IJCSE International Journal of Computer Sciences and Engineering Open Access
Survey Paper Vol-6, Special Issue-6, July 2018 E-ISSN: 2347-2693

# Automatic Generation of MCQS from Domain Ontology- A Survey

**Sahana Serin V. P.[1*], Viji Rajendran V.[2]**

[1,2]Department of Computer Science and Engineering, N.S.S College of Engineering, Palakkad, Kerala, India

[*]*Corresponding Author: sahanaserinvp@gmail.com, Tel.: +91-8281491022*

**Available online at: www.ijcseonline.org**

***Abstract—*** Ontologies are knowledge representation structures, that models domain knowledge by concepts, instances, rolesand their relationships. Assessment systems can exploit this knowledge by using multiple choice Questions (MCQs). Online assessment systems are mainly using MCQs instead of subjective questions for conducting the tests. Using MCQs for assessments has merits as well as demerits. For assessing wide range of knowledge, MCQs are used. It is because they require very less administrative overhead as well as provide instant feedback to test takers. There are several ontology based MCQ generation approaches proposed by many authors. These approaches generates different kinds of questions, in one approach the stem of all generated question remains the same, another one make use of the semantics of the domain, represented in the form of TBox axioms andABox axioms, to frame interesting MCQs. Some other methods differ in generating distractors for the questions. There are approaches which controls the difficulty level of generated MCQs. This paper gives a literature review and comparison of some of the methods for MCQ generation from ontology.

***Keywords***— Multiple Choice Questions, Distractors, Ontology

## I. INTRODUCTION

Automatic question generation systems are relatively new field. It helps the test takersto ease the burden and reduce the cost for manual assessment of the tests. MCQs are well suited for testing student’s knowledge in a subject area. It can also be implemented in online learning systems because it gives immediate feedbacks. Use of MCQs reduces the time to value the answers.MCQ design is a difficult task, either by hand or automatically. It is because,in addition to a good question and a correct answer,we need to generate suitable incorrect options.

Ontologies are a set of concepts and categories in a subject area or domain that includes their properties and the relations between them. The knowledge in a particular domain can be exploited using assessment systems by MCQs. The existing approaches [5] , thatgenerates MCQs from ontologies represented in Web Ontology Language (OWL) have limitations. They generate simple concept related questions or analogy type questions using roles. Structured domain knowledge in the form of Description Logic (DL) based ontologies can be used to generate MCQs automatically. So online assessment systems can utilize existing knowledge bases for assessing learner’s skills and knowledge in a subjectarea.

An ontology is a set of terminological or assertional axioms. Terminological axioms are related to relationships between concepts. Assertional axioms are related to relationships between individuals and rolesor between individuals and concepts. DL ontologies have formal semantics [1] .Ontology is a logical theory which indicates that explicitly stated knowledge can infer implicit knowledge. DLs are decidable fragments of first order logic with roles and concepts. An ontology can be defined as a restricted set of axioms. A domain ontology represents concepts which are belonging to part of the world. Domain ontology provides particular meanings of terms applied to that domain. Domain ontologies represent concepts in very specific and often selective ways and thus they are often conflicting [2] .

An MCQ item is made up of the following parts:

- A stem
- A key
- Some distractors

The stem is a statement or sentence that introduces a question. The key is simply the correct answer. MCQs contain a number of plausible incorrect answers. They are called distracters. The optimal number of distractors for MCQs remains varying. One way to alter the difficulty of an MCQ is by the similarity between the key and distractors. If they are more similar, then it is difficult to select the correct one because of difficulty to discriminate between them. Stem and answers derived from concepts are used for generating conceptual knowledge MCQs from an ontology. Hence if there is a proper similarity measure for concepts, then we can generate MCQs with controllable difficulty.

The remainder of the paper is organized as follows: section 2 presents relevant works on automated question generation. Section 3 presents comparison of these works and section 4 provides conclusions.

## II. Related Work

*A. Same stem for all generated MCQs*

Ontology-based question generation methods vary with the characteristics of the generated Questions. Papasalouros et al. [4] suggested some strategies based on classes, properties and terminologies of ontologies for framing MCQs and the corresponding distracters. Their MCQ generation methods have limitation that they lack proper theoretical support for when to use which strategy, and uses same stem for all generated questions*(”Choose the correct sentence” )*. Their strategies are presented with examples taken from a domain ontology in the Greek ancient history domain called ‘Eupalinos Tunnel’.

When using class-based strategies for generating question items, declarations in the form Class(*Individual*) in the ontology are converted in to sentences of the form ‘*Individual* is a(n) Class’. For example, ‘Eupalinos is an engineer’. All the choices (including key and distractors) for the questions in this strategy are in same form “Individual is a(n) Class”. For questions generated by property-based strategies, choices are generated in the form ‘*Individual* propertyName *Individual*’. Terminology-based Strategies are based on concept/sub concept relationships, instead of dealing with individuals in the ontology. For example, ‘rulers are politicians’ is the correct answer, since rulers is a subclass of Politician. ‘rulers are monks’ is a distractor, since Monk is a sibling class of Politician [4]. The same stem of all generated MCQs and same sentence format of all choices remains as a limitation of this approach.

