An Overview of Automatic Question Generation Systems

Ms. Deepshree S. Vibhandik, Prof. Mrs. Rucha C. Samant

***Abstract*: Question Generation from text is Natural Language Processing task. It involves a joint effort between Natural Language Understanding (NLU) and Natural Language Generation (NLG). Automatic Question Generation (AQG) is the task of generating reasonable questions from an input, which can be structured (e.g. a database) or unstructured (e.g. a text). An AQG system would be useful for building an automated trainer for learners to ask better questions, and for building better hint and question asking facilities in intelligent tutoring systems. Another benefit of AQG is that it can be a good tool to help and to improve the quality of the Question Answering (QA) systems. Tackling AQG in the field of computational linguistics has got immense attention from the researchers. This review paper draws on recent developments in NLP research to look at the evolution of automatic Question Generation Systems in a new light.**

***Keywords:* Natural Language Processing, Automatic question generation, Question Taxonomy, Named Entity Tagger**

I. INTRODUCTION

Over the past ten years, research in the field automated systems has moved toward the Natural Language Processing (NLP). NLP is an area of research and application that explore how computers can be used to understand and manipulate natural language text or speech to do useful things. NLP holds great promise for making computer interfaces that are easier to use for people, since people will be able to talk to the computer in their own language, rather than to learn a specialized language of computer commands. NLP researchers aim to gather knowledge on how human beings understand and use language so that appropriate tools and techniques can be developed to make computer systems understand and manipulate natural languages to perform the desired tasks. NLP research has evolved from the era of punch cards and batch processing to the era of Google and the likes of it [1].

In the ongoing advent of computer based technology there are many changes being made till now in various fields that tend to move from manual systems to automated systems. The automatic systems help us with much cost and time efficient solutions. It has affected various areas. Education is one of these areas which have gone through this evolution. In the education field, the academicians are majorly dependent on their own for generating questions for various examinations. This survey paper focusing on different techniques that are helpful for generating various types of questions mentioned in Graesser and Pearson's Taxonomy [2].

There are various approaches of generating trigger questions, generic questions or factual questions for different purposes like academic writing or reading comprehension. The rest of the paper is organized as follows: Section 2 presents the novel approach of generating trigger questions for students' learning through writing; Section 3 discusses method of Question and Answer about the contents of story for language learning; Section 4 describes question generation method from sentences; Section 5 illustrates rule based and template based approach of question generation for Technical Institution; Section 6 concludes the paper and outlines future areas of AQG research.

II. DETAILS OF G-ASK AQG SYSTEM

A good literature review to summarize and build upon current knowledge about a topic is a key part of academic writing [3]. In learning through writing, students need to consider trigger questions and monitor their understanding. Generic trigger questions are potentially forms of support for writing to learn.

Generating trigger questions to support academic writing involves three basic concepts: extract citations from students' composition, citation classification and generate template based questions [4]. The Fig. 1 depicts the basic architecture of Question Generation system.

The input to the system is a literature review paper and the output is a set of generated questions. The question generation process follows 3 steps –

Step 1- Preprocessing: This step performs extraction of citations from students' composition. By using the pattern matching techniques citations are extracted from literature review papers. There are five types of citation styles viz. textual syntactic, textual parenthetical, prosaic, pronominal and numbered. The Named Entity Tagger (NER), LBJ is used to identify citations with Prosaic style and a simple Pronoun Resolver, finding the nearest Name Entity appearing before the pronoun, is used to identify citations with Pronominal style.

ISSN: 2278 – 7798

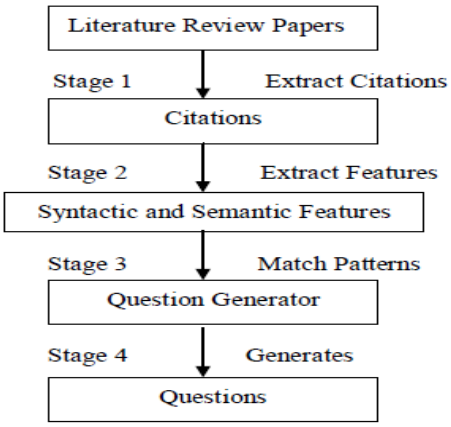

Fig. 1: System Architecture

Step 2- Extracting Syntactic and Semantic Features: The syntactic features include subject, predicate verb, auxiliary verb and predicate, voice and tense which are essential to perform subject-auxiliary inversion. To extract syntactic features the Phrase Structure Tree derived from the original citation. The Stanford Parser is used to parse a sentence into a Phrase Structure Tree. Tregex [5] is a powerful pattern matching technique which can match an individual word, regular expression, a POS tag or group of POS tags such as a Noun Phrase (NP) or Verb Phrase (VP). According to the predicate verb or auxiliary verb it determines the tense of the sentence and generates the verb lemma by using WordNet.

