Ontological Approach to Plot Analysis and Modeling

Yevhen Burov[0000-0001-6124-3995], Victoria Vysotska[0000-0001-6417-3689], Petro Kravets [0000-0001-8569-423X]

Lviv Polytechnic National University, Lviv, Ukraine
yevhen.v.burov@lpnu.ua[1], victoria.a.vysotska@lpnu.ua[2], petro.o.kravets@lpnu.ua[3]

Abstract. The paper addresses the problem of structural plot aspects analysis for both fiction and non-fiction works. For this, an approach based on ontological modeling is proposed. The requirements and the ontology for plot analysis and construction is developed. The structure of plot is represented by contextual graphs, containing scenes and action nodes. The meta-information, describing the specific scene is formalized using contextual ontology being a subset of common ontology of discourse. The proposed approach can be used for creation of new tools facilitating the detailed and multi-faceted analysis of literary works.

Keywords: plot analysis, ontology, discourse, contextual graph, scene

1 Introduction

One of the central problems of computational linguistics is the detection and analysis of meaning for both fiction and non-fiction literary works. The meaning is a complex, multidimensional concept. It hides not only in the meaning of sentences and paragraphs, but also in the structure of the text when one part references or reinforces what has been said in other part. The thoughts of author in text form the arcs of cohesion [1] which contribute to better understanding and overall logical integrity of the text. A complex text is a collection of ideas, stories, anecdotes, tropes, and allusions to real and hypothetical events, which should not contradict but rather reinforce each other within the context of a goal or the main message of literary work. The attentive reader should be aware of the soundness and relevance of research underlying the author's text and should be able to follow and study it in detail if needed. Moreover, the literary work is usually a part of larger discourse, and detecting all its meaningful components allows to establish the relations with other works within the context of discourse.

The implementation of the computational analysis of meaning requires the developing of tools for meta-analysis of text which can be used by writers and readers on all stages of literary work existence, starting from planning and writing and ending by reading and analysis.

2 Background analysis

The idea of usefulness of a common formalized framework for the analysis of texts permeates the works on literary and discourse analysis for a long time. For example, [2] states that "literary pragmatists need a taxonomical apparatus which will apply to genres of all kind, literary or otherwise". Such apparatus can be used to detect and compare common patterns of meaning in texts.

The more recent work [3] elaborates on importance of narrative analysis of meaning, introduces the modes of meaning in detailed analysis of sentences. Author specifies modes of narrative (poetical, analytic) and different types of writing (scientific, historical) each having the different requirements to elucidation and understanding of meaning. However, the [3] considers the notion of ontology in the most general, philosophical sense, not trying to apply the plethora of concepts, methods and tools developed in the knowledge engineering area for knowledge representation and analysis.

Some articles use analytical, mathematical methods to analyze meaning. In [4] is developed the method of character analysis based on graphs for systemic analysis of literary works. Graphs represent characters as nodes and vertices as relations between characters. Authors introduce the graph models for dialogues, and build dialog and opinion networks. Those networks are also used for sentiment analysis. The plot structure model formalizes the change of the narrative contents or the recipient's emotion according to the event development. The [5] applies information-theoretic approach to the evaluation of narrative complexity and uses it for support of writing process.

In order to explore the dynamics of changes in characters or plot development in [6] the temporal networks of characters are used. Authors compute the time-averaged eigenvector centralities, Freeman indices and vitalities of characters as measures for evolution of characters in a story.

However, the current research explores and formalizes only some specific aspects of meaning in a literary work. In order to study meaning in all its complexity we need a unifying platform, the common set of formalized concepts and relations which can be used to represent different facets of meaning and various its models. The best candidate for this platform is arguably the modern knowledge engineering framework based on ontologies.

