Development of a Complex Ontology of Optics

Dmitry Mouromtsev

National Research University of Information Technologies, Mechanics and Optics, Sankt-Petersburg, Russia

mouromtsev@mail.ifmo.ru

Abstract. The paper describes the construction of domain ontology of optics for educational purposes and its possible application in scientific museums.. A short description of the domain knowledge structure is given. Patterns used to resolve the knowledge engineering problem in the optics make one doubt about the sufficiency of OWL expressivity for correct representation of specific relationships pronounced in the semantics of the said domain.

Keywords: Ontology engineering, domain ontology design, semantic science

1 Introduction

The task of the optics ontology design was a part of a project concerning the development of intelligent educational system for the optics domain. The application of this system is planned in the educational center for optical technologies' of National Research University of Information Technologies, Mechanics and Optics. A part of this center is "Optimus"[1] museum that has a lot of artifacts from antiquity to the present time. The mission of the center is familiarization of the visitors with the world of science, as well as providing pupils and students an environment to study the history of the optics, and some of its subjects and topics on real artifacts. At the same time, the museum possesses not only the showpieces but also interactive facilities including special equipment to run experiments and games.

Optimus maintains relations with similar museums in Europe that also have ancient manuscripts, optical devices, instruments and objets d'art that present the world of knowledge about the optics and science in general. Genetically, this cultural heritage includes individual facts of history of science, results of experiments and inventions, demonstration of optical phenomena and so forth. To have a clever eye for the collections, one may require a description of many other concepts such as scientific theories, models, physical processes and the like. In this context, ontology of optics can provide added value in the sense of education and research opportunities. On one hand, it can combine references to items from various museums and centers, and on

[1] http://www.optimus.edu.ru/en

the other, using it, one can create new exhibitions focusing on particular issues of the optics.

In addition, a domain ontology is a cognitive tool to create an educational technique of a new type, being a sort of an intelligent guide to the world of knowledge, based on semantic search and reasoning.

Unlike textbooks and other teaching materials, ontology provides not only a structuring of contents but explicitly shows interrelationships of domain subjects, their mutual influence on each other and evolution through time. Furthermore, the ontology can represent objects of the real world exhibited in the science museums or used in educational laboratories. From this point of view the use the ontology supposed two cases, as an interactive tool for navigation in a knowledge portal and a base for an intelligent guide in the museum of science.

Representation of scientific knowledge in the form of OWL ontology is developed in the approach pursued by the semantic science, defined as application of semantic technology and reasoning to the practice of science [1]. This paper describes the initial step of the development of the optics ontology and discusses some problems of knowledge engineering and representation in this domain.

2 Motivation

Until now, speaking of educational web-based applications, they have in mind primarily the different systems of distance education. In recent times this term was also referred to e-learning 2.0 [2], [3] implying the use of learning resources based on Web 2.0 [4]. Unlike the traditional approach where the student is invited to study a kit of materials, perform the test tasks are then be checked by a teacher, a new form of e-learning 2.0 becomes more social and includes web 2.0 software, such as blogs, wikis, podcasts and virtual worlds (like Second Life). The last phenomenon has also been referred to as "Long Tail Learning" [5], because the process of teaching involves not only teachers, but also a large number of students, exchanging experience in dealing with specific problems and thereby educating each other.

At first glance, such technology platform is able to satisfy the requirements of a virtual space for the museum of science. However, solely the use of Web 2.0 technology does not provide a solution to the aforementioned problem of «intelligent guidance» to the user in searching and handling the information. Exactly for this reason the project was focused on the development of the domain ontology.

Implementation of the active learning technology on the basis of ontological modeling was already discussed by the author [6]. The idea to use ontologies for e-learning tasks is well known and there is a lot of literature on this issue. One the most developed mindmaps describing ontologies for education is suggested in [7]. From the application perspective, ontologies are used as a cognitive tool for such tasks as knowledge construction, knowledge externalization, knowledge communication and knowledge assessment.