The semantic features include the name of author and citation category based on taxonomy of conceptual citation categories proposed by Lehnert et al [6]. The LBJ Tagger is used to detect authors' names and a rule-based approach to classify the citations.

Step 3- Question Generation: In this final step the template based questions are generated. Once the syntactic and semantic features are extracted from a citation, match the predefined patterns in repository of templates to generate corresponding questions.

To explore the ability of this system, the Bystander Turing Test conducted by Pearson Graesser. For performance evaluation the pilot study conducted on dataset of six literature reviews. The dataset contains 1,088 sentences including 221 citations. Table 1 show that 145 citations have been extracted and the recall is 0.66 in average.

Table 1: Citation Extraction Result

	Rule 1	Rule 2	Rule 3	Rule 4	Rule 5	Rule 6	Rule 7	Total
Number of Citations	18	22	12	50	16	29	74	221
Number of retrieved Citations	10	12	7	27	10	17	62	145
Recall	0.56	0.55	0.58	0.54	0.63	0.59	0.84	0.66

Table 2 illustrates 161 questions generated and the average semantic correct-ness: 60%.

Table 2: Question Generation Result

	Rule 1	Rule 2	Rule 3	Rule 4	Rule 5	Rule 6	Rule 7	Total
Number of Generates Questions	10	9	14	7	6	17	98	161
Number of Correct Questions	6	9	9	4	4	10	56	97
Precision	0.6	1	0.64	0.57	0.67	0.59	0.57	0.6

III. Details Of QA System

Some computer assisted language learning systems are equipped with test functions to ask about the contents of a story. The objective of this system is to realize a QA function which provides suitable questions for each learner and tailored advice according to the answer [7]. The QA function gives questions about a story in English to a learner who has studied the contents of the story by hearing or reading, and the learner answers the questions in English.

The foremost task of this system is to extract the syntactic and semantic information of stories based on Definite Clause Grammar. Syntactic information consists of a syntactic tree and feature structure. The syntactic tree expresses parts of speech and modification relationships of words and phrases. The feature structure expresses grammatical functions of words and phrases and grammatical information such as sentence structure and idioms. Semantic information consists of time and space information and information about verbs, nouns and modifiers. The next task is question generation by using syntactic and semantic information of an original story and dictionaries of synonyms and antonyms.
The QA system achieves the task of question generation from original story by preparing various methods.

1. *To ask about the content of one sentence:*
Firstly to extract the subject, predicate, object, complement and modifiers of a sentence from its syntactic tree and feature structure. Then to extract attributes of sentence such as tense, voice, kind of sentence, idioms and auxiliary verbs from feature structure. Also replace pronouns with their referent by referring semantic information. Finally question is prepared by using this information. For example, the method generates “Who is busy?” from “Sam is busy”.

2. *To use synonyms and antonyms:*
This method replaces a word with its synonym or antonym by referring to dictionaries of synonyms or antonyms and generates a question sentence by using one

ISSN: 2278 – 7798

of the methods to ask about the content of one sentence. For example, the method generates "Is Sam free?" from "Sam is busy" by using dictionary of antonyms.

3. *To use modifiers appeared in plural sentences:*
Semantic information of each noun is stored separately. Modifiers of the same object are stored into the same entry of an instance. Because information of modifiers which are expressed in different sentences is gathered into the same entry of an instance, it is easy to combine the modifiers of an object, which appeared separately, and generates question sentences. For example, the method generates "Did Sam sit on a small white bench?" from the two sentences "Sam sat on a white bench." and "There was a red bicycle near the small bench." which have the same object "bench".

4. To ask the contents represented by plural sentences by using a relative pronoun:

Firstly, the method selects a main sentence and a subordinate sentence which have an antecedent, and selects a relative pronoun by referring to the case and the semantic category of the antecedent in the subordinate sentence. Secondly, the method takes the antecedent away from the subordinate sentence. Finally, the method combines two sentences and the relative pronoun and generates a question sentence. For example, "Was there a red bicycle near the small bench which Sam sat on?" is generated from the two sentences "Sam sat on a white bench." and "There was a red bicycle near the small bench."

5. *To ask relationship of time and space:*
This method generates questions which ask when an event occurs or what event occurs in a place by using inclusion relationship. Also this method generates what event occurs after an event by referring to partial ordering in time information. For example, "What did happen this morning?" is generated from sentence "On that day, at early in the morning the Malaysian plane was crashed."

