INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS, 2024, VOL. 70, NO. 1, PP. 191-197
Manuscript received October 2, 2023; revised March, 2024. doi: 10.24425/ijet.2024.149530

Optimization of computer ontologies for e-courses in information and communication technologies

Nazym Sabitova, Yuriy Tikhonov, Valerii Lakhno, Kariyrbek Makulov, Olena Kryvoruchko, Vitalyi Chubaievskyi, Alona Desiatko, and Mereke Zhumadilova

***Abstract*—A methodology is proposed for modifying computer ontologies (CO) for electronic courses (EC) in the field of information and communication technologies (ICT) for universities, schools, extracurricular institutions, as well as for the professional retraining of specialists. The methodology includes the modification of CO by representing the formal ontograph of CO in the form of a graph and using techniques for working with the graph to find optimal paths on the graph using applied software (SW). A genetic algorithm (GA) is involved in the search for the optimal CO. This will lead to the division of the ontograph into branches and the ability to calculate the best trajectory in a certain sense through the EC educational material, taking into account the syllabus. An example is considered for the ICT course syllabus in terms of a specific topic covering the design and use of databases. It is concluded that for the full implementation of this methodology, a tool is needed that automates this procedure for developing EC and/or electronic textbooks. An algorithm and a prototype of software tools are also proposed, integrating machine methods of working with CO and graphs.**

***Keywords*—electronic courses; electronic textbooks; computer ontologies; graph; matrix description; subject area; genetic algorithm**

I. INTRODUCTION

IN the era of education globalization and the rapid pace of information and communication technology (ICT) development, the need for innovative teaching methods and tools is undeniable. These methods can be used not only in higher education institutions but also in schools, starting from primary grades. This task is also relevant, for example, when retraining specialists in a specific subject area (SA), especially in ICT, where the content of educational material changes quite dynamically.

The relevance of conducting research in this direction is justified by the fact that electronic courses and/or electronic textbooks (hereinafter referred to as EC) of such an SA as ICT should have increased cognitive value. Moreover, some ICT knowledge, due to the pace of new information growth, may be lost when studied in print [1, 2]. This fully applies, for example, to such EC subsections as: programming, computer networks, computing systems, databases, etc.

Considering the possibility of representing the CO ontograph of the SA in the form of a graph [2], the large dimensionality of the resulting graphs, and the existing developments in computer representation and graph processing, we can state that new research is needed in the direction of transforming the formal description of the CO ontograph in the form of a graph and using graph processing techniques, in particular, the genetic algorithm (GA), as one of the most efficient algorithms for solving optimization problems. All of the above determined the relevance of the problems addressed in this research.

II. LITERATURE REVIEW

In the works [1, 2, 3, 4], the usefulness of knowledge representation methods using CO (Computer Ontologies) is demonstrated.

Ontological approaches in the process of creating EC (Electronic Courses) have also been the subject of many studies. For instance, in [4, 5], it is shown that ontological models are intended for the implementation of information technology for automated processing of CO in the SA (Subject Area).

In the study [6], a system of ontologized design of EC is described, which translates EC from the information level to a level based on ontological knowledge.

The computer ontology of the educational SA should be cognitive. This can be achieved by taking the universally significant CO of the SA as a basis and then modernizing it for a specific course. The basis is the syllabus or the working curriculum of the discipline, for example, ICT, or a separate module of the academic discipline. In this context, metrics for graphs can be used.

In the work [7], the correspondence of structure, functionality, and usability of CO is studied. However, the authors do not provide descriptions of specific already implemented practical solutions that would help practicing educators implement their ideas, for example, by resorting to specialized software.

Nazym Sabitova is with L. N. Gumilyov Eurasian National University, Astana, Kazakhstan (e-mail: sab_nazym@mail.ru).

Yuriy Tikhonov is with Luhansk Taras Shevchenko National University, Poltava, Ukraine (e-mail: uruytihonov1@gmail.com).

Valerii Lakhno is with National University of Life and Environmental Sciences of Ukraine, Kyiv, Ukraine (e-mail: lva964@nubip.edu.ua).

Kaiyrbek Makulov and Mereke Zhumadilova are with Yessenov University, Aktau, Kazakhstan (e-mail: kaiyrbek.makulov@yu.edu.kz, mereke.zhumadilova@yu.edu.kz).

Olena Kryvoruchko, Vitalyi Chubaievskyi and Alona Desiatko are with State University of Trade and Economics, Kyiv, Ukraine (e-mail: {kryvoruchko_ev, chubaievsyi_vi} @knute.edu.ua, desyatko@gmail.com).