The advantage of using ontologies as a tool of education is a systematic approach to the study of a subject area based on the logic [8]. As it stated in [9], ontologies formalize the meanings of terms used in a domain, and provide clear human-readable definitions that disambiguate their usage, along with logical axioms that allow automated reasoning. They enable consistency checking, classification, and query answering over knowledge of a particular domain, enabling intelligent computer applications to be built which support the work of scientists or teachers within the domain of interest and across interrelated neighboring domains. From this point of view, ontology [10] is a formal specification of the meaning of the vocabulary used in an information system. Ontologies are needed so that information sources can interoperate at a semantic level. [1]

There are many definition of "ontology", but in terms of software applications it is essential that the ontology is used as a cognitive tool for such tasks as knowledge formalization, knowledge externalization, knowledge transfer and knowledge assessment [11]. Ontology engineering can provide an effective model of a domain for the learning process that helps both teachers and students to acquire knowledge, use it and control the learning.

Finally, there is a recent trend to development of so-called Linked Learning systems using a distributed learning content shared in Linked Open Data cloud [12].

3 Background

3.1 Domain area description

One of the first tasks of this work was a design of the optics ontology prototype in a historical perspective. Then the historical ontological model must be integrated into the scientific ontology of the optics domain, including the following areas: theories and descriptions (noumena), the laws of mathematical models and experiments (experiments), experiments (phenomena) and the personalities, the parameters (characteristics), tools (instruments), the basic elements.

Area of theories and descriptions (noumena) includes a conceptual approach to the division of the nature of light (waves, quanta, rays), and historical aspects of the appearance and development of these ideas (chronological tree or constituting the time).

The area of the laws and the mathematical models consists of the principles of geometrical optics (Snell's law, Fermat's principle), and mathematical models used in physical optics (Stokes vector, the Jones matrix, the Brewster angle, etc.)

In the area of experiments (tests) there are examples of fundamental optical experiments and schemes of their observation: the Young's experience, Newton's rings, observation of interference by division of wavefront and others.

In the area of phenomena and personalities identified are the basic optical phenomena: absorption, refraction, diffraction, etc., as well as their variants (classes and their specific content). These phenomena are associated with the scientists discovered them. These data form the GIS component of the ontology.

The area of parameters (characteristics) is a set of measurable quantities describing the quality (properties) of various optical phenomena: the refractive index, absorption coefficient, the degree of polarization, and so forth..

The instrument's area includes a taxonomy of optical devices including partially disclosed classes of interferometers and polarization prisms and some others.

The area of the basic elements consists of optical mediums, devices and components (lenses, prisms, mirrors), partially disclosed classes of interference devices (split lens, biprism, bimirrors).

Finally, one can also highlight uncertain area uniting classes: "Personalities", "Historical hypotheses", as well as the two coordinate components — the time ("Temporal component") and GIS ("Spatial component"). Together they form the basis of historical and scientific ontological model.

3.2 Knowledge engineering

Visual modeling techniques play a very important role in the process of knowledge engineering. In the case of the knowledge base on optics the most attractive model is a concept map [13]. In contrast to the Buzan's mind maps [14], concept maps do not require a single center and allow defining various types of relationships between individuals. A fragment of the developed concept map is shown at Fig. 1.

