

Towards automatic generation of e-assessment using semantic web technologies

Marija Cubric

University of Hertfordshire

Milorad Tasic

University of Nis

Abstract:

Semantic Web technologies have been increasingly used as a tool for generating, organizing and personalizing e-learning content, including e-assessment. In this paper we discuss and extend an innovative approach to automated generation of computer-assisted assessment (CAA) from Semantic Web – based domain ontologies. We expand the work previously done in this area in two important directions: first, we add new ontology elements (annotations), to the meta-ontology used for generating questions; second, we add semantic interpretation of the mapping between the 'domain' ontology and the target 'question' ontology. The semantic interpretation is based on the notion of 'question templates' that are founded on the Bloom's taxonomy of educational objectives, but can be sourced equally well in any other pedagogical premise applicable to question design and content (e.g. Kolb's learning theory). We show with examples obtained from the prototype implementation how that works in practice. The primary application domain for this work is in automated assessment for e-learning, and in particular, development of intelligent CAA systems and question banks, but the ideas can be further generalized in the context of ontology engineering and evaluation.

Introduction

The work on ontologies in e-learning domains has been primarily concentrated on ontological formalization of learning objects, instructional processes and learning designs (Sicilia and Barriocanal, 2005; Knight, et al. 2006) as well as on data mining techniques for the discovery of ontologies from various learning corpora (e.g. Montoyo, et al. 2005).

The role of ontologies in designing learning assessment has been less studied and only recently, are techniques for ontology-based assessment design strategies starting to emerge. Chung, Niemi, and Bewley (2003) describe an assessment authoring support system for "aiding assessment authors to populate the assessment ontology with values specific to the users' purposes". Holohan et al. (2005) concentrate on semi-automatic generation of simple learning objects such as slide shows and objective tests, in the context of an adaptive learning environment. They further extend their work in Holohan et al. (2006) to include dynamic problem generation, using domain-specific algorithms: the example considered was from the domain of relational databases and the resulting problems were database queries.

Papasalouros et al. (2008) describe various ontology-based strategies for automatic generation of Multiple Choice Questions (MCQ), from arbitrary knowledge domains. The generation is based on the basic meta-ontology relations between a 'class' and 'individual', as well as between two individuals (binary 'role'). Their strategies were further optimized and implemented as a plugin for the Protégé¹ ontology editor, by Tomic and Cubric (2009).

The main focus of our current work is ontology-based automatic generation of assessment of an arbitrary knowledge domain. The main novelty is in extending the existing body of research in two important directions: First, enriching the meta-ontology used for question generation with the new elements, such as annotations. Second, adding semantic interpretation to the mapping between the domain and the MCQ ontology in terms of 'question templates' (Figure 1). The first extension helps us in defining additional question types (section 4) and in defining new strategies for creating 'distracters' i.e. incorrect answers presented as choices in MCQs. It also provides us with rich textual descriptions of the ontology elements that we use in question generation. The second extension enables us to use levels in the Bloom's taxonomy (Bloom and Krathwohl, 1956) for labeling the mapping from elements of a domain to elements of the MCQ ontology. Papasalouros et al. (2008) discuss difficulties in generating syntactically correct questions using the natural language processing (NLP) techniques. We suggest the use of question templates, as an alternative to the NLP and primary means for generating question text ('stem'). Moreover, the templates provide us with necessary pedagogical underpinning for classification of different question types in the context of Bloom's taxonomy.

The rest of the paper is organized as follows: in section 2 (Knowledge Domain Ontology) we summarize main definitions related to ontologies; in section 3 (MCQ Ontology) we define the target (MCQ) ontology; in section 4 (Question Generation) we describe annotation-based 'question templates' and strategies for question generation and provide example questions corresponding to different levels of Bloom's taxonomy; in section 5 (Conclusions) we conclude with the summary of current and future work.

¹ <http://protege.stanford.edu/>

Knowledge Domain Ontology

The term ontology has been in use for a long time in philosophy, while it recently attracted a lot of attention within the computer science community due to the growing popularity of Semantic Web. In philosophy, the term ontology has been used since the 17th century to refer both to a philosophical discipline (Ontology with a capital "O"), and as a domain-independent system of categories that can be used in the conceptualization of domain-specific scientific theories (for more in-depth discussion see Guizzardi (2007)).

In this paper, we adopt the 'computer science' use of the word ontology to mean a formal specification of a conceptualization of a knowledge domain in terms of concepts, objects, relationships, and features relevant to modeling of the domain (Gruber, 1993). This set of abstract entities may be further specified by means of the definitions from the representational vocabulary such as classes, instances, relations, and properties respectively (see Table 1). The specification provides meanings for the vocabulary and adds formal constraints for its coherent use.

