Stages of Conceptualization and Formalization in the Design of the Model of the Neuro-Fuzzy Expert System of Professional Selection of Pupils

Bogdan B. Buyak[1][0000-0003-1496-7573], Ivan M. Tsidylo[1][0000-0002-0202-348X],
Victor I. Repskyi[1][0000-0003-4297-5335] and Vitaliy P. Lyalyuk[2][0000-0001-7793-1397]

[1] Ternopil V. Hnatiuk National Pedagogical University,
2, M. Kryvonosa St., Ternopil, 46027, Ukraine
{buyak.bogdan, tsidylo, repskyi}@tnpu.edu.ua
[2] Kryvyi Rih Metallurgical Institute of the National Metallurgical Academy of Ukraine,
5, Stephana Tilhy St., Kryvyi Rih, 50006, Ukraine
vitalij.lyalyuk@gmail.com

Abstract. The article describes the problem of designing a neuro-fuzzy expert system of professional selection at the stages of conceptualization and formalization, which involves the definition of concepts, relationships and management mechanisms necessary to describe the solution of problems in the chosen subject field. The structural model of the decision making system for determining the professional selection of students for training in IT specialties is substantiated. Three subsystems are proposed as structural components for studying: psychological peculiarities, personal qualities, factual knowledge, abilities and skills of students. The quality of the system's operation is determined by the use of various techniques for acquiring knowledge on the basis of which the knowledge base of the neuro-fuzzy system and the combination of the use of fuzzy and stochastic data will be formed.

Keywords: professional orientation, professional selection, expert system, design, neuro-fuzzy modeling, conceptualization, formalization, structural model.

1 Introduction

1.1 The Problem Setting

One of the topical problems of artificial intelligence is the construction of diagnostic neuro-fuzzy expert systems, the substantiation of their structure and individual subsystems. Such systems make it possible to combine various forms of presentation and methods of processing knowledge. They are the study environment for the optimal combination of various mechanisms of knowledge processing in the process of solving the problem of effective interaction of heterogeneous information in conditions of uncertainty. The range of application of fuzzy data processing methods is expanding every year, covering areas such as the design of industrial robots and household appliances, automatic recognition of speech and images, and so on.

The social and humanitarian sphere of human activity concerning the application of intelligent technologies of automation of information processes and modeling of systems that are characterized by fuzzy and vague input values has not remained unaltered. One of the most active and perspective areas of applied research, based on the construction of information systems in terms of the theory of fuzzy sets – fuzzy simulation. The advantage of this approach is that fuzzy simulation allows you to obtain more adequate results in comparison with results based on the use of traditional analytical models and management algorithms, as well as to simulate complex systems with multiple choice and heterogeneous input data.

Professional orientation and professional selection of young people for the study of IT specialties is a complex multi-faceted process, the automation of which requires a non-standard approach and the application of artificial intelligence technologies, in particular methods and algorithms of fuzzy logic. In the process of analyzing the factors that influence the professional choice of students, in addition to the psychological and personal qualities of the individual, one must take into account the tendency to study certain disciplines, in particular, the disciplines of the professional-practical cycle of the curriculum. In addition, an external evaluation of the specialists of the chosen forecasting field of education is required as well as their independent recommendations for choosing a future profession.

1.2 Analysis of Recent Research and Publications

Problems of professional orientation and professional selection of students, in particular, practical aspects, are considered in many works by well-known teachers and scholars: Ivan D. Bekh [2], Valentyna T. Lozovetska [12], Nellia H. Nychkalo [14], Olena M. Otych [15], Olha I. Shcherbak [17], Illia O. Teplytskyi [16], Hryhorii V. Tereshchuk [9], Ivan A. Ziaziun [24] and others.