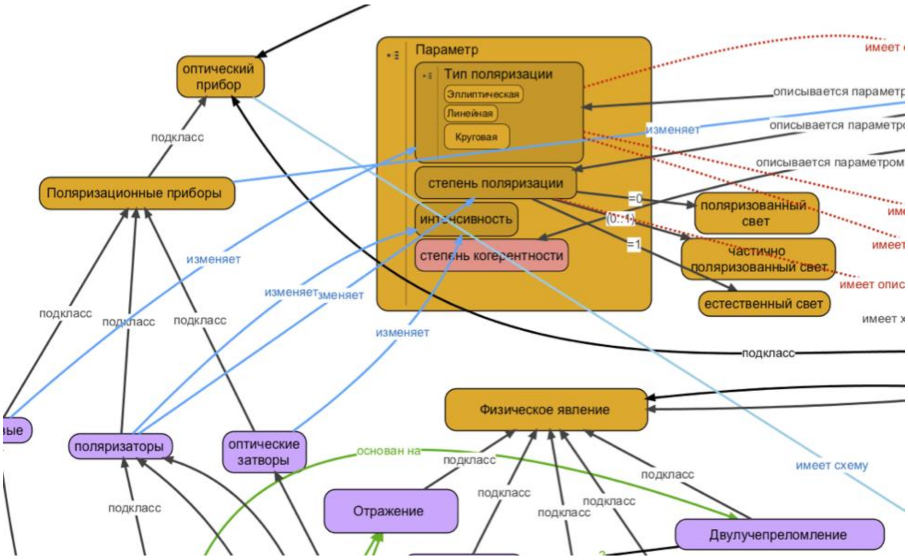

Fig. 1. A fragment of the optics domain concept-map

The main problem of the optics domain's knowledge engineering is the fact that optics is one of the oldest sciences known from ancient times. Over the centuries, many concepts have changed their scope, sometimes to the opposite. But others, on the contrary, remain unchanged for ages. For this reason, the ontological model must represent three sections of knowledge: historical, practical and theoretical. The historical section contains properties such as "explain," "carry out a test", "discover", "prove", "introduce concept". This section allows to look at a phenomenon in its historical context — to find out what scientists have studied it, what tests they carried. The practical section contains properties such as "based on", "uses". This section allows, for example, describing the application of the phenomenon — view devices, units or practical technology based on it. The theoretical section includes such properties as "a condition", "is a parameter of", "is part of". This section allows seeing the concepts explaining the considered notion.

The current structure of the domain ontology includes a set of optical ideas and personalities of ancient scholars. It gives an idea of the ancient philosophers expressing a hypothesis about the means and nature of light and the formation of the optical branches [15]. These ideas have the greatest impact. For example, Aristotle expresses ideas about the formation of rainbow and the existence of visual rays, contributes to the development of the four branches of optics, "Catoptrics", "Ray Optics", "Meteora", "Means of view, the nature of light." Relationships between classes can be ambiguous, since one idea can refer to several personalities who, in turn, may have a few ideas and a branch of antique optics can include multiple personalities and ideas.

