Automatic Generation of Question Bank Based on Pre-defined Templates

Ahmed Ezz Awad
Faculty of Science, King Abdulaziz University, Jeddah, KSA
ahmedezzaldinahmed@gmail.com

Mohamed Yehia Dahab
Faculty of Computer Sciences and Information Technology, King Abdulaziz University, Jeddah, KSA.
MDahab@kau.edu.sa

ABSTRACT

The preparation of question bank is a difficult and time consuming. This paper explains an algorithm that provides solution to automatic generation of question bank based on a set of pre-defined templates. All possible questions are generated by parameterized concepts from a set of pre-defined templates. The generated questions cover all selected topics in all level of difficulties the form of a multiple-choice question (MCQ). The outputs can be used for both examinations and asynchronous training. If the output used in asynchronous training, it augmented with a set of brief explanation. A successful automatic generation of question bank has been developed for a biology course (Bio110) in KING ABDULAZIZ UNIVERSITY (KAU).

Keywords: e-learning, question bank, meta knowledge, learning concepts, MCQ

1. INTRODUCTION

According to Bloom taxonomy [2], the paper covers the following level of knowledge: apply, understand and remember. This paper explains an algorithm that provides solution to automatic generation of question bank based on a set of pre-defined templates. The main idea behind the applied algorithm is to provide a large number of questions in question bank. The main feature of question bank is to cover all mentioned levels of knowledge for all topics in a course at different levels of difficulty. This feature enables us to provide huge number of questions in online asynchronous training course namely biology\(^1\). Also, this feature enables us to provide dissimilar paper exams in the same level.

All possible questions are generated by parameterized concepts from a set of pre-defined templates.

If the output used in asynchronous training, it augmented with a set of brief explanation.

The remainder of this paper is organized as follows: Section 2 explains the system architecture and demonstrates the algorithms. In Section 3, the evaluation. Section 4 concludes the paper with a summary.

2. Related work

In this section, we review related work on question generation. Some methods utilize ontologies, knowledge bases, and lexicons as sources for generation and other utilize unstructured text and NLP techniques.

There are numerous applications for automated test generation, including for example Random Test Generator-PRO [9] or Test Maestro II [10]. These applications as well as the majority of learning management systems employ tests generated randomly from the pool of questions in the question bank.

ExtrAns system uses a combination of robust natural

\(^1\) http://prod.kau.edu.sa/faculties/science/bio/index.aspx
language processing technology and dedicated terminology processing to create a domain-specific Knowledge Base, containing a semantic representation for the propositional content of the documents. Knowing what forms the terminology of a domain and understanding the relation between the terms is vital for the answer extraction task [5].

The NLQA system is a rule–based one that uses natural language interface. The system can be easily implemented for any domain. It can be used for building rapid language comprehension applications specific to a particular domain by using the appropriate heuristics / keywords of that domain. The NLQA is fitted with the vocabulary of the domain, its rule base file, its hierarchy, etc. created by the domain expert for a specific subject area. The framework is unaltered for any domain in which it has to work. The framework's keyword-based approach along with the shell-'information base' structure is based on principles used in ELIZA [4].

The PICO system is a Clinical question answering system which contains Knowledge extraction modules that serve as building blocks for a clinical question answering system. The PICO frame describes the structure of a well-built clinical query, and can serve as the core organizing knowledge structure of a question answering system [7].

The OntoAware system has a set of tools based on semantic technologies, knowledge representation and processing that are useful for learning content authoring, management and delivery. This authoring environment allows a semi-automatic generation of the learning objects (standard e-learning and courseware elements) by using existing ontologies or even creating a new one from scratch [5].

Another approach to generate multiple-choice questions in automatic way starts from domain specific ontologies and it is independent of lexicons [3].

Recently, [8] described a system that can automatically generate factual questions. He also discussed some of the computational and linguistic challenges related to extracting factual questions from text such as the limited ability to generate deep questions and difficulty in associating questions with their appropriate answers.

