**Australian Journal of Basic and Applied Sciences,** 10(14) September 2016, Pages: 46-50

**AUSTRALIAN JOURNAL OF BASIC AND APPLIED SCIENCES**

ISSN:1991-8178 EISSN: 2309-8414
Journal home page: www.ajbasweb.com

# Programming Exam Questions Classification Based On Bloom's Taxonomy Using Grammatical Rules

[1]Dhuha A. Abduljabbar, [1]Ghadah Al-Khafaji, [2]Nazlia Omar

*[1]Department of Computers, College of Science, University of Baghdad, Baghdad, Iraq.*
*[2]Knowledge Technology Group, Center for AI Technology, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia.*

**Address For Correspondence:**
Dhuha Abdulhadi Abduljabbar, Department of Computers, College of Science, University of Baghdad, Baghdad, Iraq.
E-mail: dhuhaftsm@yahoo.com

**ARTICLE INFO**

***Article history:***
*Received 26 July 2016*
*Accepted 21 September 2016*
*Published 30 September 2016*

***Keywords:***
*Exam questions, rule-based approach, POS, parsing.*

**ABSTRACT**

This study describes grammatical rules to classify and analyze written programming exam questions through natural language processing. However, written exam questions have always been a method for educators to assess the level of understanding of students. A good exam question should comprise of various levels of difficulties in order to increase students' thinking skills. Thus, Bloom's Taxonomy, is being used extensively by educators nowadays to frame instructional goals, classify learning assignments, drive instructions and outline evaluations. The uniqueness of the taxonomy framework helped academics to use it as a reference to carry out teaching and learning activities. The primary objective of this study was to devise a tool that would make it easier for lecturers to assess a student's cognitive level according to written examination questions. We employed a natural language processing technique to examine the cognitive levels of Bloom's taxonomy for each question by means of the development grammatical rules. These developed rules facilitated and improved the result of classification model in the programming domain. The results of this study were the point of measuring the extent to which the decision was correct, and overcome the problem of determining the cognitive category of programming questions. In addition, the results from the experiment show that the grammatical rules are a viable approach to help categorize the questions automatically according to Bloom's Taxonomy.

## INTRODUCTION

Generally, classification is an important task, which classifies and assigns an object to a class based on various attributes or rules (Madhumitha&Ilango, 2016). The goal of classification is to accurately predict the target class for each case in the data (Rajeswari*et al*., 2013).Particular, the questions classification is the main area of focus for the purposes of the current study. The purpose of questions classification (QC) is to guess the type of entity for a question, which is written in natural language, and this process was conducted by classifying the question under a category selected from a set of predetermined categories. The set of predefined categories, which are considered as question classes, are usually called question taxonomy. Questions classification is a vital element of question answering systems and a significant amount of research has been conducted with regard to it over the past ten years (Loni, 2011).

In 1956, Benjamin Bloom, together with a team of educators, came up with a system for the classification of educational objectives, and their findings were published as the Taxonomy of Educational Objectives: Book 1, Cognitive Domain. Nearly more than half a century since its first publication, the handbook is now available in more than twenty languages. These findings, known universally as Bloom's Taxonomy, is being used



**ToCite ThisArticle:**Dhuha A. Abduljabbar Ghadah Al-Khafaji, Nazlia Omar ., Programming Exam Questions Classification Based On Bloom's Taxonomy Using Grammatical Rules. ***Aust. J. Basic & Appl. Sci.,*** *10(14): 46-50, 2016*

extensively by educators nowadays to frame instructional goals, classify learning assignments, drive instructions and outline evaluations (Almerico& Baker 2004; Chang & Chung, 2009). Six ranking positions of cognitive level, ranging from the least to the most complex i.e. *Knowledge, Comprehension, Application, Analysis, Synthesis* and *Evaluation*, are defined in Bloom's Taxonomy in order to induce students to "climb to a higher level, or step, of thought".

