Modeling and Evaluation of the Mathematical Educational Ontology

Liliana Shakirova[1] [0000-0001-5758-4076], Marina Falileeva[1] [0000-0003-2228-7551], Alexander Kirillovich[1] [0000-0001-9680-449X], Evgeny Lipachev[1] [0000-0001-7789-2332], Olga Nevzorova[1,2] [0000-0001-8116-9446] and Vladimir Nevzorov[3] [0000-0002-1887-5791]

[1]Kazan Federal University
[2]Institute of Applied Semiotics of Tatarstan Academy of Sciences, Kazan, Russia
[3]Kazan National Research Technical University, Kazan, Russia

liliana008@mail.ru, mmwwff@yandex.ru, {alik.kirillovich, elipachev, onevzoro, nevzorovvn}@gmail.com

Abstract. In this paper, we discuss the current stage of development of the educational mathematical ontology OntoMath^{Edu}, firstly presented by us at INTED 2019 and CICM 2019. This ontology is intended to be used as a Linked Open Data hub for mathematical education, a linguistic resource for intelligent mathematical language processing and an end-user reference educational database. The ontology is organized in three layers: a foundational ontology layer, a domain ontology layer and a linguistic layer. The domain ontology layer contains language-independent concepts, covering secondary school mathematics curriculum. The linguistic layer provides linguistic grounding for these concepts, and the foundation ontology layer provides them with meta-ontological annotations. Our current work is dedicated to development of prerequisite relationships of the OntoMath^{Edu} ontology. We introduce these relationships by manual arrangement of the concepts of OntoMath^{Edu} by educational levels. After that, we conduct preliminary evaluation of the ontology. The ontology will be used as a foundation of the new digital educational platform of Kazan Federal University.

Keywords: Prerequisite, Ontology, Mathematical education, OntoMath^{Edu}

1 Introduction

Organization of knowledge for educational purposes requires complementing logical relations between concepts with prerequisite ones. The concept *A* is called a prerequisite for the concept *B*, if a learner must study the concept *A* before approaching the concept *B*. Prerequisite relationships are used in such tasks as automatic reading list generation [1], curriculum planning [2, 3], evaluation of educational resources [4] and prediction of academic performance [5].

While manual annotation of prerequisite relationships by expert is a time-consuming, it is still the most effective approach and can complement automatic approaches for extraction of these relationships from collections of technical documents

[6], MOOC courses [7], dependencies among university courses [8], learning paths of students [9], Wikipedia [10, 11] and Linked Open Data [12].

This work is dedicated to development of prerequisite relationships of the educational mathematical ontology OntoMath^{Edu} [13]. These relationships are introduced by manual arrangement of the concepts by educational levels.

The main contributions of this paper are two-fold: (i) developing prerequisite relationships of the OntoMath^{Edu} ontology; (ii) preliminary evaluation of this ontology.

The rest of the paper is organized as follows: In Section 2 we describe the OntoMath^{Edu} ontology. In Section 3 we introduce educational levels of OntoMath^{Edu}. And in Section 4 we conduct a preliminary evaluation of the ontology.

2 OntoMath^{Edu} description

In this section, we describe OntoMath^{Edu}, a new educational mathematical ontology [13]. This ontology is intended to be used as a Linked Open Data hub for mathematical education, a linguistic resource for intelligent mathematical language processing and an end-user reference educational database.

OntoMath^{Edu} is organized in three layers: a foundational ontology layer, a domain ontology layer and a linguistic layer.

The domain ontology layer contains language-independent math concepts from the secondary school mathematics curriculum. The description of concept contains its name in English, Russian and Tatar, axioms, and relations with other concepts. Additionally, the concepts have been semi-automatically interlinked with DBpedia [14] on the basis of the approach proposed in [15].

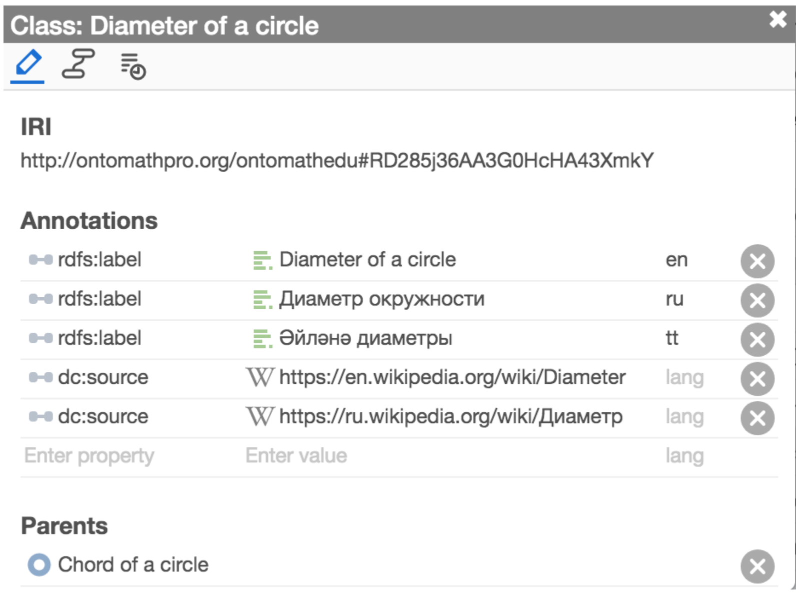

Fig. 1. *Diameter of a circle* concept in the WebProtégé editor

Fig 1 represents an example of the *Diameter of a circle* concept in the WebProtégé editor.

