Developing an ontology of mathematical logic

Russell Boyatt, russell.boyatt@warwick.ac.uk
Mike Joy, m.s.joy@warwick.ac.uk
Department of Computer Science, University of Warwick, Coventry, UK

Abstract:

An ontology provides a mechanism to formally represent a body of knowledge. Ontologies are one of the key technologies supporting the Semantic Web and the desire to add meaning to the information available on the World Wide Web. They provide the mechanism to describe a set of concepts, their properties and their relations to give a shared representation of knowledge. The MALog project are developing an ontology to support the development of high-quality learning materials in the general area of mathematical logic. This ontology of mathematical logic will form the basis of the semantic architecture allowing us to relate different learning objects and recommend appropriate learning paths. This paper reviews the technologies used to construct the ontology, the use of the ontology to support learning object development and explores the potential future use of the ontology.

Keywords: logic ontology representation education protege

1. INTRODUCTION

Mathematical logic is an area of mathematics that includes a broad range of topics such as set theory, propositional logic, predicate logic and proof (e.g. axiomatic systems, natural deduction). The knowledge and skills acquired from a familiarity with these mathematical logic topics can be applied in several disciplines, including engineering, physics, medical science and computer science. In computer science specifically, the skills developed by learning about and understanding mathematical logic form an excellent foundation to support highly technical skills such as computer programming and systems design. In an competitive and economically expanding Europe a skilled workforce is required to be capable of creative and innovative problem-solving. It would therefore be desirable to develop a high-quality set of learning material capable of developing the mathematical logic skills of students and workers throughout the European Union.

The MALog project is an EU-funded project that is designed to address the need for learning materials in the area of 'mathematical logic', which the project considers to be an umbrella term encompassing a wide variety of mathematical topics. However in broad terms it is a subfield of mathematics concerned with the application of formal logic and deductive reasoning. One of the primary aims of the MALog project is to produce pedagogically high quality learning material units to help develop mathematical logic knowledge and skills. The learning material will be targeted at the needs of both the educational and industrial learning environments. The learning material will consist of all necessary theory, example problems, exercise sheets and, where appropriate, exam questions. Also included within the learning material will be appropriate metadata including learning goals, prerequisite information and ontology data.

The material will be aimed at three learner categories: high-school students, university undergraduate students and enterprise workers. Material directed at high-school students will focus on the foundations of mathematical logic such as truth tables and set theory, whereas material directed at undergraduate students will include more advanced topics including predicate logic and propositional logic. The material for the third category, enterprise workers, will be targeted towards developing specific skills as identified by the MALog project in collobration with partner companies. The requirements in each three categories are different and the exact requirements for each learner within that category are different. The project will produce material in an innovative way by making an individual learning path available to each learner with the aim of being as flexible as possible. Using a bank of learning material the most appropriate route through the information will be selected and adapted to best suit the needs of the learner. Each learner will be able to meet their own individual learning requirements and goals. The learning path will be determined in-part from the information provided by the mathematical logic ontology developed by the project.

2. ONTOLOGIES

An ontology provides a mechanism to formally represent a body of knowledge and has its origins in philosophy. Within the information sciences, the modern use of the word ontology has a specific technical meaning and Gruber (1995) gives the definition of an ontology as a "specification of a conceptualization". In simpler terms, an ontology is description of the set of concepts, their properties and the relations between them that give a shared representation of knowledge. An ontology is used to agree a vocabulary that is consistent and shareable, where the level ambiguity of lowered as the definition of terms is constrained.

Ontologies are now widely-used as the key technology in the development of the Semantic Web and the desire to add meaning to the information available on the World Wide Web. Berners-Lee (2001) describes the Semantic Web as a vision of the World Wide Web where metadata gives added meaning and flexibility to the way that computers can process and manipulative information. The ontology is one of the key technologies supporting the development of the Semantic Web and has given rise to several ontology languages such as RDF (W3C, 2004) and the Web Ontology Language – OWL (W3C, 2009).

The technologies developed as part of the Semantic Web have been used to develop smaller and more tightly focused ontologies. Most notably, the medical science community has developed and published many different ontologies, for example, projects such as the GALEN high-level medicine ontology (Rector and Rogers, 2006) and the Gene Ontology project (Ashburner, 2001). The GALEN project provides a common reference model to ensure mediation between the terminology used in the different electronic patient record systems. These examples show how large ontologies can be developed and successfully deployed.