Considerable attention has been paid recently to the study of the possibility of automating the processes of vocational guidance and professional selection using a variety of information systems – Nailia R. Iangurazova [8], Baryi G. Iliasov [8], Elena B. Startceva [8], Andrei V. Shulepov [19] and others. In particular, [8] offers the construction of an expert system for the selection of not one, but several specialties when enrolling in a higher education institution. The basis of this approach is the modular principle of building the knowledge base of the expert system, as well as the advantages of applying this principle. To create an expert system, the authors analyzed and modeled the subject area using a structural approach.

Shulepov [19] on the basis of a review of expert systems on the choice of specialty when enrolling in the university offered a mathematical model of the system of assessment of own opportunities and the choice of a rational strategy in the labor market for a graduate of high school taking into account external socio-economic factors. The structure of such a system, the functional model and the data storage scheme are described.

The definition of students' success is one of the key determinants when choosing a profession, based on fuzzy logic: Andrei R. Aidinian [1], Aleksandr P. Kirpichnikov

[9], Irina V. Leonova [9], Stella A. Liasheva [9], Mikhail P. Shleimovich [9], Olga L. Tcvetkova [1], Ivan M. Tsidylo [20] and others.

In the paper [9] a methodology for assessing the quality of teaching of students of higher educational institutions is proposed, which is based on the use of a two-level system built on the basis of the adaptive system of neuro-fuzzy output ANFIS implemented in the package of Fuzzy Logic Toolbox MatLab system. The system for assessing the quality of training, which allows for numerical evaluation as a summary indicator characterizing the performance of an educational institution and shows how well the learning process is conducted, is proposed in [1]. Another approach to constructing a fuzzy evaluation of the results of automated testing is proposed by us in [20]. The structure, functions and mechanisms of constructing a fuzzy system model are described. The response surface of the fuzzy system reflects the dependence of the final score on the complexity of the task and the degree of correctness of the task. Testing of the controller on the test sample has been carried out, proving the functional suitability of the developed model.

To date, the scientists have made attempts to generalize conceptual studies of the problem: the application of intellectual information technologies in practical tasks – Oleh M. Berezkyi [4], Maksim I. Dli [10], Roman Iu. Golunov [10], Vladimir V. Kruglov [10], Serhiy D. Shtovba [18], F. Wasserman, and others, the construction of fuzzy expert systems – George F. Luger [13], Serhiy D. Shtovba [18], etc., decision-making in situations for which both fuzzy and stochastic data should be used – Aleksei G. Beliakov [3], Olha V. Hlon [7], Volodymyr M. Dubovoy [7], Oleh O. Kovaliuk [5], Aleksandr S. Mandel [6] and other national and international scientists.

1.3 Unresolved Aspects of the Problem

The analysis of works shows the presence of a certain structure and stage of the design of neuro-fuzzy expert systems. However, little or no attention has been paid to detailing the implementation of the most important stages – conceptualization and formalization, in particular, in the process of developing a model of the system of professional selection of students.

1.4 The Purpose

The purpose of the article is to substantiate the essential features, concepts, relationships and management mechanisms necessary to describe the solution of the problem of designing a neuro-fuzzy expert system for determining the professional suitability of future students for the training of IT specialties.

2 The Results of the Research

Designing a neuro-fuzzy expert system to determine the professional suitability of students to learn future IT skills is based on the most commonly used approach using fuzzy

knowledge bases and the hybrid learning algorithm ANFIS. Peculiarities of the implementation of such systems, as stated in the paper [8, p. 120] are as follows: processing of fuzzy statements, that is, when the prerequisite is fuzzy variables, and the output machine – a mechanism for obtaining data from them; the use of a matrix of fuzzy relations, when a set of factors and set of prerequisites are determined; the matrix contains fuzzy variables whose measure is represented as a real number within [0, 1], and to determine the causes of the state, the transformation of the matrix and factors into the form of equations of fuzzy relations is carried out, and then the resulting system is solved by the method of composition of the minimum-maximum; use of fuzzy excerpts.