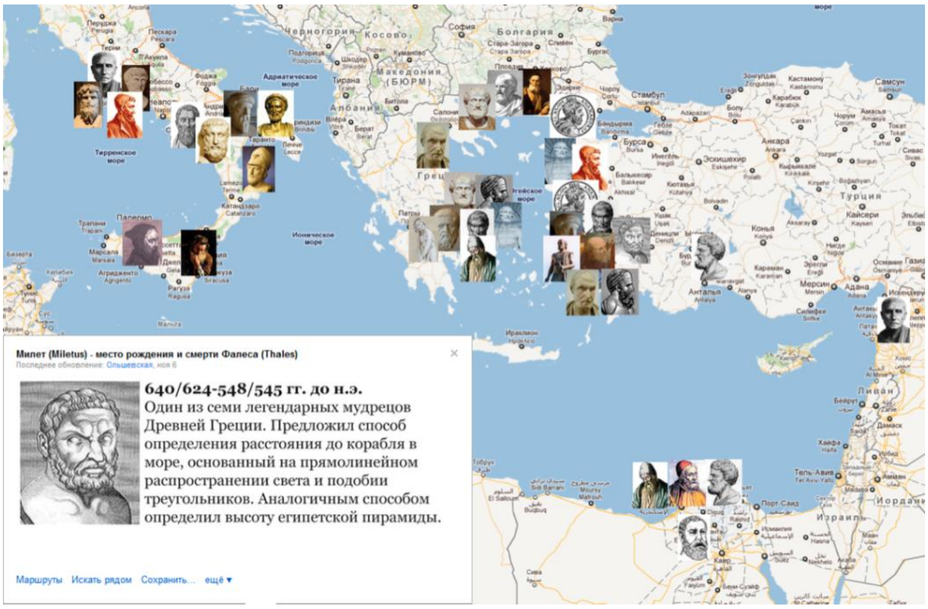

Fig. 2. Representation of “Personalities” class individuals on the map

The class "Geoinformation component" is the base class for "Region" having a subclass "City", consisting currently of the names of places of birth and death of opticians, but planned to add new categories of geographic information. Geoinformation component of the domain has been created by means of Google Map. It covers the period of a classical Greek heritage in the field of optics and its development in the

Hellenistic period until the 2nd-3rd centuries when the ancient tradition of optics was completed by mathematical works of Ptolemy and the writings of Galen on the physiological optics.

The map contains places of ancient scientists. Instead of the conventional tags provided in the tools of Google Map, there have been used portraits of each personality and a brief description was created for each, including the dates of birth and death, major achievements and so on as shown in Fig. 2.