The concepts are organized in two main hierarchies: the hierarchy of objects and the hierarchy of reified relationships (also there are three temporary hierarchies that will be dissolved). Fig. 2 represents the top-level hierarchies and the top-level concepts of the hierarchy of objects.

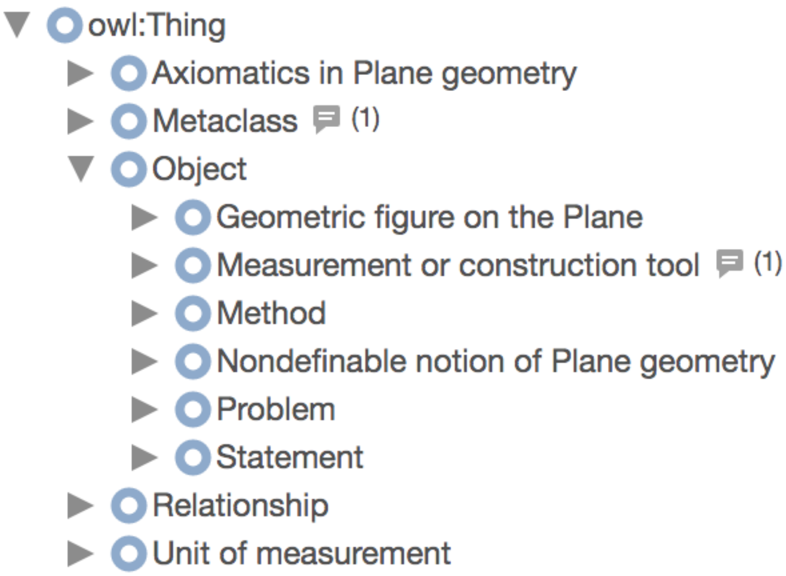

Fig. 2. The hierarchies of concepts

Fig 3 represents a fragment of the hierarchy of objects, containing the *Diagonal of a trapezoid* concept and its parents. There are four paths from this concept to the top concept *Object*, including the following: *Diagonal of a trapezoid* → *Diagonal of a quadrilateral* → *Diagonal of a polygon* → *Line segment of a polygon* → *Line segment* → *Curve* → *Geometric figure on the Plane* → *Object*.

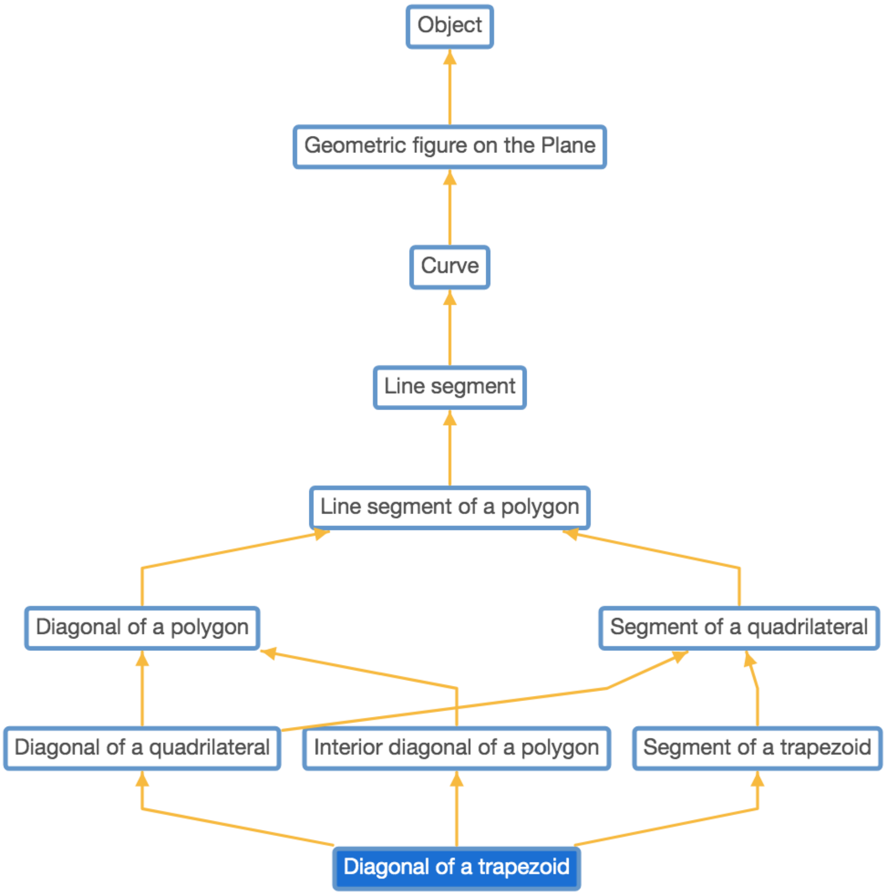

Fig. 3. The *Diagonal of a trapezoid* concept in the hierarchy of objects

There are two meta-ontological types of the concepts: kinds and roles.