In [8, 9, 10], the authors primarily focused on syntactic, semantic, and pragmatic metrics used in the compilation of EC.

In [11], aspects of the metric of the schema and the entire knowledge base (hereinafter KB) of the ontology are considered. It should be noted that this study is more theoretical than practice-oriented, as it does not offer a specific toolkit for optimizing CO (Computer Ontologies).

In [12], the authors analyze the quality criteria of CO. According to the authors [6, 7, 12, 13, 14], such criteria include: completeness and accuracy of the vocabulary of the considered SA (Subject Area); adequacy of the CO structure in terms of relationships; performance when implemented in applications; selection of the best CO from a set of available ones; perception of CO from a cognitive point of view.

As shown in [6, 7, 13, 14], methods for evaluating the quality of graphs can be used to modernize the CO of a study course when applied to ontographs.

For ontological engineering and methods of computer processing of subject knowledge, there is a gap between well-developed methods and tools for individual stages and solving applied tasks in specialized SAs. However, the tasks of automation and optimization of CO are not fully resolved. As a result, there is a large proportion of manual work in building CO for EC (Electronic Courses). Hence, the relevance and importance of the problem of developing new methods for modifying CO for EC, by using computer graph processing techniques when representing the formal ontograph of CO in the form of a graph.

The methodology should ultimately provide the integration of mathematical methods and mechanisms, algorithms, and tools for effective processing of SA and tasks for a given SA (in our case, using the discipline of ICT as an example); automated optimization of computer ontologies for EC in ICT.

III. Problem Statement

Given the need to modify the CO (Computer Ontologies) of the SA (Subject Area), using the example of ICT (Information and Communication Technologies), by employing computer graph processing techniques when representing the formal ontograph of CO in the form of a graph, this study examines the application of computer graph processing techniques, specifically the use of GA (Genetic Algorithms) and the development of tools that integrate machine methods of working with CO and graphs.

IV. Main Methods Used in the Research

During the research, the following methods were used: matrix representation of graphs corresponding to the ontograph of CO SA; computer description of matrices; a genetic algorithm for finding minimal paths on a graph.

V. Main Part of the Research

It is necessary to develop an automated CO SA design process, divided into separate branches to increase the degree of cognitive EC (Electronic Courses) and/or EU (Electronic Units). An example of the ICT discipline is considered, and more specifically, we will look at the topic - "Databases (DB)". The example is chosen as it is closely related to accompanying topics, such as programming, networks, algorithmization, etc.

Let's consider a fragment of the CO DB, see Fig. 1.