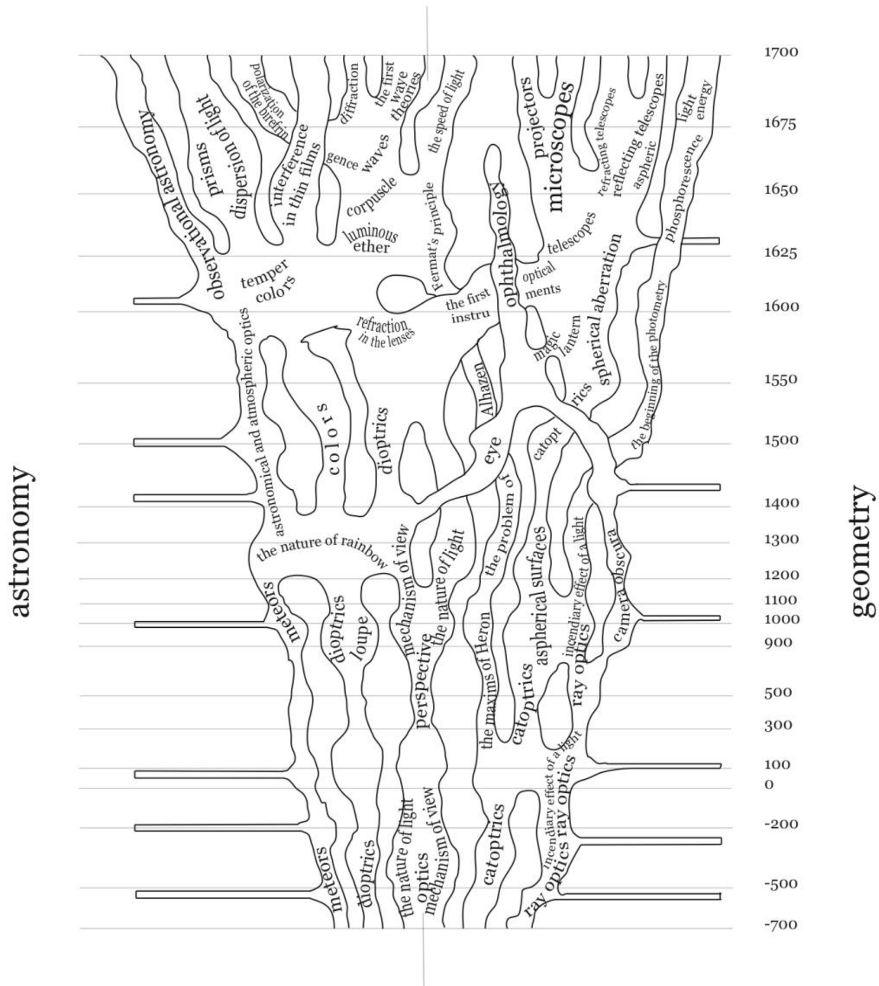

Fig. 3. Tree of optics

The class "Temporary component" is the base class for the "Age" having a subclass of "periods" including "Antiquity", "Middle Ages", "Renaissance", "the 17th Century", "The era enlightenment" [16]. In the ontology currently developed only one subclass is filled, "Antiquity." It consists of five branches of classical optics formed later the science "optics": "Catoptrics", "Ray Optics", "Meteora", "Dioptrics", "Means of view, the nature of light."

Temporary component of the domain ontology illustrated in Fig. 3 is represented as a styled tree of optics, which contains not only a time component and an extended set of ideas but also demonstrates the contribution of each scientist to particular branches of optics.

This specific graph has a connection with two domains not shown in the figure, "Geometry" and "Astronomy." Antique optical knowledge is largely determined by the methods and problems solved by these two related sciences. The tree covers the time period much beyond the antiquity and continued until the 17th century. One can see the waxing and waning "branches" as well as crossings showing the links between the various optical paths of evolution of ideas.

3.3 OWL Representation

To create an optics knowledge base as an upper level ontology there were used DOLCE Lite (a simplified version of DOLCE 2.0) [17]. This upper level ontology is suited well for creating domain ontologies of different areas, as it contains abstract classes and relationships common to all sciences such as "phenomenon", "concept", "method", "experience", "principle" etc. The use of the full version of DOLCE Lite as a basis for ontology optics seemed inappropriate and redundant, since a number of abstract classes would never have been used and would not have derived concepts. Therefore, unnecessary classes were deleted and then it was enriched with specific optics relations.

One of the first steps was the creation of four basic classes of ontology: "Personalities", "Historical hypothesis", "Temporary component", "Geoinformation component". Individuals for cities, historical hypotheses, personalities, ancient branches of optics were also created. Specific relationships between these concepts were defined. The structure created as a result of the optics domain decomposition is shown in Fig. 4.

Representation of the ontology in the OWL requires to address a number of conceptual issues, including:

1. When do we have to consider the concept as class, and when the individual?
2. What are the most well-defined relations between concepts and how to ensure the inheritance of relationships in case of the dynamic classification?
3. What structures in the ontology will minimize computational complexity of the reasoning?

