International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

# Development of a Generic Pre-processing Module for Automatic Question Generation

Dr.Nagamani CH
*Department of Computer Science and Engineering*
*MALLA REDDY ENGINEERING COLLEGE FOR WOMEN*
Maisammaguda, Dhulapally, Post via Kompally Secunderabad-500100, Telangana, INDIA
nagamani502@gmail.com

Dhawaleswar Rao CH
*Department of Computer Science and Engineering*
*Centurion University of Technology and Management*
Paralakhemundi Campus, Gajapati District, Odisha, INDIA
dhawaleswarrao@gmail.com

***Abstract*—Knowledge discovery in text has emerged as a challenge to the researchers due to the large number of textual documents available from different sources. A critical step of this process is the pre-processing phase. Several pre-processing steps have been performed for automatic generation of cloze questions from text. The task necessitates identifying the facts that can be asked to the examinee. Moreover, we need the sentences that contain a single questionable fact. Complex and compound sentences often contain more than one facts; therefore, pre-processing is necessary to handle such sentences. Also, identification of the keywords, terms or topic words becomes necessary to get an idea regarding the theme or content of the text. In this paper, we discuss various pre-processing tasks and proposed a sentence selection module for task of generating questions from text automatically.**

***Index Terms*—Automatic Question Generation, Multiple-Choice Questions, NLP, Pre-processing, Text Analysis**

## I. Introduction

Question is a fundamental technique for evaluating a learner's knowledge or comprehension. Evaluation is necessary for learning and the evaluation requires questioning. Carefully generated questions can assist the learners to facilitate their comprehension and decomposition of a problem. These can also encourage them to plan a solution before implementation and to reveal the gaps or misconceptions in knowledge [1]. Where as, the questions are been created manually takes a lot of time and requires expertise on the subject. Both are costly in many learning environments. Additionally, in the modern era of education e-learning, online learning, active learning, computer-aided learning, intelligence tutorial system etc. gained huge popularity. These smart learning platforms also demands smart assessment. Automatic question generation (AQG) plays a big role here in making such learning environments smarter. Therefore, AQG from text has turned into an important component of advanced learning technologies, help systems, instructional games, inquiry-based environments etc. [2]. As a result, AQG from text has caught the substantial attention of the research community in the last two decades. Although the term AQG is generic and has been defined as a system capable of accepting multiple input forms including text and non-text [3]; the majority of the research in this area primarily focus on question generation from a text.

Mutiple-Choice Question (MCQ) is a popular tool which is widely used in multiple levels of educational assessment. MCQ is supportive of the evaluation of well-defined knowledge and ideas included within the respective text. MCQ type has several advantages like, it can handle multiple levels of learning outcomes including recall, application, analysis, quick and accurate evaluation; reliable assessment; objective and consistent scoring (unlike the scoring of essay questions); and higher validity. With the assistance of the MCQs, the academic development units may assist academics to enhance student performance and learning outcomes about in every practical sense [4]. MCQs are a widely used response format and a popular method of assessing knowledge of various domains and applications. MCQs implemented through computerized applications are widely used globally in many standardized objective tests and also in India. The tests like SAT (Scholastic Assessment Test), GATE (The Graduate Aptitude Test in Engineering), All India Institute of Medical Sciences (AIIMS), CAT (Common Admission Test), Joint Entrance Examination (JEE), most part of the TOEFL (Test of English as a Foreign Language), GRE (The Graduate Record Examination) and many other examinations use MCQs.

MCQ is the most popular sub-area of AQG where maximum research effort have been given. Effort has been given for the development of MCQ systems in various languages and domains. In the literature we found that MCQ system is developed for English, Basque, European Portuguese, Russian, China languages. In numerous fields research has also been conducted on automated MCQ generation, including, educational domain [5], [6], language learning [7]–[17], general science [18], physics [19], medical and biology [20]–[26], sports domain [27]–[30] etc.

## II. Motivation for Automatic Multiple-Choice Question Generation

In the past two decades, the research community devoted substantial effort to find the techniques for automatic generation of MCQs from text. For automated MCQ creation in diverse fields and applications, a significant variety of solutions have been created. The development of an automated MCQ generator system is separated into many sections from the literature. Every system has a little variation in the number of stages and the general strategy, but most of the systems have a general working flow. The job may be widely separated into many subtasks such as pre-processing, sentencing, key section, question building, distractor creation and post-processing. However, system each system changes the amount of subtasks and the overall approach. The literature still lacking in various ways, despite a huge amount of study towards automated MCQ production. We could not identify a general procedure throughout our literature investigation, which may enable new researchers to start with. We were unable to identify an independent domain and application system that uses text input and generates MCQs. The majority of the systems we examined are domain-specific. Domain issues are dependent on unique domain characteristics and restricted in scope. While the question-generation activity by its very nature needs domain-specific configuration, there is still some universal sense.

In this article, as we focus on computer-assisted assessment in school education domain in India, we first targeted to generate MCQs automatically from textbook contents. However, we did not find any system that fulfills our purpose. So, we planned to develop a system for automatic MCQ generation from school textbook.

We have begun to investigate the literature where a lot of systems and technology have been discovered. Because there are several MCQ systems and various methodologies have been developed, it is necessary to study these systems. To develop a new strategy, knowledge of current ways is needed. A basic article may offer a summary of the facts and helps to create a workable vision. To this end, there has been a methodically review of the existing literature on MCQ creation.

