to this notion [2, 3, 5, 6, 7, 9, 10, 14, 15]. However, there is no neurobiological evidence to support this premise. In fact, there is a growing body of evidence that says that extrasynaptic neuromodulators influence synaptic adjustments “directly” [7]. The neurobiological evidence shows that there are many different neurotransmitters and receptors and many different cellular pathways for them to affect cellular changes. Cellular mechanisms within the cell are used to convert these “extracellular” signals into long-lasting changes in cellular properties. So the connectionist conjecture that no other physical entity directly signals changes to a cell’s behavior is a major misconception about the brain. Beyond the neurobiological evidence, this conjecture is also logically inconsistent, as discussed later.

3.2 Another Misconception - The Brain Does Not Collect and Store Any Information about the Problem Prior to Actual Learning

In connectionism, brain-like learning algorithms cannot store any training examples (or any other information, for that matter) explicitly in its memory - in some kind of working memory, that is, that can be readily accessed by the learning system in order to learn [2, 3, 5, 6, 7, 9, 10, 14, 15]. The learning system can use any particular training example presented to it to adjust whatever network it is learning in, but must forget that example before examining others. This is the so-called “memoryless learning” property, where no storage of facts/information is allowed. The idea is to obviate the need for large amounts of memory to store a large number of training examples or other information. Although this process of learning is very memory efficient, it can be very slow and time-consuming, requiring lots of training examples, as demonstrated in [11,12]. However, the main problem with this notion of memoryless learning is that it is completely inconsistent with the way humans actually learn; it violates very basic behavioral facts. Remembering relevant facts and examples is very much a part of the human learning process; it facilitates mental examination of facts and information that is the basis for all human learning. And in order to examine facts and information and learn from it, humans need memory, they need to remember facts. But connectionism has no provision for it.

There are other logical problems with the idea of memoryless learning. First, one cannot learn (generalize, that is) unless one knows what is there to learn (generalize). And one can find out what is there to learn "only by" collecting and storing some information about the problem at hand. In other words, **no system, biological or otherwise**, can "**prepare**" itself to learn without having some information about what is there to learn (generalize). And in order to generalize well, one has to look at a whole body of information relevant to the problem, not just bits and pieces of information at a time as is allowed in memoryless learning. So the notion of “memoryless learning” is a very serious misconception in these fields, and is totally inconsistent with external observations of the human learning process.

3.3 A Third Misconception - The Brain Learns Instantly from Each and Every Learning Example Presented to it

A major dilemma for this field is explaining the fact that sometimes human learning is not instantaneous, but may occur much later, perhaps at a distant point in time, based on information already collected and stored in the brain. The problem lies with the fundamental belief in the connectionist school that the brain learns “instantaneously.” Instantaneous, that is, in the sense that it learns promptly from each and every learning example presented to it by adjusting the relevant synaptic strengths or connection weights in the network. And it even learns from the very first example presented to it! The learning, as usual, is accomplished by individual neurons using some kind of a "local learning law." Note that “instantaneous learning” is simply a reflection of “memoryless learning;” just the opposite side of the same coin.

3.4 A Fourth Misconception - The Networks Are Predesigned and Externally Supplied to the Brain; And the Learning Parameters Are Externally Supplied Too

Another major dilemma for this field is explaining the fact that a network design, and other types of algorithmic information, has to be externally supplied to some of their learning systems, whereas no such information is externally supplied to the human brain. In fact, not just one, but many different network designs (and other parameter values) are often supplied to these learning systems on a trial and error basis in order for them to learn [5, 9, 10, 14, 15]. However, as far as is known, no one has been able to supply any network design or learning parameter values to a human brain. Plus, the whole idea of “instantaneous and memoryless learning” is completely inconsistent with their trial and error learning processes; there is supposed to be no storage of learning examples in these systems for such a trial and error process to take place. In other words, no such trial and error process can take place unless there is memory in the system, which they disallow.