Other systems do not require any ontology or WordNet. Such as: [6] Question bank has been generated based on parse annotated corpus and [1] uses the Wikipedia, as the source of information. The later technique uses NLP to generate named entity and distractors.

3. SYSTEM ARCHITECTURE

The proposed system contains three subsystems: Knowledge Descriptor, Questions Generator, and E-learning Executer. The Knowledge Descriptor subsystem allows the instructor to describe the learning contents. The Questions Generator subsystem receives the learning contents and generates the corresponding multiple questions. The E-learning Executer subsystem uses and allows the students to use the generated questions in education process. Error! Reference source not found., shows the system architecture. This paper describes the knowledge descriptor and Questions Generator sub systems.

![System Architecture](image)

**Figure 1: System Architecture**

The role of knowledge descriptor is to store the questions representation forms. The question generator handles the different representations of the questions and then generates the verities of each question. The role of E-learning executer is to run the generated questions bank in internet site.

Three types of questions: Concept Definition, Correct or wrong options, and Problem. The main purpose of Concept Definition of question is to measure how the students understand the concepts and its terminologies. The main purpose of Correct or wrong options of is to measure how the students can recognize the similarity and differently between terminologies. The Problem type is a complicated and it is varied. The bottleneck of this type is how to define the solution and the verities options.

3.1 Correct and Wrong Options Question
In this type of question there are three parts: question text, correct options, and wrong options. The item of Correct or wrong options question can be defined as:

\[(\text{question text}, \text{CO}, \text{WO})\]

Where:

**Question text:** represents the head of question

**CO:** Correct Option, is a finite set of correct options.

**WO:** Wrong Option, distracters, is a finite set of wrong options.

Length of correct option list = CL

Length of wrong option list = WL

The general form of the question is:

<table>
<thead>
<tr>
<th>Question Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>A) term1</td>
</tr>
<tr>
<td>c) term3</td>
</tr>
</tbody>
</table>

Where term1, term2, term3, and term4 \(\in (\text{CO} \cup \text{WO})\) and <standard word> is one of three words: “all of the above”, “both first and second options”, “none of the above”.

In case of term4 equal “all of the above”, term1, term2, and term3 belong to CO. In case of term4 equal “none of the above”, term1, term2, and term3 belong to WO. In case of term4 equal “both first and second options”, term1, and term2 belong to CO, and term3 belongs to WO. Otherwise term1 belong to CO, and term2 and term3 belong to WO.

Figure 2 shows the shows CWOQ algorithm that is used in questions generation of Correct and Wrong Options Question(CWOQ).

The algorithm uses four sub algorithms: *onlyOneSolution*, *twoSolutions*, *allOfTheAboveSolutions*, and *noneOfTheAboveSolutions* algorithms.

In addition to the correct options and wrong options, the algorithms use list of reserved words that is used in questions generations process. The words are: “First and second option”, “All of the above”, and “None of the above”.

The “**onlyOneSolution**” algorithm uses one correct answer and three wrong options or one correct answer, two wrong options. And one of reserved words. Figure 3 shows the algorithm. The number of generated questions of “onlyOneSolution” is

\[\text{CL} \times \text{WL}^3 + 3 \times \text{CL} \times \text{WL}^2\]

**Algorithm: onlyOneSolution**

1. Put this item as the first option
   1.1 Get all combinations of three items from WO.
   
   For each combination construct question
   1.2 Get all combinations of two items from WO and one word from reserved words
   
   For each combination construct question

The “**twoSolutions**” algorithm uses the word: “First and second option”, two correct options, and one wrong option. Figure 3 Shows the algorithm. The number generated questions of “twoSolutions” algorithm is:
The “allOfTheAboveSolutions” algorithm uses the word: “all of the above”, and three correct options. Figure 5 shows the algorithm. The number of generated questions of “all of the above solutions” is $\binom{CL}{3}$.