The researchers in the area of knowledge engineering were from the very beginning fascinated by formal representation of meaning conveyed by texts. For this, they developed semantic maps as graph-like structures which were later enriched by J. Sowa by adding formal logic constructs [7-8]. Conceptual graphs [8] were used to formalize the meaning to simple sentences. Later, on the basis of contextual graphs research the RDF ontology representation language was developed, which is the one of the mainstay elements in ontology engineering area [9]. The RDF formalizes meaning of simple sentences, having basic subject-predicate-object structure. The processing of texts from the specific area remains the preferred method of ontology construction [10].

Another area which can contribute to better understanding if meaning is situational awareness and decision support area. The analysis of scenes, understood as a collections of multimedia artifacts, objects and people present in specific time and location, requires the capturing and formalization of concepts and relations relevant to the scene. This is not unlike to the formalization of literary scene and can use the same approaches.

Thus [11] proposes to build a scene model based on simple qualitative descriptors (shape, color, location, topology) and use it for reasoning in home automation. The article [12] uses multimedia ontology for video scene interpretation and reasoning. This approach can be applied to understanding the meaning of scene in the movie.

However, current research results in the domain of ontological analysis of scenes or texts don't take into consideration the forms of meaning specific to literary works such as story, plot, character, references to research, and arcs of cohesion [13-14].

This paper elaborates the idea of using ontological engineering approach for plot and story analysis focusing on the aspects of meaning and structure specific to literary works.

3 Ontological modeling for plot analysis and construction

Ontological modeling reflects a specification-based approach to knowledge management, where initial specification of knowledge is built by domain experts and later processed by information systems [15-19]. It presents an alternative to machine learning, where knowledge specification is derived from processing huge datasets, evaluating the quality of obtained result and updating knowledge specification using feedback mechanisms [20-27]. The main advantage of ontological modeling compared to machine learning is the absence of learning datasets (which are not always available) [28-32]. Moreover, machine-learning models don't explain the rationale of obtained solution, while in specification-based approach the initial specification elements always has some justification [33-39].

In [13], ontology is defined as a formal model of a certain conceptualization of a subject area. Such a model contains the definition of the entities of the subject domain *En* and the relations *Rl* between them, the constraints and axioms *Ax*.

$$On = (En, Rl, Ax) \tag{1}$$

Ontology provides the shared conceptualization, containing concepts and relations, which can be reused in different application across the selected subject area [40-43]. Even more, it formalizes logical rules and dependencies existing in subject area and either automatically apply them or checks their conformance in specification created by expert [44].

In case of the plot analysis the availability of ontology will not only allow to analyze literary work, but also create the library of topics, ideas, tropes common in literary discourse and track their usage in multiple works [45-48].

Let's elucidate the main requirements for plot analysis ontology which stem from tasks performed [49-51]. Thus, the ontology

1. Should allow to formalize typical plot elements, such as characters, artifacts, ideas, research, moods, plots, roles etc.
2. Allow to represent the structure of narrative and also the structure of specific threads of narrative and their interactions and crossings.
3. Represent scene as basic element of a plot or thread in narrative.
4. Support timelines, which are events and scenes ordered according time of occurrence, allowing to check time-related consistency of work.
5. Allow to focus on development of selected characters or artifacts across the work
6. Analyze scenes selecting relevant elements and building references, describing dependencies to other scenes or external research sources.
7. Place markup in a text allowing to highlight particular text for a specified reference.
8. Use markup to specify specific repeating pattern in text. For example definition, classification, known trope or cliché.
9. Allow for logical reasoning, which is especially helpful, for example in building genealogies and defining family relationships between characters.
10. Support the specification and verification of external formal requirements for literary text, such as the number of characters, structural requirements. For example, the scientific article usually has text specifying authors and their affiliation, abstract, introduction, background analysis, method description, discussion and the list of references.

4 The ontology for plot analysis

The proposed ontology structure for plot analysis represented as a taxonomy of objects is shown on fig. 1. First of all, each literary work is always a part of discourse, sharing with other works common tropes, ideas, roles for characters. This part of ontology stores common elements, occurring in multiple works. Next comes the list of works included in discourse. For each work we specify such basic components as ideas and messages, characters, artifacts, events, research sources, text fragments. This list can be further enlarged with illustrations, multimedia components and even critical reviews. Those elements represent the basic elements for meta-analysis and discovering relationships and structures within analyzed work.