The MALog project has been developing an ontology to provide a description of the topics in the mathematical sub-field of mathematical logic. This ontology will be used as the semantic architecture allowing learning material development to be properly managed. The information provided by the ontology will give learning object authors access to better information about the individual topics and the relationships between them. The information delivered by the ontology will be used by authors to help plan paths through the learning material and also be capable of adapting the learning path to the requirements of the individual learner. The terms identified by the ontology will ensure the project develops material using a common vocabulary of mathematical terms. This will help to ensure consistency across a large number of learning objects and reducing the likelihood of confusing a learner with different names for the identical mathematical concepts. Measuring the reuse of terms will also be possible, for example, a learner would be able to identify all instances of universal quantification allowing

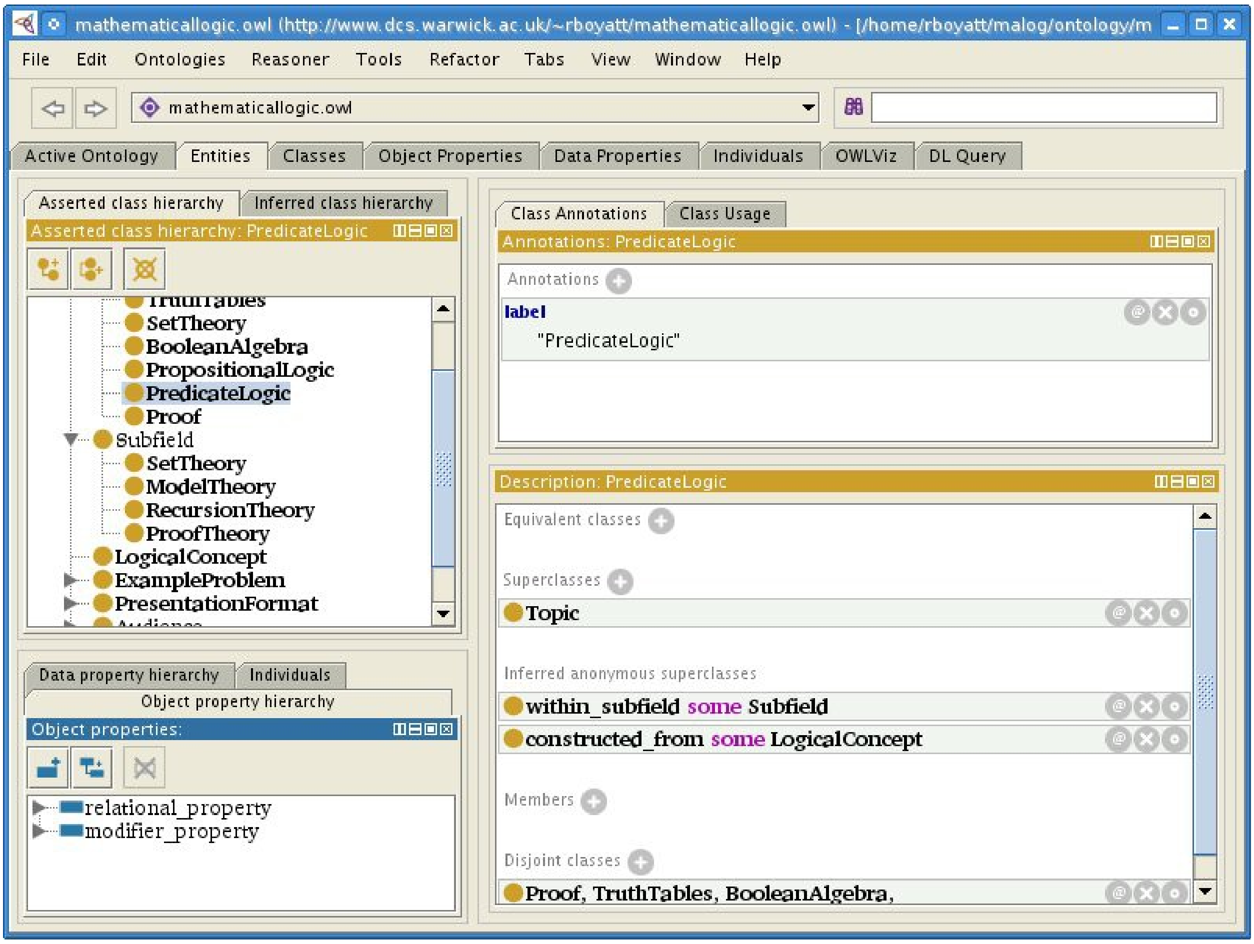

Figure 1: The Protégé Ontology Editor displaying part of an ontology.

