Artificial Neural Networks and Genetic Algorithms: An Overview

MICHAEL GR. VOSKOGLOU
Mathematical Sciences, School of Technological Applications
University of Peloponnese (ex T.E.I. of Western Greece)
Meg. Alexandrou 1, 26334 Patras
GREECE

Abstract: - In contrast to the conventional hard computing, which is based on symbolic logic reasoning and numerical modelling, soft computing (SC) deals with approximate reasoning and processes that give solutions to complex real-life problems, which cannot be modelled or is too difficult to be modelled mathematically. SC is a synthesis of several computing paradigms mainly including probabilistic reasoning (PR), fuzzy logic (FL), artificial neural networks (ANNs) and genetic algorithms (GAs). The techniques connected to the previous paradigms are not competitive, but complementary to each other and can be used together for solving a given problem. Despite the fact that SC appeared just during the 1980s, its techniques are nowadays being used successfully in many domestic, commercial and industrial applications becoming a major research object in automatic control engineering. The present paper reviews the techniques and applications of ANNs and GAs, and their relationships with FL and PR pointing out their advantages and disadvantages with respect to the traditional techniques of hard computing.

Key-Words: - Soft Computing (SC), Probabilistic Reasoning (PR), Fuzzy Logic (FL), Artificial Neural Network (GA), Genetic Algorithm (GA), Evolutionary Computing

Received: July 21, 2021. Revised: October 14, 2022. Accepted: November 21, 2022. Published: December 23, 2022.

1 Introduction

The combined effects of the three industrial revolutions that took place from the end of the 18^{th} century until recently replaced the manpower and the power of animals as means of production with machines, facilitated the mass production of goods, improved the social services and led our society to the digital era [1]. During the *fourth industrial revolution (4IR)*, which is currently in the beginning of its appearance, the advanced Internet of Things (IoT), the renewable energy, the 3D-printing and the development of the cyber-physical systems (e.g. robots, autonomous vehicles and control systems, distance medicine, etc.) are expected to merge leading humanity to a new era of progress and well-being [2], provided of course that no wrong human actions will interfere.

It is generally accepted nowadays that computers, with their speedy calculations and the wealth of information that they provide through the web, have become a valuable tool for our everyday activities affecting in a great deal the way of our life and our behaviour

In contrast to the conventional hard computing, however, which is based on symbolic logic reasoning and numerical modelling, *soft computing (SC)* deals with approximate reasoning and processes that give solutions to complex real-life problems, which cannot be modelled or is too difficult to be modelled mathematically. Its features such as intelligent control, decision making support, non-linear programming, etc., have made SC popular for people of a wide variety of disciplines and scientific backgrounds, like engineers, computer scientists and mathematicians, other natural and positive scientists, etc.

The theory and techniques of SC were introduced during the 1980's [3-6]. The term SC is due to the pioneer of fuzzy logic L. Zadeh, who defined it as "an emerging approach to computing, which parallels the remarkable ability of the human mind to reason and learn in an environment of uncertainty and imprecision" [6].

SC is a synthesis of several computing paradigms mainly including *probabilistic reasoning (PR)*, *fuzzy logic (FL), artificial neural networks (ANNs) and genetic algorithms (GAs).* The techniques connected to the previous paradigms are not competitive, but complementary to each other and can be used together for solving a given problem. For example, FL can be used for knowledge representation via fuzzy IF – THEN rules, ANNs for learning and adaptation, GAs for evolutionary computation, etc. The target of SC is to exploit the tolerance for imprecision, uncertainty, approximate reasoning, and partial truths, being inherent to the previously mentioned computing paradigms, in order to achieve close resemblance to human like thinking and decision making.

