EXPLORING SEMANTIC QUESTION GENERATION METHODOLOGY AND A CASE STUDY FOR ALGORITHMIC QUESTION POOL

Sumayyah Alamoudi and Amany Alnahdi

Department of Computer Science, King Abdulaziz University, Jeddah, KSA

ABSTRACT

Assessment of student performance is one of the most important tasks in the educational process. Thus, formulating questions and creating tests takes the instructor a lot of time and effort. However, the time spent for learning acquisition and on exam preparation could be utilized in better ways. With the technical development in representing and linking data, ontologies have been used in academic fields to represent the terms in a field by defining concepts and categories classifies the subject. Also, the emergence of such methods that represent the data and link it logically contributed to the creation of methods and tools for creating questions. These tools can be used in existing learning systems to provide effective solutions to assist the teacher in creating test questions. This research paper introduces a semantic methodology for automating question generation in the domain of Algorithms. The primary objective of this approach is to support instructors in effectively incorporating automatically generated questions into their instructional practice, thereby enhancing the teaching and learning experience.

KEYWORDS

Ontology-based approach, Automatic question generation, Education, Algorithms, E-learning, assessment

1. INTRODUCTION

The current popularity of online knowledge artifacts and formalisms such as ontologies, are good representations of domain knowledge. The Semantic Web and Natural Language Processing have all proved that ontologies based on Description Logics (DLs) are useful for explicitly transferring information across semantic technologies. Ontology is a graph-based knowledge representation that shows the properties of concepts and the relationships which connect. As a result, these ontologies can be utilized in education since they enable formulation of a specific domain by identifying all relevant concepts, their relationships, and all of the attributes and constraints that exist. Therefore, ontologies have aided the development of knowledge assessment tools [1]. Furthermore, these tools can be used in interactive learning systems to give solution for generating questions automatically and immediate feedback, and greater learner engagement.

Students performance assessment is mandatory in the education process. Traditionally, instructors perform several steps to assess students. It starts from selecting the objectives of the exam, forming the exam's questions and answers, with taking into consideration the difficulty level of the questions to measure different cognitive skills [2]. However, the creation of manual assessments takes up a large portion of an instructor's time. Therefore, it is anticipated that the time spent on exam preparation could be utilized in better ways.

Bibhu Dash et al: NLPAI, IOTBC -2024
pp. 61-44, 2024. IJCI – 2024 DOI:10.5121/ijci.2024.130205

However, the current Learning Management Systems (LMSs) that most universities use such as blackboard provides seamless learning [3]. Also, blueprint and question mark are online assessment tools that aided the blackboard in assessment creation that maps specification of assessment items based on educational [4]. While those LMSs are efficient in the technical pedagogical capabilities to assist the instructor in creating and storing questions and grading Multiple Choice Questions, True/False and short answers. However, current LMSs have limitations in automating content question generation [5].

This paper presents a methodology for automatic creation for questions based on semantic ontologies. The proposed methodology consists of several steps that include Ontology Determination, Ontology Analysis, Stem Creation, Question Generation, and Validation. To illustrate the methodology a case study was applied and presented in the Algorithm domain.

The rest of this paper is presented as follows. Section 2 indicates relevant work. Section 3 presents the methodology. Section 4 presents a case study. The conclusion and future research directions are presented in section 5.

2. RELATED WORK

Ontology is philosophical displine that studies the meaning of existence [6]. Several ontologies have been developed to represent subfields in science. ChemOnto is an implemented chemistry ontology that addresses chemical elements and their interaction [7]. Also, in chemistry, [8] developed OntoKin ontology for capturing data and semantics of chemical kinetic reaction mechanisms. The ontology was created using a formal representation language (OWL) and validated by domain experts. Moreover, the authors of [9], introduced a natural science ontology. In physics, [10] presents the Physics-based Simulation Ontology (PSO) for accurate modeling of physics-based simulations in engineering design. There are few ontologies in the computer science field. For example, researchers from Aljouf university [5] built a computer-related ontology from the NSC document database that contains 520 documents in computer science related fields. Also, the authors [11], presented Pasonto: An Ontology for learning pascal. Additionally, the authors in [12] created an SQL ontology to represent the semantics of databases. Although most studies use monolingual ontologies, researchers in [13], used bilingual ontology in French and English. Ontologies can be used to integrate concepts and present them in different languages. Also, the authors in [14] used Arabic and English to create an ontology for algorithms concepts. The constructed ontology organizes concepts in the Algorithms field and presents them in both Arabic and English languages.

