# Question Level based Opinion Generation in Web based Interactive Systems

Rajimol R
Department of Computer Science and Engineering
Mangalam College of Engineering
Kottayam, India

Vinodh P Vijayan
Department of Computer Science and Engineering
Mangalam College of Engineering
Kottayam, India

**Abstract**: The greatest challenge in opinion generation system is to understand the meaning of the question asked and answering the question correctly based on the knowledge level of users. The solution for this is the question level based opinion generation system, which generate opinions by understanding the in-depth meaning of the question. Opinions are extracted from reviews on different websites. System uses lexical based algorithm for finding the semantic relation between words and thus obtains the keyword used to retrieve the opinions. System also uses transformation based classification to learn the system. Transformation classification is based on rules and is ends up when no more transformation is possible for the data.

**Keywords**: Opinion generation system, Lexical based algorithm, Machine learning, Transformation based classification, Keyword identification.

## 1. INTRODUCTION

Opinion generated for multiple levels of users are a challenging issue in an interactive opinion generation system. The question level based opinion generation system helps to generate opinions with intelligence. Some of the related works generates opinions directly from the set of reviews. But here the system generates opinions according to the level of question asked by the user. It uses lexical driven algorithm to find the semantic relation between the sentences and also make use of knowledge bases to make the system intelligent by learning.

The system collects reviews from different sites as similar to other opinion generation system but it also learns the system in order to answering different levels of questions. It also perform sentiment analysis and ranking to the collected reviews [1]. System performs parsing technique in order to identify the keyword that strongly represents the question. Then it performs depth value calculation. Based on the value, system retrieves related opinions and presents it to the user.

System first analyzes the question asked by the user. Then it performs appropriate keyword identification and depth calculation. This function can be accomplished by using the lexical driven algorithm. Lexical analysis classifies the sentence in to tokens or group of characters, which helps to identify the keywords easily from a sentence.

After question analysis and keyword identification, the system generates appropriate opinions for the user.

System uses transformation based classification for learning the system. Thus it helps to make the system intelligent. Usually the transformation classification starts with simple solutions and then its process like a cycle. It ends up when no more transformation is possible.

## 2. RELATED WORKS

Automatic expressive opinion generation system helps to generate expressive opinions rather than simple and common opinions [1]. This system generates expressive opinions for enjoyable conversation. It takes opinions from reviews and ranked them. It also performs sentiment analysis to classify the opinions into positive and negative opinions. Uniqueness of generated opinion is based on adjective frequency.

Topic relevance based opinion generation [2] focuses on the problem of searching opinions over general ideas. It applies ranking to documents which contains individual opinions. From the ranked documents opinions are retrieved.

Micro opinion generation is an unsupervised approach [3]. It deals with generating concise summary of opinions with maximum of 2 to 5 words. Since the opinions provided as summaries or small content, it helps the user to easily understand the generated opinions. Major idea is to use existing words in original text to compose meaningful summaries.

OPINE is a novel method for mining reviews in order to build a model with product features. It is an unsupervised information retrieval system and it uses relaxation labeling as the key concept [4]. It extracts the noun phrases from reviews and retains the nouns whose frequency greater than threshold value.

The talent to detect expected or valuable things is called serendipity. Wikipedia forms a graph like structure consists of article nodes and category nodes [5] called triplets. This uses supervised data for processing.

MONEA is efficient development platform architecture for multi functional robots. Firstly, it embodies the meta architecture for networked-robots. Secondly, it provides some development models. Finally, it doesn't require heavy weight middleware [6].

Speech based interactive information system is basically a question answering system. System works in two modes [7]. When user asks a question, it switches to retrieval/QA mode and generate answers. Otherwise it works in system recommendation mode.

Schema is a multi party interaction human robot [9]. System consists of multi user tracking and fusion module, multi party dialogue manager and virtual human and robot control module.

Opinion mining deals with extracting required data or information from large dataset, it can be a web. Sentiment analysis classifies the opinions into positive or negative polarities [8]

## 3. OPINION GENERATION SYSTEM

Automatic expressive opinion generation system is used to produce enjoyable conversations rather than simple or common sentences. In order to make deeper knowledge about the question, the system uses lexical driven algorithm and machine learning concepts in it. Opinion generation is a classic Natural Language Processing (NLP) problem. The opinion generation system also needs to be intelligent in order to realize the question and to generate suitable opinions for the question.