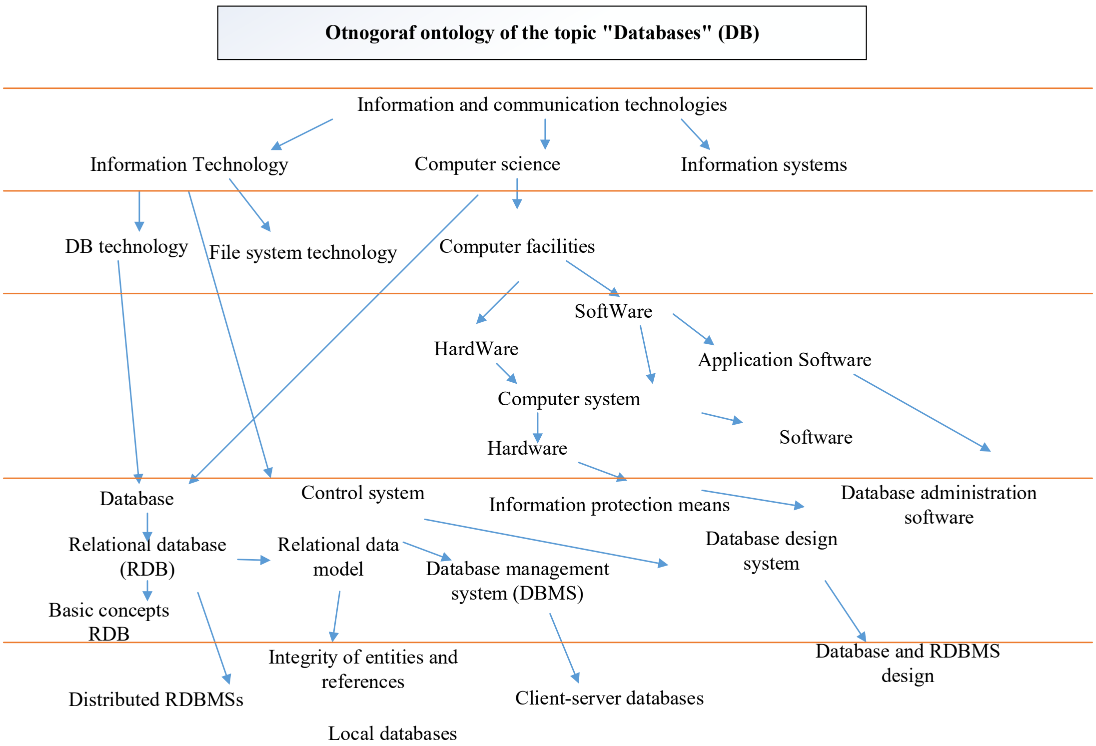


Fig. 1. Fragment of the ontograph of the CO PDO ICT (topic Databases)

In Figure 1, the root concept of the ontograph is divided into levels of depth of knowledge about PDO (horizontal red lines). Moreover:

- The depth, height of the concept (tier), and width of the graph are large;
- It's not a tree, and it's not balanced;
- The presence of cycles also hinders perception;
- The graph is branched;
- Edge density, characterizing the graph's closeness to a fully connected graph, is irregular (this metric will automatically detect errors when constructing the CO by an expert);
- The variety of the number of connections is irregular;
- The graph's complexity (including the analysis of the ratio of nodes with multiple inheritances to all nodes of the graph and the analysis of the average number of parent nodes at a node) is high.

Overall, if all the fragment's concepts and their connections are included in the ICT EC, the EC will be relevant to modern knowledge about PDO but will be difficult to perceive and weakly cognitive.

To improve perception, the expert (in this case, the ICT discipline teacher preparing the corresponding EC) outlines, based on the syllabus and their experience, the depth level of studying the PDO, indicating on the ontograph the concepts up to which the student should study the PDO. All higher concepts up to the root are automatically included in the EC based on the CO description. To improve perception and for the correct operation of the EC, the ontograph is divided into branches. Each branch leads from the root concept to the final one specified by the expert. Graph processing methods are used.

The procedure for selecting the final concepts of a branch is poorly formalized. For example, one could set the depth of PDO study (and the final branch concepts) by the levels of the ontograph. However, as noted above, the heights of the concepts (tier) of this graph are large, i.e., there may be several concepts at the level of one branch, on the other hand, the expert may believe that the final concepts do not necessarily have to be located on the same level. Therefore, it is advisable to manually select the final branch concepts.

For example, for the EC on the topic of Databases, 200 concepts were selected. If you display them all on the screen (Fig. 2), it is impossible to work with it.

Therefore, the ontograph is displayed on the screen with navigation tools that allow you to view it in fragments, see Fig. 3. On it, the final vertices of the branches are selected.