In accordance with the stages of the development of intelligent systems, let us consider the most time-consuming stages – conceptualization and formalization, which, according to the authors [16]: "is, indeed, "a stumbling block" at the very first stage of the development of models of intelligent systems". At these two stages, one can say, the feasibility of using one or another type of knowledge for the presentation of information is determined. At the stage of obtaining knowledge, we analyze the problem area and determine the methods of solving problems.

First of all, you need to pay attention to the development environment of the future system. The software environment for designing will be MATLAB – one of the oldest, carefully elaborated and time-tested automation systems for mathematical and scientific and technical calculations, built on the extended presentation and application of matrix operations [20]. The specific environment for the development of a neuro-fuzzy expert system is fuzzy output editor Fuzzy Inference System Editor. This editor consists of the following programs: Membership Function Editor, Rule Editor, Rule Viewer, and Surface Viewer [23, p. 18].

The problem of professional selection of students is the correspondence of psychophysiological characteristics of a student with the requirements of the IT specialty. Based on this, the goals of the system development include: to determine the main criteria for professional review; to show their significance and validity; to simulate the work of the expert system and evaluate its effectiveness.

Accordingly, professional selection involves a system of activities that can identify people who, by their individual personal qualities, are most suitable for studying and continuing professional activities in a particular profession. Proficiency is preceded by students' inclination and abilities, quality of training, interests, and environment. The stages of professional selection include: psychological study of the profession in order to identify the requirements for a person; selection of psycho-diagnostic methods of research, including tests that most characterize those mental processes and occupational actions, concerning which the professional suitability is to be assessed; forecasting the success of training and further activities on the basis of comparison of information: a) about the requirements of the profession to the person and received psychodiagnostic data with an emphasis on the assessment of personality characteristics; b) the possibility of purposeful improvement and compensation of professionally important qualities, as well as the probability of adaptation to the profession, the possibility of emergence of extreme situations and influences.

In the final version, the system should determine the degree of professional competence of the student for further training on the IT specialty based on the prescribed

criteria on the basis of the programmed knowledge. Taking into account the complexity of the task of professional selection, we divided all the input data into three subsystems: psychological peculiarities of the student, student's personal qualities, factual knowledge, abilities and skills (KAS's) in possession of the appropriate programming languages. Each of these subsystems will have separate input variables that will be responsible for the formation of this knowledge base with subsequent fuzzy logic output for a certain type of criteria. The outputs of the three subsystems that will set the condition will be inputs to the overall resultant system, the output of which will be the final result of the student's professional suitability. As a result, we obtain a hierarchical expert system that will operate as inbound variables both with fuzzy linguistic and statistical information. Thus, we have completed the stage of development of intelligent systems – conceptualization.

At the formalization stage, we fill in the system with input data to form a knowledge base. The general system will have three subsystems described above, as well as output variables that will eventually show the result. The first subsystem will be a subsystem called "psychological peculiarities of students". In each subsystem, first of all, it is necessary to provide input data or they can also be called input variables. The input data of the first subsystem include: memory, attention, imagination, thinking and temperament. We got five input variables. But this is not enough for the validity of the subsystem. Each input variable will contain its varieties, so that during the examination we will be able to determine which type of psychological features is most developed. In the language of fuzzy logic they are called fuzzy linguistic variables. For example, what kind of attention or thinking is dominant. Based on available varieties of memory, the first input variable of this subsystem will contain the following variables: figurative, verbal-logical (in content), short-term, long-term, operational (in duration), involuntary and arbitrary (by the way of memorization). The second input variable is attention. Accordingly, we introduce its types, namely: external, internal, involuntary, arbitrary, post-satisfactory. The third parameter is imagination. Its varieties, and accordingly, linguistic variables: artistic, technical, and scientific. The fourth input variable is thinking: visually-effective, visually-figurative, verbal-logical. The last, fifth input element is the student's temperament. Its varieties are as follows: choleric, sanguine, phlegmatic, melancholic.