### 3.1 Classes, processing and context of questions

Question classes can be of different types such as factoid typed questions, non-factoid typed questions and some questions may need deeper knowledge in order to answer it. Question processing deals with the understanding of the questions or semantics of the questions given by the user. It also needs to understand the context of the questions before processing it.

### 3.2 Opinion extraction and formulation

Opinion extraction is basically depending on the level of question asked by the user. Simple extraction may be enough for certain questions. It may want the partial answers to be extracted from various sources and combine them. At the same time, also needs to make the results of the opinion generation system to be as natural as possible.

System uses a Transformation-Based Learning (TBL) algorithm to induce rules from the training data. TBL is a technique used to learn the system by providing set of rules that converts simple solutions into all possible set. Rules are expanding by using rule sets. The algorithm greedily selects the rule that reduces the error rate the most.

The system uses a parser, which shows the structure of the English Sentence and also performs deeper logical analysis. It can be used for relation extraction, where semantic relationships between the words are extracted. It is also responsible for keyword extraction, which extracts the important

## 4. ARCHITECTURE

The system receives user questions as input. Then it performs question and topic analysis in order to produce appropriate answer for the question.

System uses pattern based extraction of words for deeper knowledge about the question with the help of lexical driven algorithm and transformation based learning.

It also uses certain hypothesis to generate rules and to make the system more intelligent. Finally the system generates appropriate opinion for the question.

System can also perform additional opinion generation by using the reviews and the uniqueness of the opinions. These can be determined by the adjective frequency of the sentence as in automatic opinion generation systems [1].

Knowledge bases and data sources are used for generating opinions. Knowledge base is trained by using some set of rules, so that the system can give intelligent answers for different level of questions.

Architecture mainly consists of three modules: question analysis module, keyword identification and depth calculation module and opinion generation module.

### 4.1 Question analysis module

In the question analysis module, system analyzes the question asked by the user. Then it splits the question into different parts in order to identify the keywords in the questions.

Question analysis module also identifies whether the question belongs to factoid or non factoid type[1].

### 4.2 Keyword identification module

Keyword identification and depth calculation module first analyze the words in the question. Then by using lexical driven algorithm, it finds the semantic relation between the words and thereby identify the keywords from the question. Then it performs the depth calculation of the corresponding keyword and stored this value in the database for opinion retrieval.

### 4.3 Opinion generation module

Opinion generation module first search the opinions for corresponding questions in database and knowledge base based on the depth value of the keywords, So that the system can give opinions for different level of questions. It also uses transformation based classification for training or learning the system. Based on the depth of question, system generates appropriate opinions and output it to the user.
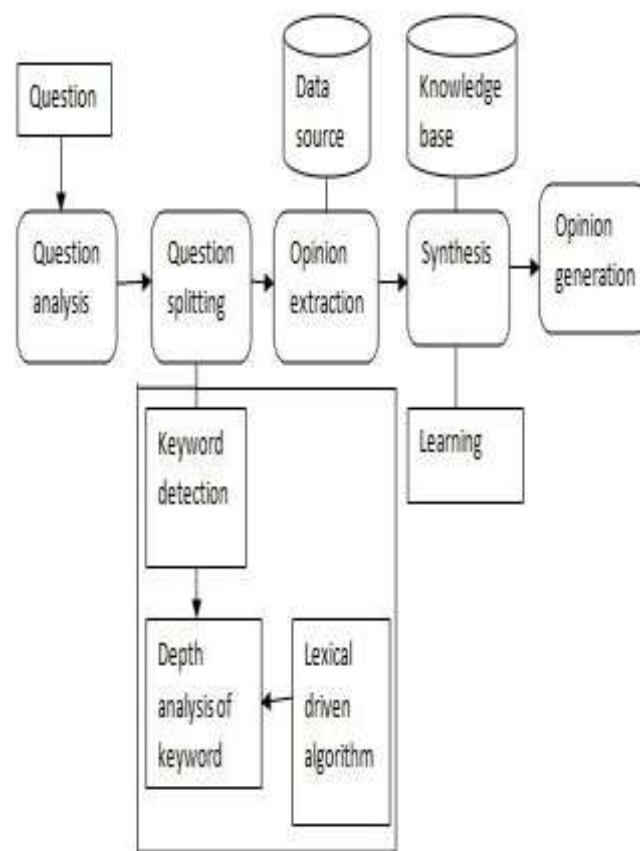


Figure 1. Architecture of question level based opinion generation system.