Gruber (1995) gives five design criteria for ontologies: clarity (conveys intended meaning), coherence (ontology should be logically consistent), extendiblity (capable of being extended within given framework), minimal encoding bias (does not rely on particular notation or implementation) and minimal ontological commitment (only terms essential, not adding unneeded assumptions). These criteria form a good basis for the evaluation of the ontology throughout the lifetime of the project and widely adopted throughout the ontology community.

3. ONTOLOGY TOOLS

Clearly developing an ontology by hand with thousands of definitions is completely impracticable. Many different software tools are available for creating and manipulating ontologies. For example, Hozo Ontology Editor[1] developed by in-part by Osaka University, KAON[2] primarily by the University of Karlsruhe and Protégé Ontology Editor primarily at Stanford University. The MALog project uses the Protégé editor due to its flexiblity and well-supported development community[3]. The tool is available freely and as open-source allowing users easy access to the technology and ensuring the future availability of the tools. Development with Protégé also benefits from a large library of pre-existing ontologies available to be examined and used where appropriate. Figure 1 shows Protégé being used to develop an ontology.

---

1 Hozo Ontology Editor is available from http://www.hozo.jp/
2 KOAN2 is available from http://kaon2.semanticweb.org/
3 Protégé is available from http://protege.stanford.edu

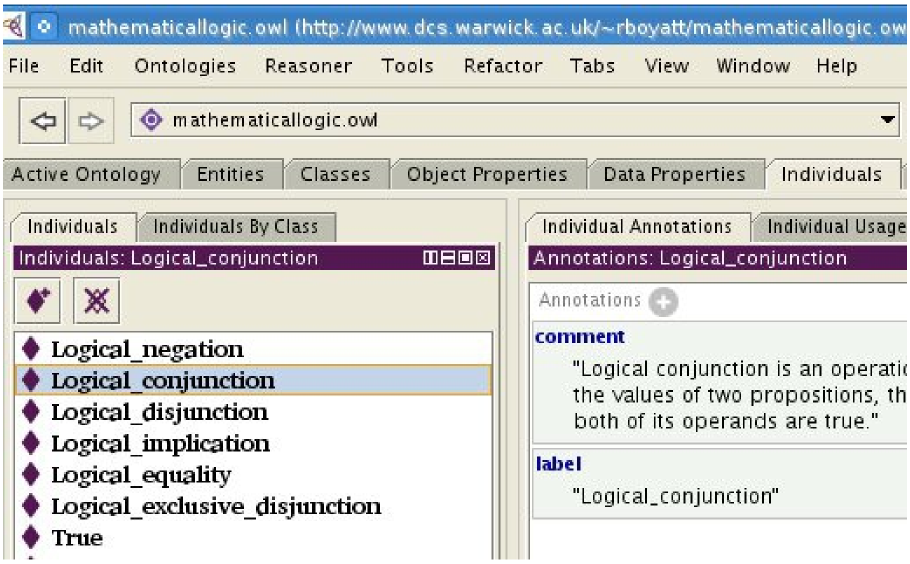

Figure 2: Using Protégé to examine ontology individuals representing mathematical logic topics

Ontology languages are formal languages used to represent ontology information. Domain knowledge is encoded using the ontology language which can also include inference rules to support processing of the domain knowledge. Examples of ontology languages include RDF, RDFS, OIL, DAML+OIL and OWL. The Web Ontology Language (OWL) is used as part of the Semantic Web and used for knowledge representation across the World Wide Web. The OWL language is serialized using an RDF/XML syntax capable of representing the classes, properties and individuals within an ontology. Due to the prominent use of the language in the Semantic Web, it has attracted attention from medical, industrial and academic fields. The language and supporting technologies are as a result undergoing a period of intensive research and development suggesting even wider use in the future. The OWL family of ontology languages are supported within the Protégé tool.

4. DEVELOPING THE ONTOLOGY

Mathematical logic is a subfield of mathematics primarily concerned with logic and reasoning. Constructing an ontology for such a large domain has required a careful staged development. The construction of the ontology began at the early stages of the MALog project itself and has now progressed through several stages of development. The sections below explain each stage of the ontology development.

