Journal of Engineering and Applied Sciences 12 (18): 4693-4696, 2017
ISSN: 1816-949X

Automated Generation of Test Questions from Knowledge-Base

Assel Omarbekova, Alma Zakirova, Bayan Abduraimova and Sailaubek Duzelbayev
L.N. Gumilyov Eurasian National University, Astana City, Republic of Kazakhstan

Abstract: Kazakhstan has a necessity for artificial intelligence application technologies in the development of educational resources and education process. The actual study describes the system development of test questions automated generation in terms of a knowledge-base. This research has an application attitude, provides detailed examples of ontology development and SPARQL-queries in RDF-documents. Additionally, Python programming language based question generating program's implementation is described with involvement of all required libraries required for RDF-files operation.

Key words: Knowledge-base, RDF-documents, SPARQL-queries, intelligent systems, implementation, operation

INTRODUCTION

The purpose of the State Program for Education Development in the Republic of Kazakhstan for 2011-2020 is the increase of educational competitiveness, development of human capital by providing qualitative education for sustainable economic growth. One of the main directions of the program is the e-Learning. The following task is formulated within such direction's implementation program interactive and intellectual digital educational resources will be developed for each subject studied in primary and specializing schools.

Here with, now a days in the republic, there is a necessity for artificial intelligence application in the development of educational resources and education process. At the moment "artificial intellect" Research Studies Institution of L.N. Gumilyov Eurasian National University runs investigations in this field, various subjects related knowledge-base as well as control and training systems in terms of a knowledge-base are being developed.

Since, 2015 the research institution became a member of International Project TEMPUS PROMIS "professional network of master's degrees in informatics as a second competence". The goal of the program is a professional training of universal persons which will combine the skills of their first profession (that which they had obtained during their bachelor course) with theoretical and technical skills in computer science, allowing them to make designs and manage the software tools (Omarbekova *et al*., 2017). The project is being implemented in association with:

- About 5 universities in the European Union (Germany, Lithuania, Poland, Finland, France)
- About 10 universities from the central Asian countries (Kazakhstan, Kyrgyzstan, Uzbekistan
- Tajikistan, Turkmenistan)
- European enterprises

MATERIALS AND METHODS

Knowledge-base development: The systems development capable to infer from the knowledge base are on the rise worldwide. Removing the existing and generation new knowledge and conformation from the existing knowledge base are realized by means of specialized engines designed to discuss and infer. In the study (Baolin *et al*., 2004), the researchers propose to use the RDF for knowledge representation and to use query language for knowledge extraction. By Wang *et al*. (2016), Xie *et al*. (2016), Pah *et al*. (2007) and Buraga and Cioca (2005) the developers describe methods of knowledge base handling, triplet's retrieval, demand assignment.

A task has been set to develop knowledge-bases for all compulsory disciplines within the project TEMPUS PROMIS. At the moment the sauri and knowledge-bases are implemented in the disciplines of "software engineering", "programming", "algorithms and data structures". During the works were composed definitions thesauri, discipline field ontology were composed during the works. Ontology describes basic concepts of the discipline field and determines relationship between them (Nurgazinova *et al*., 2015). The domain ontologies are developed via. protege.

Furthermore, the ontology files have been verified by their validity via. "RDF validator" web service at "www.w3.org". The system deploys triples list and data model's graph. A fragment of generated graph after successful validation of "algorithms and data structures. rdf".

Corresponding Author: Assel Omarbekova, L.N. Gumilyov Eurasian National University, Astana City, Republic of Kazakhstan

Table 1: The examples of questions of designed system

Cognitive process	Examples of questions of designed systems
Knowledge	
Replay	Define the term
Remembering,	What classes they contain?
recognition	Classify the following terms
	To what class they belong?
Understanding	
Interpretation	Continue with definition
Instantiation	Give an example
-	Find an extra item of a series
Classification	Find the items that do not belong to the class
Comparison	To which class belongs items

RESULTS AND DISCUSSION

Test questions generation: The intellectual system which automatically generates test questions in terms of knowledge-base has been developed. The development of test questions for student's knowledge assessment requires no-nonsense approach. The test tasks of high quality are developed thoroughly in a way that they would not be ambiguous or unclear for the tested students. Various methodologies were investigated under test tasks development as well as deeply studied Bloom's taxonomy. In accordance with taxonomy, the student's skills are divided into six categories, knowledge, comprehension, application, analysis, synthesis and assessment. For taxonomic categories belong the following: knowledge (information), understanding, application, analysis, synthesis, materials evaluation and methods together with adopted purposes. The most primitive of the skills are knowledge and understanding and the most advanced: the application; analysis; synthesis evaluation. In the process of question development, it was decided to generate tasks in evaluation of knowledge and understanding as for the question formation of the issues for the application, synthesis, analysis and evaluation it is necessary a system which is able to make reasoning in full. The following are examples of questions of the designed system for the level of "knowledge" and "understanding" (Table 1).

