БЪЛГАРСКА АКАДЕМИЯ НА НАУКИТЕ . BULGARIAN ACADEMY OF SCIENCES

ПРОБЛЕМИ НА ТЕХНИЧЕСКАТА КИБЕРНЕТИКА И РОБОТИКАТА, 52
PROBLEMS OF ENGINEERING CYBERNETICS AND ROBOTICS, 52

София . 2001 . Sofia

The Artificial Neural Networks in the Context of the Modern Tendencies in Evolutionary Computing

Stefan Koynov, Chavdar Korsemov, Hristo Toshev

Institute of Information Technologies, 1113 Sofia

1. Introduction

1.1. Why do the artificial neural networks borrow means from other domains of the artificial intelligence? AI domains related to the design and tuning of the artificial neural networks

A key shortcoming of the current state of neural network technology is the lack of any effective design methodology. Neural network technology is becoming widely accepted in various industries and numerous operating applications have demonstrated significant performance improvements over prior methods. But current approaches to developing neural network applications are a critical barrier to realizing the full potential of this technology. Optimizing neural network applications is a formidable design task requiring myriad choices in the number of neurons and layers, their interconnections and the training algorithm to be employed as well as all its parameters.

The rest of this chapter introduces the following aspects: the relations between the artificial neural networks (ANN) and the AI as the greatest domain of the information technologies; the relations between the ANN and the evolutionary computing (EC) as the domain directly responsible for the design, the optimization and the training of the ANN. (Section 1.2 presents an abstract of the ways for system modeling which are directly related to the ANN domain).

Chapter 2 introduces the ANN properties thus proving the necessity to borrow approaches and techniques from other AI domains and chapter 3 views the modern achievements of applying the EC (including the genetic synthesis) for the design, the optimization and the training of the ANN.

1.2. Approaches to system modeling

Fig. 1 views the basic aspects of the approaches to system modeling ([1]; the figures and the tables are placed in the Appendix).

Section 1.3 presents a comparative analysis of the similarities and of the differences between the AI and the ANN as two different domains of the information technologies.

1.3. AI and ANN

According to [2] the AI systems possess the following general properties: knowledge representation, reasoning and training.

According to [3] the ANN possess the following main characteristics: level of explanation, processing modes and structure of the information presentation. Table 1 shows a comparison between the AI as the environment of the ANN and the ANN themselves.

Section 1.4 summarizes the role and the place of the EC in the cycle of the design, the optimization and the tuning of the ANN.

1.4. The evolutionary computing and the ANN

The speculation that the ANN could be optimized using simulated evolution goes back at least to [4]. Chapter 2 introduces an information about the efforts dedicated to the application of the EC for the design, the optimization and the training of the ANN. Fig. 2 summarizes the aspects of the cooperation between the EC and the ANN with remarks of the correspondent references:

2. ANN

2.1. ANN and the time scale (substantial ideas)

The history of the fundamental hypotheses and models concerning the ANN ([13]) show that some ideas have originated independently from more than a single scientist, but not all the scientists enjoy merited popularity.

2.2. Properties and applications of the ANN pointing to the necessity to borrow methods and techniques from other AI domains

Figs. 3 and 4 illustrate the principle properties and applications of the ANN (they follow [14]).

It is evident that the set of properties including the wide range of the applications is a source of principal troubles during the processes of the design, the choice of parameters, the mathematical description, the optimization and the initialization of the ANN. Therefore there is a necessity to borrow means from other AI domains including the EC approaches (also see section 1.1).

The next chapter is an introduction to the concrete projects with ANN based on EC approaches which include the design, the optimization and the learning of the ANN.

3. Evolutionary computing and the ANN

3.1. The POE model

The POE model is introduced in [15]. It is aimed at the most significant features of the living creatures; also it is possible to make a comparison with the goals during the process of constructing an ANN (Table 2).

[16–19] comprise a review of the modern ANN applications which are accessible for the POE-model (Fig. 5).

Table 2 and Fig. 5 naturally fix the connections between the ANN and the EC. The latter is an object of section 3.2.

3.2. Modern solutions with ANN based on approaches with EC

The most common network involved in this studies has been the multilayer perceptron. If the network architecture is predefined, the problem becomes one of trying to find optimal settings for the weight and bias terms. A natural representation for this is a real-valued vector where each component corresponds to a weight or bias. For such continuous-valued representations the common method of mutation is to add a multivariate zero mean Gaussian random variable to all components, thereby in this case changing all of the weights and biases simultaneously. The behavior of each offspring network is strongly related to its parent and there is a continuous range of possible new behaviors.