1. First we identified the domain and scope of the ontology to be constructed. The ontology should include as much information about the broad field of mathematical logic as possible. There is clearly the potential to include many different topics but the ontology should focus on the knowledge relevant to the educational needs of the MALog project. There would also be the potential to build a full knowledge representation of mathematical logic. However, this project limits itself to building an ontology and describing the shared vocabulary to be used. It will not be necessary to use the information collected to perform any kind of mathematical calculation. In the early stages of the MALog project a needs-analysis survey was conducted to investigate the areas of mathematical logic already taught in schools and universities, and identify topics where more teaching materials were needed. The data collected helped to identify broad topics which should be included

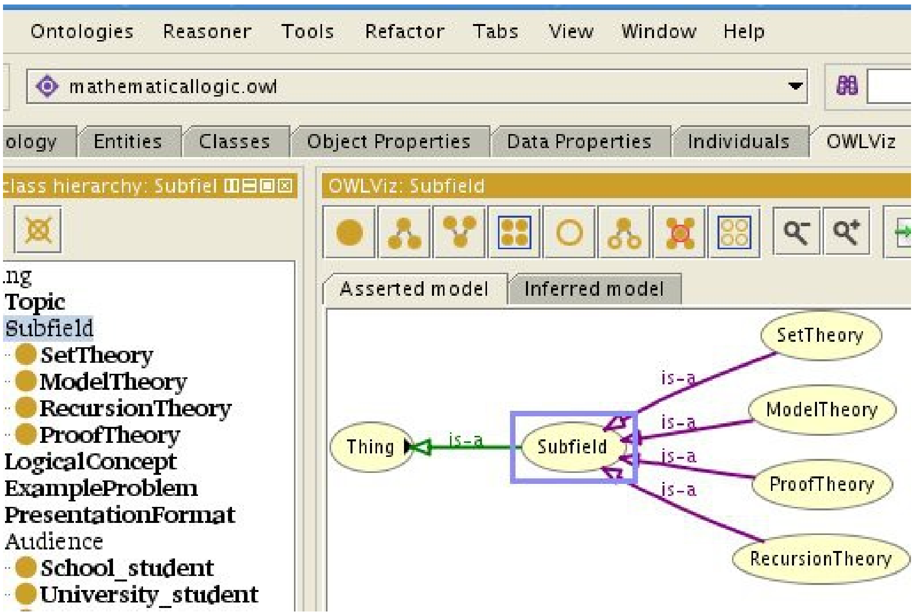

Figure 3: Visualising aspects of ontology class hierarchy with Protégé.

2. The development of any ontology should explore existing resources to ensure that nothing is available that could be adapted or extended. While some mathematical ontologies do exist, the project had established in early project proposals that none existed suitable to represent the required domain. Other mathematical ontologies such as GeoSkills, built by the I2Geo project, were nevertheless helpful in terms of identifying construction style and general approach to ontology construction (InterGeo, 2008).

3. At the next stage of the ontology development, collection of relevant information began to allow many different terms to be identified. The resources used at this stage include mathematical text books, mathematical logic teaching resources, websites and published mathematical literature. As the resources were found and consulted, identification of the vocabulary began - ready to be entered into the ontology.

4. Ontology construction began by constructing the class hierarchy. Classes were initially created by high-level topic areas, e.g. 'set theory', 'predicate logic', 'propositional logic'. The class hierarchy only provides the very basic structure. The relational and modifier properties add detail to the classes and their relationships. For example, a relational property 'within_subfield' is used to establish the relationship between topics and the subfield of mathematical logic (e.g. 'axiomatic set theory' within subfield of 'set theory'). Figure 3 shows part of the class hierarchy in graph form (the Protégé tool includes visualisation functionality).

5. When the basic ontology structure had been constructed individual instances could be created from the research conducted at stage 4. An individual instance of a class should fit within the definition of that class. Creation of individual instances with the Protégé tool is shown in Figure 2. For example, 'venn diagrams' is an instance of the 'axiomatic set theory' class. The information added with the relational and modifier properties helped add additional information from the domain as the ontology expands.