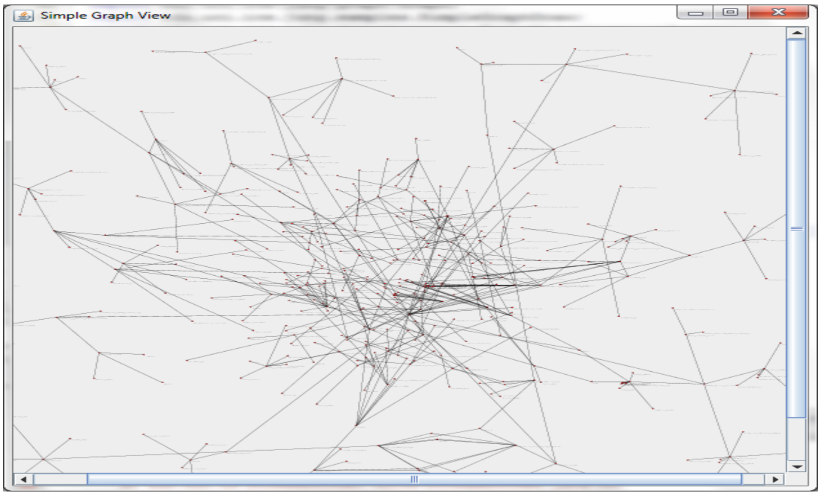


Fig. 2. Displaying the entire database ontograph on the screen

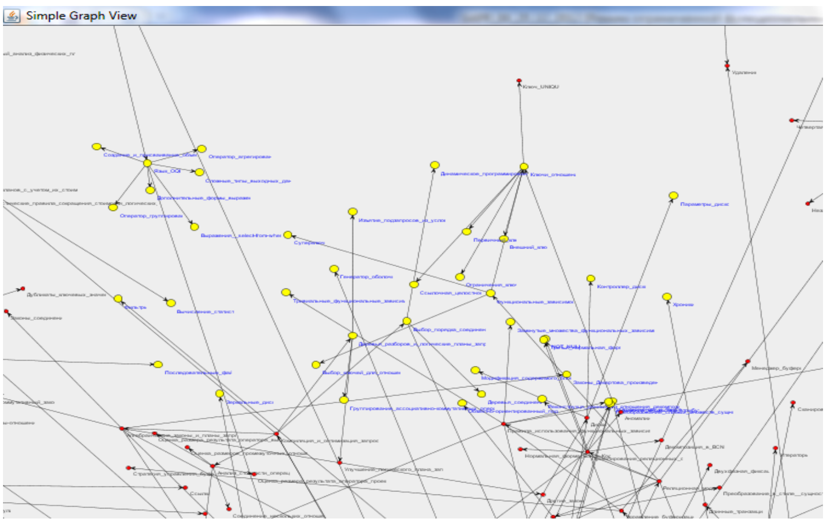


Fig. 3. Displaying a fragment of the CO (Cognitive Object) on the screen for expert work

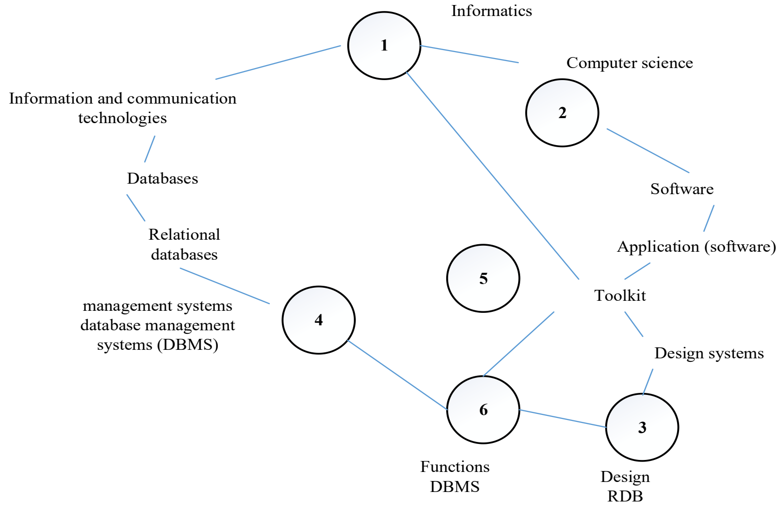


Fig. 4. Vertices selected by the expert (subjectively selected for illustration)

Next, it is necessary to create a matrix description of the CO fragment selected by the expert. Let's provide an example. As mentioned earlier, the educational CO of the PDO (Professional Development Object) database includes about 200 concepts. Working with a matrix of dimension 200, especially manually performing genetic algorithm procedures, is not rational. Therefore, for illustration, we take a small fragment of the ontograph. We will select the vertices chosen by the expert from the fragment (see Fig. 1) (see Fig. 4) and illustrate the matrix description on them and then the steps of the GA (Genetic Algorithm) operation.

In Figure 4, the selected vertices are numbered. Note that the selection of vertices was done subjectively, with the aim of demonstrating the general concept of computer optimization capabilities of the ontograph.

Let's simplify the graph obtained, as a result of which we get the following image, see Fig. 5.

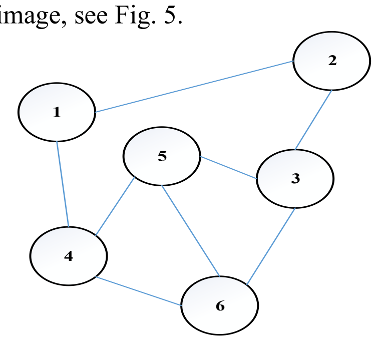


Fig. 5. Initial graph for the GA (Genetic Algorithm)

Then, the adjacency matrix for the graph shown in Figure 5 will look as shown in Table I.

TABLE I
FRAGMENT OF THE KB FORMED FOR DETECTING CYBER THREATS TO PAC

		Numbers of graph vertices					
		1	2	3	4	5	6
Numbers of graph vertices	1	0	1	0	1	0	0
	2	1	0	1	0	0	0
	3	0	1	0	0	1	1
	4	1	0	0	0	1	1
	5	0	0	1	1	0	1
	6	0	0	1	1	1	0

The initial population of the GA (Genetic Algorithm) [15, 16] is generated using corresponding algorithms when solved by machine. The most well-known algorithm selects the attributes of a new object randomly. The initial population is created once at the beginning of the GA's operation, and then it is replaced by more adapted individuals. In the example, the initial gene is chosen in the form of a random set. An example of such a set is shown in Figure 6.