## 5. EXPERIMENTAL RESULTS

Experiments are conducted on Intel Core i3 processor with CPU of 2.40GHz. Data are extracted from different review sites. 14000 sentences are extracted from 20000 reviews. Comments to posts are not considered. To avoid mistakes done by sentence tokenizer, the length of the sentence must be < 200 characters.

Opinion generation systems are meant for generating opinions to users for their question. But most of the systems don't care about the satisfaction of the user. Satisfaction of the user can be

measured in many ways. Most preferable way is to measure the acceptance of opinion and response time of opinions.

The most important consideration for question level based opinion generation system is that whether they capable of providing opinions for complex or intelligent question also.

Question level based opinion generation system shows better performance in terms of acceptance of sentence and response time.

## 5.1 Acceptance of opinions

Most of the opinion generation system faces a problem for providing acceptable opinions that requires some knowledge. The problem is how to find the exact meaning of the question. This system solves this problem by identifying the keywords from the question. It also calculates the depth value for the keywords, so the system can easily understand the meaning and can provide more acceptable opinions.

Question level based opinion generation system produces more acceptable sentences. It uses learning techniques in order to retrieve and update the knowledge. So the system can produce more appropriate and context relevant sentence as output
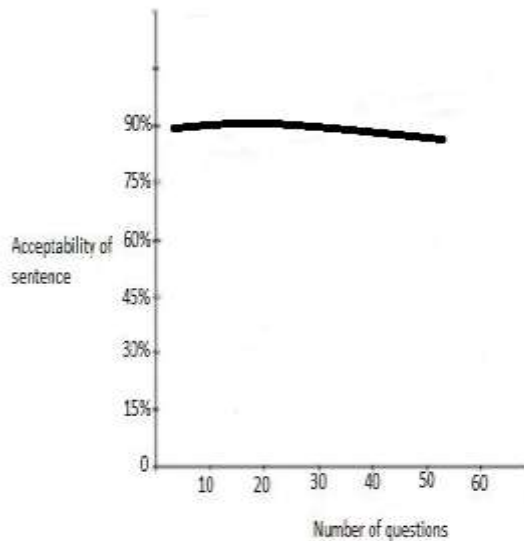


Figure 2: Acceptability of opinion

For measuring the performance of acceptability, plot a graph with number of questions in x-axis and acceptability of sentence in y-axis. Number of questions asked by the user can vary from 0 to any number. System shows better performance in any type of question, that is for factoid and non factoid typed questions.

System also checks whether the opinions given to the user is satisfied or not. It also provides better acceptable sentence even for intelligent questions.

Thus the system provides more acceptable opinions for any type of questions and any number of questions.

## 5.2 Response time

The response time of the system is directly related to the complexity of the sentence. As level of question becomes difficult, it may take more time to generate the opinions. Question level based opinion generation system shows better results even for complex questions.

In general opinion generation systems, it feels difficulty to provide opinions within specified time. So this may reduces the efficiency of the system and thus affect the users satisfaction.

These problems are happened due to the difficulty to understand the meaning of the question and also due to the unavailability of the data.

But in question level based opinion generation system, it uses data source as well as knowledge base for generating appropriate opinions. Knowledge base is periodically updated by using the transformation based classification. This classification or learning continues until no more selection is required.