Generation of test questions to estimate student's low level skills. The program implementation starts with installation of IDLE Python 3.5 and "rdflib" library and import into "tkinter" library to develop graphical interface. The result of program is shown on Fig. 1. The system allows to open any file of rdf and owl type. Choose the file computer science.rdf and click "Generate". The questions based on a knowledge base will be shown. Generation of test questions to estimate student's higher level skills (Algorithm 1):

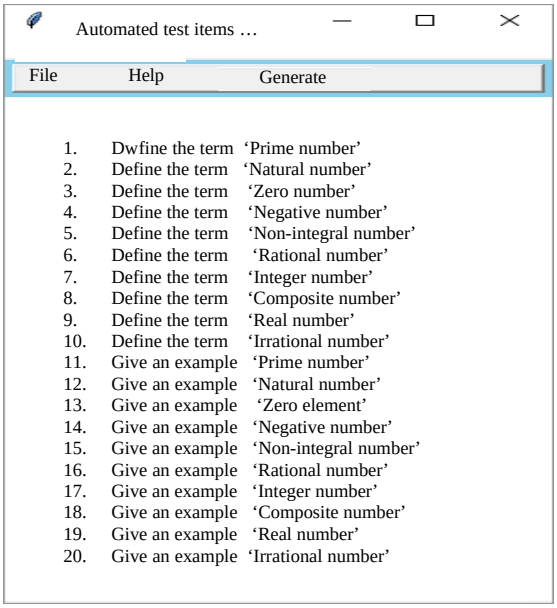

Fig. 1: A form with questions testing the knowledge of terms

Algorithm 1: Test question

```
from tkinter import
import tkinter.messagebox
import rdflib
root = Tk()
def generate():
 g = rdflib.Graph()
 g.parse(file_path)# the path to the file
 # SPARQL inquiray
 qres = g.query("""
 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
 PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
 PREFIX plants: <http: //www. linked data tools. com/plants#>
 PREFIX  dc: <http://purl.org/dc/elements/1.1/#>
 PREFIX owl: <http://www.w3.org/2002/07/owl#>
 SELECT ?name
 WHERE
{?entity rdf:type/rdfs:subClassOf* numbers:integer;
  rdfs:label ?name.} LIMIT 5""")
 qres2 = g.query("""
 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
 PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
 PREFIX plants: <http: // www. linked data tools. com/plants#>
 PREFIX  dc: <http://purl.org/dc/elements/1.1/#>
 PREFIX owl: <http://www.w3.org/2002/07/owl#>
 SELECT ?name
 WHERE
{?entity rdf:type/rdfs:subClassOf* numbers:real;
  rdfs:label ?name.} LIMIT 5""")
 global item_arr
 item_arr = []
 for row in qres:
  items = str(row[0])
  item_arr.append(items)
 global label
 label = Label(frame, text = "I. Classify the following terms. To what class
they belong?" width = "200", height ="2", bg = "grey", fg = "white", font
= ("Helvetica", 11), justify = LEFT, anchor = W)
 label.pack()
```

```
 for item in range(len(item_arr)):
  str_number = str(item+1)
  the_item = str_number+". "+item_arr[item]
   label = Label(frame, text = the_item, width = "200",height = "2",
bg = "white", fg = "grey", font = ("Helvetica", 11), justify = LEFT,
anchor = W)
  label.pack()
 global item_arr2
 item_arr2 = []
 for row in qres2:
  items = str(row[0])
  item_arr2.append(items)
 global label
  label = Label(frame, text = "II. What classes they contain?", width =
"200",height = "2", bg = "grey", fg = "white", font = ("Helvetica", 11),
justify = LEFT, anchor = W)
 label.pack()
 for item in range (len(item_arr2)):
  str_number = str (item+1)
  the_item = str_number+". "+item_arr2[item]
   label = Label (frame, text = the_item, width = "200",height = "2", bg =
"white", fg = "grey", font = ("Helvetica", 11), justify = LEFT, anchor = W)
  label.pack()
def openFile():
 root.fileName = filedialog. askopen file name (file types = (("RDF files",
".rdf"), ("OWL files", ".owl"), ("All files", "*.*")))
 global file_path
 file_path = root.fileName
 return file_path
def quit_program():
 exit()
def show_info():
  tkinter.messagebox.showinfo ('Window title', 'This application generates
low order thinking skills questions')
root.geometry ("600x400+300+200")
root.title ("Automated test items generator")
frame = Frame (root, bg = "sky blue",)
frame.pack()
menu = Menu (root)
root.config (menu = menu)
subMenu = Menu(menu)
menu.add_cascade(label = "File", menu = subMenu)
subMenu.add_command (label = "Open file", command = openFile)
subMenu.add_separator()
subMenu.add_command (label = "Exit", command = quit_program)
helpMenu = Menu (menu)
menu.add_cascade (label = "Help", menu = helpMenu)
helpMenu.add_command (label = "Info", command = show_info)
button1 = Button(frame, text = "GENERATE", fg = "blue", command =
generate)
button1.pack(side = "top", fill = 'both', expand = True, padx = 4, pady =
4) root.mainloop()
```

Here with Fig. 2 shows a program with graphical interface generating test tasks for objects classification. After system's development system errors have not been detected according to the testing results and system was duly generating test tasks from the knowledge-base for different discipline fields. Several RDF file were used under system's testing such formats as OWL and TTL. The rdflib library for Python programming language, involved to process RDF files, during the testing process turned out to be unable to operate with TTL files. However the system duly processed RDF and OWL files with query language SPARQL for RDF documents.