1	2	3	4	5	6	№
5	6	4	1	3	2	Gene

Fig. 6. Encoding of a chromosome consisting of vertices of the original graph

Next, a pair of parent chromosomes is selected, see Figure 7.

1	2	3	4	5	6	№ vertex
1	2	3	4	4	6	Parent 1
6	5	4	3	2	1	Parent 2

Fig. 7. Parent chromosomes

The crossover operator is applied to the parent chromosomes, the result is shown in Figure 8.

1	2	3	4	5	6	№
1	2	3	6	5	4	Descendant 1
6	5	4	1	2	3	Descendant 2

Fig. 8. Offspring chromosomes

Next, the mutation operator is applied. In the considered example, a simple method of swapping two genes was used for gene mutation, see Figure 9. Two genes are randomly selected and swapped.

1	2	3	4	5	6	№
5	6	4	1	3	2	Gene

a) Chromosome before mutation

1	2	3	4	5	6	№
5	3	4	1	6	2	Gene

b) Chromosome after mutation

Fig. 9. Example of the mutation operator's work during the optimization of CO concepts

The "tournament selection" [15] was used as the selection mechanism. That is, from a population containing individuals *N*, individuals were randomly selected *T*. The best individual is then recorded in the intermediate population. In essence, a tournament was held between the selected individuals. This operation was repeated *N* times. Then, the individuals in the obtained intermediate population were randomly crossed. With the growth of parameter *T*, the selection between individuals becomes stricter. This is due to the fact that if an individual has a low fitness indicator, it has no chance to "produce offspring".

For the small example considered, after 15 generations, the following result was obtained, see Figure 10.

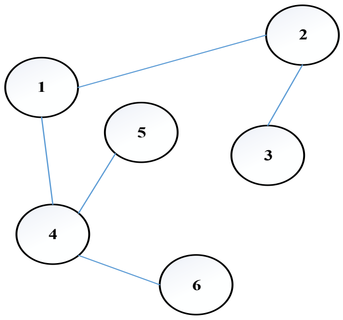


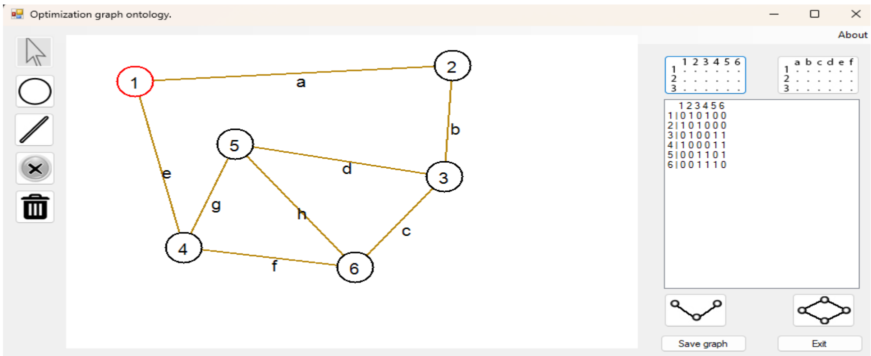
Fig. 10. Graph scheme with the procedure for dividing the vertices of the CO graph

This algorithm was implemented in the Python language in the Visual Studio 2022 programming environment, see Figure 11. As a result, an optimized representation of the ontograph for the considered part of the EC "ICT" was obtained, see Fig. 12.

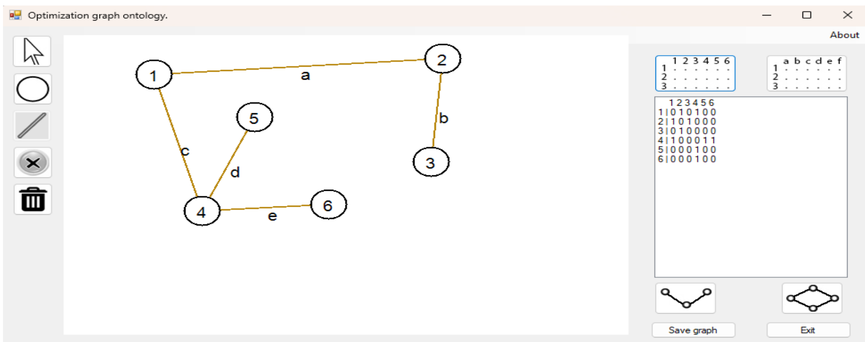
In Figure 12 a) the initial GA graph and its adjacency matrix for the vertices of the graph, displaying the concepts of the educational discipline selected by the expert, are shown.