The current W3C Semantic Web standards² suggest specific formalisms for encoding ontologies, such as Resource Description Framework (RDF), Resource Description Framework Schema (RDFS), and Web Ontology Language (OWL) stack, which all vary in their expressive power (McGuinness and van Harmelen, 2004). The RDF language is a basis for encoding and it is usually interpreted as a set of statements about ontology resources in a form of subject-predicate-object triples (similar to Entity-Relationship model). RDFS extends RDF by providing a framework for description of application-specific classes and properties, while OWL adds more vocabulary for describing classes and properties and combined with a reasoning tool, provides logical facilities for reasoning and inference (Pahl and Holohan, 2009).

Table 1: Standard ontology elements at different levels of abstraction

Conceptualization	Specification	Relevant language constructs
Concept	Class	rdf:Class, rdfs:Class, owl:Class
Individual	Instance	rdfs:Resource, owl:Thing
Relationship	Relation	rdf:Property, owl:ObjectProperty
Feature	Property	rdf:Description, owl:DatatypeProperty
Textual description	Annotation	rdfs:label, rdfs:comment

Table 1 lists the most common ontology elements that may be exploited from domain ontology. Note, however, that the third column presents only few of the possible ontology language elements that can be used for the question generation.

In addition to the standard elements (first four items in Table 1), many of the currently developed ontologies incorporate 'annotations' i.e. textual descriptions of the underlying

² <http://www.w3.org/RDF/>

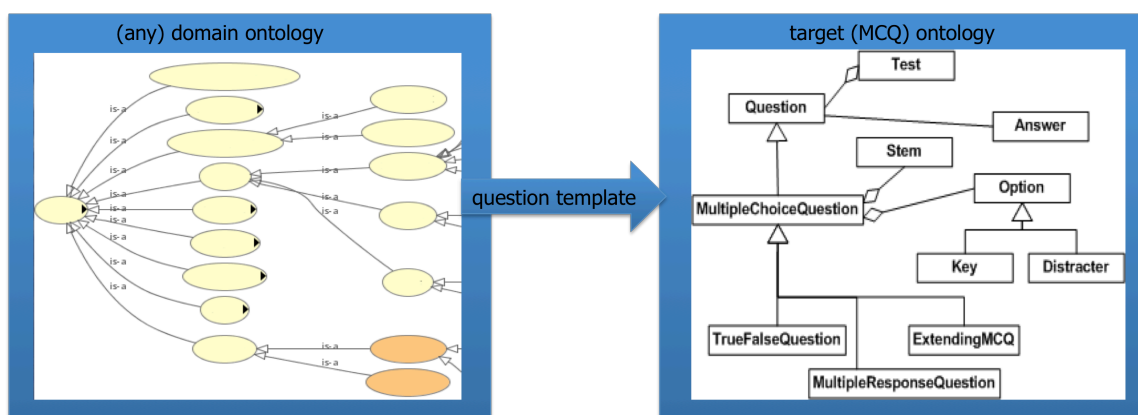
ontology elements (e.g. Gene ontology as described by Harris, et al. 2004). The role and usage of annotations in ontologies is currently topic of many research studies, for example, image annotations used in image retrieval (Liu et al., 2007), automating annotations with WordNet (Sanfilippo et al., 2006), to name a few.

There are other ontology elements not discussed in this paper, but potentially useful for future work. This is particularly true for the ontology elements defined at logic level, related to restrictions, rules, function terms, axioms, etc. It should be noted that the 'logic level' is not an exactly specified term i.e. its definition is context-dependent. Usually, the logic level is referred to within the OWL. As recently indicated by Rodriguez (2009) in the context of Web of Data, "it is necessary to separate RDF from its logic language legacy and frame it simply as a data model". Keeping that in mind and in order to avoid the trap of 'vagueness', we constrain our focus in this paper on the data model (RDF) aspect of the domain ontology, and we leave reasoning component and OWL representation for the purpose of automatic assessment for future research.

MCQ Ontology

We define MCQ ontology as a basis for development of the MCQ format specification. The ontology is based on the standard IMS (Instructional Management Systems) Test and Question information model (IMS, 2002) and it gets populated with question instances during the process of question generation. A Unified Modeling Language representation of the MCQ ontology is shown in Figure 1.