*B. Semantic based stem generation*

Constructing stems for MCQ is a challenging task. It is more time consuming and requires experience in question generation. The OntoQueengine[15], is an ontology based MCQ generation engine that generates a set of stems from a given ontological domain. The stems contains MCQ-type, true/false, and fill-in items. MCQ stems are generatedusing statements from the ontology. OntoQue uses three strategies for generation of stem: individuals,class membership, and property as in [4]. Class membership gives stems of the type “what is the kind of.” For generating MCQ stems using this strategy, all classes that are defined and their instance members are collected from a group of RDF statements in the ontology.

In individual-based strategy, first list all individuals from the domain ontology. Then for each individual, collect all assertions in which the individual is a subject or object. These statements give the basic form for generating the stem. There are properties that always relate subjects and objects. The engine contains algorithms for construction of stems from property axiomssuch as transitive properties. Before removing subjects or objects,statements are extracted and converted to natural language. Since the statement is true, this removed element will then use as the key [14]. Distracters are generated randomly from a set of classes except the object class in class-based strategies.

Distracter quality of seeming reasonable is an important aspect of an MCQ item. It is considered as one of the major challenges faced by test takers when creating MCQs for the tests. Although the previous systems [4][14] uses algorithms and techniques to identify options that are similar to the key as distracters, results shows that this is not sufficient. A plausible distracter creates confusion and makes students think. One good approach for creating quality distracters is to consider student’s repeated areas of misunderstanding or their common errors.

*C. Varying difficulty level of MCQs*

The similarity measure of concepts in an ontology can be used to control the difficulty of generated MCQs. Good quality MCQs can be generated by improving distractor generation methods based on similarity. A method by Alsubait et al. [11] presented the calculation of similarity between concepts. These concepts can be used as distractors by considering similarity between them. The complication of an MCQ can alter with the similarity between the key and distractors. The more similar they are, the more knowledge is needed to distinguish between them and it is difficult to select the correct one.

They presented a similarity measurer that will compute the similarity between concepts w.r.t. ontology*O*. The similarity measurer is enclosed with many similarity measures with different precision, ambiguity and computational costs. All measures inspireJaccard’s similarity coefficient and the similarity *Sim(C, D, L, O)* of two concepts *C*, *D* in a concept language *L* w.r.t. an ontology *O* is defined as follows:

$$Sim(C,D,L,O) = \left|\frac{Subs(C,L,O) \cap Subs(D,L,O)}{Subs(C,L,O) \cup Subs(D,L,O)}\right|$$

Where *Subs(C, L, O)* is the set of all concepts which subsumes *C* in *L* w.r.t. *O*.Using this approach the same stem and key can be used to generate different questions with different levels of difficulty based on the similarity between the key and the selected distractors.

*D. Label-set based MCQ generation*

Vinu et al. [10] introduced two new approaches for MCQ generation based on label-sets. Label-set is a set containing the constraints satisfied by the instances. They associated each instances with the label-set. Label-set of an individual instance is called node label-set and label set of a pair of instances is called edge label-set. The overall architecture of their system is given in figure.1.

 

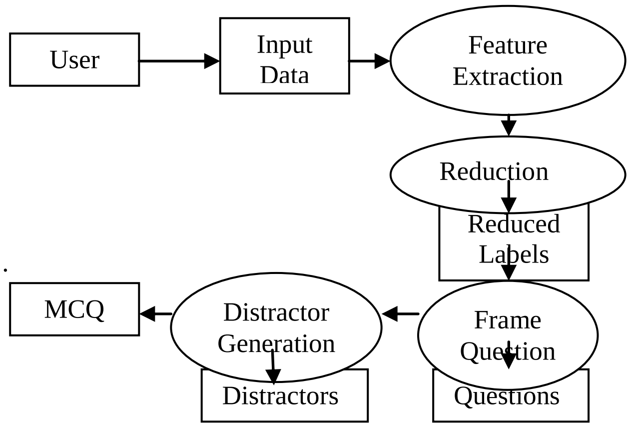


Figure 1: Label-set based MCQ generation system

All the logical expressions in a label-set are not useful for framing a stem, so they can be further reduced. Redundant entities are removed and combine two or more expressions to form a human understandable sentence for a meaningful stem generation. It is achieved by label-set reduction process.
The system uses some reduction rules to obtain reduced label-sets. Their system uses approach introduced in [11] to control difficulty level of MCQs. When used with semantically-rich ontologies, this system can generate good MCQs.