The knowledge of the second subsystem of "student's personal qualities" is presented in the following way. The input elements of the subsystem will be: language proficiency, inclinations and abilities, overall capacity, success. We also use varieties of input elements in the form of linguistic variables, in order to classify each of them in more detail. The first input element is to speak the language, both native and foreign. Its varieties are as follows: Ukrainian, English, Russian, German, Polish. The second input element is the student's personal inclinations and abilities. Here we will include the following: educational, creative (in terms of application), theoretical and practical (by type of thinking) [11]. The third input element is the total capacity. Its varieties are as follows: depth and breadth of mind, self-discipline, self-control, activity, initiative, responsibility. The fourth element of this subsystem is the success – high, low, and average. This subsystem will output a result that describes the input variables for the general system, the level of personal qualities of the person who was tested.

The third subsystem will be a subsystem called "KAS's" (knowledge, abilities, skills). This will be the largest subsystem and will include two more subsystems both of which will be affiliated with the IT specialty "Web programmer". The first is the knowledge of programming languages, and the second is a software package.

Subsystem "Knowledge of programming languages". Here it is necessary to foresee types of programming languages and the level of proficiency in them. It should be noted that in the knowledge base we will introduce the types of programming languages in two directions: client programming languages and server programming languages. The first direction will include: HTML5, XTML, CSS3, Java-Script along with the JQuery library. The server languages include: PHP, C#, ASP.NET, Java. Each input variable will provide the level of proficiency: high, sufficient, average, low, zero. It is these two areas of programming languages that play an important role, since the expert system, when deducing the result, will not only be able to show whether a professionally qualified student will study in this specialty, but will also output the result of the direction of student learning in the IT industry. Whether this will be the direction of preparation closer to web design or it will be a server programmer.

The second subsystem is a "Software product". The input of its knowledge base will be those software products that the future web programmer should be directly proficient in. Therefore, the input variables of the subsystem will be the following software applications and packages: Notepad, Sublime Text, Php Storm, Aptana, DreamWeaver, MySQL, Microsoft Visual Studio, Photoshop (for designers). Each of the above input variables will have linguistic variables corresponding to the level of proficiency: high, sufficient, average, low, zero.

The two subsystems described above will provide a fuzzy logical output, the initial results of which will be input data for the subsystem "KAS's" whose output values, in turn, will be input values for the general neuro-fuzzy expert system for determining the professional suitability of future students to study IT skills.