Figure 1 Mapping between domain and target (MCQ) ontology



The root concept is a composite class Test consisting of an arbitrary number of questions. At the next level there are classes Question and Answer which are in an n-to-m relation i.e. every question may be assigned zero or more answers while an answer can be assigned to zero or more questions. Every question has an instance of the Stem class as an attribute for textual representation of the question. The class Question is further specialized into the MultipleChoiceQuestion subclass. The MultipleChoiceQuestion has one correct answer and one or more wrong answers, corresponding to instances of Key class and Distracter class respectively. Depending on the number of Options, Keys and Distracters. It can be further specialized into TrueFalseQuestion, MultipleResponseQuestion etc.

Question Generation

We define 'question generation' to be a process of populating Test class instances (part of MCQ ontology shown in Figure 1), with a list of questions, that are based on the knowledge described in the underlying domain ontology. Each individual question consists of exactly one instance of the Stem, exactly one instance of the CorrectAnswer, and one or more instances of the Distracter. We extend notation used in Papasalouros et al., (2008), with some annotation-specific symbols i.e. $A:x$, for: annotation x is a description of class A and $x(A)$ for: annotation text x contains the concept A .

Annotation-based stems and strategies

Examples of stems that correspond to the basic meta-ontology relations between classes and individuals, were introduced in Holohan et al, (2005) e.g. "Which one of the following items is an example of the concept A ?" or "Which one of the following items is not an example of the concept A ?" etc. We extend their approach by introducing new stems that are using annotated information from the domain ontology. For example:

- a) Which one of the following definitions describes the concept A ?
- b) Read the text x below and decide which one of the following concepts is a correct replacement for the blank space in the text³
- c) Read the paragraph x below and decide which one of the following concepts it defines.

We also add new annotation-based strategies for generating question distracters based on the assumption that the more similar distracters are, the more difficult question becomes (Mitkov et al, 2008). As the 'semantic similarity' measure we combine a text similarity measure with an ontology elements similarity measure (Bach and Dieng-Kuntz, 2005). To start with, we sort a list of annotations according to the similarity measure, and then narrow the selection down to a number of items from the top of the list. We say that the selected items make 'high similarity' pairs with the original annotation.

The following strategies are used in the current prototype implementation for generating distracters for the type of questions described above:

- (1) If the correct answer is an annotation x , that describes the concept A i.e. if $A:x$, then, for a distracter choose any annotation y , such that for some $B \neq A$, $B:y$ and the pair (x,y) has 'high similarity' measure.
- (2) If the correct answer is a concept A , that occurs in annotation x i.e. if $x(A)$ and $A:y$ for some y , then, for a distracter, choose any concept $B \neq A$, such that for some z , $B:z$ and the pair (y,z) has 'high similarity' measure.
- (3) If the correct answer is a concept A , described by the annotation x i.e. if $A:x$ then, for a distracter, choose any concept $B \neq A$, such that for some annotation y , $B:y$ and the pair (x,y) has 'high similarity' measure.

³ 'Blank space' is generated by removing an arbitrary pre-defined concept from the text.

There are other different techniques for computing the text similarity measure. For example, Cohen et al. (2003) provide a comprehensive overview of similarity measures for short text segments that can be used for the purpose, and we continue experimenting with other combinations of similarity measures as a part of our ongoing research.

Semantic interpretation

Bloom’s taxonomy represents one of the most influential models of learning objectives and educational competencies. (Anderson and Sosniak, 1994). While the critiques of Bloom have been mainly focused on the sequential nature of the model and questioned its usefulness in describing the process of learning (Moore, 1982) its use in assessment authoring has been widely accepted in practice and has become a part of the educational ‘folklore’.

We use ‘question templates’ described earlier to assign semantics to the mapping between the domain and the target MCQ ontology. In this way, we are able to use levels in the Bloom’s taxonomy (knowledge, comprehension, application, etc.) for labeling the mapping from elements of a domain to elements of the MCQ ontology and to use individual levels’ action verbs, such as define, relate, analyze etc (Felder and Brent, 1997; CAA centre resources (2002)) in forming the ‘stems’ for question templates. This approach we call ‘semantic interpretation’.

In Table 2 we present some illustrative examples for the knowledge, comprehension, application and analysis level questions together with corresponding strategies for question generation. Strategies are numbered as in the previous section i.e. (1), (2) or (3), and combined with property-based and class-based strategies (Papasalouros et al., 2008).