In order for humans to generalize well in a learning situation, the brain has to be able to design different networks for different problems - different number of layers, number of neurons per layer, connection weights and so on - and adjust its own learning parameters. The networks required for different problems are different, it is not a "same size fits all" type situation. So the networks cannot come "pre-designed" in the brain; they cannot be inherited for every possible "unknown" learning problem faced by the brain on a regular basis. Since no information about the design of a network is ever supplied to the brain, it implies that network design is performed internally by the brain. Thus, it is expected that any brain-like learning system must also demonstrate the same ability to design networks and adjust its own learning parameters without any outside assistance. But the so-called autonomous learning systems of connectionism depend on external sources to provide the network design to them [14, 15]; hence they are inherently incapable of generalizing without external assistance. This implies again that connectionist learning is not brain-like at all.

3.5 Other Logical Problems with Connectionist Learning

There are other logical problems with these connectionist ideas. Strict autonomous local learning implies pre-definition of a network "by the learning system" without having seen a single training example and without having any knowledge at all of the complexity of the problem. There is **no system, biological or otherwise**, that can do that in a meaningful way; it is **not a "feasible idea" for any system**. The other fallacy of the autonomous local learning idea is that it acknowledges the existence of a "master system" that provides the network design and adjusts the learning parameters so that the autonomous learners can learn. So connectionism's autonomous learners, in the end, are directed and **controlled by other sources after all**! So these connectionist ideas (instantaneous learning, memoryless learning and autonomous local learning) are completely illogical, misconceived and incompatible with what can be externally observed of the human learning process.

4. CONCLUSIONS

One of the "large" missing pieces in the existing theories of artificial neural networks and connectionism is the characterization of an autonomous learning system such as the brain. Although Rumelhart [15] and others have clearly defined (conjectured) the "internal mechanisms" of the brain, no one has characterized in a similar manner the external behavioral characteristics that they are supposed to produce. As a result, the field pursued algorithm development largely from an "internal mechanisms" point of view (local, autonomous learning, memoryless learning, and instantaneous learning) rather than from the point of view of "external behavioral characteristics" of human learning. **That flaw is partly responsible for its current troubles**. It is essential that the development of brain-like learning algorithms be guided primarily by the need to reproduce a set of sensible, well-accepted external characteristics. If that set of external characteristics cannot be reproduced by a certain conjecture about the internal mechanisms, than that conjecture should not be valid.

This article essentially described some of the prevailing notions of connectionism and showed their logical inconsistencies and how they fail to properly account for some very basic aspects of human learning. So there is definitely a need for some new ideas about the internal mechanisms of the brain. From the last three debates and based on evidence from experimental psychology and neurobiology, it appears that a very convincing argument can be made that **there are subsystems within the brain that control other subsystems**. This "**control theoretic**" notion, which allows external sources to directly control a cell's behavior and perform other tasks, is finding growing acceptance among scientists [13]. This notion has many different labels at this point: non-local means of learning, global learning and so on. **It would not be fair if it is not acknowledged that such control theoretic notions are already used, in one form or another, in almost all connectionist learning systems**. For example, all constructive learning algorithms [5, 9, 10] use non-local means to "decide" when to expand the size of the network. And the back-propagation algorithm itself [14, 15] depends on a non-local, external source to provide it with the design of a network in which to learn. So connectionist systems inadvertently acknowledge this "control theoretic" idea, by using a "master or controlling subsystem" that designs networks and sets learning parameters for them. **In other words, as baffling as it may sound, the control theoretic ideas have been in use all along; they are nothing new**. Only recently has such non-local means of learning been used effectively to develop robust and powerful learning algorithms that can design and train networks in polynomial time complexity [1, 4, 11, 12]. Polynomial time complexity, by the way, is an essential computational property for brain-like autonomous learning systems.

In addition, a control theoretic framework resolves many of the problems and dilemmas of connectionism. Under such a framework, learning no longer needs to be instantaneous, but can wait until some information is collected about the problem. Learning can always be invoked by a controlling subsystem at a later point in time. This would also facilitate understanding the complexity of the problem from the information that has been collected and stored already. Such a framework would also resolve the network design dilemma and the problems of algorithmic efficiency that have plagued this field for so long [1, 4, 11, 12]. So one can argue very strongly for such a theory of the brain both from a computational point of view and from the point of view of being consistent with externally observed human learning behavior.