The scenes, which are the basic components of a plot are also presented as a component of ontology tree. This tree includes references to structures which can be perceived in the analyzed work. One of these structures is a plot, represented as contextual graph containing scenes. There are also the linear work structure representing the order of text fragments in the final published variant. Multiple timeline views help to focus on analysis of character development, artifact usage, or how the most important ideas or messages are conveyed across the work. The timeline and ideas views are also used to check for inconsistencies and logical conflicts. The ontology can be easily enriched by other elements reflecting, for example, the critical evaluations of particular elements or work as a whole.

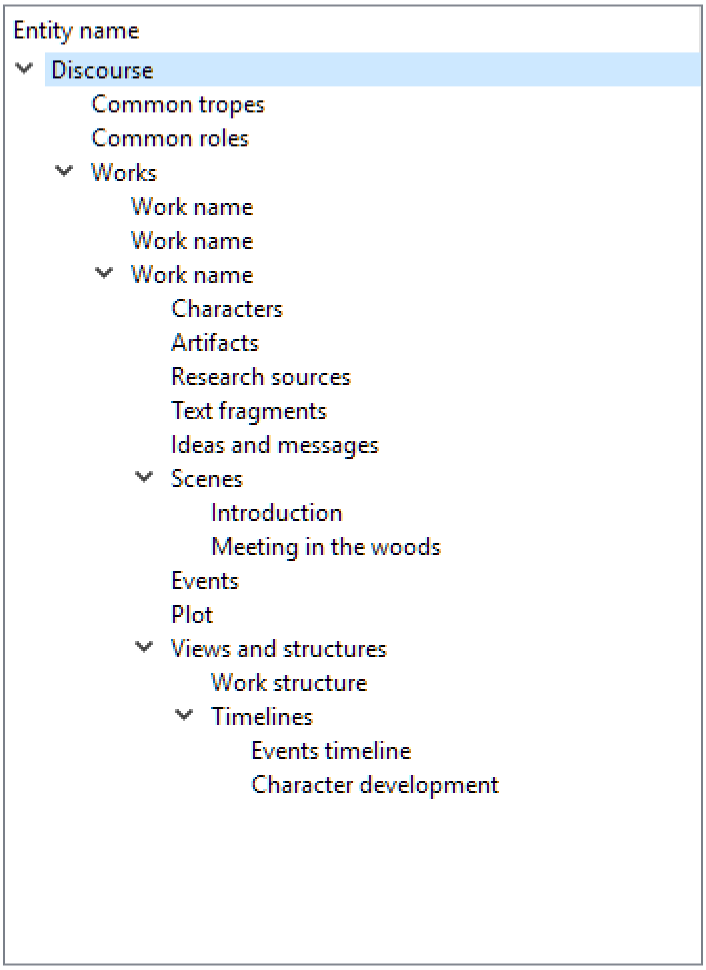

Fig. 1. The structure of ontology for plot analysis

5 Using contextual graphs to represent plot structure

Let us represent the structure of plot as a set of connected scenes. The connections between scenes can reflect time order or sequential order presented to reader or just possible variants of plot development as seen by author. The analysis of possible plot developments requires focusing on all relevant elements for each scene (scene context) which is helped by contextual graph model.

The model of context graphs was initially developed for analysis of emergencies by metro operators. [14]. The authors of this model note that further actions taken in particular situations (scenes) are hard to formalize and are influenced by human operator's experience. In order to make a correct decision the operator should have all relevant information and not be overwhelmed by unnecessary detail. The contextual graph not only represent the structure of possible solutions depending on contextual information, but allows the operator to focus on small set of relevant data in order to make better decision.

Contextual graph is an acyclic graph with one input node and one output node (fig 2).