The developed structural model of the system (Fig. 1) allows for obtaining complete and visual information about the subject domain, justify the structure of knowledge bases and select and develop the means of its filling, determine the requirements for the software and hardware of the system. The organization of the structure of the model was made on the basis of the modular approach, which allowed us in a separate module (subsystem) to provide means and methods for obtaining, processing and preservation specific to it, the knowledge base in separate units of knowledge and their use by several subsystems in accordance with the hierarchy of acceptance processes solution to the task. The knowledge base will be formed on the basis of the creation of so-called heuristic rules that will consist of a different combination of the principle "if (condition) and (condition) then (result)", which will provide a separate result. Validity of any expert system depends on the number of heuristic rules. The more rules the expert provides, the better will be the examination and the more accurate results at the output. However, not always a greater number of rules will lead to a better result. Therefore, to optimize the system, we propose the use of the ANFIS algorithm.
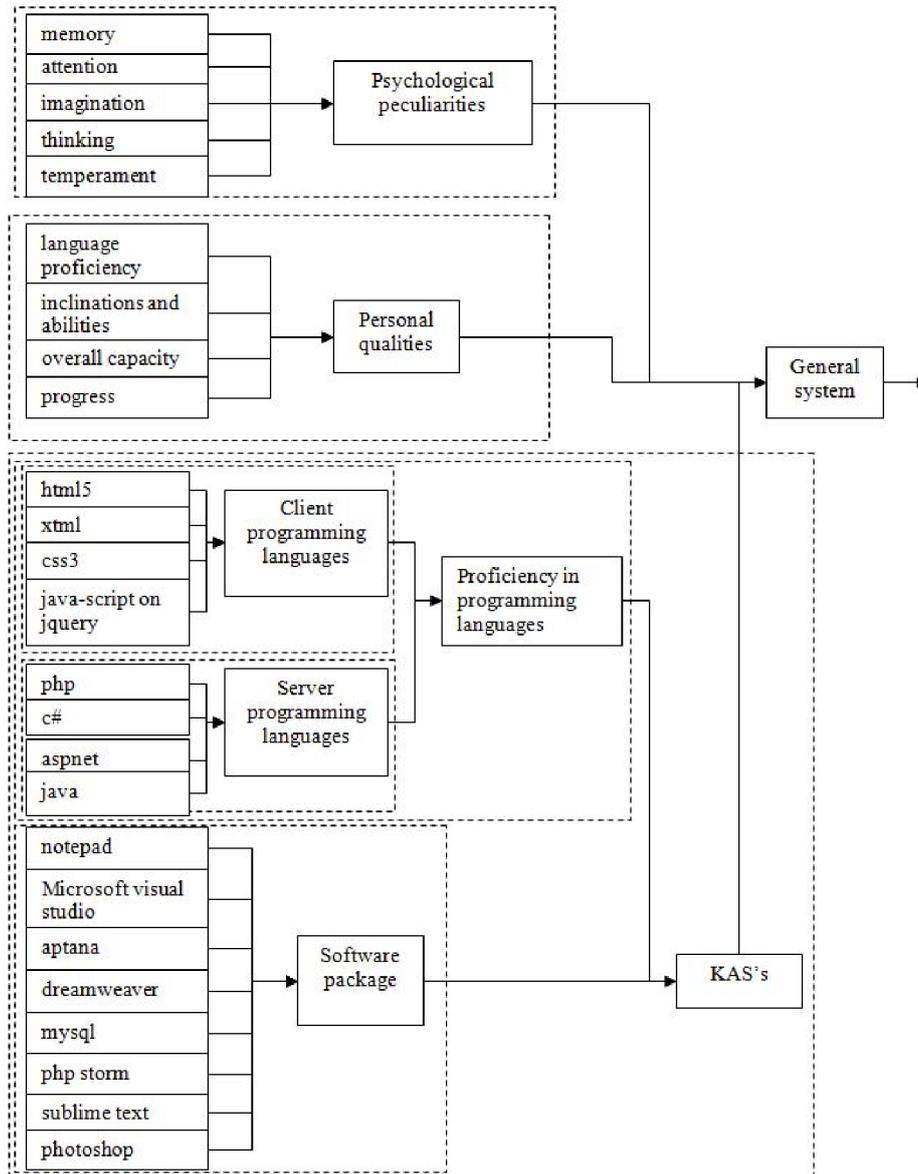
Fig. 1. Structural model of the general neuro-fuzzy expert system for determining the professional suitability

3 Conclusions and Prospects for Further Research

Thus, for the automated expert assessment of the professional competence of students for the training of IT specialties, subsystems of decision-making have been allocated

and their hierarchy has been determined, on the basis of which a general model is created that describes all the decision-making processes. This approach can be attributed to the IDEF methodology, which is defined and used to solve problems of modeling complex systems. Modeling of processes of professional orientation and professional definition can be attributed precisely to such systems, which allow to represent and analyze the models of activity of a diverse spectrum. As systems models we propose the following subsystems: psychological peculiarities, personal qualities, factual knowledge, skills and abilities. Hence, the width and depth of the study of processes in the system are being met. Namely, both by the use of various techniques for obtaining knowledge on the basis of which the knowledge base of the neuro-fuzzy system will be formed, and by the combination of the use of fuzzy and stochastic data.