In this article, we provide a general six-phase pipeline for the production of automated MCQs. This includes (1) Pre-processing, (2) Sentence selection, (3) Key selection, (4) Question formation, (5) Distractor generation, and (6) Post-processing. We wanted to combine the strategies examined by the literature review for the construction of the MCQ system from the school textbook. However, many of these policies have not been applied to the field. Various tactics have also been used to implement certain phases. A good methodology for the creation of a new system is thus difficult to define. This research gap has prompted us to develop a generic pipeline. We have thus tried to create a general MCQ pipeline that will function well on many topics. We offer a generic MCQ generating pipeline from the school book in the first half of the essay.

### A. *Question Taxonomy*

The first step to generate questions from a text is to determine the knowledge included. The kind of knowledge that is integrated into or to be requested in text differs from one level to another. Various taxonomy questions have also been offered [31]–[35] in the literature. we have also discovered. But the scope is the main parameter to create taxonomy if we concentrate on the issue generating activity.

The authors categorised the assignment in two more broad groups on the basis of its extent. These are questions of the objective kind and questions of the subjective kind. The target questions have a specified scope where the information buried in a single phrase is addressed in the questions. On the other hand, the queries of the subjective or narrative kind have as their scope the whole text or a paragraph. Subjective and objective characteristics have relatively good benefits and in particular areas demonstrate dominance over others. For example, questioning of target type is one of the most effective techniques of an e-learner to get feedback.

Again, complicated descriptive issues offer greater educational advantages for reading understanding than simple objective inquiries. [5], [36], [37]. The subjective and objective kinds may be divided into many subclasses. The questions are classified from a question generation point of view in Figure.1. However, most of the present AQG systems concentrate on producing inquiries, i.e. objective matters, from a single sentence.

From Figure.1 we observe that, objective type questions are classified into two types: *Wh*-questions and *fill-in-the-blank (FIB)* questions [38]. In FIB questions two type of questions are possible: *Cloze* and *Open-cloze* questions [38], [39]. Cloze questions are the question sentences (sentence with one or more blanks) with a set of alternatives. Question sentences without alternatives are known as Open-cloze questions [39]. Questions generated for assessing the content knowledge of the learner are called *factual* questions [38], [39].

Wh-calls are the matter that starts with Wh-word. A question may have a number of alternatives or possibilities, one of which is the right response. A Questiona and Answer containing a series of options is referred to as the multi-choice question (MCQ) [**?**], [7]. These MCQs are mostly Wh-questions but include alternatives, and some writers also include cloze questions [**?**], [7], [8], [38]–[40]. These are also called Cloze questions. Therefore Wh-questions are divided into four varieties depending on the intended complexity: *Shallow, Intermediate, Deep* and *Cloze*. [35].

Variations in objective inquiries were represented by many related terms in the literature. The new researchers in this field frequently become confused with these variances. Now we are debating these questions more thoroughly to clarify them.

The questions are the fact or know-how in [**?**], [8], [41]–[43]. For the production of factual questions, non-fiction books that provide factual information instead of views are necessary. Recent questions are employed at some time after the student

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

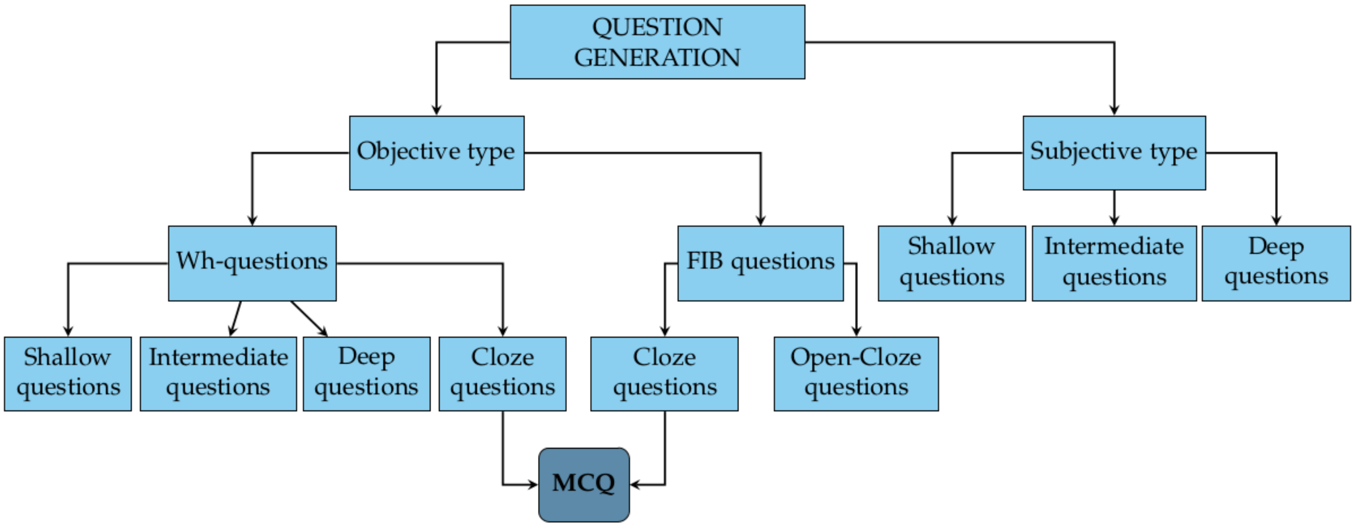


Fig. 1. Question Taxonomy based on Question Generation

read the texts [5] to evaluate the recall of certain information. It is also used to evaluate the content in the text of a student's knowledge.