Despite the fact that SC has appeared just during the 1980s, its techniques are nowadays being used successfully in many domestic, commercial and industrial applications becoming a major research object in automatic control engineering. In the current era of the 4IR and the growth of the advanced IoT, as the processing by computer devices is increasing and their cost is reduced, the need of using SC methods has become too important and has the potential to expand further. SC appears as a new multidisciplinary field, to construct a new branch of *Artificial Intelligence (AI)*, known as *Computational Intelligence (CI),* a term which is sometimes considered a synonym to SC [7]. It is recalled that AI is the branch of Computer Science that focuses on the theory and practice of creating “smart” devices mimicking the human reasoning and behavior [8, 9]. AI aims at making computers able to perform autonomous improvements with the help of given data without needing the commands of a program created by humans.

The present paper reviews the techniques and applications of ANNs and GAs, and points out their advantages and disadvantages with respect to the traditional techniques of hard computing. The rest of the paper is organized as follows: Section 2 is devoted to FL and PR, with emphasis to *Bayesian Reasoning (BR)*, which appears today as a link between BL and FL. Section 3 presents the basics of AANs and their relationships with FL and PR. Section 4 examines the structure of GAs and the article closes with a discussion and the general conclusions presented in Section 5.

2. Fuzzy Logic and Probabilistic Reasoning

The Aristotle’s (384-322 BC) *bivalent logic (BL),* based on the “principle of the excluded middle” (everything is either true or false), used to be for more than 23 centuries the basis for the development of science and human civilization. Despite to the fact that opposite views about the existence of a third area between true and false appeared early in the human history, integrated propositions for multi-valued logics introduced only in the early 1900s, mainly by Lukasiewicz and Tarski [10, 11].

Zadeh, based on the concept of *fuzzy set (FS)* [12], introduced during the 1970’s the infinite-valued FL [13], where the truth values are represented by numbers in the unit interval [0, 1]. FL satisfies the Lukasiewicz’s “principle of valence”, according to which propositions can have intermediate truth values (partial truths) between true or false. FL extends and completes the traditional BL by examining what happens in the “area” between true and false. In addition, FL and its generalizations (e.g. see [14]) treat effectively all the types of the existing in the real world uncertainty.

A *fuzzy system* is a collection of FSs related and bound together. Fuzzy systems are not only able to use their own knowledge to represent and explain phenomena of the real world, but can also increase it with the help of given data, i.e. they emulate the way in which humans learn, corresponding to the “software” of the human brain.

During the 1990s, E. Jaynes argued that *probability theory* can be viewed as an extension of BL reducing to it when something is absolutely certain or absolutely impossible [15]. Probability theory, however, the development of which was based on the principles of BL, has been proved suitable for tackling only the cases of uncertainty which are due to randomness (e.g. games of chance) [16]. As a result, the Jaynes’ probabilistic logic is subordinate to FL.

Bayesian reasoning (BR) appears today as a link between BL and FL [17, Section 5]. It is recalled that the Bayes’ rule, which is a straightforward consequence of the well-known formula calculating the traditional probabilities, is expressed in the form

$$P(A/B) = \frac{P(B/A)P(A)}{P(B)} \quad (1)$$

Formula (1) calculates the conditional probability P(A/B) in terms of the *prior probability* P(A) which is fixed before the experiment, the conditional probability P(B/A) of the inverse process, and the *posterior probability* P(B). The last two probabilities are calculated with the help of the experiment’s data.

BR has been traced in many everyday life and science situations [18]. Recent researches have shown that most of the mechanisms of the human brain’s function are Bayesian [19]. This makes BR a very useful tool for AI. Note that fears have been already expressed that the Bayesian machines of AI could become too smart in future leaving to humans a second role only [20]! Sir H. Jeffrey's characterization, therefore, of the BR as the "Pythagorean theorem of probability theory" [21] becomes absolutely justified.

3. Artificial Neural Networks

The human brain consists of more than a billion neural cells each one of them working as a simple processor of information. The interaction between all cells and their parallel processing makes the brain's function and abilities possible.