Algorithm: allOfTheAboveSolutions
1- Get all combinations of three items from CO
2- For each combination do
   2.1 Put “all of the above” as fourth option.
   2.2 construct question

Figure 5: “allOfTheAboveSolutions” algorithm

The total generated questions = $CL \times \binom{WL}{3} + 3 \times CL \times \binom{WL}{3} + WL \times \binom{CL}{2} + \binom{CL}{3} + \binom{WL}{3}$

Example of Correct and Wrong Options Question

In subject Bioenergetics, there are some properties of Enzymes. Table 1 shows this example; the question text is ask about the enzyme: Enzyme _____. There are four correct answers and five wrong answers, so $CL = 4$ and $WL = 5$. The result of applying the CWOQ algorithm is the generating of 270 questions.

Table 2 shows samples of the generated questions.

Table 1: CWOQ example

<table>
<thead>
<tr>
<th>Questions: Enzyme</th>
<th>Correct answer list</th>
<th>Wrong answers list</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) speeds up the cell chemical reactions</td>
<td>1) is consumed in the reaction</td>
<td></td>
</tr>
<tr>
<td>2) has a particular target molecule called the substrate</td>
<td>2) is not specific for substrate</td>
<td></td>
</tr>
<tr>
<td>3) is a protein</td>
<td>3) is a carbohydrate</td>
<td></td>
</tr>
<tr>
<td>4) is specific for substrate</td>
<td>4) is a lipid</td>
<td></td>
</tr>
<tr>
<td>5) slows down the cell chemical reactions</td>
<td>5) is a carbohydrate</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: samples of CWOQ generated questions

<table>
<thead>
<tr>
<th>Question</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enzyme _____</td>
<td>Only one solution uses three wrong options</td>
</tr>
<tr>
<td>A) speeds up the cell chemical reactions</td>
<td></td>
</tr>
<tr>
<td>B) is consumed in the reaction</td>
<td></td>
</tr>
<tr>
<td>C) is not specific for substrate</td>
<td></td>
</tr>
<tr>
<td>D) is a carbohydrate</td>
<td></td>
</tr>
</tbody>
</table>
Enzyme _______  
A) speeds up the cell chemical reactions  
B) is consumed in the reaction  
C) is not specific for substrate  
D) First and second choice

Only one solution uses two wrong options and first and second word

Enzyme _______  
A) speeds up the cell chemical reactions  
B) is consumed in the reaction  
C) is not specific for substrate  
D) all of the above

Only one solution uses two wrong options and all of the above.

Enzyme _______  
A) speeds up the cell chemical reactions  
B) has a particular target molecule called the substrate  
C) is a libid  
D) First and second choice

Two solutions

Enzyme _______  
A) has a particular target molecule called the substrate  
B) is a protein  
C) is specific for substrate  
D) all of the above

All of the above solution

Enzyme _______  
A) is a carbohydrate  
B) is a libid  
C) slows down the cell chemical reactions  
D) none of the above

None of the above solution.

3.2 Concept Definition (CD) Question

The item of CD can be defined as:

CDQ = (Concept Name, [question text], CT, TD, UT)

Where:
CT: Concept’s Terms, is a finite set that contains the terms names of the concept
TD: Term Definition, is a finite set that contains the definition of the terms
UT: unrelated Terms, is a finite set that contains the unrelated terms (distractors).

Each element of the concept’s terms list is mapping to item of the terms definition list. Length of concept’s terms list = Length of terms definition list = TL. Length of unrelated list = OL. We can construct three forms of multiple questions from the CD: "is a concept" question, "terms definitions" question, and "determine term" questions.