*Shallow* questions also known as *simple* questions are the questions that focus more on facts, such as: yes/no, what?, who?, when? and where?. *Intermediate* questions deals with how?, how much?, how many?, what does? questions etc. *Deep* questions also called as *complex* questions are defined as the questions that involve more on logical thinking, such as: why/how, What next?, What if?, why?, how did?, what do, why didn't? etc. Deep or high-level questions prompt to better learning than shallow or low-level questions that attention on recognition and recall [5]. Deep questions are the kind of open-ended queries. It covers a large amount of content instead of a single sentence. Therefore, requires recall and deep thinking to answer such questions [44].

Open-cloze question contains a sentence with one or more blanks in it and does not have alternatives. A cloze question contains a sentence with one or more blanks and a set of alternatives. The question sentence with one or more blanks is known as *stem* [39]. The correct answer is also called as *key* or target word. The wrong alternatives are referred as *distractors* [9]. For example, consider the following cloze question.

*stem:* The nearest planet to the sun is ______.
(a) Earth *(distractor-1)*
(b) Jupiter *(distractor-2)*
(c) Mercury *(key)*
(d) Venus *(distractor-3)*

The automated creation of MCQ is an important field of study. Researchers have developed or focused on the unresolved issues of MCQ systems in new and upcoming applications. The automated MCQ generating literature is thus fairly robust. Because several systems have been designed for the development of MCQ and various methodologies. To develop a new strategy, knowledge of current ways is needed.

### B. Various approaches of MCQ Generation

We are now discussing the methods to the creation of the automated MCQ systems. We have determined that the primary motivation for the researchers was the approach anticipated to follow to prepare MCQs manually from a book. The user initially has to access the information in the input text for manual production of MCQs. Since an MCQ requires mainly an information sentence, it also indicates the sentences containing any facts or information which is doubtful. Identifying the word or sentence that works as the key is the next challenge. He then asked from the statement, where the answer is the key. The ultimate aim is for the analysis of the input text or a wider context to choose certain distractors.

A similar method might be used to construct an automated MCQ creation system. The assignment consists of many steps. Every system has a little variation in the number of stages and the general strategy, but most of the systems have a general working flow. The workflow of a system with six steps may also be shown in Figure.2: (1) Pre-processing, (2) Sentence selection, (3) Key selection, (4) Question creation, (5) Generating distrators and (6) Post-processing. Below are the approaches utilised in the literature for developing these stages.

### C. Pre-processing

Knowledge discovery in text has emerged as a challenge to the researchers due to the large number of textual documents available from different sources [45]. A critical step of this process is the pre-processing phase. Several pre-processing steps have been performed for automatic generation of cloze questions from text. The task necessitates identifying the facts

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

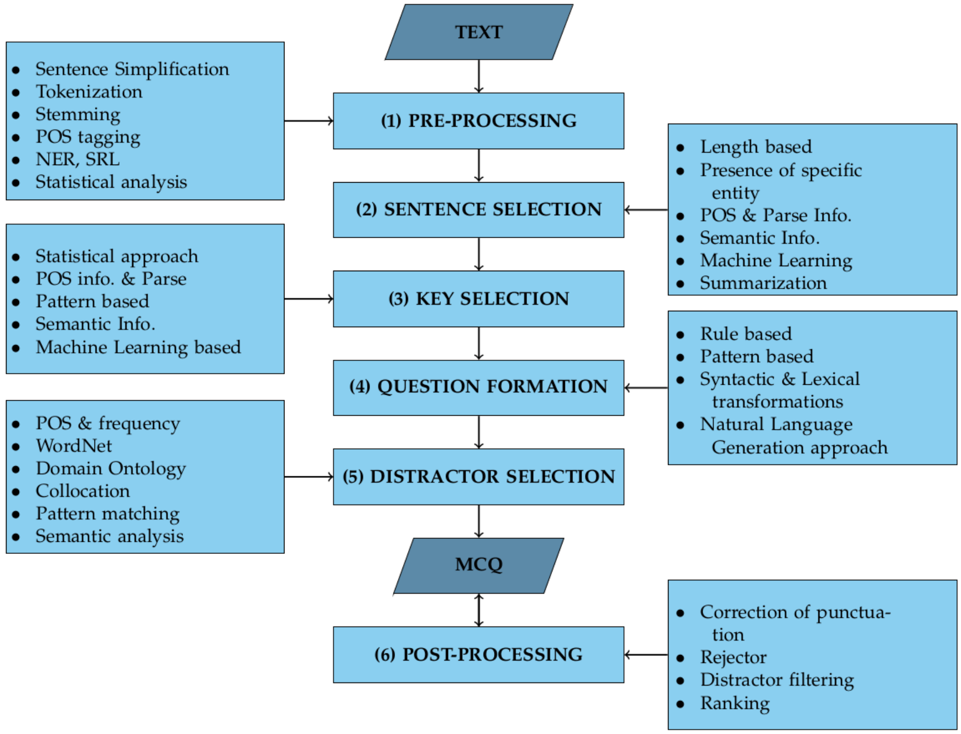


Fig. 2. The flowchart of the question generating system

that can be asked to the examinee. Moreover, we need the sentences that contain a single questionable fact. Complex and compound sentences often contain more than one facts; therefore, pre-processing is necessary to handle such sentences. Also, identification of the keywords, terms or topic words becomes necessary to get an idea regarding the theme or content of the text. We discuss below various pre-processing tasks used in the existing systems.

Text **normalization** normally refers to the input text conversion in required formats and to the elimination from the text of [46] of unneeded content. These processes are dependent on the domain and applications in significant part. The methodology to be used in this case thus relies on the need for this specific activity. In [5], [47]–[49], many forms of text normalisation and sentence normalisation were utilised.