The variability between each parent and its offspring can be controlled directly by specifying the individual variances for each Gaussian perturbation (or covariances). One common method for accomplishing this is to set the step size (standard deviation) to be proportional to the mean squared error of the parent network. In this manner, as better solutions are discovered, the step size is reduced and the search effort is concentrated around the parent and conversely, a larger variance is used for parents with relatively poor performance.

Some of the first efforts at applying EC to optimizing ANN can be found in [20–21]. More recent research has involved simultaneously evolving both the structure and weights of feed forward and recurrent ANN ([22, 23, 24]). Some attention has also being given to evolving fuzzy ANN in which classification of input patterns are made with respect to their fuzzy membership in evolved clusters ([25]).

3.3. Genetic synthesis and optimization of the ANN

EPRI has been sponsoring research by Tariq Samad and Steven A. Harp at Honeywell Technology Center in the use of genetic algorithms to synthesize and optimize the design of the ANN. Initial investigation was directed at validating the approach by optimizing an already trained neural network developed manually in a realistic application of heat rate modeling performed for EPRI by Robert Uhrig at the University of Tenessee. The same data sets from TVA's Sequoyah nuclear power plant were used for the ANN training and testing. Appropriate criteria for optimization were determined to include accuracy (e.g. low error prediction of plant gross heat rate), learning speed and the ANN simplicity (e.g. low number and density of connections). The ANN design experiments and evaluations began with a review of various ANN designs appropriate for the thermodynamic modeling application. Experiments were conducted that showed the genetically optimized ANN had a significant performance improvement over manually developed ANN. This research also demonstrated that the choice of input variables is critical in the ANN applications. The genetic algorithm's ability to simultaneously optimize input selections along with the ANN structures and learning parameters was vital to the accurate modeling. Another result was that genetic optimization need not be limited to evolving individual trained ANN. The ANN architectures can be designed for classes of applications and then later trained on specific data for different applications within a given class ([11, 12]).

Experiments on two additional applications have been completed: engine NO_x emissions and ozone levels in New York City. Both resulted in the automatic production of models competitive with the best conventional non-linear statistical models developed for these problems by the exercise of considerable statistical expertise and manual adjustment.

Demonstration software has been developed that runs under Microsoft Windows 3.x on a PC. The software features an easy-to-use graphical interface and incorporates tutorial material on ANN, genetic algorithms and the neurogenetic design technique. Two problems are included – heat rate modeling and engine NO_x emissions prediction – along with optimized ANN models for them. The software also contains an ANN specification and training facility, allowing users to test and compare their own hand-crafted designs with the genetically optimized ones and to have the GA optimize their design within the constraints they establish for it.

Research is continuing with the objective of demonstrating conclusively the power of genetic synthesis of the ANN for modeling and analysis applications throughout the electric power industry. An in-depth assessment of this design technique requires applying it to a variety of additional problems relevant to EPRI members. The first of these is the real time pricing (RTP) of electric power in a competitive market. A GA is being used to design and optimize an ANN model using data from the Marriot Marquis Hotel in New York City. The data includes two yearrs of hourly loads, rates, weather information, etc. obtained in a joint EPRI and Consolidated Edison RTP experiment. Accurately forecasting the loads and the rates would be of value to any electricity customer who had energy storage facilities or co-generation capability.

The investigations are also exploring ways to extend the state of the art for this general approach by exploring extensions to both the genetic and neural components of the system.

APPENDIX

System modeling	Math equations	+ : Simplicity of the description
		Vectors of big dimensions
		Difficult rewriting the system of the equations (if possible)
		Difficult re-set up of the system (following its description)
		Limited computing resources for multivariable systems along the time axis
	ANN	+ Convenient for ill-defined tasks or for implicit algorithm formulations
		Convenient for iterative modeling (structural and algorithmic iterativity)
		Complicated design with respect to the fuzzy systems
		Unpredictable behavior
		Asymptotic convergence
		Small input changes lead to a similar, but new training
	Linguistic rules	Non-mathematical (logical) formalism
		+ Easy system re-set up
		Languages with ill-defined grammars admitting contradictory conclusions from facts

