Book Reviews

ROJAS, Raul: **Theorie der neuronalen Netze. Eine systematische Einführung.** (Theory of neural networks. A systematic introduction). Berlin-Heidelberg-New York, etc.: Springer Verlag 1993. XVIII,446p., ISBN 3-540-56353-9

Books on neural networks are no rarity today. The author of the book under review does not hesitate to speak of a publication explosion in the 80s. In his opinion, these books, not excluding the well-known work by Rumelhart and McClelland(1986), have the setback of being, in part, highly pragmatic and of frequently "resembling a hodgepodge of the various models proposed in the course of time". To remedy this shortcoming is the purpose of this handbook, which according to its title and subtitle professes to be a systematic introduction to the theory of neural networks. This means that the various models already existing are not only listed and described, but also related to one another. The author admits that despite this theory-oriented systematic approach we are still confronted with a puzzle requiring the various models to be fitted together to an overall picture. To this extent this book, too, resembles the depictions of neural networks it criticizes.

Beginning, like almost all other descriptions, with an introduction to the biology of natural neural networks, it then, in its first part devoted to the forward-oriented networks, goes on to describe the classical model of McCulloch and Pitts (1943) and the 'perceptron' of Rosenblatt (1958). Inevitably following next is a critical analysis of the perceptrons of Minsky and Papert, which, because of the host of negative results regarding the properties of individual perceptrons hindered for a long time the development of neural networks in the competition with symbol-processing AI. Likewise emphasized, however, are the positive achievements of Minsky and Papert, who carried out a classification of perceptrons according to their limitations and already then singled out important elements of the differences between sequential and parallel data processing. Then there follows a description of the perceptron learning algorithm, which was the key to the initial success of this model, and the difference between monitored and non-monitored learning is explained. This difference is of cardinal importance inasmuch as monitored learning, which always needs a 'teacher' who checks the output and changes the process accordingly, is biologically implausible in that it means an interference from outside. In non-monitored learning, on the other hand, it is the network itself which determines whether or not a change of the weights is necessary, thus coming much closer to the self-organizing interconnectivity in the human brain. A further successful step, one which brought about the revivification of the long-neglected neural networks, was the construction of multilayer networks permitting a wider application spectrum, but also requiring greater training efforts. One of the most important learning methods is the so-called backpropagation algorithm, which seeks the minimum of the error function of a given learning problem by descending in the direction of the gradient. The combination of those weights of a network which minimize the computation error is regarded as the solution of the learning problem. With the description and deduction of the backpropagation algorithm the first, voluminous part of the book ends.

The second, far shorter part is devoted to a theoretical analysis of the backpropagation algorithm, which was developed as a generalization of statistical regression methods. In this connection it also becomes clear why this learning method, which became popular because of its numerous application fields (robotics, language and pattern recognition tasks, coding problems, etc.), is looked at critically. The reason: backpropagation can discover and exploit statistical regularities and also, by permanent learning, continually change the network parameters - yet it is an illusion to believe that backpropagation can replace a more exact knowledge of the problems. Of this method it can be said - as it can be said of statistics in general - that even the best methods cannot change insufficient or incoherent data into reliable results (cf.p.196). This theoretically well-founded warning applies above all to those speculations which have been indulged in in financial circles on using neural network systems for prognosticating stock-market quotations or raw material prices without calling attention to their merits and shortcomings. This is due first and foremost to the complexity of the problem.

The complexity of learning is also the theme of the theoretical part. What happens when the network becomes bigger and the number of unknwon weights increases? Independent of the backpropagation methods, which, though the most popular one, is not the only learning method in multilayered networks, it is therefore necessary to investigate both the complexity of a problem itself and of the learning process involved in solving this problem. At this point the author, both historically and systematically, reverts far back to the study of the fundamentals of mathematics. Hilbert's thirteenth problem and, with it, the question of the solvability of higher-degree algebraic equations, the further development and the answering of this question by Kolmogorov are discussed, as well as the solving of the general learning problem for neural networks within the framework of modern complexity theory, which has found this problem to be intractable, i.e. not efficiently calculable. These highly abstract theoretical considerations, which may seem superfluous to many a practitioner, are justified with the argument that in a field such as the neural networks, where numerical methods are applied intensively, the fundamental limits to efficient calculations should be known as well. It is only on this basis that network structures can be designed that permit faster solutions to learning problems (p.224). At this point the overall conception of the book becomes clear. For the author regards the theory of neural networks a priori as one of the five possible models of calculability (mathematical, logical-operational, machine = computer, cellular automatism, and biological = neural networks). As a biological model, artificial neural networks lay claim to being an alternative calculability paradigm capable of supplying those sophisticated solutions which a million years of evolution have producedby self-organization.