## III. COMPARISON

Table 1: Comparison of various ontology based MCQ generation methods

| No | Method | Methodology | Advantages | Disadvantages |
|---|---|---|---|---|
| 1 | Same stem for all generated MCQs | Use some strategies based on class, properties and terminologies of ontology | Easy to generate syntactically correct stems and distractors | Generates poor MCQs , All MCQs have same stem |
| 2 | Semantic based stem generation | Use strategies based on class, properties and individual of ontology | Generates meaningful stems , Different kinds of stems using different strategies | Generates poor MCQs |
| 3 | Varying difficulty level of MCQs | Compute similarity between two concepts to control difficulty level of MCQs | Meaningful stems, generates MCQs of different difficulty levels | Questions are not syntactically perfect |
| 4 | Label-set based MCQ generation | Associates each instances with label set and label set is used to form a stem | Generates good quality MCQS | Questions are not syntactically perfect |

## IV. CONCLUSION

.There exist many researches on automatic MCQ generation approaches from ontologies. It is proven that ontology based MCQ generation methods are very useful for generating good quality MCQs. These methods work well for defining the semantics of questions to generate. But the problem of syntactically correct question generation is not completely solved. In order to overcome syntactic problems, more advanced natural language generation techniques should be utilized. Generated distractors determine the hardness and quality of an MCQ. So that distractor generation process has to be given more importance.

## REFERENCES

[1] F. Baader, D. Calvanese, D. L. McGuinness, D. Nardi, and P. F. (eds.) PatelSchneider. "*The Description Logic Handbook: Theory, Implementation and Applications*" , Cambridge University Press, second edition, 2007.

[2] Muhammad Aun Abbas, " *A Unified Approach for Dealing with Ontology Mappings and their Defects*" , Universite de Bretagne Sud, 2016.

[3] T. Alsubait, B. Parsia, U. Sattler, " *A similarity-based theory of controlling mcq difficulty*" , 2013 Second International Conference on e-Learning and e- Technologies in Education (ICEEE), pp. 283-288, 2013.

[4] A. Papasalouros, K. Kanaris, K. Kotis, " *Automatic Generation of Multiple Choice Questions from Domain Ontologies*" , International Association for Development of the Information Society, pp. 427-434, 2008.

[5] M. Cubric, M. Tosic, " *Towards automatic generation of e-assessment using semantic web technologies*", Proceedings of the 2010 International Computer Assisted Assessment Conference, ISSN:2045-9432, Vol.1, 2010.

[6] E. Holohan, M. Melia, D. McMullen, C. Pahl, " *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, 2005.
[7] M. Tosic, M. Cubric, " *SeMCQ-protege plugin for automatic ontology-driven multiple choice question tests generation*" , Proceedings of the 11th International Protege Conference, 2009.
[8] B. Aitko, S. Stankov,M. Rosi, A. GrubiAi, " *Dynamic test generation over ontology-based knowledge representation in authoring shell*" , Syst. Appl. 36(4), pp. 8185-8196, 2009.
[9] Ian Horrocks, Peter F.Patel- Schneider, Frank van Harmelen, "*From SHIQ and RDF to OWL: the making of a Web Ontology Language*" , Web Semantics: Science, Services and Agents on the World Wide Web, Vol. 1, Issue.1, pp. 7-26, December 2003.
[10] Vinu E.V., Sreenivasa Kumar P. , " *A novel approach to generate MCQs from domain ontology: Considering DL semantics and open-world assumption*" , International journal on Web Semantics: Science, Services and Agents on the World Wide Web, 2015.
[11] Tahani Alsubait, Bijan Parsia, Ulrike Sattler, " *Ontology-Based Multiple Choice Question Generation*" , German Journal on Artificial Intelligence,Vol.30, Issue.2, pp. 183-188, June2016.
[12] T. Alsubait, B. Parsia, and U. Sattler, " *Measuring similarity in ontologies: How bad is a cheap measure?*" ,27th Inernational Workshop on Description Logics (DL-2014), 2014.
[13] T. Alsubait, B. Parsia, U. Sattler, " *Generating multiple choice questions from ontologies: Lessons learnt*" , Proceedings of the 11th International Workshop on OWL: Experiences and Directions, Vol.1265, pp. 73-84, 2014.
[14] M. Al-Yahya, " *OntoQue: a question generation engine for educational assesment based on domain ontologies*" , Proceedings of the 11th IEEE International Conference on Advanced Learning Technologies (ICALT '11), Athens, Ga, USA, pp. 393-395, July 2011.

## Authors Profile

*Mrs. Sahana Serin V. P.* pursed Bachelor of Technology in Computer Science and Engineering fromUniversity of Calicut, India in 2016. She is currently pursuing Master of Technology in Computer Science and Engineering, NSS College of Engineering, Palakkad, Kerala, India.

*Mrs.Viji Rajendran V.* pursued Ph.D in Cloud Services from Anna University, Chennai,India in 2017. She is currently working as Associate Professor in Department of Computer Science and Engineering, NSS College of Engineering, Palakkad, Kerala, India. She has published more than 4 research papers in reputed international journals and conferences including IEEE and it's also available online. Her main research work focuses on Cloud service discovery and Semantic web.