Construct "is a concept" question

The purpose of this question is to measure whether the student knows the terms of each concept or not. The construction of “is a concept” question process is similar to the Correct and Wrong Options Question. The question text set to be ____ is a concept. The correct options is setting by the concept terms (CT), and the wrong options is setting by the unrelated Terms (UT). The process uses the same four algorithms; Figure 7 shows “is a concept” algorithm.

```
Head of Question: ___ is <Definition> Options
a) Term1  b) Term2  c) Term3  
d) Term4 or “none of the above”
```

Algorithm: "is a concept" Algorithm

1 -question text ← “____ is a <concept name>”.
2- CO ← CT , WO ← UT
3- call onlyOneSolution
4- call twoSolutions
5- call allOfTheAboveSolutions
6- call noneOfTheAboveSolutions

Figure 7: “is a concept” Algorithm

Construct "terms definitions" question
The "term definition" questions ask the student about the definition of term, so the construction of these questions uses the two lists CT, and TD. Also we use the option "none of the above" only. "term definition" question can be defined as:

**Head of Question:** <term name> is

**Options**

- a) Define1
- b) Define2
- c) Define3
- d) Define4 or "none of the above"

Where <term name> CT, Define1, Define2, Define3, and Define 4 TDs. Where each term has only one definition, so there no need of “all of the above solutions” and “two Solutions” options; Figure 8 shows “terms definitions” Algorithm.

**Algorithm:** "terms definitions"

1- for each term of CT do
   1.5 CO ← get definition of term from TD
   1.6 KO ← other definitions
   1.7 call onlyOneSlioution
   1.8 call noneOfTheAboveSolutions

Figure 8: "terms definitions" Algorithm

**Construct "determine term" question**

The question: "determine term" ask the student to get the term of the given definition. So the construction of these questions uses the two lists CT, and TD. Also we use the option "none of the above". The question: "determine term" can be defined as:

Where <Definition> belongs to TD, Term1, Term2, Term3, and Term4 belong to CT.

Where definition is refereed to only one item so there no need of “all of the above solutions” and “two Solutions” options; Figure 9 shows "determine term" Algorithm.

**Table 3:** "is a concept" question example

<table>
<thead>
<tr>
<th>Concept:</th>
<th>Exceptions to Mendel Laws</th>
</tr>
</thead>
<tbody>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>Incomplete dominance</td>
<td>Heterozygote has intermediate phenotype</td>
</tr>
<tr>
<td>Codominance</td>
<td>Heterozygote expresses phenotypes of both homozygotes.</td>
</tr>
<tr>
<td>Multiple alleles</td>
<td>Three or more alleles in a population for the same locus.</td>
</tr>
<tr>
<td>Pleiotropy</td>
<td>The phenomenon of one gene mutation being responsible for or affecting more than one phenotypic characteristic.</td>
</tr>
<tr>
<td>Polygenes</td>
<td>Multiple independent pairs of genes may have similar and additive effects on the phenotype.</td>
</tr>
</tbody>
</table>

| Unrelated terms (distractors): dominance recessiveness |

Table 4 shows samples of the generated questions.

**Table 4:** Sample of generated questions

<table>
<thead>
<tr>
<th>Question</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Which of the following is an exceptions to Mendel Laws</td>
<td></td>
</tr>
<tr>
<td>A) Incomplete dominance</td>
<td></td>
</tr>
<tr>
<td>B) dominance</td>
<td></td>
</tr>
<tr>
<td>C) recessiveness</td>
<td></td>
</tr>
<tr>
<td>D) all of the above</td>
<td></td>
</tr>
<tr>
<td>Which of the following is an exceptions to Mendel Laws</td>
<td></td>
</tr>
<tr>
<td>A) Pleiotropy</td>
<td></td>
</tr>
<tr>
<td>B) dominance</td>
<td></td>
</tr>
<tr>
<td>C) recessiveness</td>
<td></td>
</tr>
<tr>
<td>D) none of the above</td>
<td></td>
</tr>
<tr>
<td>two wrong options and all of the above</td>
<td></td>
</tr>
<tr>
<td>two wrong options and none of the above.</td>
<td></td>
</tr>
</tbody>
</table>
Which of the following is an exceptions to Mendels Laws
A) Pleiotropy
B) dominance
C) recessiveness
D) First and second choice