This claim is honored in this book first and foremost by the description of the self-organizing networks as designed by Kohonen (1982) and, particularly, by the description of the so-called genetic algorithms which simulate an evolution process in the computer. First, however, there follows, in the third part of the book, a highly detailed description of the recursive networks to which the learning process in associative memories according to Hebb's rule, as well as the Hopefield model with its variants (e.g. Boltzmann machines) belong.

The book closes with a survey of the hardware for neural networks. Deserving of special mention are the detailed bibliography, as well as the historical notes accompanying each chapter. They considerably contribute to the readability of the book, highly abstract-theoretical and brimming with mathematical formulae as it is otherwise. If one asks, in summing up, whether the author has achieved his purpose of not only describing the various models but also linking them together systematically, this question must be answered unreservedly in the affirmative. By looking at the host of different network topologies and learning strategies from the - admittedly highly abstract and formal - point of view of the calculability paradigm, the chaotic tangle they intitially presented has made way for a readily overseeable overall picture. The sober, maybe all too sober description thus achieved, which also consistently emphasizes the limitations of such artificial neural networks in comparison with the unsurpassable complexity and efficiency of natural neural networks, contributes to the demystification of the artificial ones. Therefore, over and beyond its intrinsic didactic function, this book can be recommended to all those seriously interested in the theoretical foundations of an information processing technology which has already succeeded in closing major gaps which conventional methods have so far been unable to fill. Erhard Oeser

Prof.Dr.E.Oeser, University of Vienna,
Institute for Science Theory and Science Research
A-1090 Wien, Sensengasse 8, Austria.

FUGMANN, Robert: **Subject Analysis and Indexing. Theoretical Foundation and Practical Advice.** Frankfurt/M: INDEKS Verlag 1993. XVI,250p., ISBN 3-88672-500-6 = Textbooks for Knowledge Organization, Vol.1

This work is a compilation of lectures given by the author in 1992 and 1993 at the School of Library and Information Science, Indiana University, Bloomington, Indiana, USA. These lectures were quite evidently based on the author's book "Theoretische Grundlagen der Indexierungspraxis", Frankfurt/M, 1992, likewise published by INDEKS Verlag. For the didactic purposes of the lectures, the English version was revised, shortened and partly also expanded. The core of Fugmann's theoretical foundation is formed by the so-called "five-axiom theory of indexing and information supply". This theory is found in the chapter on information storage and retrieval. The point of departure for these five axioms is to be seen in the survival power or endurance of information retrieval systems. We will briefly introduce the five axioms here:

1. *Axiom of Definability:* Every document collection needs to be ordered. This order must, however, be definable, on the one hand in the sense of undirected or unfocused information supply (serendipity effect in browsing) and on the other hand for focused or directed searching. From this demand the distinction between pertinence (subjective need) and relevance (objective need) is derived.

2. *Axiom of Order:* Order is defined by the author as the "meaningful proximity of the parts of a whole at a foreseeable place", thus rendering the compilation of information relevant to a given subject into an order-creating process. This axiom dominates the entire classification theory.

3. *Axiom of Sufficient Order:* This axiom pertains to the degree of order as depending on the size of the document collection, the frequency of searches and the specificity of the enquiries, as well as to the type of order, which should be both loss-avoiding and noise-avoiding.

4. *Axiom of Representational Predictability and of Loss-Avoiding Order:* The completeness of the information to be retrieved depends on the predictability of its description. The prerequisite for this property is a so-called index language (i.e. information retrieval language) whose lexics and syntax are, through standardization, better "predictable" than those of natural languages, with - in particular - synonymy needing to be excluded.

5. *Axiom of Representational Fidelity and of Noise-Avoiding Order:* An index language must be capable of representing concepts and concept connections in a semantically adequate and unambiguous fashion, this is what the author means by "representational fidelity".

When presented in coherence, these five axioms are trivial in themselves, but they constitute basic demands on information retrieval languages and thus on a classification type of pragmatic nature for the ordering of documents or subjects. Such information retrieval languages are artificial languages which, just like natural languages, must fulfill two basic functions - namely the cognitive function (language as carrier of objects of human cognition) and the communicative function, with this latter function needing to be interpreted both technologically and socially. In addition, however, information retrieval languages also must fulfill the following func-