A kind is a concept that is rigid and ontologically independent [16]. So, for example, the *Triangle* concept is a kind, because any triangle is always a triangle, regardless of its relationship with other figures. Fig. 4 represents an example of a kind concept (namely, the *Triangle* concept).

A role is a concept that is anti-rigid and ontologically dependent [16]. An object can be an instance of a role class only by virtue of its relationship with another object. So, for example, the *Side of a triangle* concepts is a role, since a line segment is a side of a triangle not by itself, but only in relation to a certain triangle. Fig 5 represents an example of one of the role concepts, namely the *Side of a triangle* concept. Each instance of this concept is related to an instance of the *Triangle* kind concept by the relation of ontological dependence.

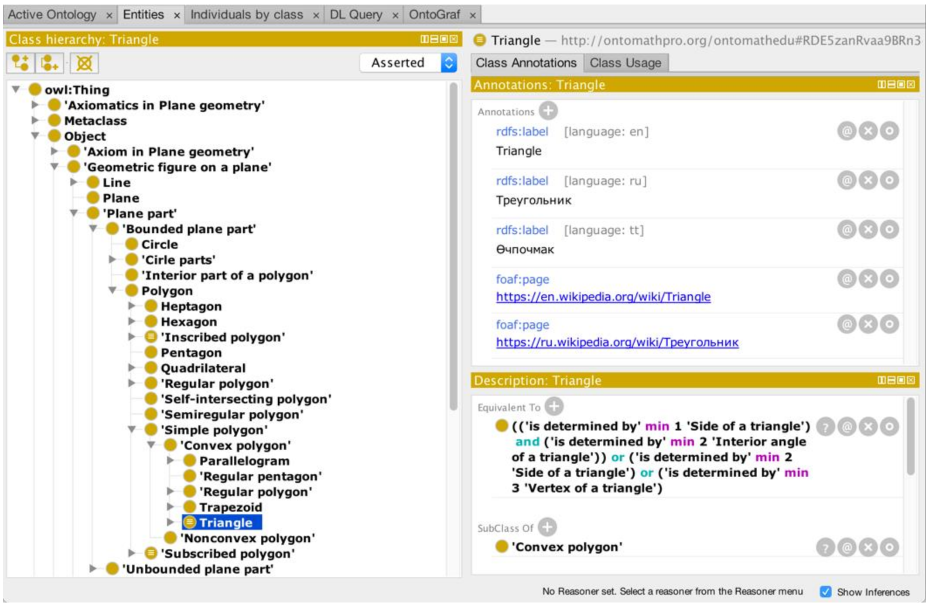

Fig. 4. The *Triangle* kind concept in the Protégé editor

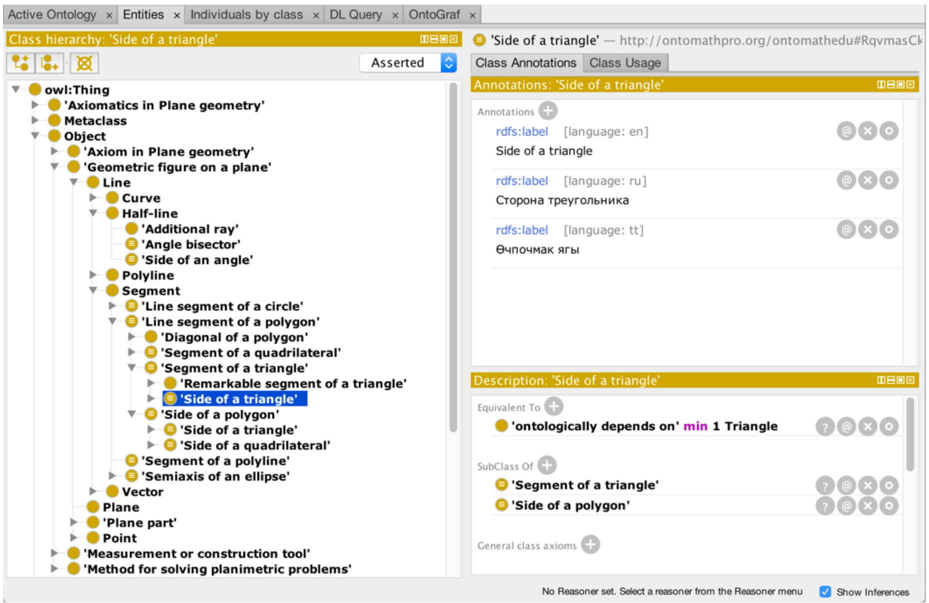

Fig.5. The *Side of a triangle* role concept in the Protégé editor

Properties of concepts are defined by the axioms, expressed by the formalism of description logics. For example, the description of the *Triangle* concept at Fig 4 contains axioms, stating that any instance of this concept determined by 1 side and 2 angles, or by 2 sides and 3 points.

Relations between concepts are represented in the ontology in a reified form, i.e. as ontological concepts, not as ontological properties. Thus, the relationships between concepts are first-order entities, and can be a subject of a statement. All instances of a relationship are linked to its participants by object properties.