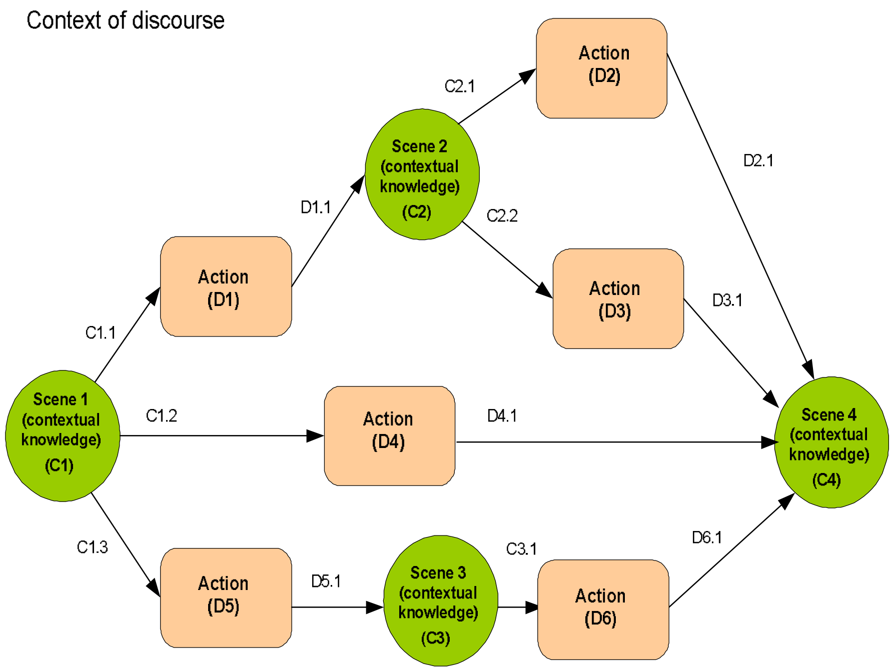

Fig. 2. Contextual graph representing a fragment of plot

The contextual graph has two types of components: contextual elements and actions. Contextual elements reflect contextual knowledge and actions – the possible actions (decisions) taken depending on the values of contextual element. The arcs coming from contextual elements to actions have marks showing which part of contextual knowledge is used in action. In more than twenty years after their development, the contextual graphs were successfully used to solve tasks in the areas of healthcare, military, transportation, information security [14].

With the purpose of plot analysis, we will interpret the contextual elements as scenes-nodes. Such nodes combine not only actual scene description as literary text, but also references to events, position in general story timeline, links to relevant research, ideas. The action node, which usually comes after scene will be construed as a change in the plot which occurred after action. It will contain information about how basic plot elements (characters, events, artifacts) were changed after scene and thus will reflect the evolution of plot. The action node coming after the specific scene can be split in several parts describing dependencies between scenes. In the process of plot construction action nodes can describe the alternative paths for plot development and can be used by author for the evaluation and selection of the best path.

Formally, the contextual graph used for plot analysis and construction is represented by tuple:

$$Pl = (Sc, Ac, Ar), \tag{2}$$

where *Sc* is a set of all scene descriptions, *Ac* is the set of all actions descriptions and *Ar* is a set of arcs between them.

6 Scene analysis and representation

A scene node is the important part of plot structure. It represents the context in which this scene is created. In knowledge-based systems the context is often used as a filter limiting the number of parameters to be tracked. In order to represent the knowledge pertaining to the scene it is advisable to use contextual ontology [15] $Con \subseteq On$, that is a part of general ontology On. The definitions of contextual ontologies are stored in knowledge base. In practice it is often handy to track the change of contextual ontologies between scenes by splitting the contextual ontology in parts, according to important descriptive parameters of the story. Those parts are considered as separate contexts. For example, such parts may be the contexts of locations, characters, time, and narrative ideas.

$$Con = Con_{char} \cup Con_{loc} \cup Con_{time} \cup Con_{ideas} . \tag{3}$$

The changes of those contexts are tracked across the story and for each scene *i* the contextual scene ontology Sc_i is built and instantiated.