In Figure 12 b) the final graph and its adjacency matrix for the vertices of the graph, displaying the concepts of the educational discipline selected by the expert after optimization using the genetic algorithm, are shown.

Such cybernetic modeling and visualization of some stages of EC design allow us to talk about the prospects of conducting a wider range of work in the direction of automating the compilation of EC with the possibility of optimizing the compiled CO. In fact, the results obtained in comparison with Figure 4 allow the expert to receive specific recommendations for branching out, see Figure 12.



a) The initial GA graph and its adjacency matrix for the vertices of the graph, displaying the concepts of the educational discipline selected by the expert



b) the final graph and its adjacency matrix for the vertices of the graph, displaying the concepts of the educational discipline selected by the expert after optimization using the genetic algorithm

Fig. 11. An example of screenshots of the prototype system for automated selection of final branch ontograph concepts for an electronic textbook on TKT

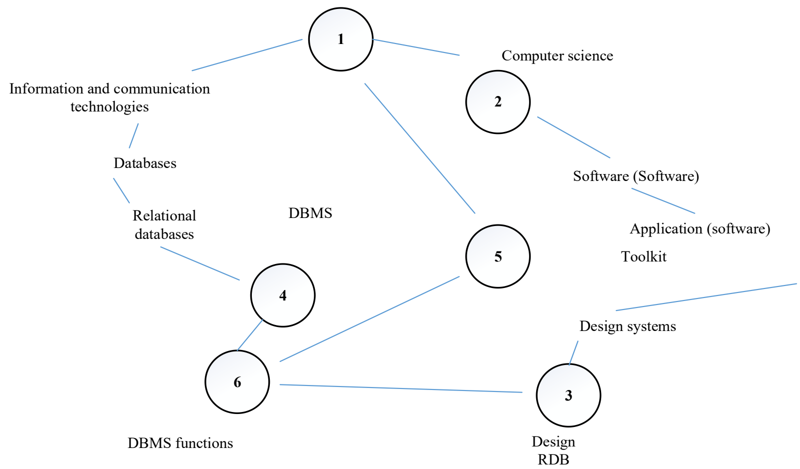


Fig. 12. Optimized division of branches of the ontograph fragment for the EC (Electronic Textbook)

VI. Methods and Models

Based on the conducted research, the following algorithm for preparing an ontograph for a specific PDO (Personalized Digital Object) or EC (Electronic Textbook) is proposed, see Figure 13.

In block 1, the CO (Conceptual Object) description in the OWL language is entered into the PC's RAM. The description of a specific educational discipline in the OWL language is obtained in advance, for example, through the Protege editor.

In block 2, the ontograph is displayed on the screen.

In block 3, the selection of final concepts is done manually due to the weak formalization of this task.

In block 4, the CO description in OWL is transformed into a computer matrix description of the corresponding graph (the adjacency matrix is automatically filled).

In block 5, the chromosome definition is determined, and the attributes of the new object are selected randomly.

In block 6, a pair of parent chromosomes is selected.

In block 7, the crossover operator is applied to the parent chromosomes.

In block 8, the mutation operator is applied.

In block 9, "tournament selection" [15] is used as the selection mechanism.

As mentioned earlier, this algorithm is implemented in the C# programming language in the Visual Studio 2022 programming environment. The proposed algorithm can be used by a teacher to prepare an ontograph for a specific PDO or EC (Electronic Textbook).

VII. Discussion of the Obtained Results

The proposed methodology for modifying CO for EC, using the example of one of the ICT PDO topics, by representing computer ontologies as a graph and using graph processing techniques and finding optimal paths on the graph, in particular, the genetic algorithm, can be used for the automated preparation of EC. However, at this stage, there are no corresponding computer tools. In particular, for converting the OWL description of computer ontologies into a computer description of the graph corresponding to the CO ontograph. Such a transformation manually is not feasible due to the large dimensionality of the task.

Therefore, for a full-fledged implementation of this methodology, it is necessary to create a tool that automates this procedure for developing EC and/or electronic textbooks.