Fig 6 represent an example of a reified relationship concept, namely, *Mutual arrangement between a circumscribed triangle and an inscribed circle*. Each instance of this concept is linked to its participants, namely to an instance of the *Circumscribed triangle* role concept and an instance of the *Inscribed circle* role concept.

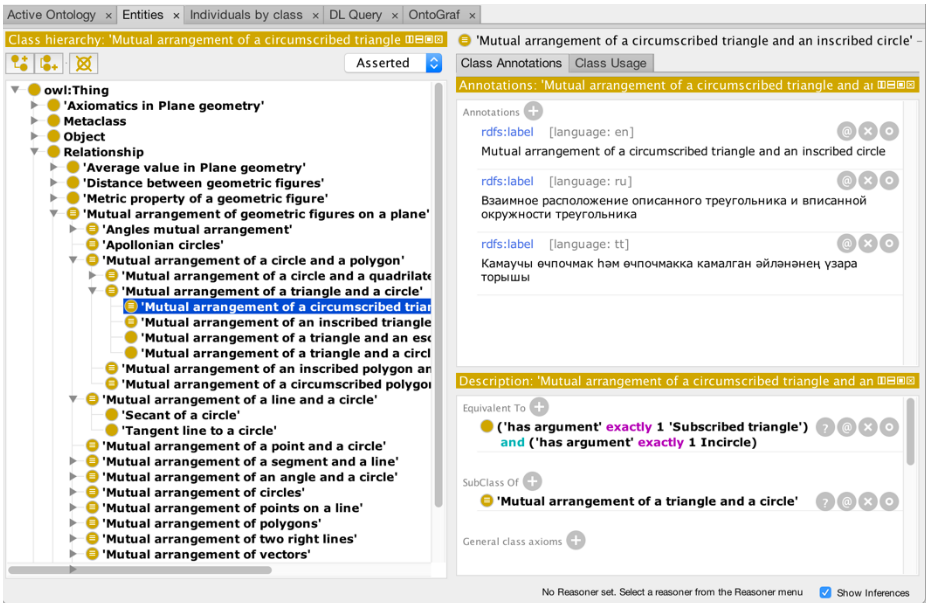

Fig.6. An example of a reified relationship concept in the Protégé editor

The linguistic layer contains multilingual lexicons under development, providing linguistic grounding for the concepts from the domain ontology layer.

A lexicon consists in (a) lexical entries, denoting mathematical concepts; (b) forms of lexical entries; (c) syntactic trees of multi-word lexical entries, (d) and syntactic frames. A syntactic frame contains a subcategorization model for a particular lexical entry and its mapping to parameters of a corresponding math concept Fig 7 represents an example of the "Riemann integral of *f* over *x* from *a* to *b*" lexical entry, where the "from *a*" dependent constituent expresses the lower limit of integration, "to *b*" express the upper limit, and "of *f*" express the integrated function.

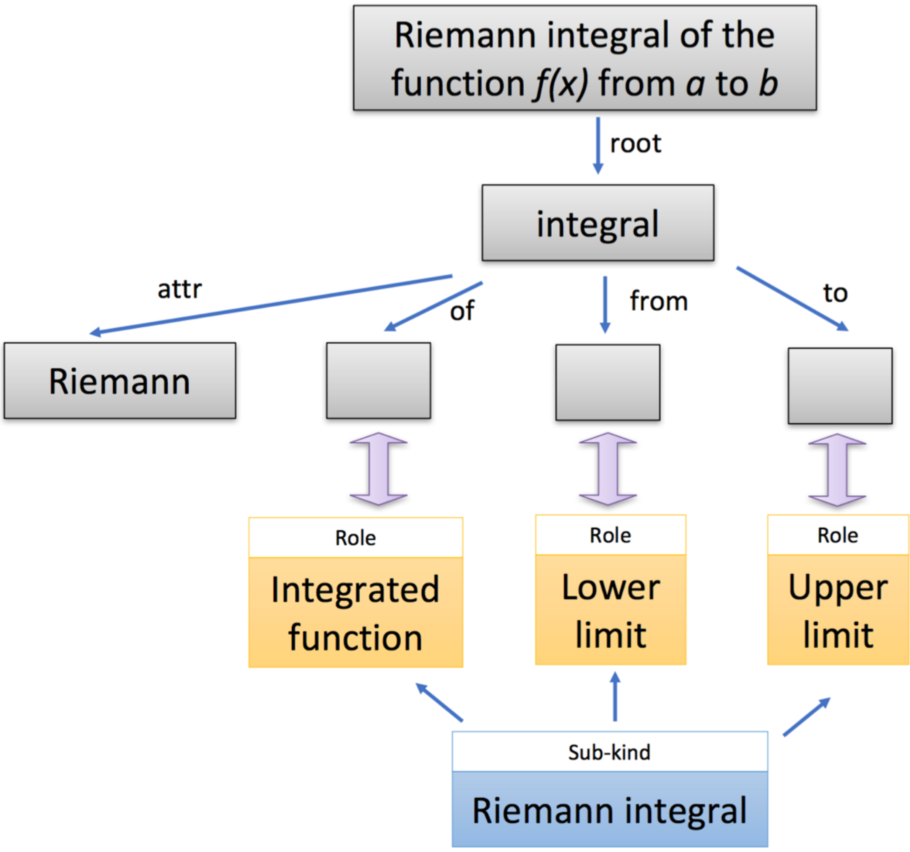

Fig.7. Syntactic frame for the "Riemann integral of *f* over *x* from *a* to *b*" lexical entry

The lexicons are expressed in terms of Lemon [17], LexInfo, OLiA [18] and PREMON [19] ontologies. According to the project, the lexicons will be interlinked with the external lexical resources from the Linguistic Linked Open Data (LLOD) cloud [20], first of all in English [21, 22], Russian [23] and Tatar [24].