5. ACKNOWLEDGMENTS

This research was supported, in part, by grants from the College of Business, Arizona State University.

6. REFERENCES

[1] Bennett, K.P. and Mangasarian, O.L. Neural Network Training via Linear Programming. In P.M. Pardalos (ed), Advances in Optimization and Parallel Computing, North Holland, Amsterdam, 1992.

[2] Churchland, P. On the Nature of Theories: A Neurocomputational Perspective. Reprinted as chapter 10 in Haugeland, J. (ed), Mind Design II, 1997, MIT Press, 251-292.

[3] Churchland, P. and Sejnowski, T. The Computational Brain. MIT Press, Cambridge, MA, 1992.

[4] Glover, F. Improved Linear Programming Models for Discriminant Analysis. Decision Sciences, 21 (1990), 4, 771-785.

[5] Grossberg, S. Nonlinear neural networks: principles, mechanisms, and architectures. Neural Networks, 1 (1988), 17-61.

[6] Hebb, D. O. The Organization of Behavior, a Neuropsychological Theory. New York: John Wiley, 1949.

[7] Levine, D. S. Introduction to Neural and Cognitive Modeling, Hillsdale, NJ: Lawrence Erlbaum, 1998.

[8] Minsky, M. and Papert, S. *Perceptrons*. The MIT Press, Cambridge, MA, 1988.

[9] Moody, J. & Darken, C. Fast Learning in Networks of Locally-Tuned Processing Units, Neural Computation. 1 (1989), 2, 281-294.

[10] Reilly, D.L., Cooper, L.N. and Elbaum, C. A Neural Model for Category Learning. Biological Cybernetics, 45 (1982), 35-41.

[11] Roy, A., Govil, S. & Miranda, R. An Algorithm to Generate Radial Basis Function (RBF)-like Nets for Classification Problems. Neural Networks, 8 (1995), 2, 179-202.

[12] Roy, A., Govil, S. & Miranda, R. A Neural Network Learning Theory and a Polynomial Time RBF Algorithm. IEEE Transactions on Neural Networks, 8 (1997), 6, 1301-1313.

[13] Roy, A. Summary of panel discussion at ICNN'97 on connectionist learning. "Connectionist Learning: Is it Time to Reconsider the Foundations." INNS/ENNS/JNNS Newsletter, appearing with Neural Networks, 11 (1998), 2.

[14] Rumelhart, D.E., and McClelland, J.L.(eds.) Parallel Distributed Processing: Explorations in Microstructure of Cognition, Vol. 1: Foundations. MIT Press, Cambridge, MA., 1986, 318-362.

[15] Rumelhart, D.E. The Architecture of Mind: A Connectionist Approach. Chapter 8 in Haugeland, J. (ed), Mind Design II, 1997, MIT Press, 205-232.

About the author

Asim Roy is a Professor of Computer Information Systems at Arizona State University. He received his B.E. in Mechanical Engineering from Calcutta University, India, M.S. in Operations Research from Case Western Reserve University, Cleveland, Ohio, and Ph.D. in Operations Research from University of Texas at Austin. He was a Visiting Scholar with the neural network group at Stanford University. He has been an Associate Editor of *IEEE Transactions on Neural Networks* and is on the Review Board of *Applied Intelligence*. His research interests include neural networks, theory of brain-like systems, pattern recognition, prediction and forecasting, computational learning theory and nonlinear multiple objective optimization. His research has been published in *Management Science, Decision Sciences, Mathematical Programming, Financial Management, Neural Networks, Neural Computation, Naval Research Logistics, The ORSA Journal on Computing, IEEE Transactions on Neural Networks, IEEE Transactions on Fuzzy Systems* and others. He has recently published a new theory for brain-like learning and computing. This new theory challenges the classical learning ideas that have dominated the field of brain-like computing for the last fifty years.