6. Evaluation of the ontology will be conducted alongside the construction of the mathematical logic learning materials. The clarity and coherence (from Gruber's criteria)

will be examined as the information is used to structure the learning materials and identify appropriate learning paths. The extendiblity and minimal ontological commitment will also be tested and measured as the ontology is modified during the next stage of the MALog project. Issues we may expect to encounter are miscategorized topics, whether topic group has occurred appropriately and whether a good vocabulary has been chosen.

5. FUTURE DEVELOPMENTS

The current version of the mathematical logic ontology cannot be considered the final version and further changes are inevitable. The ontology will continue to evolve throughout the lifetime of the project as information is updated and errors or omissions are discovered requiring the addition of new material. The ontology will not simply influence the development of the learning objects with a one-way flow of information. As the learning objects are developed, it will be possible to improve the ontology, for example, refining the vocabulary used, adding additional connections between topics or adding new information about the difficulty level of a topic.

As the ontology continues to evolve and is published it will be possible for other people to make use of the information. It hoped that the use of OWL will allow other projects to make use of the ontology information for on-line learning materials. The current version of the ontology only exists in English but it is hoped to expand this to encompass several other languages. The MALog project involves partners from several different countries and the learning material will be published in multiple languages. Some authors such as Espinoza et. al (2008) and Peters et. al (2007) have shown how ontologies can be adapted to cope with multiple languages. With the eventual publication of the learning materials in English, Finnish, French, German and Romania, it is hoped to expand the ontology to include additional language information. Authors such as Kerremans (2008) have presented approaches to the problem of culture-specific terminology in ontologies. As the MALog project is interacting with schools, universities and workplaces in multiple countries we expect to discover culture-specidesirablefic terms that must also be incorporated into the ontology to ensure coherence.

6. CONCLUSIONS

The ontology has been developed to represent a body of knowledge – in this particular case, mathematical logic. The ontology will now be used to aid in the creation of sixty pedagogically high-quality learning units. Once completed, both the learning materials and the ontology will be published and freely available to all interested European educational institutions and enterprises.

Acknowledgements

This work is supported in part by the Lifelong Learning Programme of the European Union, Project MALog, 505326-LLP-1-2009-1-FI-KA3-KA3MP.

References.

Ashburner, M., Ball, C. A., Blake, J. A. et. al (2000). “Gene Ontology: tool for unification of biology”, Nature Genetics, Vol. 25, pages 25-29.

Berners-Lee, T., Hendler, J. and Lassila O. (2001). "The Semantic Web", *Scientific American,* Vol. 284, No. 5, pages 34-43.

Espinoza, M., Gómez-Pérez A. and Mena, E. (2008) "Enriching an Ontology with Multilingual Information", *The Semantic Web: Research and Applications,* Lecture Notes in Computer Science, Vol. 5021, p. 333-347.

Gruber, Thomas (1995). "Toward Principles for the Design of Ontologies Used for Knowledge Sharing". *International Journal Human-Computer Studies* Vol. 43, Issues 5-6, p. 907-928

InterGeo Project (2008), "GeoSkills Ontology", Available from: *http://www.inter2geo.eu/2008/ontology/GeoSkills. Accessed 8th October 2010.*

Kerremans, K., Temmerman, R. and Tummers, J. (2003), "Representing Multilingual and Culture-Specific Knowledge in a VAT Regulatory Ontology: Support from the Termontography Method". In *On The Move to Meaningful Internet Systems 2003: OTM 2003*, Lecture Notes in Computer Science, Vol 2889, p. 662-674.

Peters, W., Montiel-Ponsoda, E., Aguado de Cea, G. and Gómez-Pérez, A. (2007). "Localizing Ontologies in OWL", In *Proceedings of the OntoLex07 Workshop at the 6th International Semantic Web Conference*, p. 13-22.

Noy, Natalya F. and McGuinness, Deborah L. (2001). "Ontology Development 101: A Guide to Creating Your First Ontology". Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880.

Rector, A. and Rogers, J. (2006). "Ontological and Practical Issues in Using a Description Logic to Represent Medical Concept Systems: Experience from GALEN", *Reasoning Web*, Lecture Notes in Computer Science, Vol. 4126, p. 197-231.

W3C World Wide Web Consortium (2009), "OWL 2 Web Ontology Language Guide", http://www.w3.org/TR/owl2-guide/. Accessed 12th October 2010.

W3C World Wide Web Consortium (2004), "RDF/XML Syntax Specification (Revised)", http://www.w3.org/standards/tech/rdf/. Accessed 12th October 2010.