This ontology typically contains references to instances of such ontology types as Characters, Artifacts, Location, Ideas, Time, Research, Text fragments. Therefore, the ontology of scene can be represented as a tuple:

$$Sci = (T_{char}, T_{art}, T_{loc}, T_{idea}, T_{time}, T_{res}, T_{txt}) . \tag{4}$$

Similarly, information about action Ac_j can be represented using ontology types instances, inherited from scene description.

7 Conclusion

This article describes an approach which can be used in the processes of plot analysis and construction for both fiction and non-fiction literary works. Conversely to the most used approach of formalizing meaning using concepts, relations and axioms, the proposed approach focuses on formalizing and studying larger concepts such as common tropes, ideas and patterns. The representation and analysis of such objects is required in the study of narrative and discourse being an important part of computational linguistics. Its usefulness transcends the notion of just another author's and analyst's toolbox. It can be used to construct or analyze the work focusing on specific facets such as stories, characters, and threads. Ontological approach facilitates detecting and tracking common tropes or other repeating patterns in multiple works across the discourse.

Moreover, the usage of ontologies for plot analysis allows using logical or rule-based reasoning enforcing structural requirements and other detectable patterns, detecting logical inconsistencies. Contextual graphs could help to construct multiple variants of alternate plots developments within the same literary work. Contextual ontologies and contextual graphs can constitute a basis for detecting and formalizing repeating patterns of meaning across multiple works.

References

1. Pinker, S.: The sense of style: The thinking person's guide to writing in the 21st century. New York, Penguin Books (2015)
2. Ventola, E.: Approaches to the Analysis of Literary Discourse. In: Tidningsbokhandeln, PB33, SF-21601, Pargas, Finland. (1991)
3. Champigny, R.: Ontology of the narrative: An analysis. Walter de Gruyter GmbH & Co KG. (2018)
4. Kwon, H. C., Shim, K. H.: An improved method of character network analysis for literary criticism: a case study of< hamlet>. In: International Journal of Contents, 13(3), 43-48. (2017)
5. Kwon, H., Kwon, H. T., Yoon, W. C.: An information-theoretic evaluation of narrative complexity for interactive writing support. In: Expert Systems with Applications, 53, 219-230. (2016).
6. Prado, S. D., Dahmen, S. R., Bazzan, A. L., Carron, P. M., Kenna, R.: Temporal network analysis of literary texts. In: Advances in Complex Systems, 19(03), 1650005. (2016)
7. Sowa, J. F.: Concept mapping. In: Talk Presented at the AERA Conference, San Francisco. (2006)
8. Sowa, J. F.: Conceptual graphs for representing conceptual structures. In: Conceptual Structures in Practice. Chapman and Hall/CRC, 112-148. (2016).
9. Rdf 1.1. primer, https://www.w3.org/TR/2014/NOTE-rdf11-primer-20140624/ (2018)
10. Buitelaar, P., Cimiano, P., Magnini, B.: Ontology learning from text: An overview. In: Ontology learning from text: Methods, evaluation and applications, 123, 3-12. (2005)
11. Falomir, Z., Olteţeanu, A. M.: Logics based on qualitative descriptors for scene understanding. In: Neurocomputing, 161, 3-16. (2015).
12. Sikos, L. F.: Utilizing multimedia ontologies in video scene interpretation via information fusion and automated reasoning. In: Federated Conference on Computer Science and Information Systems (FedCSIS), 91-98. (2017).
13. Gruber, T. R.: A translation approach to portable ontology specifications. In: Knowledge acquisition, 5(2), 199-220. (1993).
14. Brézillon, P.: Elaboration of the Contextual Graphs representation: From a conceptual framework to an operational software. In: ISTE OpenScience, Published by ISTE Ltd. London, UK, 1-26. (2017).
15. Cabrera, O., Franch, X., Marco, J.: A context ontology for service provisioning and consumption. In: Eighth International Conference on Research Challenges in Information Science (RCIS), 1-12. (2014)