**Sentence simplification** may also be taken into account as part of the normalisation of text. It consists of transforming complicated and composite phrases into simple phrases, usually by utilising the external system and [50]. The simplification of sentence in many systems has been conducted; like, [**?**], [**?**], [**?**], [**?**], [20], [49], [51]. **Structural and tokenization** were also done in some circumstances. Tokenization is dividing the text of the document into a stream of words, symbols and numbers and so on. The chapters, sections, sub-sections, paragraphs and other pertinent tags are identified as a [**?**] and structural analysis. The structural analysis and filtering of [52] methods may overlook the superfluous parts of the input text. Such pre-processing is seen in various systems; like, [8]–[12], [53]–[56].

Some **static analysis** [43] was done on a text for the extraction or further processing in succeeding modules. In the following modules. Selection of keywords typically requires a certain statistical study, for example. Most are word level analysis such as word frequency, n-gram, term frequency*inverse document frequency (Tf*Idf), co-occurrence statistics, etc. These analyses are most frequent. Although they are not specifically described in many studies as pre-processing steps, they may be regarded as pre-processing. In different modules of numerous systems, statistical analysis is used [5], [8], [10], [12], [23]–[25], [30], [53], [54], [57]–[60].

**Lexical analysis** is the study of words. And it is the process of taking an Input Character string and creating a series of symbols called the [61] lexical tokens. Inflexions may influence several task modules, especially in selecting key and generating distractors. The stemming [62] process was accomplished in many Systems [5], [48], [63], [64].

**Syntactic analysis** covers the function of words in the creation of sentences. Because of the fact that the information to be requested belongs to a certain word or sentence category (the target category depending on the application and field) and the distractors must also belong to the same category, the sentence should be syntactically analysed. Many systems have been utilised with various degrees of syntactic analysis such as [65], [66], or shallow syntactic parsing or deep parsing. For example, [8], [9], [11]–[13], [21], [24], [28], [29], [49], [53], [55], [58], [67]. Names also play a vital role in text. And in many domains, like, history, sports, health, entertainment, names of something are commonly asked as

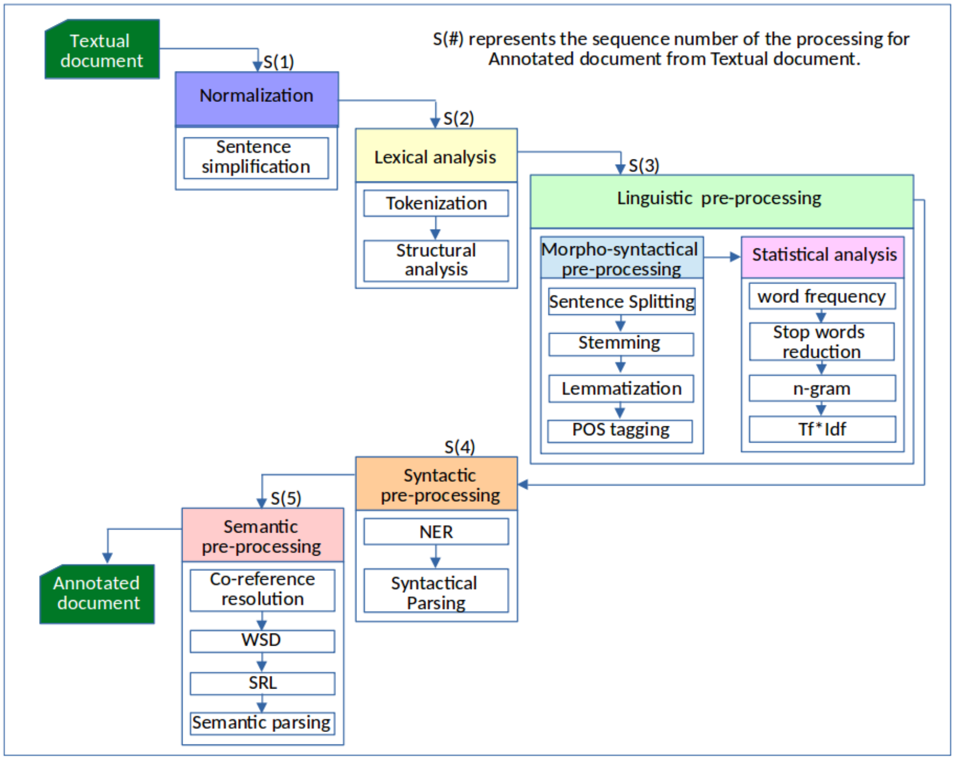


Fig. 3. Pre-processing general framework module

question. Therefore, identification of named entities is required in such scenario. **Named entity recognition** (NER) [68] is performed in a number of existing systems including [20], [23], [27], [28], [69]–[71]. Here also we would like to point out that many of these writers have not explicitly called syntactic analysis a pre-processing phase, but before the main modules of the systems, these analyses may be carried out, and they may be considered a pre-processing phase.

**Semantic processing** or contextual meaning of the words, phrases, sentences or the complete text of the phrase [72]. Different semanticized workloads in NLP have been defined. Many of these semantic processing are used by AQG because the meaning or information in the input text are understood. **Coreference resolution** is a crucial pre-processing step to map the pronouns of their respective substantives, in particular, the pronoun resolution. A pronoun cannot be used as a question in a majority of applications, hence it is vital to identify the related noun. Various authors have performed the task Coreference resolution like; [5], [27], [69], [73], [74] and many others. Words in natural language text are ambiguous, a word might carry multiple meanings depending on the context. **Word sense disambiguation** (WSD) [75] identify the exact sense of a word given in a text according to the context. We find the use of WSD in [5], [8], [12], [56], [58], [70], [76], **discourse analysis** [24], **semantic role labeling** (SRL) [77] or **semantic parsing** [23], [70], [73], [74], [78], [79] have been used as pre-processing technique in the literature.