The structural order of the inseparable, in our opinion, elements and characteristics of the student's personality reflects our vision of the automation of the process of professional orientation and professional decision. The interdisciplinary nature of the task is beyond doubt. Failure to take into account some features took place at both the conscious and unconscious levels in order to avoid overloading the created model with unnecessary data.

The prospect of further inquiries into this issue is seen in the implementation of the designed system in the MATLAB software environment, in particular in the Fuzzy Logic Toolbox, and the involvement of specialists in pedagogy, psychology, physiology, programming, atomization of information processes, etc., at the stages of conceptualization and formalization. After all, only in cooperation of experts from various industries one can get high-quality and effective software product, including a computer-aided neuro-fuzzy expert system for determining the professional suitability of students for the training in IT specialties.

References

1. Aidinian, A.R., Tcvetkova, O.L.: Metodika otcenki kachestva obucheniia studentov vuza s ispolzovaniem neiro-nechetkogo podkhoda (Methods of estimation of the quality of students' education with the use of neuro-fuzzy approach). Programmnye produkty i sistemy. **29**(4), 189–193 (2016)
2. Bekh, I.D.: Stanovlennia profesionala v suchasnykh sotsialnykh umovakh (Becoming a professional in modern social conditions). Pedahohika tolerantnosti. 3–4, 157–162 (2001)
3. Beliakov, A.G., Mandel, A.S.: Analiz dostovernosti vyvodov, formiruemykh s pomoshchiu ekspertno-statisticheskikh sistem (The analysis of reliability of the conclusions formed with the help of expert statistical systems). Ros. AN. In-t probl. upr. im. V.A. Trapeznikova, Moscow (2002)
4. Berezsky, O., Melnyk, G., Batko, Yu.: Modern Trends in Biomedical Image Analysis System Design. In: Laskovski, A. (ed.) Biomedical Engineering, Trends in Electronics, Communications and Software, pp. 461–480. InTech, Rijeka (2011)
5. Dubovoi, V.M., Kovaliuk, O.O.: Modeli pryiniattia rishen v upravlinni rozpodilenymy dynamichnymy systemamy (Models of decision-making in the management of distributed dynamic systems). UNIVERSUM-Vinnytsia, Vinnytsia (2008)
6. Granin, S.S., Mandel, A.S.: Stationary Inventory Control Policies in Supply Systems under Inflation Condition. Automation and Remote Control. **77**(8), 1453–1460 (2016)

7. Hlon, O.V., Dubovoi, V.M.: Modeliuvannia system keruvannia v umovakh nevyznachenosti (Modeling of control systems under uncertainty). UNIVERSUM-Vinnytsia, Vinnytsia (2004)

8. Iliasov, B.G., Startceva, E.B., Iangurazova, N.R.: Modulnaia struktura bazy znanii ekspertnoi sistemy "Vybor spetcialnosti pri postuplenii v vuz" (Modular structure of the knowledge base of the expert system "The choice of specialty at admission to an institution of higher education"). Neftegazovoe delo. 2. http://ogbus.ru/authors/Ilyasov/Ilyasov_1.pdf (2006). Accessed 29 Jan 2017

9. Kirpichnikov, A.P., Liasheva, S.A., Shleimovich, M.P., Leonova, I.V.: Proektirovanie avtomatizirovannoi sistemy obrabotki dannykh uspevaemosti studentov na osnove nechetkoi logiki (Designing an automated system for data processing of students' progress based on fuzzy logic). Vestnik Kazanskogo tekhnologicheskogo universiteta. 372–376 (2014)

10. Kruglov, V.V., Dli, M.I., Golunov, R.Iu.: Nechetkaia logika i iskusstvennye neironnye seti (Fuzzy logic and artificial neural networks). Izdatelstvo fiziko-matematicheskoi literatury, Moscow (2001)