To evaluate the system performance there are two kinds. One is to investigate how method question generation covers collection of books of problems on the market. The other is to investigate how the question sentences generated by the module are correct semantically. In the investigation, it is confirmed that this methods cover 160 of 201 question sentences (80%) in the collection books for novices. The question generation module generated 1263 question sentences from the sentences in the textbook for second grade and 714 questions from the sentences for third grade. 1179 (93%) and 669 (94%) question sentences are correct semantically respectively.

IV. Details Of AQG From Sentence

In Automatic Sentence-to-Question generation system, where for a given sentence, the Question Generation (QG) system generates a set of questions for which the sentence contains, implies, or needs answers [8]. It is difficult to generate accurate questions from the complex sentences. Hence to facilitate the question generation task, it needs to build elementary sentences from the input complex sentences using a syntactic parser. The basic architecture of this QG system is shown in Fig. 2.

The system task is mainly divided into 3 steps – data processing, elementary sentence construction and sentence classification and question generation.

1. *Data Processing:*
The input text file given in the dataset has the information about actual answers for the questions. Hence firstly it needs to draw out relevant sentences from input text file. For extracting such relevant sentences that contain the answer from input file, it is necessary to remove redundant text from document. Then pass the relevant sentences to the Named Entity tagger and Part-of-Speech tagger. The POS tagged sentences provide information about the verbs and their tenses. NE tagged sentences provide information about person's name, location, organization, facility, time etc.

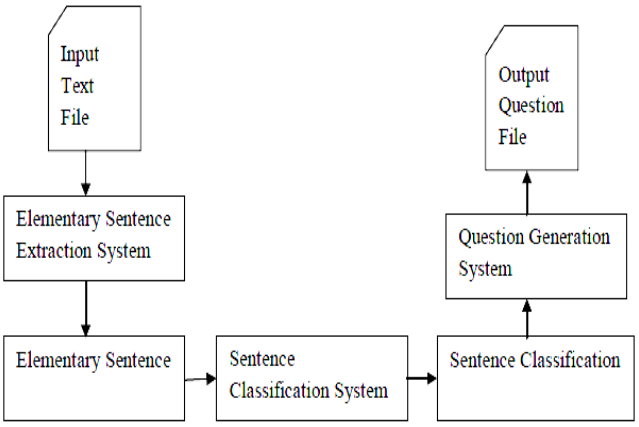

Fig.2: Architecture of QG System

2. *Elementary Sentence Extraction:*
To generate more accurate questions, the elementary sentences are extracted from the complex sentence. By using Charniak Parser, each complex sentence is parsed and then constructs the syntactic representation. It also constructs three arrays, one for the Noun Phrases (NPs), second for the Verb Phrases (VPs) and third for the Prepositions (PPs) with their location in the tree. Finally elementary sentences are constructed by combining NPs with the VPs and PPs by reading the NPs till the scope of the VPs and PPs that are in the VPs scope.

ISSN: 2278 – 7798

3. *Sentence Classification and Question Generation:*
Elementary sentences are the inputs of this module. Based on the associated POS and NE tagged information, the sentences are classified by using two simple classifiers – Fine Classifier and Coarse Classifier. Fine classifier determines the fine classes. Them the set of classes reduced to a coarse class determined by the class hierarchy. This method defines five coarse classes as – Human, Entity, Location, Time and Count. If a sentence has the structure – "Human Verb Human" then it will be classified as "whom and who" type of question. For example, the system generates "Mr. John teaches whom?" and "Who teaches to Sam?" from "Mr. John teaches to Sam."

For system performance evaluation, the Precision and Recall measures are calculated.

V. Details Of Multi Agent AQG

The proposed framework helps in question generation by deploying agents; the agents will perform various operations like document processing, information classification and question generation [8]. The architecture of this multi agent system is shown in Fig. 3.

In this multi agent system the Document Processing Agent (DPA) extracts the words from ranked list of words extracted from input text. The Information Classification Agent (ICA) classifies the input on the basis Bloom's taxonomy. The Question Generation (QG) module generating questions with the help of rules and template based database.

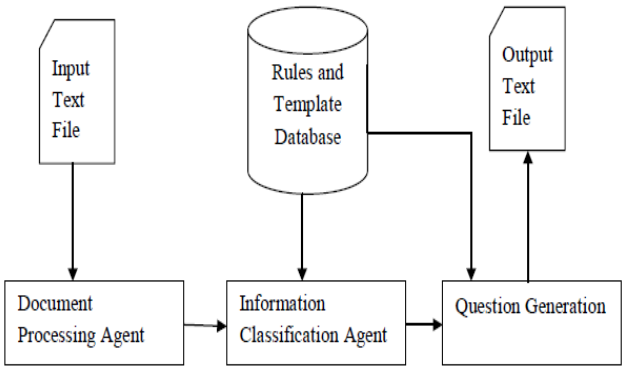

Fig. 3: System Architecture

1. *Document Processing:*
The DPA takes the input text document, processes it by Tree Tagger Tool and produces output one word per line along with the tag and lemma of that token. Then the processed output is ranked according to word count i.e frequency of occurrence of each word. Finally stemming process is performs to get proper keyword.