The foundation ontology layer provides the concepts with meta-ontological annotations, defined by the foundation ontology UFO [16].

3 Educational Levels

In addition to universal statements about mathematical concepts, the ontology contains the statements that are linked to special concepts named viewpoints. The following types of points of view are currently being developed:

- Definitions. From different points of view, the same concept can be defined differently. These different definitions can determine some concepts through different systems of other concepts.
- Educational levels. To implement the principle of consistency and continuity in teaching concepts in the field of geometry, we introduced the notion of educational level and applied this to the presentation of ontology concepts.

Let us consider consistency in the study of the *Triangle* topic. This topic is studied in grades 7–9, including grades with advanced math program.

Table 1 presents the first level of studying definitions of the *Triangle* concept in a grade 7 (this is basic level). This level includes four stages of studying this topic in grade 7. At the second level (in a grade 8), the *Triangle* concept is expanded by the two new concepts (*Inscribed triangle* and *Subscribed triangle*). At the third level (in advanced course), other types of triangles defined in the ontology are also considered.

Table 1. Educational levels for the *Triangle* topic in the OntoMath^{Edu} ontology

Educational levels			**Stages of studying**	**Concepts**
3rd	2nd	1st		
+	+	+	1	Triangle
+	+	+	2	Acute triangle, Obtuse triangle,
+	+	+	3	Isosceles triangle, Equilateral triangle
+	+	+	4	Right triangle
+	+	+	1	Inscribed triangle, Subscribed triangle
+	+		1	Medial triangle
+	+		2	Orthogonal triangle, Triangle with vertices at Euler points

This means the possibility of a parallel study of these pairs of concepts that can be arranged in any sequence and it will be better to study these concepts simultaneously by comparing their properties. The second level includes concepts studied in grades 7-8. The third level includes concepts studied in grades 8–9 and in grades with advanced math program and also the concepts of previous levels. To take into account the methodological features of teaching mathematics, it is necessary to determine object properties in the OntoMath^{Edu} ontology, which we shall conditionally name didactic relations.

In the current version of the OntoMath^{Edu} ontology the following didactic relations are defined:

1. The *Studied simultaneously* relation connects the concepts that should be studied together, for example, the *Line* and *Ray* concepts;
2. The *Studied later* relation (the inverse relation of the *Studied earlier*). For example, the *Isosceles triangle* concept is studied later than the *Acute triangle* concept. The *Studied later* relation as well as its inverse relation, are transitive, therefore we can build the sequences of the *Studied later* relations, which form a certain sequence of concepts in learning;
3. The *Concept-level* relation determines the relevance of the concept to the educational level, for example, the concept *Triangle* is connected by the *Concept-level* relation with a stage 1 of the first educational level (see Table 1). The *Concept-level* relation is used as a criterion for building a learning sequence of concepts.

4 Analysis of the OntoMath^{Edu} ontology

In this section, we report the results of a preliminary evaluation of the OntoMath^{Edu} ontology.

The structural properties of this ontology were analyzed using the analytical software tools of the OntoIntegrator system [25]. The OntoIntegrator system is a development tool focused on the tasks of automatic text processing using various ontological models. The main functional capabilities of this system are:

- designing ontological models of arbitrary structure with wide data visualization capabilities;
- development of scientific applications related to text processing;
- natural language processing based on ontological and linguistic models.

The analytical tools of the OntoIntegrator system allow us to explore various structural properties of ontologies. When using these tools for the analysis of the OntoMath^{Edu} ontology, quantitative and qualitative results were obtained that made it possible to identify some structural features, as well as to identify specific steps for improving the ontology.

In total, 776 concepts, 5 hierarchies, 2338 text inputs of concepts, 836 class-subclass relations were defined in the OntoMath^{Edu} ontology.

The Fig. 8 represents a diagram of the distribution of objects by subclasses in the *Object* hierarchy, here 1 is the *Assertion* subclass, 2 is the *Geometric figure on a plane* subclass, 3 is the *Task* subclass, 4 is the *Tool for measuring or drawing geometry shapes* subclass, 5 is the *Method* subclass, 6 is the *Undetectable concepts of plane geometry* subclass.