## III. Proposed System

Rao and Saha [80] outlined the steps to be followed for the automatic generation of MCQs from a text. We have primarily adopted those steps. However, for the implementation of the individual steps, we have used our own approach. We discuss below the techniques in the literature for the development of the individual phases of MCQ generation.

The strategies we adopted for developing the system for automatic MCQ generation are discussed here. The system focuses on preparing questions from school level subjects. So it takes a chapter of a textbook as input and prepares questions from that. The proposed system contains four core modules, namely, pre-processing, sentence selection, key selection and distractor generation. The individual modules use a hybrid approach. These are summarized below.

The pdf version of the textbook is given as input to the system. The system first converts the pdf document into readable text format and cleans the text. Students might face difficulty in understanding or memorizing long sentences. Compound, complex and compound-complex sentences often contain multiple facts. However, one MCQ normally deals with a single fact. The occurrence of multiple facts in a sentence might make it over-informative and the question itself might contain certain clues that help in guessing the correct answer. Therefore, from the questioning point of view also, simple sentences are better. So, the system converts the compound, complex and compound-complex sentences into simple sentences.

| TABLE 1 | | | |
|---|---|---|---|
| EVALUATION OF THE PRE-PROCESSING MODULE | | | |
| EVALUATION METRIC | Evaluator 1 | Evaluator 2 | Evaluator 3 |
| Utilizable | 92.05 | 95.12 | 95.16 |
| Informative Sentence Quality | 95.01 | 95.19 | 94.40 |
| Informative Sentence Readability | 93.12 | 95.03 | 95.05 |

*1) Pre-processor Module:* The pre-processing module performs text extraction, text normalization, lexical analysis, linguistic, syntactic and semantic processing on the input text and makes the text ready for the subsequent modules. The module is generic and expected to work well on a variety of inputs. In Figure 3 we present the pre-processing framework.

For performing sentence simplification, first we extract the sentences having entities from annotated text using Stanford-NERTagger[1]. Second, sentences like compound, complex and compound-complex have been extracted using Stanford dependency parser[2]. Then these sentences are converted into simple sentences using their context-free grammar(CFG) structure and POS tagging based rules. Here the extracted compound, complex, compound-complex sentences, their corresponding CFG structures, and the POS tagging rules are stored in three different files named keywords contained sentences, CFG, and sentence simplification rules.

The CFG approach is based on the computation of CFG structure of an input sentence with CFG structure of a set of simplified sentences. Therefore, for the task, we need a set of sentences that act as a simplified set. In order to create the simplified set, we collect a number of the existing pre-processed named entity containing sentences from the corpus. CFG structure matching approach primarily focuses on the structure of the sentence CFG, rather fact embedded in the sentence. This module matches the CFG structure between the input sentence and the named entity or keyword contained sentences. If both CFG structures match, then the input sentence incorporates the structure of the particular named entity rule-based simplified sentence. Otherwise, the input sentence is computed as a simple sentence.

*2) Evaluation of the Pre-processor Module:* The score we got after manually evaluating 122 system-generated informative sentences is shown in Table 1. We attain an average accuracy of 94.46 percentile based on the scores.

## IV. Conclusion and Future Work

We have presented a pre-processing framework and Sentence Selection module for automatic generation of MCQ questions from school textbooks. The MCQ system contains four key modules, namely, preprocessing, sentence selection, key selection and distractor generation. Hybrid approaches have been used for implementation of the individual modules. Manual evaluation results demonstrate that the proposed system is capable of generating accurate MCQs. Individual modules also show good accuracy. Additionally, the number of MCQs generated by the system is quite high. A high recall is necessary for the system to utilize it in a real application scenario. Although the system contains some domain (or, subject) specific features and resources, it can be ported to other domains (or, subjects) with minor effort. directions to extend the present work. Here Key selection and Distractor generation phases are been considered as future work.

[1] https://nlp.stanford.edu/ner/

[2] https://nlp.stanford.edu/software/nndep.html

## References

[1] K. E. Boyer, W. Lahti, R. Phillips, M. D. Wallis, M. A. Vouk, and J. C. Lester, "Principles of asking effective questions during student problem solving," in *Proceedings of the 41st ACM Technical Symposium on Computer Science Education*. ACM, 2010, pp. 460–464.

[2] K. Boyer and P. Piwek, "Qg2010: The third workshop on question generation," in *International Conference on Intelligent Tutoring Systems. Pittsburgh, PA: Carnegie Mellon University*, 2010.

[3] V. Rus, Z. Cai, and A. Graesser, "Question generation: Example of a multi-year evaluation campaign," *Proc WS on the QGSTEC*, 2008.

[4] G. T. Crisp and E. J. Palmer, "Engaging academics with a simplified analysis of their multiple-choice question (mcq) assessment results." *Journal of University Teaching and Learning Practice*, vol. 4, no. 2, pp. 88–106, 2007.

[5] M. Heilman, "Automatic factual question generation from text," Ph.D. dissertation, Language Technologies Institute, School of Computer Science, Pittsburgh, PA, USA, 2011, aAI3528179.

[6] A. Kurtasov, "A system for generating cloze test items from russian-language text." in *RANLP*, 2013, pp. 107–112.

[7] T. Goto, T. Kojiri, T. Watanabe, T. Iwata, and T. Yamada, "Automatic generation system of multiple-choice cloze questions and its evaluation," *Knowledge Management & E-Learning: An International Journal (KM&EL)*, vol. 2, no. 3, pp. 210–224, 2010.