Fig. 1. Approaches to system modeling

Table 1. Comparison between the AI and the ANN

Presentation	AI	ANN
Knowledge	Symbolic, digital, successive processing	Structurally-coded knowledge; its expansion depends on the neurons
Manner of processing	Depends on the successive nature of the languages and on the approaches to reasoning	Parallel, distributed, strong in biological prototypes
Structure	Quasi-linguistic of symbolic expressions	Decisive, dependent on the concrete task, the solution is difficult to expand, the ANN model must be generalized

EC & ANN	ANN design ([5]–[8]) ANN training ([9], [5]–[8], [10]) Genetic synthesis and ANN optimization ([11]–12])

Fig. 2. EC, its role and its place in the cycle of the design, the optimization and the tuning of the ANN

ANN (properties)	Advantages	Adaptivity		
		Stability (of a fuzzy and noisy input)		
		Robustness		
	Disadvantages	Curse of dimensionality		
		Design problems		
		Energy function (choice)		
		Convergence (classification problems)		
		Time for learning		
	Classification	Feedforward ANN	One-layer perceptron	
			Multilayer perceptron	
			ANN with RBF	
		Recurrent ANN	Competitive learning	
			Kohonen's SOM	
			Hopfield's ANN	
			ART networks	
	Learning	Paradigm		
		Rule	Error correction	
		Architecture	Competitive learning	
			Combination (error correction & competitive learning)	
			According to physical principles	
		Algorithm	Supervised	
			Unsupervised	Associative memories
				Categorization
			Popular algorithms	Perceptrons
				Competitive learning

Fig. 3. ANN description

ANN (tasks)	Basic	Function approximation Coding & decoding with ANN Data compression Image compression Modeling		
	Dynamic systems	Identification Modeling Optimization Control	Time series	Generating Prediciton
	Recognition	Images	Principal applications Typical applications Concrete solutions	
		Speech	Basic concepts	General scheme Difficulties
			Recognition with ANN	

Fig. 4. ANN applications

Table 2. Goals for the ANN design and their counterparts in the POE model

Goal	ANN	Analogy with the POE model
Model	New ANN models / evolution of popular models	ANN Phylogeny (new species / evolution of living creatures)
Design	Specifying the ANN model for practical applications	Ontogeny (growth of living creatures)
Learning	Supervised / unsupervised	Epigenesis (*basically* unsupervised)

ANN (modern applica-tions)	Language learning (Steels, 1995)	Inborn abilities (phylogeny) Acquired abilties (epigenesis)
	ANN populations (Yao 1993, Nolfi *et al.* 1994, Liu & Yao, 1996)	Evolution on the global level ("phylogeny") Learning on an individual ("epigenesis")

Fig. 5. Modern POE-analogical ANN applications

4. Conclusion

This paper is a review of the interrelations between the ANN and the state of the art in the field of the EC. The analysis starts with the place of the ANN in the approaches to system modeling followed by a comparative analysis of the ANN as a private case of the AI which then is restricted to the field of the EC. The second part of the paper presents the realized approaches to animating the chosen ANN models (including the design-, the optimization- and the learning- phases) starting from an EC background, including genetic synthesis. The material is illustrated with schemes, tables and a representative bibliography list.