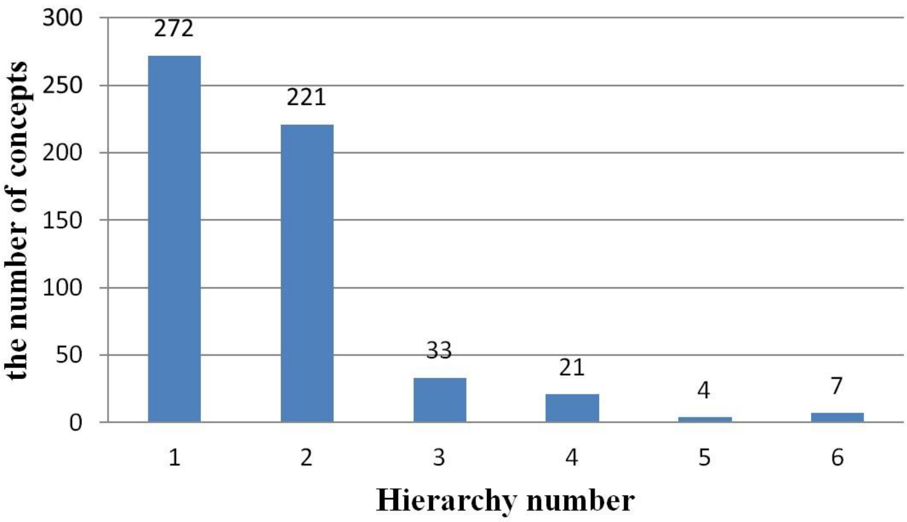

Fig. 8. The diagram of the distribution of objects by subclasses in the *Object* hierarchy

As already noted, the OntoMath^{Edu} ontology was built manually based on school textbooks. The general names were used to denote the names of important concepts (problems, theorems, methods, etc.). Below the results of linguistic analysis of the names of ontological concepts were carried out. Fig. 9 shows the frequency distribution of concept names by the number of words in their names.

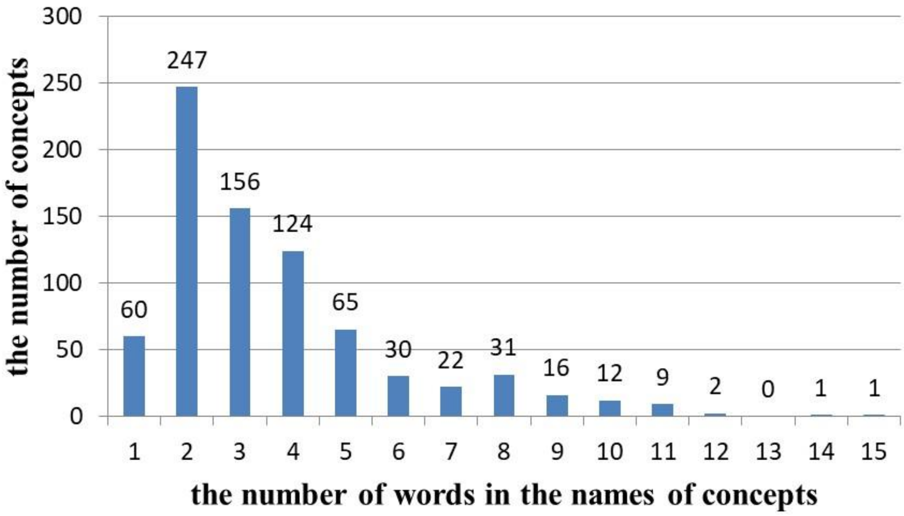

Fig. 9. The frequency distribution of concept names by the number of words in their names

The most frequent classes are two- and three-words concept names which are related to the main objects of the subject area. More longer names (more than 5 words) actually refer to the formulations of standard problems and theorems of plane geometry. Thus, a feature of the OntoMath^{Edu} ontology is not only the systematization of elementary geometry objects, but also the systematization of typical problems, theorems, and drawing methods, which is important for application in the education.

Examples of concept names are given in the Table 2.

Table 1. Examples of concept names

Length of name (in words)	Concept name (English translation)
1	Astrolabe; Vector; Hyperbola; Hypotenuse; Homothetic transformation
2	Axiom of congruence; Vertex of a square
3	Semimajor axis of an ellipse; Interior part of an angle
4	Interior part of an angle
	Axiom of a zero-vector postponement
5	Tangent line to a circle
	Mutual arrangement of points on a line
6	Cutting square into unequal squares
	Tangent segment from a point to a circle
7	Theorem about product of segments of intersecting chords
	Axiom of uniqueness of a vector postponement from given point
8	Rule of finding the coordinates of the product of a vector by a number;
	Axiom about scalar product of a vector into an equal vector
9	Inversion of angles between straight lines and circles property
10	Axiom of distributivity of multiplying vector by a real number related vector addition;
	Problem of costructing a triangle given three sides
11	Axiom of distributivity of multiplying vector by a real number related numbers addition
	Theorem about area of a parallelogram with given two sides and angle between them
12	Rule allow to find the coordinates of sum, difference and product by a number using coordinates of vectors
	Theorem about area of a parallelogram with given side and the altitude drawn to this side
14	Problem of constructing a triangle given two sides and the included angle
15	Problem of constructing a triangle given two angles and the included side

When developing an ontology for education, it would be useful to have data about the significance of concepts in the training course. Data on the frequency of concepts in the textbooks by Sharygin [26] and Atanasyan [27], the relationships of high-frequency concepts (the contextual environment of high-frequency concepts) contributes to the identification of the most important concepts of academic discipline. Subsequent ranking concepts in terms of their significance may be useful for testing. High-frequency concepts (with frequency of occurrence) for two school geometry textbooks are given in the Table 3 and the Table 4, and low-frequency concepts are given in the Table 5 and the Table 6.