[8] R. Mitkov, L. An Ha, and N. Karamanis, "A computer-aided environment for generating multiple-choice test items," *Nat. Lang. Eng.*, vol. 12, no. 2, pp. 177–194, Jun. 2006.

[9] J. Pino, M. Heilman, and M. Eskenazi, "A selection strategy to improve cloze question quality," in *Proceedings of the Workshop on Intelligent Tutoring Systems for Ill-Defined Domains. 9th International Conference on Intelligent Tutoring Systems, Montreal, Canada*, 2008, pp. 22–32.

[10] E. Sumita, F. Sugaya, and S. Yamamoto, "Measuring non-native speakers' proficiency of english by using a test with automatically-generated fill-in-the-blank questions," in *Proceedings of the Second Workshop on Building Educational Applications Using NLP*, ser. EdAppsNLP 05. Stroudsburg, PA, USA: Association for Computational Linguistics, 2005, pp. 61–68.

[11] A. Hoshino and H. Nakagawa, "Assisting cloze test making with a web application," *TECHNOLOGY AND TEACHER EDUCATION ANNUAL*, vol. 18, no. 5, pp. 2807–2814, 2007.

[12] C.-L. Liu, C.-H. Wang, Z.-M. Gao, and S.-M. Huang, "Applications of lexical information for algorithmically composing multiple-choice cloze items," in *Proceedings of the Second Workshop on Building Educational*

*Applications Using NLP*, ser. EdAppsNLP 05. Stroudsburg, PA, USA: Association for Computational Linguistics, 2005, pp. 1–8.

[13] Y.-C. Lin, L.-C. Sung, and M. C. Chen, "An automatic multiple-choice question generation scheme for english adjective understanding," in *Workshop on Modeling, Management and Generation of Problems/Questions in eLearning, the 15th International Conference on Computers in Education (ICCE 2007)*, 2007, pp. 137–142.

[14] S. Smith, P. Avinesh, and A. Kilgarriff, "Gap-fill tests for language learners: Corpus-driven item generation," in *Proceedings of ICON-2010: 8th International Conference on Natural Language Processing*, 2010, pp. 1–6.

[15] J. Lee and S. Seneff, "Automatic generation of cloze items for prepositions," in *Eighth Annual Conference of the International Speech Communication Association*, 2007, pp. 2173–2176.

[16] D. Higgins, "Item distiller: Text retrieval for computer-assisted test item creation," *Educational Testing Service Research Memorandum (RM-07-05). Princeton, NJ: Educational Testing Service*, 2007.

[17] Y. Susanti, T. Tokunaga, H. Nishikawa, and H. Obari, "Evaluation of automatically generated english vocabulary questions," *Research and Practice in Technology Enhanced Learning*, vol. 12, no. 1, p. 11, 2017.

[18] I. Aldabe and M. Maritxalar, "Automatic distractor generation for domain specific texts." in *IceTAL*. Springer, 2010, pp. 27–38.

[19] R. Shah, D. Shah, and L. Kurup, "Automatic question generation for intelligent tutoring systems," in *Communication Systems, Computing and IT Applications (CSCITA), 2017 2nd International Conference on*. IEEE, 2017, pp. 127–132.

[20] N. Afzal and R. Mitkov, "Automatic generation of multiple choice questions using dependency-based semantic relations," *Soft Comput.*, vol. 18, no. 7, pp. 1269–1281, Jul. 2014.

[21] N. Karamanis, L. A. Ha, and R. Mitkov, "Generating multiple-choice test items from medical text: A pilot study," in *Proceedings of the Fourth International Natural Language Generation Conference*, ser. INLG '06. Stroudsburg, PA, USA: Association for Computational Linguistics, 2006, pp. 111–113.

[22] W. Wang, T. Hao, and W. Liu, "Automatic question generation for learning evaluation in medicine," in *International Conference on Web-Based Learning*. Springer, 2007, pp. 242–251.

[23] I. E. Fattoh, "Semantic based automatic question generation using artificial immune system," *Computer Engineering and Intelligent Systems*, vol. 5, no. 8, pp. 74–82, 2014.

[24] M. Agarwal and P. Mannem, "Automatic gap-fill question generation from text books," in *Proceedings of the 6th Workshop on Innovative Use of NLP for Building Educational Applications*, ser. IUNLPBEA '11. Stroudsburg, PA, USA: Association for Computational Linguistics, 2011, pp. 56–64.

[25] A. E. Awad and M. Y. Dahab, "Automatic generation of question bank based on pre-defined templates," *International Journal of Innovations Advancement in Computer Science*, vol. 3, no. 1, pp. 80–87, 2014.

[26] G. Kumar, R. E. Banchs, and L. F. D'Haro, "Automatic fill-the-blank question generator for student self-assessment," in *Frontiers in Education Conference (FIE), 2015 IEEE*. IEEE, 2015, pp. 1–3.

[27] A. Narendra, M. Agarwal, and R. shah, "Automatic cloze-questions generation," in *Proceedings of the International Conference Recent Advances in Natural Language Processing RANLP 2013*. INCOMA Ltd. Shoumen, BULGARIA, 2013, pp. 511–515.

[28] M. Majumder and S. K. Saha, "Automatic selection of informative sentences: The sentences that can generate multiple choice questions," *Knowledge Management & E-Learning: An International Journal (KM&EL)*, vol. 6, no. 4, pp. 377–391, 2014.

[29] ——, "A system for generating multiple choice questions: With a novel approach for sentence selection," *ACL-IJCNLP Workshop on NLPTEA 2015*, pp. 64–72, 2015.