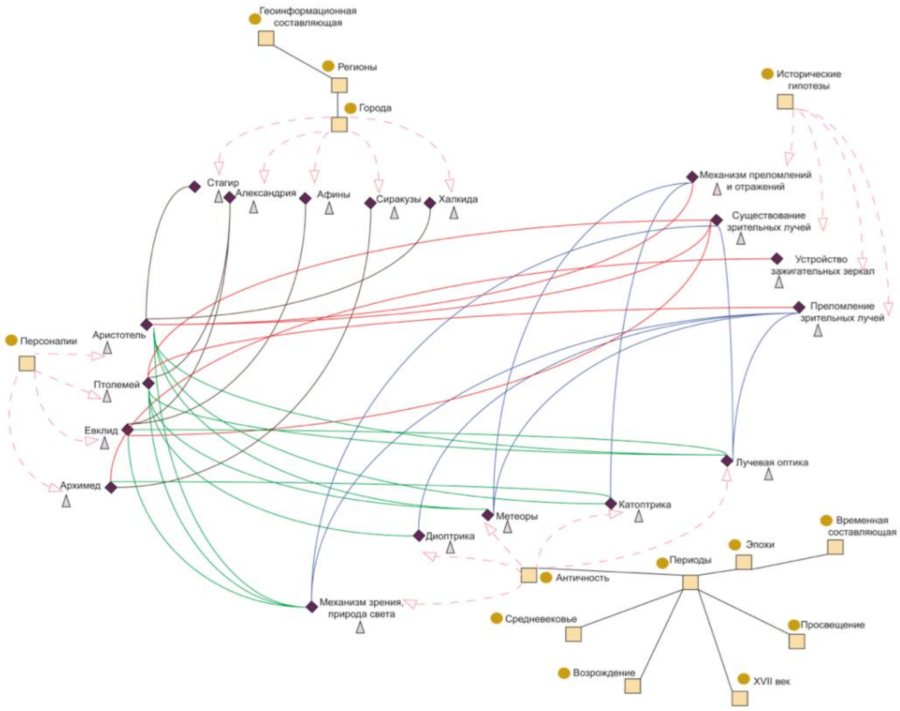

Fig. 4. Top level structure of optics domain ontology

The first issue was solved by a criteria of an existing object in real world. In other words, when defining the concept to be a class or an individual, the principle of "uniqueness" was chosen. That is, any device or subject of cultural heritage is considered as individual. The same situation is with the scientists’ profiles considered as individuals of the class "personalities".

The laws of optics, formulas and various optical schemes and models should also be considered as individuals, as they exist in a single copy. Similar arguments are also true for the physical characteristics and parameters. Although in the latter case, the situation is not so obvious: the "intensity" of the parameter can be merely represented in ontology as an individual, but, for example, "coherence" is of two kinds, "spatial" and "temporal".

At the same time a number of concepts in the ontology correspond to "pure abstractions" of the real world, which surely must be classes without any individuals. For example, such an abstraction is the concept of "Light", “Wave” or subclasses of “Optical Phenomenon”. One of the reasons why we should not consider things like different kinds of optical phenomenon as individuals of this class is that diffraction observed of some scientist in his experiments is not the same used in the process of holography recording in some laboratory. So we cannot say that particular type of diffraction is the unique subject as we can see an infinite number of realization of some phenomenon. A view of a part of optics ontology is shown on a Fig. 5.

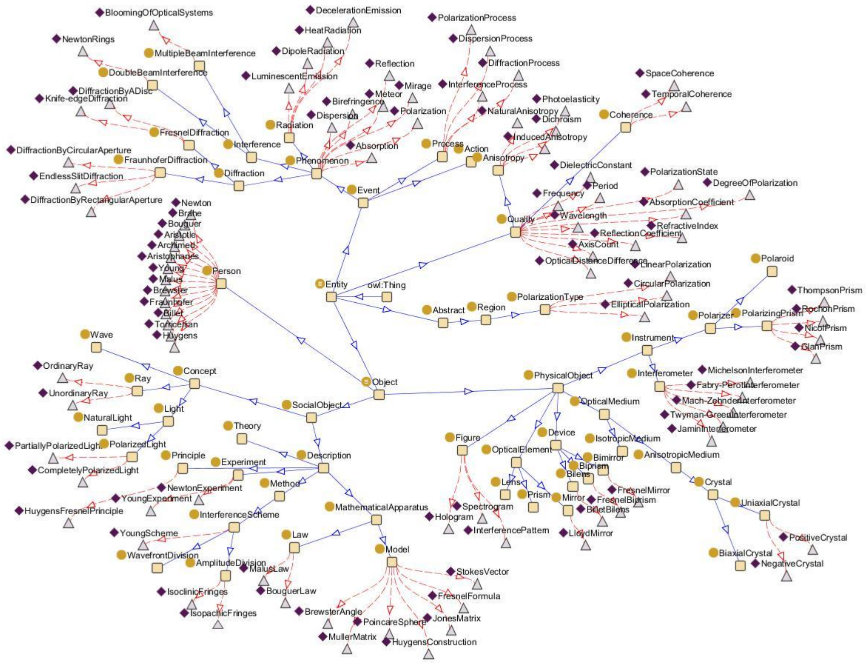

Fig. 5. A part of optics ontology view

Here we are faced with the second issue to be solved in this work: "What are the most well-defined relations between concepts?" In most cases it is only problem of knowledge engineering. For example, the relationship between "Diffraction grating" and "Fraunhofer diffraction" is "based on" and this fact is easily represented by the next expression:

```
DiffractionGrating SubClassOf basedOn some
FraunhoferDiffraction
```

where `basedOn` is an ObjectProperty with Domain "OpticalDevice" and range "PhysicalPhenomenon".