Through this research, a rule-based approach was adopted for determining the category of an examination question based on Bloom's Taxonomy cognitive level. In general, question classification by using rule-based approach depends on using rules determined manually by knowledgeable engineers, with the help of domain experts (Panicker *et al*., 2012). The rule-based approach tries to match the questions with some manually handcrafted rules. This approach however, suffers from the need to define too many rules (Rahman, 2015; Li & Roth, 2006). Furthermore, rule-based approach performs well on a particular dataset. The rules that were present in the rules database would attempt to identify the syntactic structure and match it with the appropriate and similar form. Once the appropriate rules had been found, a special category would be allocated to the particular question with the identified syntactic structure. The rule-based text classification has also seen a lot of research, since researchers can easily create rules based on a particular aim, such as classifying any documents that contain any information on medicine into a medically-related class. Costagliola and Fuccella (2009), Hermjakob (2001) and Duch, (2011) agreed that rules are built according to identified situations.

Thus this study developed rules, by using Natural Language Processing, which consists of combination of tagging, parsing and extracting keywords. These developed rules facilitated and improved the result of classification of exam questions to the cognitive level of Bloom's Taxonomy in the programming domain.

## MATERIALS AND METHODS

The main aim of this paper is to develop rules for classifying exam questions based on the grammatical structure of the questions. The rules were developed from a training set of examination questions in programming subjects. There were two reasons for applying the rules (Omar *et al*., 2012):

• The rules will distinguish the suitable keywords for each question depending on its category.

• The rules will help choose the correct category if a keyword shares more than one category. For example, List may fall under the Knowledge or Analysis category.

To classify questions using the rule-based approach, several rules were identified by utilising syntactic patterns from questions. The rule-based approach consists of several phases. The overall process of the rule-based approach in determining the Bloom's Taxonomy category of a given question was explicitly described in the following.

### 1. *Data Set Planning:*

The empirical evaluation of the question classification was implemented on the data set of the programming exam questions in C++ and Java. This data set was previously used by Haris(2013) and research experiment was performed on it. Programming exam questions from the data set were collected from bank exam FTSM/UKM.

### 2. *NLP-Processing Phase:*

In this step, the questions are pre-processed to be transformed into a simpler form and act as an input to the rule-based system. The processing phase begins with:

• POS tagging: this module is responsible for tagging words in a given question with their part-of-speech so that each word is annotated with its part-of-speech (POS) tag which is a grammatical tag; e.g., verb, noun or adjective. The Stanford tagger was used for identifying the POS tags for question words. Figure 1 displayed the output obtained from the Stanford tagger on the sample question "Explain the structure of a method in the program?" from the comprehension category.

Explain/VB, the/DT, structure/NN, of/IN, a/DT, method/NN, in/IN, the/DT, program/NN, ?/.

**Fig. 1:**The POS output obtained from the Stanford tagger

Where VB is Verb ( base form), DT is Determiner, NN is Noun (singular or mass), NNS is Noun-plural, IN is Preposition or subordinating conjunction, JJ is Adjective, TO is "to", and CC is Coordinating conjunction.

- Shallow parsing: this module divides the sentence and groups its words together based on their POS tags to make more meaningful phrases. It identifies noun phrases, verb phrases and simple adjectival and adverbial phrases. Shallow parsers represent the task of recovering only a partial amount of syntactic information to identify phrases from natural language sentences. On the other hand, partial parsing can be much faster, more robust and sufficient for many natural language processing applications (Patrick, 2009). To generate classification patterns, shallow parsing was exhaustively applied to training questions. Figure 2 showed the shallow parsing output of the same given question above obtained from the Stanford parser.

(VP (VB Explain))
(NP (DT the) (NN structure))
(NP (DT a) (NN method))
(NP (DT the) (NN program))

**Fig. 2:** Shallow parsing for question sample from comprehension obtained by Stanford parser
Where **VP** is Verb Phrase, **NP** is Noun Phrase, **ADVP** is Adverb Phrase, and **PP** is Prepositional Phrase.

### *3. Developing Rules:*

Beside the rules designed by (Haris, 2013) to classify questions into their categories, this work designed new rules by utilising keywords and syntactic structures of each question. As defined in (Haris& Omar, 2012), each Bloom's cognitive level has its keywords that exemplify the level and represent intellectual activity. Table 1 presented the samples of keywords from each Bloom's cognitive level. Table 2 showed Sample of rules for each cognitive level.