16. Alipanah, N., Parveen, P., Khan, L., Thuraisingham, B.: Ontology-driven query expansion using map/reduce framework to facilitate federated queries. In: International Conference on Web Services, 712-713. (2011).
17. Lytvyn, V., Sharonova, N., Hamon, T., Vysotska, V., Grabar, N., Kowalska-Styczen, A.: Computational linguistics and intelligent systems. In: CEUR Workshop Proceedings, Vol-2136 (2018)
18. Vysotska, V., Fernandes, V.B., Emmerich, M.: Web content support method in electronic business systems. In: CEUR Workshop Proceedings, Vol-2136, 20-41. (2018)
19. Kanishcheva, O., Vysotska, V., Chyrun, L., Gozhyj, A.: Method of Integration and Content Management of the Information Resources Network. In: Advances in Intelligent Systems and Computing, 689, Springer, 204-216. (2018)
20. Lytvyn, V., Vysotska, V., Chyrun, L., Chyrun, L.: Distance Learning Method for Modern Youth Promotion and Involvement in Independent Scientific Researches. In: First International Conference on Data Stream Mining & Processing (DSMP), 269-274. (2016)
21. Korobchinsky, M., Vysotska, V., Chyrun, L., Chyrun, L.: Peculiarities of Content Forming and Analysis in Internet Newspaper Covering Music News, In: International Conference on Computer Science and Information Technologies, CSIT, 52-57. (2017).
22. Naum, O., Chyrun, L., Kanishcheva, O., Vysotska, V.: Intellectual System Design for Content Formation. In: International Conference on Computer Science and Information Technologies, CSIT, 131-138. (2017)
23. Euzenat, J., Shvaiko P.: Ontology Matching. In: Springer, Heidelberg, Germany. (2007).
24. Maedche, A., Staab, S.: Measuring Similarity between Ontologies. In: Knowledge Engineering and Knowledge Management, 251-263. (2002).
25. Xue, X., Wang, Y., Hao, W.: Optimizing Ontology Alignments by using NSGA-II. In: The International Arab Journal of Information Technology, 12(2), 176-182. (2015).
26. Martinez-Gil, J., Alba, E., Aldana-Montes, J.F.: Optimizing ontology alignments by using genetic algorithms. In: The workshop on nature based reasoning for the semantic Web, Karlsruhe, Germany. (2008).
27. Vysotska, V., Lytvyn, V., Burov, Y., Gozhyj, A., Makara, S.: The consolidated information web-resource about pharmacy networks in city. In: CEUR Workshop Proceedings, Vol-2255, 239-255. (2018).
28. Lytvyn, V., Vysotska, V., Burov, Y., Veres, O., Rishnyak, I.: The Contextual Search Method Based on Domain Thesaurus. In: Advances in Intelligent Systems and Computing, 689, 310-319. (2018)
29. Lytvyn, V., Vysotska, V.: Designing architecture of electronic content commerce system. In: Computer Science and Information Technologies, Proc. of the X-th Int. Conf. CSIT’2015, 115-119. (2015)
30. Rashkevych, Y., Peleshko, D., Vynokurova, O., Izonin, I., Lotoshynska, N.: Single-frame image super-resolution based on singular square matrix operator. In: IEEE 1th Ukraine Conference on Electrical and Computer Engineering (UKRCON), 944-948. (2017)
31. Tkachenko, R., Tkachenko, P., Izonin, I., Tsymbal, Y.: Learning-based image scaling using neural-like structure of geometric transformation paradigm. In: Studies in Computational Intelligence, 730, Springer Verlag, 537–565. (2018)
32. Lytvyn, V., Vysotska, V., Uhryn, D., Hrendus, M., Naum, O.: Analysis of statistical methods for stable combinations determination of keywords identification. In: Eastern-European Journal of Enterprise Technologies, 2/2(92), 23-37. (2018)
33. Lytvyn, V., Vysotska, V., Pukach, P., Bobyk, I., Uhryn, D.: Development of a method for the recognition of author’s style in the Ukrainian language texts based on linguome-

