An Automatic Answering System Using Template Matching For Natural Language Questions

Pachpind Priyanka P¹, Bornare Harshita N², Kshirsagar Rutumbhara B³, Malve Ashish D⁴

¹,²,³,⁴BE Comp S.N.D COE & RC, YEOLA

Abstract- Using computers to reply normal language question is an good-looking plus hard trouble. Usually such trouble are handle under two group Open Domain trouble and Close Domain trouble. In this paper we are present a scheme that try to solve close domain trouble.During a close domain, answers to questions are not alive in the public area and so they cannot be search using a search engine. Consequently answers have to be store in a record by an manager . Then, the challenge is to distinguish the normal language question so that the solution could be matched to the particular answer in the record. We employ a pattern matching procedure to perform this matching. In this system our aim is to exercise this scheme with non-native English speakers, we housing a method to beat the mismatches we might meet due to spelling mistakes. The scheme is developed so as to the questions can be put using small messages as of a cellular phone and so the system is designed to appreciate SMS language in addition to English.

Keywords—Mobile, GSM Modem, SMS, FAQ, Automatic Answer, Template Matcher, NLP.

I. INTRODUCTION

At the present days using computers to respond User’s questions is becoming an exciting new impression with the increased use of computers. In this mechanisms allow users to raised questions in a normal talking and give a true answer. Accepting user questions in normal languages requires normal Language Processing. Normal Language Processing acting a huge role in the ICT and Question Answering (QA) systems. Normal language processing is the come close to analyzing text based on both a set of theories and a set of technologies. It will become vital to be able to ask queries and obtain answers, using usual language terms, rather than the keyword based needs of users as they will provide an accurate, faster, suitable and helpful way of giving answers to user questions.

We have adopted in this project an automated FAQ (Frequently Asked Question) answering system that replies with pre-stored answers to user questions asked in regular English, rather than keyword or sentence structure based retrieval mechanisms.

This is achieved using a pattern matching technique with some other mechanisms like disemvoweling, matching synonyms, etc. The normal language processing technique developed for FAQ retrieval does not analyze client query. As an alternative study is applied to FAQs in the database. Thus, the work of FAQ retrieval is reduced to keyword matching creating an fantasy of cleverness. The system is both growing and handy. Growing because its question answering ability improves as more questions are asked and new FAQ entries are formed. It is portable because the system could be used for any problem domain (closed) by changing the knowledge base. Typically, there are two types of question answering systems: (1) closed-domain question answering that deals with questions under a specific domain, and can be seen as an easier task on one hand as the NLP systems can exploit domain-specific knowledge frequently formalized in ontology but harder on the other as the information is not generally available in the public domain; and (2) open domain question answering that deals with questions about almost everything, and can rely only on general ontology and world knowledge. On the other hand, as mentioned earlier these systems usually have much more data available in the public domain from which to extract the answer.

As depicted in Figure 1, there exist two methods [1][2], for coming up with an suitable answer for a user question and they are AI method[2] and FAQ search method. The AI method focuses on answer generation by analyzing questions and creating an “understanding” of the question. This requires difficult and higher linguistic study programs.

II. LITERATURE SURVEY

The artificial intellect approach uses an ontology based knowledge base in order to get the user question and then query the FAQ database. The techniques consider the similarities in work, sentence length, word order or distance of identical work of the user question to decide whether it is equivalent to an FAQ. Values are assigned to reflect the similarity and then the sum of these values is used for comparison.
This approach works rather badly for short questions and when the query and FAQ use different wording to carry the same meaning. A closed domain question answering system based on the FAQ search method joined with template matching has been modified to service users in our system.

The QA system can better satisfy the needs of users as they will provide an accurate, faster, suitable and effective way of giving answers to user questions. The approach we have adopted in this project is an automated FAQ (Frequently Asked Question) answering system that replies with pre-stored answers to user questions asked in usual English, rather than keyword or grammar base retrieval mechanism. This is achieved using a pattern identical method with some additional mechanisms like disemvowelling, identical synonyms, etc.

III. RELEVENCE OF WORK

In this section we depict the working of our system. The system can be subdivided into three main modules: (1) pre-processing, (2) question template matching, and (3) answering. Figure 2 shows the system architecture of the question and answering system. Each module is described in detail in the following subsections.

In this project the following modules are described:

1) Preprocessing,
2) Question template matching
3) Answering

IV. WORKING

A. Pre-Processing Module

Pre-processing module mostly consists of three operations: (1) converting SMS abbreviations into common English words, (2) removing stop words, and (3) removing vowels. Because the system is normal to process texts with both normal and SMS languages it is needed to change the SMS abbreviations with the parallel English words before processing user questions advance. This is complete by referring to pre-stored usually use SMS abbreviations. Stop words are the language that put in no result to the connotation of a condemnation even if they are removed. Removing stop words is done to raise the use of the system by save time and disk space. Examples of stop words are the, a, and, etc. Next step in this module is to remove vowels from the text to feel spelling mistakes. This process is called disemvoweling which will be discuss in details in coming section.
B. Question-Template Matching Module

The pre-processed text is coordinated against each and every pre stored template awaiting it finds the best coordinated template with the received text. In order to do this, templates are shaped according to a specific syntax and the facts are describe in part IV. more in this module, words that are careful to have synonyms are referred in a synonym file. This synonym file can be adapted according to the related domain and are efficient from a criterion database [6]. It is worth note that the templates here are for question and not intended for answers. The major target of this system is to see the closest pattern that equal the query we include received from the user.

C. Answering Module

Each and every template expressive a question are pre stored in a database with its answer, now when the best coordinated template for the question is found, the matching answer will be returned to the end user.

V. SOFTWARE REQUIREMENT SPECIFICATION

A) Operating Environment

- Operating System- Windows XP.

B) Design & Implementation Constraints

- as long as such answering system that can attract more and more customer.
- Reducing cost for system.

Graphical user interface provides administrator for easy interface with software.

C) Assumptions And Dependencies

- Server ask for that allow the official user to act together with the scheme [4].
- Answering system that includes textual sms data .
- Security for authorized administrator.

D) Hardware Interface

- RAM- Min 256 MB.
- Space on HDD- Min 10GB.
- USB GSM Modem with good specification .
- Mobile phone.

E) Software Interface

- Java 1.5 on words
- Net Beans
- My SQL
- Wamp Server

F) Communication Interface

- Mobile
- Computer

VI. CONCLUSION

The proposed system provides user the new technique of information assembly the final result is a elegant easy to use automatic answering system with the facility of detecting and answering questions asked in English or SMS language. You can basically use the system to augment customer alertness of yourself and your product or service; Give your existing user offers and promotions without expenses time.

REFERENCES