[30] A. S. Bhatia, M. Kirti, and S. K. Saha, "Automatic generation of multiple choice questions using wikipedia," in *International Conference on Pattern Recognition and Machine Intelligence*. Springer, 2013, pp. 733–738.

[31] B. Bloom, "Bloom taxonomy of educational objectives: Book 1," 1956.

[32] R. D. Nielsen, J. Buckingham, G. Knoll, B. Marsh, and L. Palen, "A taxonomy of questions for question generation," in *Proceedings of the Workshop on the Question Generation Shared Task and Evaluation Challenge*, 2008.

[33] K. E. Boyer, W. J. Lahti, R. Phillips, M. Wallis, M. A. Vouk, and J. C. Lester, "An empirically-derived question taxonomy for task-oriented tutorial dialogue," in *Proceedings of the Second Workshop on Question Generation*, 2009, pp. 9–16.

[34] C. Forăscu and I. Drăghici, "Question generation: Taxonomies and data," in *AIED*, vol. 2009, 2009, p. 14th.

[35] A. Graesser, V. Rus, and Z. Cai, "Question classification schemes," in *Proc. of the Workshop on Question Generation*, 2008.

[36] R. C. Anderson and W. B. Biddle, "On asking people questions about what they are reading," *Psychology of learning and motivation*, vol. 9, pp. 89–132, 1975.

[37] C. Hamaker, "The effects of adjunct questions on prose learning," *Review of Educational Research*, vol. 56, no. 2, pp. 212–242, 1986.

[38] B. Das and M. Majumder, "Factual open cloze question generation for assessment of learner's knowledge," *International Journal of Educational Technology in Higher Education*, vol. 14, no. 1, p. 24, 2017.

[39] M. Agarwal, "Cloze and open cloze question generation systems and their evaluation guidelines," *International Institute of Information Technology, Hyderabad*, 2012.

[40] T. Goto, T. Kojiri, T. Watanabe, T. Iwata, and T. Yamada, "An automatic generation of multiple-choice cloze questions based on statistical learning," in *Proceedings of the 17th International Conference on Computers in Education*. Asia-Pacific Society for Computers in Education, 2009, pp. 415–422.

[41] M. Liu, V. Rus, and L. Liu, "Automatic chinese factual question generation," *IEEE Transactions on Learning Technologies*, vol. 10, no. 2, pp. 194–204, 2017.

[42] M. Heilman and N. A. Smith, "Question generation via overgenerating transformations and ranking," Carnegie-Mellon Univ Pittsburgh PA Language Technologies Inst, Tech. Rep., 2009.

[43] D. Metzler and W. B. Croft, "Analysis of statistical question classification for fact-based questions," *Inf. Retr.*, vol. 8, no. 3, pp. 481–504, May 2005.

[44] I. Labutov, S. Basu, and L. Vanderwende, "Deep questions without deep understanding." in *ACL (1)*, 2015, pp. 889–898.

[45] R. Curia, M. Ettorre, L. Gallucci, S. Iiritano, and P. Rullo, "Textual document pre-processing and feature extraction in olex," *WIT Transactions on Information and Communication Technologies*, vol. 35, 2005.

[46] R. Sproat, A. W. Black, S. Chen, S. Kumar, M. Ostendorf, and C. Richards, "Normalization of non-standard words," *Comput. Speech Lang.*, vol. 15, no. 3, pp. 287–333, Jul. 2001.

[47] M. Heilman and N. A. Smith, "Good question! statistical ranking for question generation," in *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, ser. HLT '10. Stroudsburg, PA, USA: Association for Computational Linguistics, 2010, pp. 609–617.

[48] L. Bednarik and L. Kovacs, "Implementation and assessment of the automatic question generation module," in *Cognitive Infocommunications (CogInfoCom), 2012 IEEE 3rd International Conference on*. IEEE, 2012, pp. 687–690.

[49] H. D. A. D. Ali, "Automatic question generation: a syntactical approach to the sentence-to-question generation case," 2012.

[50] D. Vickrey and D. Koller, "Sentence simplification for semantic role labeling." in *ACL*, 2008, pp. 344–352.

[51] M. Heilman and M. Eskenazi, "Application of automatic thesaurus extraction for computer generation of vocabulary questions," in *Workshop on Speech and Language Technology in Education*, 2007.

[52] N. J. Belkin and W. B. Croft, "Information filtering and information retrieval: Two sides of the same coin?" *Commun. ACM*, vol. 35, no. 12, pp. 29–38, Dec. 1992.

[53] D. Coniam, "A preliminary inquiry into using corpus word frequency data in the automatic generation of english language cloze tests," *Calico Journal*, vol. 14, no. 2-4, pp. 15–33, 1997.

[54] A. Hoshino and H. Nakagawa, "A real-time multiple-choice question generation for language testing: A preliminary study," in *Proceedings of the Second Workshop on Building Educational Applications Using NLP*, ser. EdAppsNLP 05. Stroudsburg, PA, USA: Association for Computational Linguistics, 2005, pp. 17–20.

[55] C.-Y. Chen, H.-C. Liou, and J. S. Chang, "Fast: An automatic generation system for grammar tests," in *Proceedings of the COLING/ACL on Interactive Presentation Sessions*, ser. COLING-ACL '06. Stroudsburg, PA, USA: Association for Computational Linguistics, 2006, pp. 1–4.

[56] S. d. S. L. Curto, "Automatic generation of multiple-choice tests geraçao automatica de testes de escolha m ultipla," Master's thesis, Instituto Superior Tecnico, Information Systems and Computer Engineering, 2010.