Question generation is a multidisciplinary technique that includes natural language processing and artificial intelligence/knowledge representation. The approaches described in the associated literature may be categorized into different categories such as approaches that depend on Natural Language Processing (NLP) methods generate questions using automated analysis for text [15], while they may also utilize other resources like DBpedia and Wordnet. Second, Template-based approaches can also be used for question generation [16]. Authors of [17] discussed several studies that used predefined patterns/templates for inquiries, filled with certain values from a vast number of permissible options. Finally, semantic-based systems [5], automate examination creation by utilizing semantic web standards, namely ontologies. Semantics-based techniques are thought to have several benefits over natural language resources [15].

Automation of generating questions for practice and self-assessments has recently been a trend in research field. Authors of [18] provided an overview of semantic web applications for automatic

question generation. The study classifies the reviewed studies based on their technological and pedagogical features. An automated question generation based on biology ontology is proposed in [19] to evaluate student's understanding of the biological concepts and the relationship between them. Also, it proposed a way to generate multiple-choice distractors (incorrect alternatives) using the same structure of the ontology.

3. METHODOLOGY

To create questions based on the Algorithm ontology, a systematic approach is proposed. The approach starts by determining the ontology of interest. Then the ontology is to be studied and analyzed for identifying properties used for creating question stems. Following that, generic question stems are to be create based on the analysis for ontology axioms especially properties. Then questions are to be automatically generated based on the stems and axioms associated to these questions. The last step is to validate the automatically generated questions. The steps of the process are depicted in Figure 1.

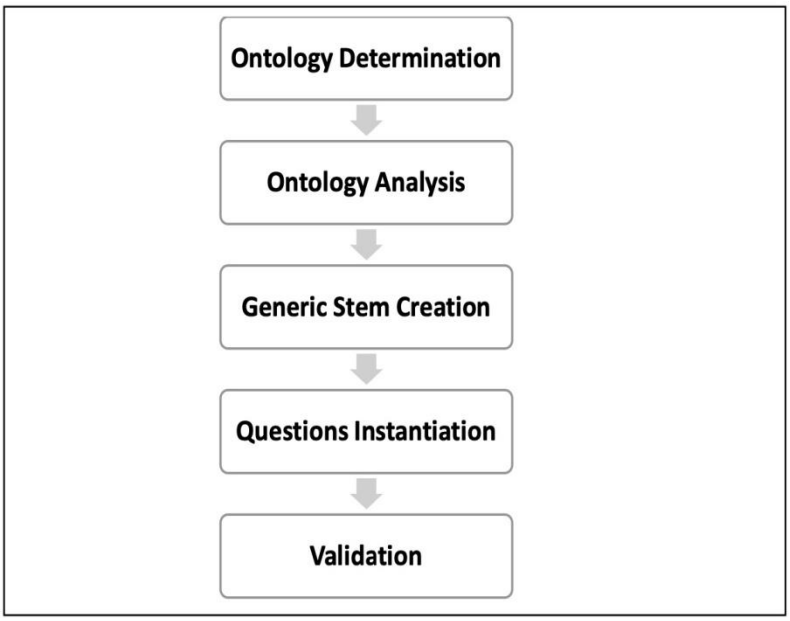

Figure 1 Question generation based on ontologies

The methodology employed in this study involves extracting relevant information from the ontology, generating questions based on the extracted information, and evaluating the generated questions. The ontology is analyzed to extract concepts based on predefined question categories, which are determined by established question models from previous research [16]. Information retrieval from the ontology is performed using SPARQL queries, a query language for extracting data from RDF graphs. The extracted information, in the form of data triplets, serves as input for the question generation algorithm. The question generation algorithm utilizes question templates aligned with different knowledge levels in Bloom's taxonomy. These templates guide the generation of multiple-choice questions with various types, based on the information derived from the ontology. The generated questions are integrated into an interactive learning system, where they are administered to students as assessments. The students' responses are then evaluated using a scoring module, which assigns scores and provides feedback. To validate the generated questions, experts assess their validity, while also comparing the difficulty levels of the questions with the categories defined in Bloom's taxonomy. The agreement among experts

regarding the evaluation outcomes is measured using Kappa calculations, a statistical method for assessing inter-rater agreement. By following this methodology, the study aims to automate the process of question generation by leveraging ontological analysis, question templates, and expert evaluation. The generated questions are designed to align with appropriate difficulty levels and provide effective assessments for students.

5. CASE STUDY

To exemplify the proposed methodology for generating questions automatically, a case study for applying the proposed methodology is presented in this section. For the Ontology Determination step, the ontology used in the case study is Algorithm ontology. The ontology current edit of the ontology [4] consists of 114 classes, 23 instances, and 6 properties. We assume that an expert curated ontology is correct and complete, as opposed to one that is automatically generated. The ontology contains 403 axioms. The analysis of the ontology indicated the that main properties used for stem creation include *has_running_time* and *belongs_to_algorithmic_pardigm*. These properties are illustrated in Figure 2. For example, *has_running_time* used for generating questions enquire about the amortized cost of the algorithm.

