**Question Generator for Narrative Stories**

**Chapter I**

**Introduction**

Technology has been playing a great role in human lives and new developments in technology continue to arise. And in the broad field of automation, one of these is Natural Language Processing (NLP) or simply the process of making computers understand text or information in the form of human readable language.

Reading is a part of the four macro skills and isn’t the easiest to learn. And it is even harder comprehending what you’ve read. If the reader is a motivated reader, the case would be easier. However, reading, for most of us, is never really that fun. But reading can be an art in which most people should master. Assessing, what you’ve understood from what you’ve read, on your own can be a hard task. Hiring teachers, instructors and professors for comprehension assessment is rather a timely task. Thus, people thought of automating this process, Automated Question Generation. It is under the field of NLP since it involves text processing via computer technology.

Question Generation (QG) is the task of automatically generating question from various inputs such as raw text, database or semantic representation according to Paul in 2011. The generation of question needs information from a text or sentences. The goal of the researchers is to create an automated system that can take a narrative story as an input and produce an output questions for assessing a reader’s knowledge about the story. The system will also give the correct answer on the generated questions.

**Background of the Study**

In the past years, many researches had conducted about question generation. This researches give some hint and methodologies in generating question however, despite of this researches some problems are not yet been solve.

According to Simkin M. and Kuechler W. (2005), many instructors appear to use Multiple Choice (MC) tests as a preferred assessment tool. In teaching economics, for example, investigation estimate MC test usage for student assessments between 45% and 65 % (Becker & Watts, 2001; Siegfried & Kennedy, 1995). They sense is that the actual number is probably higher today because MC tests are so ubiquitous across the other disciplines (even in teaching compositional writing) and also because they are now used so frequently on entrance examinations, online tests, mass lecture courses and certification tests.

In the study conducted by Afzal and Pekar in 2010, they worked on generation of multiple-choice questions. They presented an unsupervised approach for Relation Extraction from surface-based patterns intended to be deployed in an e-Learning system for automatic generation of multiple choice questions. The researchers experimented with three different surface-based approaches and showed that PoS-based and verb-centred patterns achieve higher precision compared to untagged word patterns. They explored different ranking methods and found that the Chi-Square ranking method obtained higher precision than the other ranking methods. Afzal and Pekar employed two techniques: the rank-thresholding method and score-thresholding method and found that thresholding scores perform better.

In the year of 2010, Michael Heilman created another Question Generator (Automatic Factual Question Generation from Text). The system has three stages in generating questions; the first stage is to transform set of sentences into simpler declarative statement. The implementation includes operations for extraction and simplifying complex sentences and for resolving pronouns. The declarative sentence is turn into a set of questions by executing a series of well-defined syntactic transformations (WH-movement, subject-auxiliary inversion, etc.). Last process involved in Heilman’s system is to score and rank questions according to features of the source sentences, input sentences, question, and transformations used in generation. As a result in his system, factual statement extractor provided more fluent and correct outputs and better coverage. The supersense tagger that was use was more accurate at predicting semantic type labels of nouns than name entity recognizer, statistical ranking improved the acceptability of top-ranked questions, and user study spent less time and mental effort in creating factual questions. Heilman recommended having an alternative representation for QG, information extraction for QG and other NLP transformations such as sentence compression, sentence fusion, or paraphrase generation in transforming set of sentences into simpler declarative statement.

Agarwal and Mannem in 2011 also developed a QG that can generate gap-fill questions for content in a document. Gap-fill questions are fill-in-the-blank questions with multiple choices (one correct answer and three distracters) provided. In generating gap-fill questions, Agarwal and Mannem used sentence selection that involves identifying informative sentences in the document, which can be use to generate gap-fill questions. Those sentences are, and then processed in the key selection stage to identify the key on which to ask the question. In the final process, the distracters for the selected key are identified from the given chapter by searching for words with the same context as that of the key. They evaluated the system on two chapters of a standard biology textbook.

Despite of these wonderful researches, there still a need for improvement of QG. There still some problems encountered from the past researchers that Question Generation for narrative stories may solve.

**Statement of the Problem**

How will the system be effective and efficient in generating questions and answers?

* What are the possible solutions for the problems of related works:
* Handling of quoted statements and words
* Question difficulty control
* Optimization
* Anaphora resolution
* Optimized and accurate parser to use
* Generating a correct and accurate answer-key

**Objective**

The main goal of the research is to develop an effective and efficient system that generates questions and answers from a narrative story.