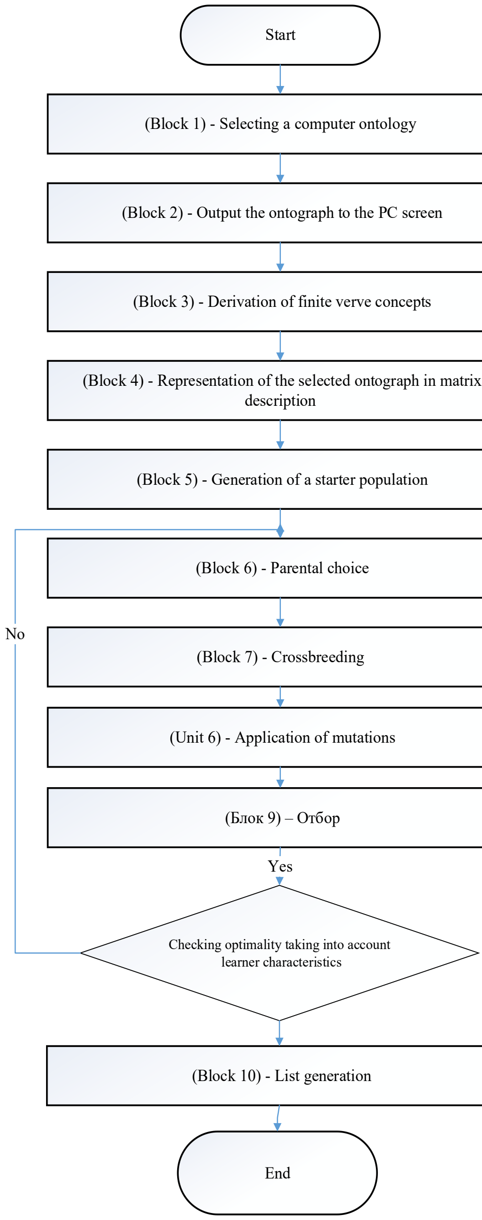


Fig. 13. Flowchart for preparing CO (Conceptual Object) for EC (Electronic Textbook) based on the use of GA (Genetic Algorithm) for ontograph optimization

Conclusion

As a result of the research carried out within the framework of this article, the following main results were obtained:

1. The methodology for modifying computer ontologies (CO) for electronic courses or electronic textbooks was considered. This methodology is implemented by representing the formal ontograph of CO as a graph and using graph processing techniques to eliminate cycles and find minimal paths on the graph. For the latter task, it is proposed to use a genetic algorithm. The application of genetic algorithm at this stage will lead to the division of the ontograph into branches and the possibility to calculate the best trajectory for the educational material of electronic courses and/or electronic textbooks in a certain sense, taking into account the syllabus (or the working program of the discipline).

2. Within the framework of the mentioned methodology, a technique and an algorithm for preparing the CO ontograph for use in automated EC development have been proposed, reducing costs and preparation time.

3. It was concluded that for more effective use of the proposed methodology, it is necessary to develop a full-fledged software toolkit. In the developing toolkit, it is necessary to take into account the transformation of the OWL description of computer ontologies into a computer description of the graph corresponding to the CO ontograph.

4. An algorithm for the operation of such a toolkit has been proposed, integrating machine methods for working with CO and graphs.

5. The obtained results are aimed at increasing the efficiency of preparing electronic courses and/or electronic textbooks for higher educational institutions, schools, or extracurricular institutions.

6. For the successful application of the proposed methodology during subsequent research, it is planned to develop an original software tool to implement all stages of the methodology for modifying computer ontologies.

References

[1] Monika Rani, Riju Nayak, O.P. Vyas, "An ontology-based adaptive personalized e-learning system, assisted by software agents on cloud storage," *Knowledge-Based Systems*, Volume 90, 2015, Pages 33-48, ISSN 0950-7051, doi:10.1016/j.knosys.2015.10.002, 33-48

[2] A. Ameen, K. U. R. Khan and B. P. Rani, "Creation of Ontology in Education Domain," 2012 *IEEE Fourth International Conference on Technology for Education*, Hyderabad, India, 2012, pp. 237-238, doi:10.1109/T4E.2012.50

[3] Uschold, M., & Gruninger, M., "Ontologies: Principles, methods and applications," *The Knowledge Engineering Review, 11*(2), 93-136, doi:10.1017/S0269888900007797

[4] Sabitova, B.SH. Razakhova, L.U. Taimuratova, al etc., "The use of ontological modeling in the preparation of electronic courses in the field of information and communication technologies", *Journal of Theoretical and Applied Information Technology*, Vol.101, No 15, PP. 5999-6013.