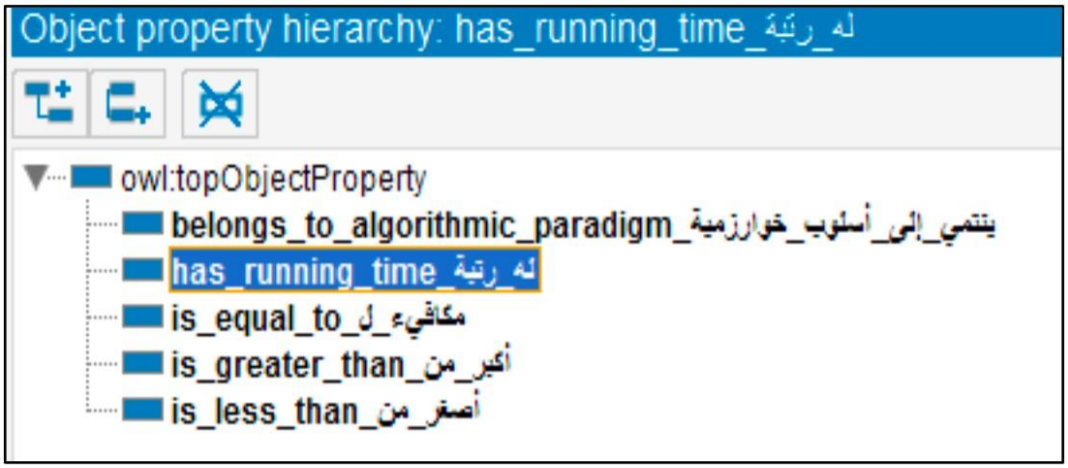

Figure 2 Properties in Algorithms Ontology presented in protege editor

For stem creation step, a generic question stem is composed and the variables that are to be substituted are to be identified. Figure 3 illustrates sample of the generic stems that can instantiated. Ontology libraries are used for parsing the ontology. To implement the automatic question generation, a python ontology library was used. Both closed questions and open-ended questions are automatically created. Sample of the questions generated are depicted in Figure 4.

Is <subclass> a/an superclass?

Does <Algorithm> has <algorithm_running_time> running time?

What is <concept>?

Figure 3 Generic Stem creation

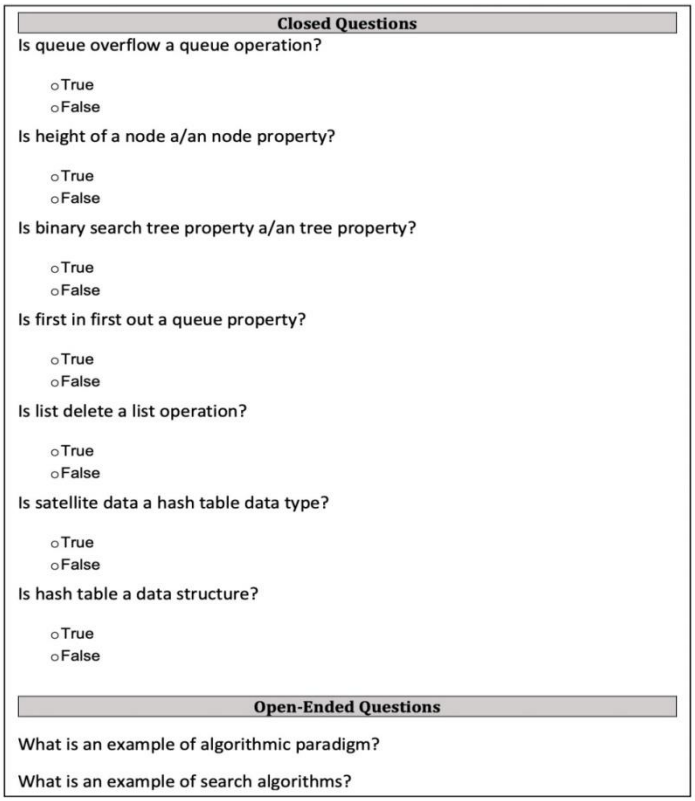

Figure 4 Sample from question pool in the Algorithm domain generated using the proposed ontology-based

6. CONCLUSIONS

This research paper addresses the automatic question generation based semantic ontologies by proposing a systematic approach for question generation. Automatic question generations facilitate the rapidity of creating question pool relevant to academic subjects. Ontologies are semantic representations based on logical relations that can be exploited for automatic question generation. The proposed methodology consists of sequence of steps include Ontology Determination, Ontology Analysis, Stem Creation, Question Generation, and Validation. A case study is also presented for applying the proposed methodology was presented in this paper for automatic question generation in the domain of Algorithms. Additionally, the proposed methodology holds potential for various practical applications. It can be utilized to create flashcards for in-class activities, establish a comprehensive question bank, enable self-assessment opportunities for learners, and be customized for different academic domains. By further exploring these research directions, the field of automated question generation from ontologies can advance, empowering educators with powerful tools to enhance assessment practices, support personalized learning experiences, and foster a deeper understanding of algorithmic concepts among students.

