# Mathematical Modeling of Artificial Neural Networks

**Radu Mutihac**
*University of Bucharest, Romania*

## INTRODUCTION

Models and algorithms have been designed to mimic information processing and knowledge acquisition of the human brain generically called artificial or formal neural networks (ANNs), parallel distributed processing (PDP), neuromorphic or connectionist models. The term network is common today: computer networks exist, communications are referred to as networking, corporations and markets are structured in networks. The concept of ANN was initially coined as a hopeful vision of anticipating artificial intelligence (AI) synthesis by emulating the biological brain.

ANNs are alternative means to symbol programming aiming to implement neural-inspired concepts in AI environments (neural computing) (Hertz, Krogh, & Palmer, 1991), whereas cognitive systems attempt to mimic the actual biological nervous systems (computational neuroscience). All conceivable neuromorphic models lie in between and supposed to be a simplified but meaningful representation of some reality. In order to establish a unifying theory of neural computing and computational neuroscience, mathematical theories should be developed along with specific methods of analysis (Amari, 1989) (Amit, 1990). The following outlines a tentatively mathematical-closed framework in neural modeling.

## BACKGROUND

ANNs may be regarded as dynamic systems (discrete or continuous), whose states are the activity patterns, and whose controls are the synaptic weights, which control the flux of information between the processing units (adaptive systems controlled by synaptic matrices). ANNs are parallel in the sense that most neurons process data at the same time. This process can be synchronous, if the processing time of an input neuron is the same for all units of the net, and asynchronous otherwise. Synchronous models may be regarded as discrete models. As biological neurons are asynchronous, they require a continuous time treatment by differential equations.

Alternatively, ANNs can recognize the state of environment and act on the environment to adapt to given viability constraints (cognitive systems controlled by conceptual controls). Knowledge is stored in conceptual controls rather than encoded in synaptic matrices, whereas learning rules describe the dynamics of conceptual controls in terms of state evolution in adapting to viability constraints.

The concept of paradigm referring to ANNs typically comprises a description of the form and functions of the processing unit (neuron, node), a network topology that describes the pattern of weighted interconnections among the units, and a learning rule to establish the values of the weights (Domany, 1988). Although paradigms differ in details, they still have a common subset of selected attributes (Jansson, 1991) like simple processing units, high connectivity, parallel processing, nonlinear transfer function, feedback paths, non-algorithmic data processing, self-organization, adaptation (learning) and fault tolerance. Some extra features might be: generalization, useful outputs from fuzzy inputs, energy saving, and potential overall high speed operation.

The digital paradigm dominating computer science assumes that information must be digitized to avoid noise interference and signal degradation. In contrast, a neuron is highly analog in the sense that its computations are based on spatiotemporal integrative processes of smoothly varying ion currents at the trigger zone rather than on bits. Yet neural systems are highly efficient and reliable information processors.

### Memory and Learning

The specificity of neural processes consists in their distributive and collective nature. The phenomenon by biological neural networks (NNs) are changing in response to extrinsic stimuli is called self-organization. The flexible nature of the human brain, represented by self-organization, seems to be responsible for the learning function which is specific to living organisms. Essentially, learning is an adaptive self-organizing process. From the training assistance point of view, there are supervised and unsupervised neural classifiers. Supervised classifiers seek to characterize predefined classes by defining measures that maximize in-class similarity and out-class dissimilarity. Supervision may be conducted either by direct comparison of output with the desired target and estimating error, or by specifying whether the output is correct or not (reinforcement learning). The measure of success in both cases is given by the ability to recover the original classes for similar but not identical input data. Unsupervised classifiers seek similarity measures without any predefined classes performing cluster analysis or vector quantization. Neural classifiers organize themselves according to their initial state, types and frequency of the presented patterns, and correlations in the input patterns by setting up some criteria for classification (Fukushima, 1975) reflecting causal mechanisms. There is no general agreement on the measure of their success since likelihood optimization always tends to favor single instance classes.

Classification as performed by ANNs has essentially a dual interpretation reflected by machine learning too. It could mean either the assignment of input patterns to predefined classes, or the construction of new classes from a previously undifferentiated instance set (Stutz & Cheesman, 1994). However, the assignment of instances to predefined classes can produce either the class that best represents the input pattern as in the classical decision theory, or the classifier can be used as a content-addressable or associative memory, where the class representative is desired and the input pattern is used to determine which exemplar to produce. While the first task assumes that inputs were corrupted by some processes, the second one deals with incomplete input patterns when retrieval of full information is the goal. Most neural classifiers do not require simultaneous availability of all training data and frequently yield error rates comparable to Bayesian methods without needing prior information. An efficient memory might store and retrieve many patterns, so its dynamics must allow for as many states of activity which are stable against small perturbations as possible. Several approaches dealing with uncertainty such as fuzzy logic, probabilistic, hyperplane, kernel, and exemplar-based classifiers can be incorporated into ANN classifiers in applications where only few data are available (Ng & Lippmann, 1991).

The capacity of analog neural systems to operate in unpredictable environments depends on their ability to represent information in context. The context of a signal may be some complex collections of neural patterns, including those that constitute learning. The interplay of context and adaptation is a fundamental principle of the neural paradigm. As only variations and differences convey information, permanent change is a necessity for neural systems rather than a source of difficulty as it is for digital systems.

## MATHEMATICAL FRAMEWORK OF NEURONS AND ANNS MODELING

An approach to investigate neural systems in a general frame is the mean field theory (Cooper & Scofield, 1988) from statistical physics suited for highly interconnected systems as cortical regions are. However, there is a big gap between the formal model level of description in associative memory levels and the complexity of neural dynamics in biological nets. Neural modeling need no information concerning correlations of input data, rather nonlinear processing units and a sufficiently large number of variable parameters ensure the flexibility to adapt to any relationship between input and output data. Models can be altered externally, by adopting a different axiomatic structure, and internally, by revealing new inside structural or functional relationships. Ranking several neuromorphic models is ultimately carried out based on some measure of performance.

### Neuron Modeling

Central problems in any artificial system designed to mimic NNs arise from (*i*) biological features to be preserved, (*ii*) connectivity matrix of the processing units, whose size increases with the square of their number, and (*iii*) processing time, which has to be independent of the network size. Biologically realistic models of neurons might minimally include:

- Continuous-valued transfer functions (graded response), as many neurons respond to their