Which of the following is an exceptions to Mendels Laws
A) Incomplete dominance
B) Codominance
C) dominance
D) First and second choice

36) Which of the following is an exceptions to Mendels Laws
A) Incomplete dominance
B) Codominance
C) Multiple alleles
D) all of the above

Incomplete dominance is referred to ___
A) Three or more alleles in a population for the same locus.
B) Heterozygote has intermediate phenotype
C) Heterozygote expresses phenotypes of both homozygotes
D) The phenomenon of one gene mutation being responsible for or affecting more than one phenotypic characteristic

__ is referred to as Multiple independent pairs of genes may have similar and additive effects on the phenotype.
A) Polygenes
B) Incomplete dominance
C) Multiple alleles
D) First and second choice

4. EXPERIMENTAL RESULTS

A question bank is developed for course General Biology course (Bio110) at Faculty of Science, King Abdulaziz University (KAU): Table 5 shows the summary of the questions bank of Bio110. The bank contains 12 chapters with 239 sub topics and total of 46345 questions.

Table 5: Questions bank of Biology course at KAU

<table>
<thead>
<tr>
<th>Section Number</th>
<th>Chapter Name</th>
<th>subtopics</th>
<th>Generated questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Exploring Life</td>
<td>9</td>
<td>5096</td>
</tr>
<tr>
<td>2</td>
<td>The Chemical Basis Of Life</td>
<td>11</td>
<td>1233</td>
</tr>
<tr>
<td>3</td>
<td>The Molecules Of Cell</td>
<td>14</td>
<td>1915</td>
</tr>
<tr>
<td>4</td>
<td>Cells</td>
<td>30</td>
<td>10212</td>
</tr>
<tr>
<td>5</td>
<td>Tissues</td>
<td>16</td>
<td>454</td>
</tr>
<tr>
<td>6</td>
<td>Bioenergetics</td>
<td>35</td>
<td>3962</td>
</tr>
<tr>
<td>7</td>
<td>Biodiversity of Life</td>
<td>22</td>
<td>2700</td>
</tr>
<tr>
<td>8</td>
<td>Nutrition</td>
<td>22</td>
<td>5753</td>
</tr>
<tr>
<td>9</td>
<td>GAS exchange</td>
<td>22</td>
<td>8515</td>
</tr>
<tr>
<td>10</td>
<td>Excretion</td>
<td>20</td>
<td>1007</td>
</tr>
<tr>
<td>11</td>
<td>Reproduction and Embryonic Development</td>
<td>14</td>
<td>611</td>
</tr>
<tr>
<td>12</td>
<td>Genetics</td>
<td>24</td>
<td>4887</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>239</strong></td>
<td><strong>46345</strong></td>
</tr>
</tbody>
</table>

A training system for Bio110 based on the developed questions bank is published; Figure 10 shows the site of Bio110 training system.

Figure 10: The site of Bio110 training system

The system allows the student to solve the training of each chapter and then display his answer and the correct answer; Figure 11 shows the correctness the student answers.
Figure 11: The correctness the student answers

5. CONCLUSION

A systematic template and its algorithms are designed for automatically generated a questions bank. A successful automatic generation of question banks is developed for General Biology course (Bio110), Faculty of Science, KAU. This approach overcomes the problems if manual methods that ask the instructor to enter questions directly. With the template, the instructor describes the question using template of the system, and then the system generates the questions. In case of any mistake of the generated questions, the instructor modifies the description and the system re-generates the questions.

We assumed that all questions generated from a specific template have the same level of difficulty. This assumption need to be approved by statistical analysis.

Also the research assumed that all possible questions are covered.

6. REFERENCES


Ahmed Ezz Awad, Mohamed Yehia Dahab