But there are cases when we have to relate an individual to a class and vice versa. For example, suppose a relationship "Any optical shutter changes the intensity". As we stated before, parameters of physical characteristics are individuals because they are parts of some optical schemas or models. At the same time we relate the concrete parameter "Intensity" to all individuals of the class "Optical shutter". So one of the possible expression for such a relation is

```
OpticalShutter SubClassOf changes some ({Intensity}).
```

This formalization cannot be considered satisfactory because in such case we force two problems:

- possible increase of computational complexity and
- losing concept's semantics when using nominals.

4 Discussion

The possibility to set relationships between classes and individuals likely is not an exact match to the situation in the domain area. Strictly speaking, relating the individual with the class, we can say on the one hand, for example, that someone observes not a general optical phenomenon but its concrete manifestation in particular place and time. On the other hand, the relation of the class to the individual does not necessarily implies a particular individual and not another. In other words, one rather wants to specify a very specific description of the object than the object itself. For example, referring to physical parameters we mean a specific description of these parameters rather than their embodiment in the real world (which in reality may not to be at all, because these entities are imaginary).

An elegant solution is representation of objects being neither classes nor individuals proposed in the theory of prototypes. This theory allows to operate not with the objects themselves but with their prototypes — abstract structures with the most characteristic properties of the described object. Thus the prototype is not a class but the most typical representative of this class, with generic but quite specific values of its properties. The formation of prototypes can be carried out in two ways: based on the general tendency or frequency of occurrence of features [18]. In the first approach, the prototype represents the average of all existing instances of some class thereby reflecting the central tendency of a category. For example, the prototype of a living room includes a door and windows. The second model assumes that the prototype is a combination of the most common features or an intersection of unique set of attributes relating to the various instances of a given class. For example, the prototype of a cat refers to a "mew" sound.

In the field of knowledge engineering a frame-based approach of the theory of prototypes was developed by M. Minsky [19]. From the definition of Minsky it follows that frame is a kind of record where information describing the stereotypical situation is stored, and in every particular situation this record can be adapted to the real situation. The theory of frames uses the assumption that the representation of concepts in the human brain does not require a clear definition of a set of properties, but is based on the concept of a prototype combining the most common properties. Thus prototypes can be interpreted as classes or as individuals depending on the context.

In this work we can consider the prototype of partially polarized light. We are interested in the very specific properties of this concept and its relationship to other entities in the domain, for example, to polarization devices. But by the reasons described above, there is no sense to create an individual of this class.

The OWL semantics does not allow to operate with entities like prototypes. But the attempt to solve this problem, for example, by defining a special kind of abstraction for dealing with prototypes will substantially complicate the complexity of the inference. Therefore, it is interesting to raise the issue of extending the semantics of one of the OWL dialects to use prototypes in ontological engineering in the same way as in the frame theory.

5 Conclusion

The structure of the educational ontology described in this paper differs significantly from most of the ontologies developed for other purposes. This unusual approach poses a number of conceptual problems in the knowledge engineering and ontology design. By putting these issues the author does not provide any exact solution but only outlines one possible way of development. Therefore, the main result of the work for today should be considered as a discussion and criticism of the ideas expressed in this article.

6 Acknowledgements

The author would like to thank Anastasiya Olshevskaya, Sergey Stafeev, Yury Katkov and Cyril Pshenichny for their contribution to this work.