[57] C.-C. Shei, "Followyou!: An automatic language lesson generation system," *Computer Assisted Language Learning*, vol. 14, no. 2, pp. 129–144, 2001.

[58] R. Mitkov and L. A. Ha, "Computer-aided generation of multiple-choice tests," in *Proceedings of the HLT-NAACL 03 Workshop on Building Educational Applications Using Natural Language Processing - Volume 2*, ser. HLT-NAACL-EDUC '03. Stroudsburg, PA, USA: Association for Computational Linguistics, 2003, pp. 17–22.

[59] J. C. Brown, G. A. Frishkoff, and M. Eskenazi, "Automatic question generation for vocabulary assessment," in *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing*, ser. HLT '05. Stroudsburg, PA, USA: Association for Computational Linguistics, 2005, pp. 819–826.

[60] I. Aldabe, M. Maritxalar, and R. Mitkov, "A study on the automatic selection of candidate sentences distractors." in *AIED 2009: 14th International Conference on Artificial Intelligence in Education Workshops Proceedings*, 2009, pp. 656–658.

[61] W. Yang, C.-W. Tsay, and J.-T. Chan, "On the applicability of the longest-match rule in lexical analysis," *Comput. Lang. Syst. Struct.*, vol. 28, no. 3, pp. 273–288, Oct. 2002.

[62] J. Allan and G. Kumaran, "Stemming in the language modeling framework," in *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Informaion Retrieval*, ser. SIGIR '03. New York, NY, USA: ACM, 2003, pp. 455–456.

[63] S. Pandey and K. Rajeswari, "Automatic question generation using software agents for technical institutions," *International Journal of Advanced Computer Research*, vol. 3, no. 4, pp. 307–311, 2013.

[64] P. Pabitha, M. Mohana, S. Suganthi, and B. Sivanandhini, "Automatic question generation system," in *Recent Trends in Information Technology (ICRTIT), 2014 International Conference on*. IEEE, 2014, pp. 1–5.

[65] K. Toutanova, D. Klein, C. D. Manning, and Y. Singer, "Feature-rich part-of-speech tagging with a cyclic dependency network," in *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology - Volume 1*, ser. NAACL '03. Stroudsburg, PA, USA: Association for Computational Linguistics, 2003, pp. 173–180.

[66] J.-M. Balfourier, P. Blache, and T. van Rullen, "From shallow to deep parsing using constraint satisfaction," in *Proceedings of the 19th International Conference on Computational Linguistics - Volume 1*, ser. COLING '02. Stroudsburg, PA, USA: Association for Computational Linguistics, 2002, pp. 1–7.

[67] I. Aldabe, M. L. de Lacalle, M. Maritxalar, E. Martinez, and L. Uria, "Arikiturri: An automatic question generator based on corpora and nlp techniques," in *Proceedings of the 8th International Conference on Intelligent Tutoring Systems*, ser. ITS'06. Berlin, Heidelberg: Springer-Verlag, 2006, pp. 584–594.

[68] J. R. Finkel, T. Grenager, and C. Manning, "Incorporating non-local information into information extraction systems by gibbs sampling," in *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, ser. ACL '05. Stroudsburg, PA, USA: Association for Computational Linguistics, 2005, pp. 363–370.

[69] R. Shah, "Automatic question generation using discourse cues and distractor selection for cloze questions," Master's thesis, Language Technology and Research Center (LTRC), International Institute of Information Technology, Hyderabad, 2012.

[70] D. L. Lindberg, "Automatic question generation from text for self-directed learning," Ph.D. dissertation, Applied Sciences: School of Computing Science, 2013.

[71] R. Correia, J. Baptista, M. Eskenazi, and N. Mamede, "Automatic generation of cloze question stems," in *Proceedings of the 10th International Conference on Computational Processing of the Portuguese Language*, ser. PROPOR'12. Berlin, Heidelberg: Springer-Verlag, 2012, pp. 168–178.

[72] M. D. P. Requejo, "The role of context in word meaning construction: a case study," *International Journal of English Studies*, vol. 7, no. 1, pp. 169–179, 2009.

[73] W. Chen, G. Aist, and J. Mostow, "Generating questions automatically from informational text," in *Proceedings of the 2nd Workshop on Question Generation (AIED 2009)*, 2009, pp. 17–24.

[74] J. Araki, D. Rajagopal, S. Sankaranarayanan, S. Holm, Y. Yamakawa, and T. Mitamura, "Generating questions and multiple-choice answers using semantic analysis of texts." in *COLING*, 2016, pp. 1125–1136.

[75] S. Bhingardive, "Introduction to word sense disambiguation."

[76] L.-C. Sung, Y.-C. Lin, and M. C. Chen, "An automatic quiz generation system for english text," in *Advanced Learning Technologies, 2007. ICALT 2007. Seventh IEEE International Conference on*. IEEE, 2007, pp. 196–197.

[77] D. Gildea and D. Jurafsky, "Automatic labeling of semantic roles," *Comput. Linguist.*, vol. 28, no. 3, pp. 245–288, Sep. 2002.

[78] I. E. Fattoh, "Automatic multiple choice question generation system for semantic attributes using string similarity measures," *Computer Engineering and Intelligent Systems*, vol. 5, no. 8, pp. 66–73, 2014.

[79] T. Effenberger, "Automatic question generation and adaptive practice," Master's thesis, Masaryk University Faculty of Informatics, 2015.

[80] D. R. Ch and S. K. Saha, "Automatic multiple choice question generation from text: A survey," *IEEE Transactions on Learning Technologies*, 2018.