Table 2. High-frequency concepts in the textbook by Sharygin

Name	Count
Point	1595
Line	846
Triangle	793
Circle	765
Angle	632
Line segment	304

Table 3. High-frequency concepts in the textbook by Atanasyan

Name	Count
Point	1652
Line	1061
Angle	858
Triangle	848
Line segment	588
Circle	511

Table 4. Low-frequency concepts in the textbook by Sharygin

Name	Count
Ellipse	1
Centimetre	1
Trigonometric equality	2
Plane geometry theorem	2
Property of a triangle	2
Adjacent angles	2

Table 5. Low-frequency concepts in the textbook by Atanasyan

Name	Count
Heptagon	1
Polyline	1
Miter square	2
Roulette	2
Object	2
Perimeter of a rectangle	2

A general assessment of the frequency distribution of ontology concepts is given in the Table 7.

Table 6. Examples of concept names

Frequency of using (interval)	Number of concepts in in the textbook by Sharygin	Number of concepts in in the textbook by Atanasyan
1000–1620	1	2
500–999	4	4
100–499	10	13
50–99	25	31
10–49	91	111
5–9	33	42
1–4	70	61

The linguistic-statistical analysis of ontology concepts showed that the OntoMath^{Edu} ontology not only contains a systematization of the main objects of the subject area, but also includes a taxonomy of the main typical problems studied in the school geometry course. The latter circumstance makes this resource especially useful for use in education. Frequency analysis of educational texts allowed to identify the most important concepts of ontology, which can subsequently be used in ranking ontological concepts in the process of studying geometry.

5 Conclusion

In this paper, we describe educational levels of the OntoMath^{Edu} ontology, and conduct its preliminary evaluation.

The ontology will be used as a foundation of a new digital educational platform under development at Kazan Federal University

This work was funded by RFBR, projects #19-29-14084, and by the Government Program of Competitive Development of Kazan Federal University.

References

1. Gordon, J., Aguilar, S., Sheng, E., and Burns, G.: Structured Generation of Technical Reading Lists. In: Tetreault J., et al. (eds.) Proceedings of the 12th Workshop on Innovative Use of NLP for Building Educational Applications (BEA 2017), pp. 261–270. ACL (2017).
2. Agrawal, R., Golshan, B., and Papalexakis, E.: Data-Driven Synthesis of Study Plans: Technical Report TR-2015-003. Data Insights Laboratories (2015). https://web.archive.org/web/20160207113043/http://www.datainsightslaboratories.com/wp-content/uploads/2015/03/TR-2015-003.pdf.
3. Auvinen, T., Paavola, J., and Hartikainen, J.: STOPS: a graph-based study planning and curriculum development tool. In: Proceedings of the 14th Koli Calling International Conference on Computing Education Research (Koli Calling '14), pp. 25–34. ACM (2014). https://doi.org/10.1145/2674683.2674689.
4. Rouly, J.M., Rangwala, H., and Johri, A.: What Are We Teaching?: Automated Evaluation of CS Curricula Content Using Topic Modeling. In: Dorn B., et al. (eds.) Proceedings of the 11th annual International Conference on International Computing Education Research (ICER '15), pp. 189–197. ACM (2015). https://doi.org/10.1145/2787622.2787723.
5. Polyzou, A. and Karypis, G.: Grade prediction with models specific to students and courses. International Journal of Data Science and Analytics **2**(3–4), 159–171 (2016). https://doi.org/10.1007/s41060-016-0024-z.
6. Gordon, J., Zhu, L., and Galstyan, A., Natarajan, P., Burns, G.: Modeling Concept Dependencies in a Scientific Corpus. In: Erk K. and Smith N.A. (eds.) Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL 2016). Volume 1. Long Papers, pp. 866-875. ACL (2016).
7. Pan, L., Li, C., Li, J., and Tang, J.: Prerequisite Relation Learning for Concepts in MOOCs. In: R. Barzilay and M.-Y. Kan (eds.) Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL 2017). Volume 1: Long Papers, pp. 1447–1456. ACL (2017).
8. Liang, C., Ye, J., Wu, Z., Pursel, B., and Giles, C.L.: Recovering Concept Prerequisite Relations from University Course Dependencies. In: Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI-17), pp. 4786–4791. AAAI (2017).
9. Pang, Y., Wang, N., Zhang, Y., Jin, Y., Ji, W., and Tan, W.: Prerequisite-related MOOC recommendation on learning path locating. Computational Social Networks **6**, (2019). https://doi.org/10.1186/s40649-019-0065-2.
10. Gasparetti, F., Limongelli, C., and Sciarrone, F.: Exploiting Wikipedia for Discovering Prerequisite Relationships Among Learning Objects. In: Proceedings of the International Conference on Information Technology Based Higher Education and Training (ITHET 2015). IEEE (2015). https://doi.org/10.1109/ITHET.2015.7218038.
11. Zhou, Y. and Xiao K.: Extracting Prerequisite Relations Among Concepts in Wikipedia. In: Proceedings of the International Joint Conference on Neural Networks (IJCNN 2019). IEEE (2019). https://doi.org/10.1109/IJCNN.2019.8852275.
12. Manrique, R., Pereira, B., Marino, O., Cardozo, N., and Wolfgand S.: Towards the identification of concept prerequisites via Knowledge Graphs. In: Proceedings of the IEEE 19th International Conference on Advanced Learning Technologies (ICALT 2019), pp. 332–336. IEEE (2019). https://doi.org/10.1109/ICALT.2019.00101.
13. Kirillovich, A., Nevzorova, O., Falileeva, M., Lipachev, E., Shakirova, L.: OntoMath^{Edu}: Towards an Educational Mathematical Ontology. In: Kaliszyk, C., et al. (eds.) Workshop