The work is partially supported by grants from Russian Foundation for Basic Research and from Saint Petersburg City Administration.

7 References

1. D. Poole, C. Smyth and R. Sharma, "Semantic Science: Ontologies, Data and Probabilistic Theories," *Uncertainty Reasoning for the Semantic Web I*,P.C. da Costa et al., eds., LNAI/LNCS 5327, Springer, 2008
2. Karrer, T What is eLearning 2.0, http://elearningtech.blogspot.com/2006/02/what-is-elearning-20.html
3. Karrer, T Understanding eLearning 2.0 www.learningcircuits.org/2007/0707karrer.html
4. Downes, S E-Learning 2.0. http://www.downes.ca/post/31741
5. Karrer, T Corporate Long Tail Learning and Attention Crisis http://elearningtech.blogspot.com/2008/02/corporate-learning-long-tail-and.html
6. Mouromtsev D., Gorovoy V. Ontology-Based Approach to Build Intelligent E-Learning Systems // Proceedings of the 2010 international conference on e-learning, e-business, enterprise information systems, e-government (EEE Worldcomp'2010), Las Vegas, USA, July 12-15 2010. – pp. 102-107
7. Gavrilova, T., Dicheva D., Sosnovsky S., Brusilovsky P. Ontological Web Portal for Educational Ontologies // In Proc. of Workshop "Applications of Semantic Web Technologies for E-Learning. (SW-EL'05)" in conjunction with 12th Int.Conf. on Artificial Intelligence in Education (AI-ED'05), Amsterdam, 2005. – pp. 19–29.
8. Gavrilova T. Knowledge Mapping for Teaching and Learning // Int. Journal "The 21st Century: a scientific quarterly", Nr 2(24), Warsaw, Poland, 2007. – pp. 21–32.

9. Hastings J, Chepelev L, Willighagen E, Adams N, Steinbeck C, et al. (2011) The Chemical Information Ontology: Provenance and Disambiguation for Chemical Data on the Biological Semantic Web. PLoS ONE 6(10): e25513. doi:10.1371/journal.pone.0025513
10. The logic of biological classification and the foundations of biomedical ontology. In D. Westerståhl (Ed.), Invited Papers from the 10th International Conference in Logic Methodology and Philosophy of Sci- ence. Elsevier-North-Holland, Oviedo, Spain. http://ontology.buffalo.edu/bio/logic of classes.pdf.
11. Toshinobu Kasai A1, Haruhisa Yamaguchi A2, Kazuo Nagano A3, Riichiro Mizoguchi A Semantic Web system for supporting teachers using ontology alignment A International Journal of Metadata, Semantics and Ontologies Volume 2, Number 1 / 2007 pp.35 – 44.
12. Linked Learning 2011: 1st International Workshop on eLearning Approaches for the Linked Data Age – http://projects.kmi.open.ac.uk/meducator/linkedlearning/.
13. Novak J. & Cacas A. The Theory Underlying Concept Maps and How to Construct Them. Technical Report IHMC Cmap. 1, Florida Institute for Human and Machine. 2006.
14. Buzan, Tony. The Mind Map Book, Penguin Books, 1996. ISBN 978-0-452-27322-1
15. Smith A.M. Ptolemy and the Foundations of Ancient Mathematical Optics: A Source-Based Guided Study // American Philosophical Society. – Philadelphia: American Philosophical Society. 1999.
16. Lindberg D.C. Science in the Middle Ages // University of Chicago – Chicago: University of Chicago. 1978.
17. DOLCE Lite. Laboratory for Applied Ontology LOA – http://www.loa-cnr.it/ontologies/DUL.owl.
18. Robert L. Solso, Otto H. MacLin, M. Kimberly MacLin Cognitive Psychology. Pearson publishing. 2008.
19. Minsky M. A Framework for Representing Knowledge. MIT-AI Laboratory: Memo 306. 1974.