try, stylemetry and glottochronology. In: Eastern-European Journal of Enterprise Technologies, 4/2, 10-18. (2017)
34. Lytvyn, V., Vysotska, V., Pukach, P., Bobyk, I., Pakholok, B.: A method for constructing recruitment rules based on the analysis of a specialist's competences. In: Eastern-European Journal of Enterprise Technologies, 6/2(84), 4-14. (2016)
35. Basyuk, T.: The main reasons of attendance falling of internet resource. In: Proc. of the X-th Int. Conf. Computer Science and Information Technologies, CSIT'2015, 91-93. (2015).
36. Lytvyn, V., Vysotska, V., Peleshchak, I., Rishnyak, I., Peleshchak, R.: Time Dependence of the Output Signal Morphology for Nonlinear Oscillator Neuron Based on Van der Pol Model. In: International Journal of Intelligent Systems and Applications, 10, 8-17 (2018)
37. Rusyn, B., Lutsyk, O., Lysak, O., Lukeniuk, A., Pohreliuk, L.: Lossless Image Compression in the Remote Sensing Applications In: Proc. of the IEEE First Int. Conf. on Data Stream Mining & Processing (DSMP), 195-198. (2016)
38. Maksymiv, O., Rak, T., Peleshko, D.: Video-based Flame Detection using LBP-based Descriptor: Influences of Classifiers Variety on Detection Efficiency. In: International Journal of Intelligent Systems and Applications, 9(2), 42-48. (2017)
39. Gozhyj, A., Chyrun, L., Kowalska-Styczen, A., Lozynska, O.: Uniform Method of Operative Content Management in Web Systems. In: CEUR Workshop Proceedings (Computational linguistics and intelligent systems, 2136, 62-77. (2018).
40. Vysotska, V., Hasko, R., Kuchkovskiy, V.: Process analysis in electronic content commerce system. In: 2015 Xth International Scientific and Technical Conference Computer Sciences and Information Technologies (CSIT), 120-123. (2015).
41. Peleshko, D., Rak, T., Izonin, I.: Image Superresolution via Divergence Matrix and Automatic Detection of Crossover. In: International Journal of Intelligent Systems and Application, 8(12), 1-8. (2016)
42. Calvaneze, D.: Optimizing ontology-based data access. KRDB Research Centre for Knowledge and Data. In: Free University of Bozen-Bolzano, Italy. (2013).
43. Gottlob, G., Orsi, G., Pieris, A.: Ontological queries: Rewriting and optimization. In: Data Engineering, 2-13. (2011).
44. Li, Y., Heflin, J.: Query optimization for ontology-based information integration. In: Information and knowledge management, 1369-1372. (2010).
45. Keet, C.M., Ławrynowicz, A., d'Amato, C., Hilario, M.: Modeling issues & choices in the data mining optimization ontology. (2013).
46. Keet, C.M., Ławrynowicz, A., d'Amato, C., Kalousis, A., Nguyen, P., Palma, R., Stevens, R., Hilario, M.: The data mining Optimization ontology. In: Web Semantics: Science, Services and Agents on the World Wide Web, 32, 43-53. (2015).
47. Basyuk, T.: Popularization of website and without anchor promotion. In: Computer science and information technologies, 193-195. (2016)
48. Basyuk, T.: Innerlinking website pages and weight of links. In: Computer science and information technologies (CSIT), 12-15. (2017)
49. Basyuk T.: The Popularization Problem of Websites and Analysis of Competitors. In: Advances in Intelligent Systems and Computing, 689. Springer, Cham, 54-65. (2018)
50. Davydov, M., Lozynska, O.: Information System for Translation into Ukrainian Sign Language on Mobile Devices. In: Computer Science and Information Technologies, Proc. of the Int. Conf. CSIT, 48-51. (2017).
51. Davydov, M., Lozynska, O.: Linguistic Models of Assistive Computer Technologies for Cognition and Communication. In: Computer Science and Information Technologies, Proc. of the Int. Conf. CSIT, 171-175. (2017)