Thus the opinions corresponding to any type of questions is available all the time. So the system gives response within few seconds even for the questions that needs intelligence.
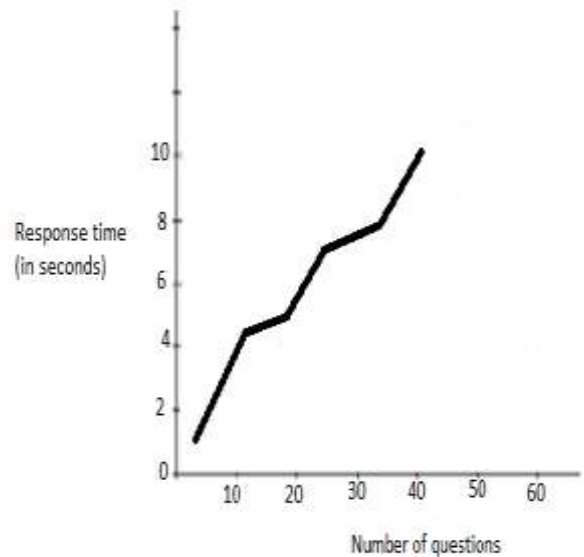


Figure 3: Response time of opinions

The response time of opinions can be measured by plotting number of questions on x-axis and response time in seconds on y-axis of a graph.

Here the number of questions can be varied from 0 to any number and to any level of question. The response time is plotted in seconds.

The obtained graph shows that the system gives answers for any number of questions within a short period of time (within seconds). System provides opinions within seconds even for different level of questions (for factoid, non factoid or complex questions).

Thus the question level based opinion generation system shows better performance in case of acceptability of opinions and response time of opinions. System also provides more accurate results as output.

## 6.  CONCLUSION AND FUTURE SCOPE

Question level based opinion generation system is a type of web based opinion generation system which answers to a user query based on the level of question. The system shows an improved level of intelligence by understanding the depth of question and answering it correctly to the expected level of user. System is tested on a varying level of sample questions normally asked by the users. The generated opinions will be

helpful for the user to obtain suitable knowledge. It also performs depth value calculation of words in order to find the semantic relation between the words. This opinion generation system generates meaningful opinions by processing the meaning of the question and related data.

In future, it can generate audio based opinions as output.

## 7. REFERENCES

[1] Yoichi Matsuyama, Akihiro Saito, Shinya Fujie, and Tetsunori Kobayashi,"Automatic Expressive Opinion Sentence for Enjoyable Conversational Systems",in IEEE/ACM transactions on audio,speech and language processing, 2015 vol 23.

[2] Min Zhang , Xingyao Ye ," A Generation Model to Unify Topic Relevance and Lexicon-based Sentiment for Opinion Retrieval", in State key lab of Intelligent Tech.& Sys,2008.

[3] D. Ferrucci, E. Brown, J. Chu-Carroll, J. Fan, D. Gondek, A. A.Kalyanpur, A. Lally, J. W. Murdock, E. Nyberg, and J. Prager *et al.*,"Building Watson: An overview of the DeepQA project," *AI Mag.*, vol.31, no. 3, pp. 59–79, 2010.

[4] Ana-Maria Popescu and Oren Etzioni, "Extracting Product Features and Opinions from Reviews", Department of Computer Science and Engineering, University of Washington.

[5] Y. Noda, Y. Kiyota, and H. Nakagawa, "Discovering serendipitous information from Wikipedia by using its network structure," in *Proc.ICWSM*, 2010.

[6] T. Nakano, S. Fujie, and T. Kobayashi, "Monea: Message-oriented networked robot architecture," in Proc. IEEE Int. Conf. IEEE Robotics Autom. (ICRA), 2006, pp. 194–199.

[7] T. Misu and T. Kawahara, "Speech-based interactive information guidance system using question-answering technique," in *Proc. IEEE* Int.Conf. Acoust., Speech, Signal Process. (ICASSP '07), 2007, vol. 4, pp.IV–145.

[8] B. Pang and L. Lee, "Opinion mining and sentiment analysis,"*Foundat.Trends Inf. Retrieval*, vol. 2, no. 1–2, pp. 1–135, 2008.

[9] Y. Matsuyama, K. Hosoya, H. Taniyama, H. Tsuboi, S. Fujie, and T.Kobayashi, "Schema: Multi-party interaction-oriented humanoid robot," in ACM SIGGRAPH ASIA Art Gallery Emerging Technol.:Adapt., 2009, pp. 82–82, ACM.