Papers at 12th Conference on Intelligent Computer Mathematics (CICM-WS 2019). CEUR Workshop Proceedings (forthcoming).

14. Lehmann, J., Isele, R., Jakob, M., Jentzsch, A., Kontokostas, D., Mendes, P. N., Hellmann, S., Morsey, M., van Kleef, P., Auer, S., and Bizer, C.: DBpedia: A Large-scale, Multilingual Knowledge Base Extracted from Wikipedia. Semantic Web Journal **6**(2), 167–195 (2015). https://doi.org/10.3233/SW-140134.
15. Kirillovich, A. and Nevzorova, O.: Ontological Analysis of the Wikipedia Category System. In: Aveiro, D., et al. (eds.) Proceedings of the 10th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2018), Seville, Spain, 18-20 September, 2018. Volume 2: KEOD, pp. 358–366. SCITEPRESS (2018). https://doi.org/10.5220/0006961803580366.
16. Guizzardi, G., Ontological Foundations for Structural Conceptual Models. CTIT (2005).
17. McCrae, J.P., Bosque-Gil, J., Gracia, J., Buitelaar, P., and Cimiano, P.: The OntoLex-Lemon Model: Development and Applications. In: Kosem I., et al. (eds.) Proceedings of the 5th biennial conference on Electronic Lexicography (eLex 2017), pp. 587–597. Lexical Computing CZ (2017).
18. Chiarcos, C.: OLiA – Ontologies of Linguistic Annotation. Semantic Web **6**(4), 379–386 (2015). https://doi.org/10.3233/SW-140167.
19. Rospocher, M., Corcoglioniti, F., and Palmero Aprosio, A.: PreMOn: LODifing linguistic predicate models. Language Resources and Evaluation **53**, 499–524 (2019). https://doi.org/10.1007/s10579-018-9437-8.
20. Cimiano, P., Chiarcos, C., McCrae, J.P., and Gracia, J.: Linguistic Linked Open Data Cloud. In: Cimiano, P., et al. (eds.) Linguistic Linked Data: Representation, Generation and Applications, pp. 29–41. Springer (2020). https://doi.org/10.1007/978-3-030-30225-2_3.
21. McCrae, J. P., Fellbaum, C., and Cimiano, P.: Publishing and Linking WordNet using lemon and RDF. In: Chiarcos C. et al. (eds.) Proceedings of the 3rd Workshop on Linked Data in Linguistics (LDL-2014), pp. 13–16. ELRA (2014).
22. Ehrmann, M., Cecconi, F., Vannella, D., McCrae, J., Cimiano, P., and Navigli, R.: Representing Multilingual Data as Linked Data: the Case of BabelNet 2.0. In: Calzolari N., et al. (eds.) Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC 2014), pp. 401–408. ELRA (2014).
23. Kirillovich, A., Nevzorova, O., Gimadiev. E., and Loukachevitch, N.: RuThes Cloud: Towards a Multilevel Linguistic Linked Open Data Resource for Russian. In: Różewski, P. and Lange, C. (eds.) Proceedings of the 8th International Conference on Knowledge Engineering and Semantic Web (KESW 2017). Communications in Computer and Information Science, vol. 786, pp. 38–52. Springer (2017). https://doi.org/10.1007/978-3-319-69548-8_4.
24. Galieva, A., Kirillovich, A., Khakimov, B., Loukachevitch, N., Nevzorova, O., and Suleymanov, D.: Toward Domain-Specific Russian-Tatar Thesaurus Construction. In: Proceedings of the International Conference IMS-2017, pp. 120–124. ACM (2017). https://doi.org/10.1145/3143699.3143716.
25. Nevzorova, O. and Nevzorov, V.: Ontology-Driven Processing of Unstructured Text. In: Kuznetsov S. and Panov A. (eds.) Proceedings of the 17th Russian Conference on Artificial Intelligence (RCAI 2019). Communications in Computer and Information Science, vol. 1093, pp. 129–142. Springer (2019). https://doi.org/10.1007/978-3-030-30763-9_11.
26. Sharygin, I.F.: Geometry, 7-9th Grades. Drofa (2018) (in Russian).
27. Atanasyan, L., Butuzov, V., and Kadomcev S.: Geometry, 7-9th Grades: Textbook for General-Education Schools. Prosveshenie (2018) (in Russian).