A *biological neural network (BNN)* [4] is a group of biological neurons connected together by *axons* and *dendrites*, their connections been called *synapses*. The *cell-body* of a BNN processes the information, the axon enables the signal conducting and the synapses control the signals. Whereas axon is the output part of a neuron, the dendrites are input elements, which receive synaptic signals from other neurons. The transmission of signals is achieved by diffusion of chemicals called *neurotransmitters*.

ANNs are inspired and based on the function of BNNs trying to simulate the learning process of the human brain [4, 5]. The term "artificial" means that ANNs are implemented in computer programs being able to handle the large number of necessary calculations during the learning process (*neural computing*) [22]. An ANN is a group of artificial neurons or nodes connected together in a way analogous to BNNs. The connections of the biological neurons are modelled in ANNs as weights between nodes. Each artificial neuron performs a particular little operation and the overall ANN's operation is the weighted sum of all these operations. The basic components of an ANN include the *input*, the *hidden layer* and the *output layer*. Fig. 1, retrieved from Wikipedia, represents graphically a simplified form of the structure of an ANN.

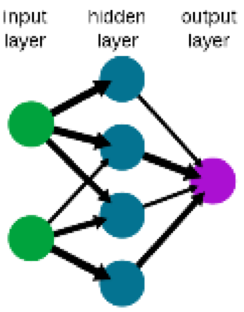

Fig. 1: Graphical representation of the structure of an ANN

An ANN must be trained to make each set of inputs to produce the desired outputs. This is done by feeding teaching patterns to the ANN and letting it to change its weighting function according to some already imposed to it learning rules. The ANNs make in general extensive use of the learning theories [23].

An ANN is usually designed for specific operations and works best if the relations between the inputs and the outputs are non-linear. ANNs are highly efficient for tackling problems in which there are no algorithms or specific rules to be followed for their solution. The applications of ANNs include analysis, classification and recognition of data, pattern recognition, control, associative memory, image processing and compression, forecasting applications, weather and stock market prediction, security and loan applications, etc. [4, 5, 24].

A disadvantage of ANNs is that there is not any general methodology for training and verifying them, therefore they cannot be used as universal tools for solving problems. Also, excessive training may be required in complex ANNs. The quality of the outcomes of an ANN depends on the accuracy of the given data.

Fuzzy systems and ANNs simulate the operations of the human mind. The ANNs, having the ability to learn and to process rapidly the information, concentrate on the "hardware" of the human brain, whereas fuzzy systems concentrate on the "software" emulating the human reasoning. Fig. 2 represents graphically the relationships among ANNs, FL and PR within the wider class of SC [2].

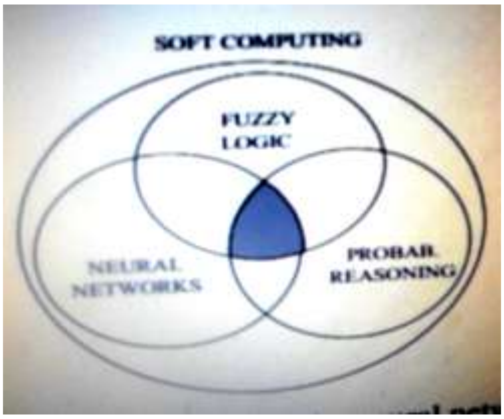

Fig. 2: Graphical representation of the relations among FL, PR and ANNs.

The intersections in Fig. 2 include probabilistic approaches to ANNs and FL systems, BR and *neuro-fuzzy systems*, which are hybrid systems using a learning algorithm from an ANN to determine their fuzzy parameters. Characteristic examples of such kind of systems are the *adaptive neuro-fuzzy inference systems (ANFIS)* providing accelerated learning capacity and adaptive interpretation capabilities to model complex patterns [24].

4. Genetic Algorithms

GAs are models of AI inspired by the process of evolution in nature, which are included in the wider class of *evolutionary computing* [5, 25]. Their basic idea is to emulate the natural selection for finding the best solution to optimization problems appearing to real-life situations. A GA, for example, can search through several designs to find the best combination resulting to the better or cheaper design of a device.