11. Krushelnytska, Ya.V.: Fiziolohiia i psykholohiia pratsi (Physiology and psychology of labor). KNEU, Kyiv (2003)

12. Lozovetska, V.T.: Profesiine stanovlennia osobystosti v suchasnykh umovakh pratsi (Psychological and pedagogical aspects of professional formation of personality in modern working condition). Naukovyi visnyk Instytutu profesiino-tekhnichnoi osvity. 5, 20–24 (2013)

13. Luger, G.F.: Artificial Intelligence: Structures and Strategies for Complex Problem Solving (6th Edition). Pearson, London (2008)

14. Nychkalo, N.: Profesiinyi rozvytok osobystosti u konteksti neperervnosti (Professional development of the individual in the context of continuity). In: Kremen, V.H., Dmytrychenko, M.F., Nychkalo, N.H. (eds.) Kontseptualni zasady profesiinoho rozvytku osobystosti v umovakh yevrointehratsiiinykh protsesiv, pp. 12–23. NTU, Kyiv (2015)

15. Otych, O.: Mystetstvo u profesiinomu rozvytku osobystosti (Art in the professional development of personality). In: Kremen, V.H., Dmytrychenko, M.F., Nychkalo, N.H. (eds.) Kontseptualni zasady profesiinoho rozvytku osobystosti v umovakh yevrointehratsiiinykh protsesiv, pp. 571–582. NTU, Kyiv (2015)

16. Polishchuk, O.P., Teplytskyi, I.O., Semerikov, S.O.: Profesiina spriamovanist fundamentalizatsii informatychnoi osvity (Professional orientation of fundamentalization of informational education). Theory and methods of learning mathematics, physics, informatics. **11**(3), 122–129 (2013)

17. Shcherbak, O.: Psykholoho-pedahohichni zasady profesiinoho rozvytku i samorozvytku pedahoha profesiinoho navchannia (Psychological-pedagogical principles of professional development and self-development of a teacher of professional training). In: Kremen, V.H., Dmytrychenko, M.F., Nychkalo, N.H. (eds.) Kontseptualni zasady profesiinoho rozvytku osobystosti v umovakh yevrointehratsiiinykh protsesiv, pp. 472–481. NTU, Kyiv (2015)

18. Shtovba, S.D., Pankevich, O.D., Nagorna, A.V.: Analyzing the criteria for fuzzy classifier learning. Automatic control and computer sciences. **49**(3), 123–132 (2015). doi:10.3103/S0146411615030098

19. Shulepov, A.V.: Postroenie ekspertnoi sistemy vybora ratcionalnoi strategii na rynke truda dlia vypusknika vuza (Constructing an expert system for selecting a rational strategy in the labor market for a graduate of the university). Vestnik kibernetiki. 8 (2009)

20. Tcidylo, I.N.: Nechetkaia otcenka rezultatov avtomatizirovannogo testirovaniia (Fuzzy evaluation of the results of automated testing). Problemy sovremennoi nauki. **11**(1), 105–114 (2014)

21. Tereshchuk, H.V.: Pedahohichna diahnostyka tsinnisnykh oriientatsii molodi v protsesi yii sotsialnoho i profesialnoho stanovlennia (Pedagogical diagnostics of value orientations of youth in the process of its social and professional formation). Pedahohika i psykholohiia. 3(12), 119–124 (1996).

22. Wasserman, Ph.D.: Neural computing: theory and practice. Van Nostrand Reinhold Co., New York (1989)

23. Zhernakov, S.V.: K voprosu o postroenii gibridnykh neiro-nechetkikh ekspertnykh sistem diagnostiki i kontrolia GTD (On the construction of hybrid neuro-fuzzy expert systems for diagnostics and control of GTE). In: Upravlenie v slozhnykh sistemakh, pp. 119–126. UGATU, Ufa (1999)

24. Ziaziun, I.A.: Pedahohichnyi profesionalizm u konteksti profesiinoi svidomosti (Pedagogical professionalism in the context of professional consciousness). In: Pedahoh profesiinoi shkoly, pp. 8–17. Naukovyi svit, Kyiv (2001)