The researchers also aim to provide solution for some of the problems of recent works regarding Question Generation by proposing an algorithm to handle quoted statements and resolve anaphora; finding a better parser that can accurately identify proper nouns and common nouns.

Proper information extraction is aimed in this study for it is a factor in order to get accurate answers for evaluating the user’s comprehension.

**Theoretical Framework**

The study’s bases of framework are the following:

Question Generation via Overgenerating Transformations and Ranking by Heilman and Smith (2009), they use techniques from summarization and sentence compression in stage 1, various simplifying transformations are performed in this stage to remove phrase types such as leading conjunctions, sentence-level modifying phrases, and appositives. In the next stage, the question transducer takes as input a declarative sentence and produces as output a set of possible questions. It identifies the answer phrase, which may be targets for WH-movement and converts them into question phrases. In the current system, answer phrases can be noun phrases or prepositional phrases, which enables who, what, where, when, and how much questions. The question transducer aims to overgenerate grammatical, though perhaps irrelevant or unimportant, questions. These rules encode a substantial amount of linguistic knowledge. They (1) mark phrases that cannot be answer phrases, due, for example, to island constraints; (2) remove each answer phrase and generate possible question phrases for it; (3) decompose the main verb; (4) invert the subject and auxiliary verb; and (5) insert one of the question phrases and post-processing. And the last stage is the question ranker, they used discriminative reranker (Collins, 2000), speciﬁcally based on a logistic regression model that deﬁnes a probability of acceptability.

**Conceptual Framework**

This part of the chapter will include an overview of the system’s processes. A narrative text will be the assumed input for the system. Before the actual question generation process, the input will first go through sentence splitting and indirect-direct speech conversion. This process is required for the system to handle quoted statements better. After these processes, the input text will proceed to the parsing process and anaphora resolution. The parsing and anaphora resolution will produce a word-sentence table. Part-of-speech (POS) tagging and Named Entity Recognition (NER) will be utilized and update the contents of the word-sentence table. A scoring method will then be applied to every word in the table to determine sentences that are relevant enough to generate questions from. After these methods, the Question Over-generation process will initiate. Before displaying output questions, a ranking method will be applied. Questions with too high or too low ranks will be disregarded.

Questions

Derived output

Narrative text

Sentence Splitting

* Conversion from indirect speech to direct speech
* Parsing
* Anaphora resolution
* POS tagging and NER
* Scoring
* Overgenerartion

Ranking Splitting

**Scope and Limitation**

This section will mainly talk about the different capabilities

To achieve desirable results there are several input constraints: input must be in the form of a narrative story, and the input must have a correct spelling and grammatical structure. The researchers limited the study to the shallow question generation domain. There are three types of shallow questions to be generated by our system. These are the following:

* Factoid Questions or WH questions (i.e. what did Ria eat today? Who killed Johnny?),
* Yes-No Questions (i.e. Is Jude the brother of Jade?), and
* Multiple Choice Questions

The system cannot generate cloze questions and other deep questions that require complex answers (e.g. inference questions, list or enumeration questions and essays).

**Significance of the Study**

Although it was mentioned earlier that question generation have been going under study for quite a time, there are still existing problems that bounds a perfect question generator. The focus of this study will be the problem of Anaphora Resolution and handling of quoted sentences. Automatic Question Generation will play an important role in the field of learning. The following are the people that may benefit in this research.

**Educators**

Through this study, educators can use the system developed as an assessment tool for their students’ comprehension about a certain literature. Educators can also save time and effort in making questions.

**Students**

The study may help the students in assessing themselves on how deep they understand a certain narrative story by answering the questions generated by the developed system.

**Researchers**

Other researchers may use the study as their basis in solving other NLP problems. The study might contribute solutions and answers for future research on the area of NLP especially in the field of Question Generation and can open another research relevant to the study.

**Bibliography**

Paul P. (2011) Question Generation, <http://www.questiongeneration.org/>

Heilman M. (2010). Automatic Factual Generation from Text [PDF Document],

<http://www.cs.cmu.edu/~mheilman/>

Agarwal, M. & Mannem, P. (2011). Automatic Gap-fill Question Generation from Text Books [PDF Document],

Simkin, M. & Kuechler W. (2005). Multiple-Choice Tests and Student Understanding: What is the Connection? [PDF Document]

Afzal, N. & Pekar, V. (2010). Unsupervised Relation Extraction for Automatic Generation of Multiple-Choice Questions [PDF Document]