[5] Tikhonov, U., Lakhno, V., Skliarenko, E., Stepanenko, O., & Dvirnyi, K., "Development of ontological approach in e-learning when studying information technologies," *Eastern-European Journal of Enterprise Technologies*, *5*(2 (83), 13–20, doi:10.15587/1729-4061.2016.79230

[6] Palagin, A.V., "An Ontological Conception of Informatization of Scientific Investigations," *Cybern Syst Anal* 52, 1–7 (2016), doi:10.1007/s10559-016-9793-6

[7] Palagin, A., Petrenko, N.G., Velychko, V., Malakhov, K., & Tikhonov, Y.L., "To the problem of "The Instrumental complex for ontological engineering purpose" software system design," *ArXiv, abs/1802.03544*..

[8] Duque-Ramos, A., Fernandez-Breis, J.T., Iniesta, M., Dumontier, M., Aranguren, M.E., Schulz, S., Aussenac-Gilles, N., Stevens, R. 2013. "Evaluation of the OQuaRE framework for ontology quality," *Expert Systems with Applications*, 2013. № 40. P. 2696–2703.

[9] Burton-Jones, A., Storey, V., Sugumaran, V., Ahluwalia, P., "A semiotic metrics suite for assessing the quality of ontologies," *Data & Knowledge Engineering*, 2005. Vol. 55. № 1. P. 84–102.

[10] Wang, Y., Bao, J., Haase, P., Qi, G., "Evaluating Formalisms for Modular Ontologies in Distributed Information Systems," In: Marchiori, M., Pan, J.Z., Marie, C.d.S. (eds) Web Reasoning and Rule Systems. RR 2007. *Lecture Notes in Computer Science*, vol 4524. Springer, Berlin, Heidelberg, doi:10.1007/978-3-540-72982-2_13

[11] Tartir, S., Arpinar, I., Moore, M., Sheth, A., Aleman-Meza, B., "OntoQA: metric-based ontology quality analysis," In: *Proceedings of the Workshop on Knowledge Acquisition from Distributed*, Autonomous, Semantically Heterogeneous Data and Knowledge Sources (KADASH), Citeseer, 2006.

[12] Gangemi, A., Catenacci, C., Ciaramita, M., Lehmann, J., "Modelling Ontology Evaluation and Validation," In: Sure, Y., Domingue, J. (eds) The Semantic Web: Research and Applications. ESWC 2006. *Lecture Notes in Computer Science*, vol 4011. Springer, Berlin, Heidelberg, doi:10.1007/11762256_13

[13] Darrel Whitley, "A Genetic Algorithm Tutorial;" November 10, 1993; *Technical Report CS-93-103* (Revised); Department of Computer Science, Colorado State University, Fort Collins, US.

[14] Eremeev, A.V., "Modeling and Analysis of Genetic Algorithm with Tournament Selection," In: Fonlupt, C., Hao, JK., Lutton, E., Schoenauer, M., Ronald, E. (eds) Artificial Evolution. AE 1999. *Lecture Notes in Computer Science*, vol 1829. Springer, Berlin, Heidelberg, doi:10.1007/10721187_6

[15] G. Alexandru-Lucian and A. Iftene, "Using a genetic algorithm for optimizing the similarity aggregation step in the process of ontology alignment," 9th *RoEduNet IEEE International Conference*, Sibiu, Romania, 2010, pp. 118-122.

[16] J. Wang, Z. Ding and C. Jiang, "GAOM: Genetic Algorithm Based Ontology Matching," 2006 *IEEE Asia-Pacific Conference on Services Computing (APSCC'06)*, Guangzhou, China, 2006, pp. 617-620, doi:10.1109/APSCC.2006.59

[17] V. Lakhno, V. Malyukov, B. Akhmetov, B. Yagaliyeva, O. Kryvoruchko and A. Desiatko, "University Distributed Computer Network Vulnerability Assessment," 2023 *IEEE International Conference on Smart Information Systems and Technologies (SIST),* Astana, Kazakhstan, 2023, pp. 141-144, doi:10.1109/SIST58284.2023.10223501

[18] Valerii Lakhno, Zhuldyz Alimseitova, Yerbolat Kalaman, Olena Kryvoruchko, Alona Desiatko, and Serhii Kaminskyi, "Development of an Information Security System Based on Modeling Distributed Computer Network Vulnerability Indicators of an Informatization Object," *International Journal of Electronics and Telecommunications* Volume 69, Issue 3, Pages 475 – 483, doi:10.24425/ijet.2023.146495

[19] Lakhno, V. et al. (2023), "The Model of Server Virtualization System Protection in the Educational Institution Local Network," In: Shakya, S., Papakostas, G., Kamel, K.A. (eds) Mobile Computing and Sustainable Informatics. *Lecture Notes on Data Engineering and Communications Technologies*, vol 166. Springer, Singapore, doi:10.1007/978-981-99-0835-6_33