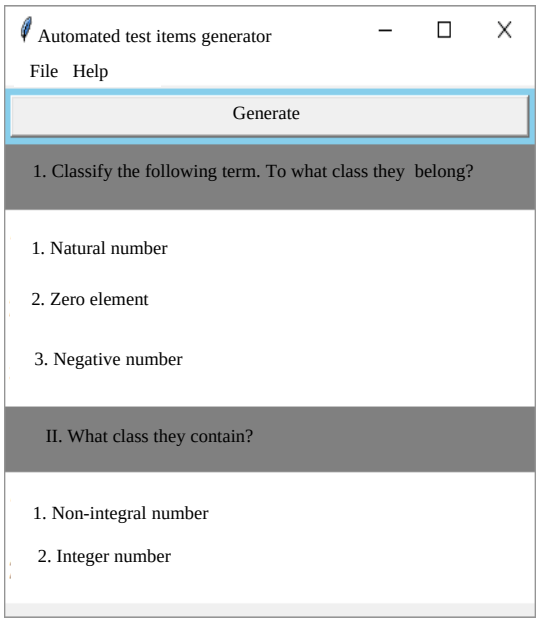

Fig. 2: Program's graphical interface with a question

Trial launches and attempts to create test questions generated in terms of argumentation were made with involvement of owlready library for Python programming language. The mentioned library includes Hermit engine which allows making conclusions related to one or another resource for definite classes.

CONCLUSION

To summarize it all, conclusions can be made in relation to relevancy and development tendency of this direction. Now a days, the whole world runs huge work related to the artificial intelligence knowledge based intellectual systems development. Made investigations in this work give new ideas and open new possibilities for further operation and investigation of knowledge based intellectual systems.

The technology will allow teachers, non-professional programmers to automate the process of formation of known valid test questions to the given questions on a particular subject field. The research results, carried out in this research open up new opportunities of intelligent systems based on knowledge bases for further work and study. As the system for the task generation from the knowledge base is one of the small steps to automation the process of knowledge evaluation of students. Further, developments are aimed at student's response handling and evaluation of answer correctness.

REFERENCES

Baolin, L., W. Xingyan and Z. LI, 2004. A rdf-based knowledgebase implementation for. Proceedings of the ICWL 3rd International Conference on New Horizon in Web-based Learning, August 8-11, 2004, World Scientific, Beijing, China, pp: 45-50.

Buraga, S.C. and M. Cioca, 2005. Modeling aspects of semantic web-based E-learning systems. Proceedings of the 3rd Balkan Region Conference on Engineering Education, September 12-15, 2005, UAIC, Sibiu, Romania, ISBN:973-739-147-0, pp: 210-213.

Nurgazinova, G., V. Golenkov, A. Sharipbay, A. Omarbekova and A. Barlybaev, 2015. Designing of intelligent reference system for algebra based on the knowledge database. Proceedings of the CAAI 2015 International Conference on Control, Automation and Artificial Intelligence, August 23-24, 2015, CAAI, Thailand, Phuket, pp: 1-7.

Omarbekova, A., A. Sharipbay and A. Barlybaev, 2017. Generation of test questions from RDF files using python and sparql. J. Phys. Conf. Ser., 806: 1-7.

Pah, I., I. Maniu, G. Maniu and S. Damian, 2007. A conceptual framework based on ontologies for knowledge management in e-learning systems. Proceedings of the 6th WSEAS International Conference on Education and Educational Technology, November 21-23, 2007, WSEAS, Venice, Italy, pp: 283-286.

Wang, L., Y. Zhang and T. Liu, 2016. A deep learning approach for question answering over knowledge base. Proceedings of the 5th CCF Conference on Natural Language Processing and Chinese Computing and 24th International Conference on Computer Processing of Oriental Languages, December 2-6, 2016, Springer, Kunming, China, pp: 885-892.

Xie, Z., Z. Zeng, G. Zhou and T. He, 2016. Knowledge base question answering based on deep learning models. Proceedings of the 5th CCF Conference on Natural Language Processing and Chinese Computing and 2nd International Conference on Computer Processing of Oriental Languages, December 2-6, 2016, Springer, Kunming, China, pp: 300-311.