Table 2: MCQ examples

Level	Question Stem	Correct answer	Strategy for generating distracters
Knowledge	Which of the following definition describes the concept <A>?	any x where A:x	(1)
	Read the paragraph <x> and decide which one of the following concepts it defines	any A, where A:x	(3)
Comprehension	Which one of the following response pairs relates in the same way as <a> and in the relation <R>?	(c,d) where R(a,b) and R(c,d)	Property-based strategies
Application	Which one of the following examples demonstrates the concept<A>?	any a, where A(a)	Class-based strategies
Analysis	Analyze the text <x> and decide which one of the following words is a correct replacement for the blank space in <x>	any A, such that x(A)	(2)
	Read the paragraph <x> and decide which one of the following concepts generalize the concept defined by<x>	any B, A<B and A:x	(3) combined with class-based strategies

Although some authors question the usefulness of MCQs for assessing higher-order skills above the knowledge level (Wood, 2003), examples of MCQs at all levels of Bloom's taxonomy can be found in the literature and practice (see for example Bull & McKenna (2003) and other CAA centre resources (2002)). The use of Bloom's taxonomy as a basis for semantic interpretation enables us to order the generated questions according to increasing educational objectives and to apply some of the ideas from adaptive computer-assisted assessment (see for example Lilley and Barker, 2002) in tailoring the tests according to the target competencies and objectives of individual learners. It is important to notice here that the Bloom's taxonomy is only one of the learning theories applicable to question design and content. Other approaches such as Kolb's experiential learning (Kolb, 1984) or Curry's 'onion model' (Curry, 1983) can be equally useful. Barker (2008) provides some good examples of assessment tasks based on Kolb's model that can be further extended to MCQs. Decoupling the semantic interpretation from the question generation process allows us to customize the question generation process by changing the underlying learning theory 'on the fly', according to different subject areas (Atherton, 2009) or different cultures (Neal and Schoenborn, 2010).

Conclusions

Creating engaging assessment strategy is one the most difficult areas in every learning design. Objective testing (including MCQs) has been extensively studied and evaluated as a method for formative and summative assessment that can speed-up the assessment process and engage students in a regular 'learning conversation'. While the results of the objective tests might not always be used to evaluate "deep learning" they certainly form a useful base and can be used as a "seed" for further assessment enhancements. Furthermore, if correctly implemented, they are addressing some of the most important student needs, such as, prompt and frequent feedback (NSS, 2009). However, creating a good objective test is not only difficult but also a very time-consuming task, which prevents its more wide-spread adoption and use. In this paper we are trying to address this problem, by extending the current research on automatic assessment (MCQ) generation with the application of annotations as well as with the semantic interpretation of the generation based on educational competencies defined by Bloom. The prototype of the proposed approach is implemented as a Protégé plugin⁴. We envision extending it in future along several prospective directions: extend and enrich the template base; add other ontology components, such as rules, axioms, restrictions, events etc. to the meta ontology used for question generations; extend the empirical dataset to include some larger 'real-life' ontologies; generalize the MCQontology to include other types of objective tests; apply integration of text similarity and ontological approaches to generation of distracters etc.

Acknowledgments

The prototype was developed with help of students on the Semantic Web course at Faculty of Electronic Engineering, University of Nis⁴. We would also like to thank the reviewers and our colleagues Trevor Barker and Amanda Jefferies for their constructive comments.

⁴<http://infosys3.elfak.ni.ac.rs/nastava/Wiki.jsp?page=SemantickiWebKurs>

References

- Anderson, L. W., Sosniak, L. A. (1994) Bloom's Taxonomy: A forty-year retrospective The National Society For The Study Of Education (June 27, 1994)
- Atherton, J. S. (2009) *Learning and Teaching; Experiential Learning* Available on-line at: <http://www.learningandteaching.info/learning/experience.htm>
- Bach, T.L. and Dieng-Kuntz, R. (2005) Measuring Similarity of Elements in OWL Ontologies. *AAAI'2005 Workshop on Contexts and Ontologies: Theory, Practice and Applications*. Pittsburgh, USA. 96-99.
- Barker, T. (2008), E-Learning 2.0 and Kolb's Learning Styles Inventory, *Proceedings of SOLSTICE Conference*, Edge Hill University, UK, June, 2008
- Bloom, B. S., Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals by a committee of college and university examiners. Handbook 1. Cognitive domain*. New York: Addison-Wesley.
- Bull, J. & McKenna, C. (2003) *Blueprint for Computer Assisted Assessment*. London: RoutledgeFalmer.
- CAA Centre Resources (2002) <http://www.caacentre.ac.uk/resources/> (accessed 12/5/2010)
- Chung, G., Niemi, D., Bewley, W. L. (2003) Assessment Applications of Ontologies. Paper presented at the Annual Meeting of the American Educational Research Association
- Cohen, W., Ravikumar, P., Fienberg, S. (2003) A comparison of string distance metrics for name matching tasks. In *IIWEB*, 73–78.
- Curry, L. (1983) An organization of learning style theory and constructs, in: L. Curry (Ed) *Learning Style in Continuing Education*. (Canada, Dalhousie University).
- Felder R.M., Brent, R. (1997) Objectively Speaking, *Chemical Engineering Education*, 31(3), 178-179
- Gruber T. R. (1993), A Translation Approach to Portable Ontology Specifications. *Knowledge Acquisition*, 5(2):199-220.
- Guizzardi G. (2007), On Ontology, ontologies, Conceptualizations, Modeling Languages, and (Meta)Models, *Frontiers in Artificial Intelligence and Applications, Databases and Information Systems IV*, Olegas Vasilecas, Johan Edler, Albertas Caplinskas (Editors), ISBN 978-1-58603-640-8, IOS Press.
- Harris M.A., et al. (2004) The Gene Ontology (GO) database and informatics resource, *Nucleic Acids Res*, 32, D258–D261
- Holohan, E. et al, (2005) Adaptive e-learning content generation based on semantic web technology. *Proceedings of Workshop on Applications of Semantic Web Technologies for e-Learning*, Amsterdam, The Netherlands, pp. 29-36.
- Holohan, E. et al, (2006) The Generation of E-Learning Exercise Problems from Subject Ontologies. *Proceedings of the Sixth IEEE International Conference on Advanced Learning Technologies*, 967-969.