Alan Turin proposed in 1950 an abstract "learning machine" which would parallel the principles of evolution, but computer simulation of evolution started in 1954 with the work of the Norwegian – Italian mathematician N. A. Barricelli who used the computer of the Institute for Advanced Study in Princeton, USA [26]. The evolution in a GA, being an iterative process, starts with a randomly created population - called *generation* - of possible solutions of the corresponding optimization problem, which are usually referred as *individuals*. The initial generation's size varies according to the form of the corresponding problem, but it typically contains several hundreds or even thousands of individuals. Often the initial generation includes the entire range of the possible solutions, referred as the feasible region or search space.

In the next step each individual is evaluated, with the help of a properly defined *fitness function*, on how well fits the desired requirements. The best fit individuals of the first generation are selected for breeding the next generation, along with a small proportion of less fit individuals, which ensures the genetic diversity of the subsequent generation (*mutation*)

A new generation is created then by combining the best aspects of the selected individuals (*crossover*). The same circle is repeated until a termination condition is reached; e.g. a solution satisfying a minimum criterion, or the completion of a pre-fixed number of generations, or the end of the existing budget, or, in the best case, when successive iterations no longer produce better results, or even a combination of the previous conditions.

GAs have found successful applications in many sectors, such as biomedical and control engineering, code-breaking, games' theory, automated manufacturing and design, climatology, etc. On the contrary, GAs cannot be used easily in problems where the number of individuals exposed to mutation is large (e.g. design of a house, of a complex engine, of a plane, etc.), due to the high possibility of an exponential increase of the search space. In such cases GAs can be applied only if the corresponding problem is broken down into the simplest possible parts (e.g. air-foils instead of whole aircraft designs). Also GAs cannot solve effectively problems in which the fitness criterion is simply a right/wrong statement, e.g. decision making problems. The suitability of a GA depends on the amount of knowledge about the problem under solution. For problems with well-known data, better, more specialized, approaches often exist for their solution.

Another limitation in the use of GAs is that the search for finding the optimal solution to complex problems often requires a very expensive fitness function evaluation. Also, the termination condition is not always clear. Further, in certain problems the GAs converge to local optima rather, than to the global one not knowing how to sacrifice the local in favor of the global fitness, etc.

6. Discussion and Conclusions

In the present work we reviewed two of the main computer paradigms applied to SC, namely AANs and GAs. These paradigms are tolerant to uncertainty, imprecision, partial truths and approximation and are used for solving complex problems which cannot, or it is very difficult to be described by mathematical models. In particular, it is becoming difficult today to control the increasing complexity of the modern machinery using traditional control systems [27].

AANs and GAs, however, have also some disadvantages with respect to the traditional and stable methods of hard computing. The ANNs cannot be used as universal tools for solving related problems, because there is not any general methodology for training and verifying them. Further, excessive training may be required in complex ANNs. The use of GAs, which are usually designed for special applications only, has also a number of serious limitations, such as the frequently very expensive evaluation needed for their fitness function, the existing in certain cases vagueness about the termination process, their tendency to converge to local optima rather than to the global one, etc.

In the current era of the 4IR and the growth of the advanced IoT, however, as the processing by computer devices is increasing and their cost is reduced, the need of using SC methods has become too important and has the potential to expand further. The attempt, therefore, to improve and expand the SC techniques is an interesting and very promising area for further research.

References:

[1] Voskoglou, M.Gr, Computers and Artificial Intelligence as Tools for Education in the Forthcoming Era of the Internet of things and energy, *WSEAS Transactions on Information Science and Applications*, 16, 185-190, 2019.

[2] Schwab, K., *The Fourth Industrial Revolution*, Crown Publishing Group: NY, USA, 2016.