**Table 1:** Examples of keywords from Bloom's cognitive level

| Bloom's cognitive level | Examples of Keywords |
|---|---|
| Knowledge | Relate, recall, repeat, reproduce, state, tell, mean, describe. |
| Comprehension | State, cite, compare, extend, generalize, gives Examples, distinguish, brief |
| Application | Solve, use, write, predict, discover, show, apply, choose, prepare, define. |
| Analysis | Experiment, list, assume, outlines, diagram, deconstruct, and differentiate. |
| Synthesis | Create, design, develop, plan, set up, rewrite, write, and improve. |
| Evaluation | Predict, rate, select, support, validate, evaluate, explain, justify, value. |

The following examples demonstrated how the patterns and rules were developed where POS tagging and shallow parsing were applied.

***Question:***
*Explain the structure of a method in the program?*

***POS tagging:***
*Explain/VB the/DT structure/NN of/IN a/DT method/NN in/IN the/DT program/NN ?/. .*

***Shallow Parsing:***
*(VP (VB Explain)) (NP (DT the) (NN structure)) (NP (DT a) (NN method)) (NP (DT the) (NN program))*
Keyword: *Explain*

***Rule**:*
*{< Comprehension keyword> (<VB>) + <NP1> +<IN> +<NP2>+<IN> +<NP3>}*

**Table 2:** Sample of rules for cognitive level of BT from programming questions

| Category | Rule | Question Example |
|---|---|---|
| Knowledge | {(VB) (Knowledge Keyword) + [CD + NN + NNS) \| (NN)$^{+}$] + (PP)?} | Define Inheritance concept. |
| Comprehension | {(VB) (Comprehension Keyword) + (DT)? + [IN + NP]$^{+}$ + (WP + NP + VB)?} | Show each of the passes of the sorting phase. |
| Application | {(VB) (Application Keyword) + (IN \| RP)? + (NP) + [IN + NP]$^{+}$ } | Define an array totalsalebyquarter with size 6 of type double. |
| Analysis | {(VB) (Analysis Keyword) + (NP) + [IN + NP]$^{*}$ } | Trace the contents of matrix from the following statements. |
| Synthesis | {(VB) (Synthesis Keyword) + (NP) + [(TO + VB) \| (IN + VBG)] + (NP) + ( IN + NP)$^{*}$ } | Write a java program to show the overloading. |
| Evaluation | {(VB) (Evaluation Keyword) + (NP) + (IN) + (NNP \| NN) + [(CC + NNP) \| (CC + VB)] + (NP)?} | Justify the concept of inheritance and give a sample code. |

## EVALUATION MEASURES

The performance of rule-based approach for each particular class *c* was evaluated using popular measures for evaluating classification systems, which are, precision, recall and f-measure ($F_\beta$) defined in the following:

$$\text{Recall} = \frac{\text{No. of correctly classified questions of category c}}{\text{No. of questions of category c}} = \frac{tp}{tp + fn}$$

$$\text{Precision} = \frac{\text{No. of correctly classified questions of category c}}{\text{No. of predicted questions of category c}} = \frac{tp}{tp + fp}$$

$$F_\beta = \frac{2 * (\text{precision} * \text{recall})}{(\text{precision} + \text{recall})}$$

Where, True Positive (TP) is the set of question that is correctly allocated to the specified category, False Positive (FP) is the set of questions that incorrectly allocated to the category, False Negative (FN) is the set of questions that is incorrectly not allocated to the category and True Negative (TN) is the set of the set of questions correctly not allocated to the category.

## RESULTS AND DISCUSSION

This section aimed to test the proposed model and the review of the experimental results that were achieved using grammatical rules for exam question classification. The main aim of this model was to develop question classification system based on Bloom’s Taxonomy to define the cognitive level of question structure. In this proposed model, the programming exam question was classified based on Bloom’s Taxonomy cognitive level with several rules. The results of the experiment were recorded; outperforming the previous study that used the same data set. The results were shown in Table 3.