- IMS Global Learning Consortium, Inc. IMS Question and Test Interoperability Version 2.1 Public Draft Specification. <http://www.imsglobal.org/question/index.html>, accessed 9 June 2010.
- Knight, C., Gašević, D., Richards, G. (2006). An Ontology-Based Framework for Bridging Learning Design and Learning Content. *Educational Technology & Society*, 9 (1), 23-37
- Kolb, D. (1984) *Experiential Learning*. Englewood Cliffs. N.J.: Prentice Hall
- Lilley, M., & Barker, T. (2002). The development and evaluation of a computer-adaptive testing application for English language. *6th Computer assisted assessment conference*, July 2002, Loughborough.
- Liu Y., Zhang D., Lu G., and Ma W.-Y. (2007) A survey of content based image retrieval with high-level semantics, *Pattern Recognition*, 40(1), 262–282.
- McGuinness D. L. and van Harmelen F. (2004), OWL Web Ontology Language. *W3C Recommendation*, <http://www.w3.org/TR/owl-features/>, accessed 12 May 2010.
- Mitkov, R., Ha, L.A., Varga, A. and Rello, L. (2009) Semantic similarity of distracters in multiple-choice tests: extrinsic evaluation. *Proceedings of the EACL 2009 Workshop on GEMS: GEometrical Models of Natural Language Semantics*, pages 49–56, Athens, Greece, 31 March 2009.
- Montoyo A, Munoz R., Metais E. (2005) " Text2Onto – A Framework for Ontology Learning and Data-driven Change Discovery" in *Proceedings of the 10th International Conference on Applications of Natural Language to Information Systems (NLDB)*, volume 3513 of Lecture Notes in Computer Science, pp. 227-238. Springer, Alicante, Spain, June 2005).
- Neal, V., Schoenborn, P. (2010) Authority and Power: Comparing pedagogies across cultures, The London SOTL, 8th International Conference
- NSS (2009) UK National Students Survey <http://www.hefce.ac.uk/learning/nss/data/2009/>, accessed 12 May 2010.
- Pahl, C., Holohan, E. (2009) Applications of Semantic Web Technology to Support Content Development, *Interdisciplinary Journal of E-Learning and Learning Objects*, 5, 1-25.
- Papasalouros A., Kotis K., Kanaris K. (2008) Automatic generation of multiple-choice questions from domain ontologies. *IADIS e-Learning 2008 conference*, Amsterdam.
- Rodriguez, M. (2009), Interpretations of the Web of Data, *arXiv:0905.3378v2*, <http://arxiv.org/abs/0905.3378>, accessed 12 May 2010.
- Sanfilippo, A., et al. (2006) Automating Ontological Annotation with WordNet. *Proceedings to the Third International WordNet Conference*, Jeju Island, Korea.
- Sicilia M.A, Barriocanal E.G. (2005) On the convergence of formal ontologies and standardized e-learning, *Int. Journal of Distance Education*.3(2), 13-29.
- Tosic M., and Cubric M. (2009), SeMCQ – Protégé Plugin for Automatic Ontology-Driven Multiple Choice Question Tests Generation, 11th Intl. Protégé Conference, Poster and Demo Session.
- Wood, E. (2003) What are Extended Matching Sets Questions? BEE-j (1,1). Available on-line at: <http://bio.ltsn.ac.uk/journal/vol1/beej-1-2.aspx>, accessed 12 May 2010.