References

1. Yamakawa, T. A fuzzy inference engine in nonlinear analog mode and its application to a fuzzy logic control. – In: IEEE Trans. Neural Networks, **4**, 1993, No 3, 496-521.
2. Sage, A. P. Concise Encyclopedia of Information Processing in Systems and Organizations. N.Y., Pergamon, 1990.
3. Memmi, D. Connectionism and artificial intelligence. – In: Neuro-Nimes'89 Int. Workshop on Neural Networks and Their Applications, Nimes, France, 1989, 17-34.
4. Bremermann, H. J. Numerical optimization procedures derived from biological evolution processes. Cybernetics Problems in Bionics. H. L. Ustereicher and D. R. Moore (eds.). – In: Gordon and Breach, London, 1966, 543-562.
5. Sebald, A. V., D. B. Fogel. Design of fault tolerant neural networks for pattern classification. – In: Proc. of the First Annual Conference on Evolutionary Programming. D. B. Fogel and W. Atmar (eds.), Evolutionary Programming Society, La Jolla, CA, 1992, 90-99.
6. Porto, V. W. Alternative methods for training neural networks. – In: Proc. of the First Annual Conference on Evolutionary Programmin. D. B. Fogel and W. Atmar (eds.). Evolutionary Programming Society, La Jolla, CA, 1992, 100-110.
7. McDonell, J. R. Training neural networks with weight constraints. – In: Proc. of the First Annual Conference on Evolutionary Programming, D. B. Fogel and W. Atmar (eds.). Evolutionary Programming Society, La Jolla, CA, 1992, 111-119.
8. Sebald, A. V., J. Schlenzig, D. B. Fogel. Minimax design of CMAC encoded neural controllers for systems with variable time delay. – In: Proc. of the First Annual Conference on Evolutionary Programming. D. B. Fogel and W. Atmar (eds.). Evolutionary Programming Society, La Jolla, CA, 1992, 120-126.
9. Fogel, D. B., L. J. Fogel, V. W. Porto. Evolving neural networks. – Biological Cybernetics, **63**, 1990, No 6, 487-493.
10. Fogel, D. B., E. C. Wasson, E. M. Boughton. Evolving neural networks for detecting breast cancer. – Cancer Letters. **96**, 1995, 49-53.
11. Electric Power Research Institute, Genetic Optimization of Neural Network Architecture TR-104074. ERPI Dist. Ctr. Pleasant Hill, CA, 1994.
12. Harp, S. A., T. Samad. Genetic algorithms and neural networks for optimized modeling and control in mission earth: modeling and simulation for a sustainable future. – In: Proc. of the 1995 Western Multiconference. A. M. Wildberger, ed., Las Vegas, Nevada, Jan. 15–18, San Diego, CA: Society for Computer Simulation, 1988, 14-20.
13. Haykin, S. Neural Networks. Macmillan Publ. Company, 1994.
14. Jain, A., Artificial neural networks: a tutorial. – In: IEEE Computer, March 1996, 31-44.
15. Sipper, M., E. Sanchez, D. Mange, M. Tomassini, A. Perez-Uribe, A. Stauffer. A phylogenetic, ontogenetic, and epigenetic view of bioinspired hardware systems. – In: IEEE Trans. Evol. Comput., **1**, 1997, No 1, 83-97.
16. Steels, L. A self-organizing spatial vocabulary. – Artificial Life, **2**, 1995, No 3, 319-332
17. Yao, X. Evolutionary artificial neural networks. – Int. J. Neural Syst. **4**, 1993, No 3, 203-222.
18. Nolfi, S., D. Parisi, J. L. Elman. Learning and evolution in neural networks. – Adaptive Behavior. **3**, 1994, No 1, 5-28.
19. Liu, Y., X. Yao. Evolutionary design of artificial neural networks with different nodes. – In: Proc. IEEE Int. Conf. Evol. Comput. (ICEC'96), 1996, 670-675.

20. Fogel, D. B., L. J. Fogel, V. W. Porto. Evolving Neural Networks. – Biological Cybernetics, **63**, 1990, No 6, 487-493.

21. Fogel, D. B. An Information Criterion for Optimal Neural Network Selection. – In: IEEE Trans. Neural Networks, **2**, 1991, 490-497.

22. McDonell, J. R., D. Waagen. Neural network structure design by evolutionary programming. – In: Proc. of the Second Annual Conference on Evolutionary Programming. D. B. Fogel and W. Atmar (eds.), Evolutionary Programming Society, La Jolla, CA, 1993, 79-89.

23. McDonell, J. R. , D. Waagen. Evolving Recurrent Perceptrons for Time-Series Modeling. – In: IEEE Trans. Neural Networks, **5**, 1994, No 1, 24-38.

24. Angeline, P. J., G. M. Saunders, J. B. Pollack. An evolutionary algorithm that constructs recurrent neural networks. – In: IEEE Trans. Neural Networks. **5**, 1994, No 1, 54–65.

25. Brotherton, T. W., P. K. Simpson. Dynamic feature set training of neural nets for classification, evolutionary programming. – In: Proc. of the Fourth Annual Conference on Evolutionary Programming. J. R. McDonell, R. G. Reynolds and D. B. Fogel (eds.). Cambridge, MA, MIT Press, 1995, 83–94.

Искуственные нейронные сети в контексте современных направлений в эволюционном вычислении

Стефан Койнов, Чавдар Корсемов, Христо Тошев

Институт информационных технологий, 1113 София

(Р е з ю м е)

Представлен обзор искуственных нейронных сетей в контексте современных направлений в эволюционном вычислении. Указана родственная связь между искуственным интеллектом и этими сетями, а также рассмотрена определяющая роль эволюционного вычисления улесняющая проэктирование, оптимизацию и настройку искуственных нейронных сетей. Материал иллюстрирован примерами взаимодействия искуственных нейронных сетей и эволюционного вычисления.