**Table 3:** Results of Bloom’s Taxonomy Cognitive Level using Rule-based Approach

| Cognitive Level | Recall | Precision | $F_1$ – Measure |
|---|---|---|---|
| Knowledge | 1 | 0.5 | 0.67 |
| Comprehension | 1 | 0.83 | 0.91 |
| Application | 0.8 | 1 | 0.89 |
| Analysis | 0.83 | 1 | 0.91 |
| Synthesis | 1 | 1 | 1 |
| Evaluation | 0.75 | 1 | 0.86 |
| Average | 0.9 | 0.89 | 0.87 |

According to the experiments in table 3, the highest result yields in synthesis, comprehension and analysis respectively, because the rules patterns covered the majority of the types of questions in these classes. And worst result yields in Knowledge class due to the appearance of vague terms in some of the classifications.

However, better results were obtained in those classes that had terms or keywords that were not found in other classes. However, poor results in some classes can be attributed to the similarity of terms for each class and occurring vague terms in some of the classification due to the relatively short length of questions.

### *Conclusion:*

With this system, instructors do not have to worry about the level of difficulties of questions. Not only does this system offer assistance for instructors from a once exhausting task, faster exam generation would further help instructors to allocate their time for other educational tasks. Nonetheless, there are more activities needed to enlarge and enrich this work. In future, studies should work to improve the model by applying other algorithms.

## REFERENCES

Almerico, G.M. and R.K. Baker, 2004. Bloom’s Taxonomy Illustrative Verbs: Developing a Comprehensive List for Educator Use. Florida Association of Teacher Educators Journal., 1(4): 1-10.

Chang, W.-C. and M.-S., Chung, 2009. Automatic applying Bloom's taxonomy to classify and analysis the cognition level of English question items. *Paper presented at the Pervasive Computing (JCPC), 2009 Joint Conferences on*.

Costagliola, G. and V. Fuccella, 2009.A Rule-Based System for Test Quality Improvement.International Journal of Distance Education Technologies,7(2): 63-82.

Duch, W., 2011.Rule-Based Methods.Encyclopedia of System Biology. www.fizyka.umk.pl/publications/kmk/11-Rule%20Discovery.pdf. [1 April 2012].

Haris, S.S., 2013. Questions Classification Based on Bloom’s Taxonomy using Rule-Based and N-Gram Approach.Master's Thesis.UniversitiKebangsaan Malaysia.

Haris, S.S. and N. Omar, 2012. A Rule-Based Approach In Bloom's Taxonomy Question Classification Through Natural Language Processing, *3rd International Conference on Advancements in Computing Technology (ICACT 2012)-, Seoul, Korea 3-5 Dec 2012 Art. No. 6530368* , pp: 410-414.

Omar, N., S. S. Haris, R. Hassan, H. Arshad, M. Rahmat, N. F. A. Zainal, &, R. Zulkifli, 2012. Automated analysis of exam questions according to Bloom's taxonomy. Procedia-Social and Behavioral Sciences, 59, 297-303.

Hermjakob, U., 2001. Parsing and Question Classification for Question Answering.*Proceeding of the Workshop on Open-Domain Question Answering at ACL*-2001. pp: 16.

Li, X., and D. Roth, 2006. Learning question classifiers: the role of semantic information. Natural Language Engineering, 12(03): 229-249.

Loni, B., 2011. A survey of state-of-the-art methods on question classification.*Literature Survey, Published on TU Delft Repository.*

Madhumitha R. and K. Dr. Ilango, 2016. Decision Tree based Classification Type Question/Answer E-Assessment System. Advances in Natural and Applied Sciences, 10(1): 22-25.

Panicker, A.D., U. Athira and S. Venkitakrishnan, 2012.Question Classification using Machine Learning Approaches.International Journal of Computer Applications, 48(13): 1-4.

Patrick, Y., 2009. Natural language understanding in controlled virtual environments. University Library.

Rahman, T.A., 2015. Question Classification Using Statistical Approach: A Complete Review. Journal of Theoretical and Applied Information Technology, 71(3).

Rajeswari, K., P. Kumari, and Dr.Vaithiyanathan, 2013. Selection of Significant Features using Decision Tree Classifiers, Journal of Applied Sciences Research, 9(4): 2566-2569.