2. *Information Classification:*
After the process of stemming and word count, the output of DPA is given to the ICA for the classification purpose. The list of ranked words extracted from input text is then classifies on the basis of Bloom's taxonomy by searching appropriate action verbs in the rule repository.

3. *Question Generation:*
The question generation module takes the output of ICA as input to generate questions. The process is a template based approach, which fits the selected keywords in the question template according to the Bloom's levels. This system is able to effectively judge the cognitive level of students' learning. The use of template based Bloom's taxonomy help to generate conceptual level questions. The work can be enhanced further for advanced level question generation based on Bloom's Taxonomy levels like analysis, synthesis and evaluation.

The performance evaluation team of this study suggests that, the template based Bloom's taxonomy generates conceptual questions. The system is efficiently works to judge the students' learning.

VI. CONCLUSION

The AQG approach depicted in section-II is useful to support literature review writing. The pilot study of this method suggests that the AQG system can produce questions that are helpful to promote students' reflection on their academic writing. They slightly outperformed the questions generated by human expert. This method only evaluates a set of specific types of questions that refer to only citation. This system can be extended to work on ranking the generated question and upgrading taxonomy of citation category. The system described in section-III employed the five ways of question generation. This system is used to support teachers to prepare the test for purpose like reading comprehension. The evaluation study of this system verifies semantic correctness of the generated questions. The experiment evaluates of the generated questions are incorrect semantically. The QG system explained in section-IV is also helpful for generating factual questions based on given input text file. For performance evaluation, Recall and Precision has been calculated by using the dataset provided by TREC 2007. The system performance can be improved by extending number of interaction rules in the future. Also focus on the system classification module to make it more robust. The multi agent AQG system in section-V generates the conceptual questions based on Bloom's taxonomy that are useful for judging the students' learning. In future, the mathematical or derivational type of questions can be generated using this approach. Also to enhance the standard of question paper, this system can be extended to select questions based upon their ranks.

ISSN: 2278 – 7798

ACKNOWLEDGMENT

We would like to express our appreciation and thanks to Prof. Dr. P. C. Kulkarni, Principal, GES's R. H. Sapat College of Engg., Nashik. We are also thankful to Prof. N. V.Alone, Head of Department, Computer Engg., GES's R. H. Sapat College of Engg., Nashik .

REFERENCES

[1] Erik Cambria, Bebo White, "Jumping NLP Curves: A Review of Natural Language Processing Research", IEEE Computational Intelligence Magazine, May 2014.

[2] Ming Liu, R. A. Calvo, "Question Taxonomy and Implications for Automatic Question Generation".

[3] G. Graswell, "Writing for Academic Success: A Postgraduate Guide", SAGE, 2008.

[4] Ming Liu. R.A. Calvo, "Automatic Question Generation for Literature Review Writing Support".

[5] Roger Levy, Galen Andrew, "Tregex and Tsuurgeon: tools for querying and manipulating tree data structures".

[6] Lehnert, W., Cardie, C., Rilo, E. "Analyzing research papers using citation sentences", Proceedings of the Twelfth Annual Conference of the Cognitive Science Society (1990) 511-518.

[7] Hidenobu Kunichika, Tomoki Katayama, "Automated Question Generation Methods for Intelligent English Learning Systems and its Evaluation".

[8] Husam Ali Yllias Chali Sadid A. Hasan, "Automatic Question Generation from Sentences".

[9] Shivank Pandey, K.C. Rajeswari, "Automatic Question Generation Using Software Agents for Technical Institutions", International Journal of Advanced Computer Research, vol. 3, Dec. 2013.

Deepshree S. Vibhandik is presently pursuing Master of Engg. degree in Computer Engg. from GES's R. H. Sapat College of Engg, Nasik, Pune University. She has completed B.E. in Computer Engg in 2008. Her area of interest is Natural Language Processing, Data Mining and Data Structures.

Prof. Mrs. Rucha C. Samant is presently working as Asst. Professor in the Department of Computer Engg. in GES's R. H. Sapat College of Engg. at Nasik. She has completed B.E. in Computer Engg. from Mumbai University in 2000. Also she has completed M.E. in Computer Engg. from Mumbai University in 2010. There is 1 National Publication and 4 International Publications of research papers on her account. Her research work mainly focuses on Data Mining and Natural Language Processing.

ISSN: 2278 – 7798