[3] Gupta, P., Kulkarni, N., An Introduction of Soft Computing Approach over Hard Computing, *Int. Journal of Latest Trends in Engineering and Technology*, 3(1), 254-258, 2013.

[4] Paplinski, A.P., Basic concepts of Neural Networks and Fuzzy Logic Systems, 2005, available on line https://www.cic.ipn.mx /~pescamilla/ContInt/Paplinski2005.pdf

[5] Ibrahim, D., An overview of soft computing, *Procedia Computer Science,* 102, 34-38, 2016.

[6] Zadeh, L.A., Fuzzy logic, neural networks and soft computing, *Communications of the ACM,* 37(3), 77-84, 1994.

[7] Siddique, N., Adeli, H., *Computational Intelligence: Synergies of Fuzzy Logic, Neural Networks and Evolutionary Computing*, John Wiley & Sons, Chichester, UK, 2013.

[8] Mitchell, M. *Artificial Intelligence: A Guide for Thinking Humans*; Parrar, Straus and Gtraux: New York, NY, USA, 2019.

[9] Kastranis, A. *Artificial Intelligence for People and Business*, O' Reily Media Inc.: Sebastopol, CA, USA, 2019.

[10] Lejewski, C., Jan Lukasiewicz, *Encycl. Philos.*, 5, 104–107, 1967.

[11] Tarski, A., *Encyclopaedia Brittanica,* 2018. Available online: www.britannica.com/ biography/Alfred-Tarski .

[12] Zadeh, L.A. Fuzzy Sets, *Inf. Control*, 8, 338–353, 1965.

[13] Zadeh, L.A., Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Trans.Syst. Man Cybern.*, *3*, 28–44, 1973.

[14] Voskoglou, M.Gr., Generalizations of Fuzzy Sets and Related Theories, in M. Voskoglou (Ed.), *An Essential Guide to Fuzzy Systems,* Nova Science Publishers, NY, pp. 345-352, 2019.

[15] Jaynes, E.T., *Probability Theory: The Logic of Science*, Cambridge University Press, UK, 8th Printing, 2011 (first published, 2003).

[16] Kosko, B., Fuzziness Vs Probability, *Int. J. of General Systems*, 17*(2-3)*, 211-240, 1990.

[17] Gentili. P.L., Establishing a New Link between Fuzzy Logic, Neuroscience and Quantum Mechanics through Bayesian Probability: Perspectives in Artificial Intelligence and Unconventional Computing, *Molecules*, 26, 5987, 2021.

[18] Voskoglou, M.Gr., Bayesian Reasoning and Artificial Intelligence, *WSEAS Transactions on Advances in Engineering Education*, 17, 92-98, 2020.

[19] Bertsch McGrayne, S., *The Theory that would not die*, Yale University Press, New Haven and London, 2012.

[20] Brockman, J., What do you think about machines that think?, 2015. Available on http://edge.org/response-detail/26871

[21] Jeffreys, H., *Scientific Inference,* 3d Edition, Cambridge University Press, UK, 1973.

[22] Buckley, J.J., Hayashi, Y., Fuzzy neural networks: A survey, *Fuzzy Sets and Systems,* 66, 1-13, 1994.

[23] Voskoglou, M.Gr., Connectivism Vs Traditional Theories of Learning, *American Journal of Educational Research*, 10(4), 257-261, 2022.

[24] Jang, J.-S.R., ANFIS: Adaptive, network-based fuzzy inference system, *IEEE Transactions on Systems, Man and Cybernetics*, 23(3), 665–685, 1997.

[25] Banzhaf, W., Nordin, P., Keller, R., Francone, F., *Genetic Programming – An Introduction,* Morgan Kaufmann, San Francisco, USA, 1998.

[26] Barricelli, N.A., Esempi numerici di processi di evoluzione, *Methodos*, 45–68, 1954.

[27] Voskoglou, M.Gr., Fuzzy Control systems, *WSEAS Transactions on Systems,* 19, 295-300, 2020.