REFERENCES

[1] Marija Cubric and Milorad Tosic (2020). "Design and evaluation of an ontology-based tool for generating multiple-choice questions." Interactive Technology and Smart Education, Vol. 17, No. 2, pp. 155-173.

[2] Nargiza Gaybullaevna Dilova (2021). "Formative assessment of students' knowledge - as a means of improving the quality of education." Scientific reports of Bukhara State University, Vol. 5, No. 3, pp. 144-155.

[3] Sultan Saleh Ahmed Almekhlafy (2020). "Online learning of English language courses via Blackboard at Saudi universities in the era of COVID-19: Perception and use." PSU Research Review.

[4] Aliaa Amr Alamoudi, Lana Adey Al Shawwa, Hoda Gad, and Ara Tekian (2021). "Team-based learning versus traditional didactic lectures in teaching clinical biochemistry at King Abdulaziz University: Learning outcomes and student satisfaction." Biochemistry and Molecular Biology Education.

[5] Saleh Almuayqil, Sameh Abdel-Ghany, and Abdulaziz Shehab (2020). "Towards an ontology-based fully integrated system for student e-assessment." Journal of Theoretical and Applied Information Technology, Vol. 98, No. 21.

[6] Chandrasekaran, B., Josephson, J.R., and Benjamins, V.R. (1999). "What are ontologies, and why do we need them?" IEEE Intelligent Systems and their applications, Vol. 14, No. 1, pp. 20-26.

[7] Baboucar Diatta, Adrien Basse, and Samuel Ouya (2018). "Ontology-based database for chemical experiments: Design and implementation." In International Conference on Interactive Collaborative Learning, pp. 389-397. Springer.

[8] Feroz Farazi, Jethro Akroyd, Sebastian Mosbach, Philipp Buerger, Daniel Nurkowski, Maurin Salamanca, and Markus Kraft (2019). "Ontokin: An ontology for chemical kinetic reaction mechanisms." Journal of Chemical Information and Modeling.

[9] Selvia Ferdiana Kusuma, Daniel O Siahaan, and Chastine Fatichah (2020). "Automatic question generation in education domain based on ontology." In 2020 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM), pp. 251-256. IEEE.

[10] Hyunmin Cheong and Adrian Butscher (2019). "Physics-based simulation ontology: an ontology to support modeling and reuse of data for physics-based simulation." Journal of Engineering Design, Vol. 30, No. 10-12, pp. 655-687.

[11] Baboucar Diatta, Adrien Basse, and Samuel Ouya (2019). "Pasonto: Ontology for learning Pascal programming language." In 2019 IEEE Global Engineering Education Conference (EDUCON), pp. 749-754.

[12] Adrien Basse, Baboucar Diatta, and Samuel Ouya (2019). "Ontology-based system for automatic SQL exercises generation." In Interactive Mobile Communication, Technologies and Learning, pp. 738-749. Springer.

[13] Baboucar Diatta, Adrien Basse, and Samuel Ouya (2019). "Bilingual ontology-based automatic question generation." In 2019 IEEE Global Engineering Education Conference (EDUCON), pp. 679-684. IEEE.

[14] Amany Alnahdi, Rania Aboalela, and Amal Babour (2017). "Building bilingual algorithm EnglishArabic ontology for cognitive applications." In 2017 IEEE International Conference on Cognitive Computing (ICCC), pp. 124-127. IEEE.

[15] Yanghoon Kim, Hwanhee Lee, Joongbo Shin, and Kyomin Jung (2019). "Improving neural question generation using answer separation." In Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 33, pp. 6602-6609.

[16] Noor H. Ibrahim Teo and Mike S. Joy (2018). "Categorized question template generation for ontology-based assessment questions." International Journal of Knowledge Engineering, Vol. 4, No. 2, pp. 72-75.

[17] Ghader Kurdi, Jared Leo, Bijan Parsia, Uli Sattler, and Salam Al-Emari (2020). "A systematic review of automatic question generation for educational purposes." International Journal of Artificial Intelligence in Education, Vol. 30, No. 1, pp. 121-204.

[18] Papasalouros, A., and Chatzigiannakou, M. (2018). "Semantic Web and Question Generation: An Overview of the State of the Art." International Association for Development of the Information Society (IADIS).

[19] Katherine Stasaski and Marti A. Hearst (2017). "Multiple choice question generation utilizing an ontology." In Proceedings of the 12th Workshop on Innovative Use of NLP for Building Educational Applications, pp. 303-312.