

УДК 62-1

Kuanyshbay D., Orynbetov Z., Kariboz D.
Suleyman Demirel University, Kazakhstan, Kaskelen

VOICE ASSISTANT FOR MOBILE SYSTEMS WITH NATURAL LANGUAGE PROCESSING

Abstract: In order to handle some tasks by voice recognition it is really important that application will understand what user said. Which means that application must take a sentence that was entered by user and analyze and handle some tasks or etc. This paper is about natural language processing and analyzing sentences.

In this article, we have briefly gone through the problems of NLP and some of the common techniques used in natural language analysis. These techniques are, however, very simple and fundamental. Many more complex and efficient approaches have been developed. However, in spite of these new developments, current state of the art is still capable of only limited tasks within restricted domains. Even though, work in the subject had began more than forty years ago, it is still in the very early stage of its development and there is definitely more to meet the eyes.

Key words: voice recognition, artificial intelligence, natural language processing.

Introduction

Language is the fundamental means of communication for human beings. Though simple and comprehensive as it may appear to human, it is in fact of utmost complexity when it comes to understanding in the part of the computer. Natural language processing (NLP) is the engineering of systems that process or analyze written or spoken natural language. Since most of human knowledge is recorded in linguistic form, enabling computers to understand natural language would allow it to access all this knowledge. However, even though this idea had came across more than forty years ago, little has been achieved. Though relatively successful systems for limited applications have been made, we have yet been able to achieve, with much success, a system with general NLP. In this article, we will investigate the problem which is hindering the progress, the general process of NLP as well as some common techniques used.

The field of computer science that deals with designing computer systems that can recognize spoken words is called voice recognition. Nowadays, voice recognition system is topical technology. Despite that voice recognition system is really complicated and complex technology, it keeps improving and already being used nowadays. Below are some good examples of where voice recognition can be used:

Automated phone systems- Many companies today use phone systems that help direct the caller to the correct department. If user have ever been asked something like "Say or press number 2 for support" and user say "2," you used voice recognition.

Google Now- Google voice is a service that allows users to search and ask questions on your computer, tablet, and phone.

Siri - Apple's Siri is another good example of voice recognition that helps answer questions on Apple devices.

Car Bluetooth - For cars with Bluetooth or Handsfree phone pairing you can use voice recognition to make commands such as "call my wife" to make calls without taking your eyes off the road.

Our purpose is to create a voice assistant that will understand the user saying and handles some tasks. Which means there should be an algorithm that based on natural language processing.

Natural Language Processing

What's the problem?

The main problem with natural language is perhaps its ambiguity. Take the following sentence for example: "Time flies like an arrow." Does it mean time flying or creatures known as

"time flies" are like an arrow? The computer would need to be able to identify whether "flies" refers to a noun or a verb in order to understand the actual meaning of the sentence. In order for a system to deal with these kinds of ambiguity, it has to know a lot about the world and be able to use that knowledge to the intended effects. Current AI technology has not yet developed to the stage where such large quantity of knowledge could be handled efficiently and effectly. However, by restricting the scope of the language to that required to handle a limited task in a limited domain, performance systems capable of useful natural language communications can be achieved.

An Overview of Natural Language Analysis Process

Natural Language Analysis can be said to consist of 3 main stages:

- a. Syntactic Analysis: where the input utterance is being checked to ensure that its syntax is correct and structured representations of the possible parses are generated.
- b. Semantic Analysis: where the initial representations of the meaning of the sentence are obtained from the possible parse.
- c. Contextual Interpretation: where the initial representations are analyzed based on contextual and world knowledge and a final meaning representation is obtained.

Syntactic Analysis. Syntactic Analysis, perhaps the most well-developed area of NLP, deals with the syntax of natural language. In Syntactic Analysis, a grammar is used to determine what sentences are legal. The grammar is being applied using a parsing algorithm to produce a structure representation, or parse tree. Many different grammars and parsing algorithms have been developed, but we will consider only the following 2 simple methods:

- 1) Context-Free Grammar
- 2) Top-Down Parser

Context-Free Grammar. The simplest style of grammar, and a widely used one, is perhaps the Context-Free Grammar (CFG). Consider the following sentence: "The cat eats the rice". A parse tree showing the composition can be constructed as shown in the figure below.

The parse tree breaks down the sentence into structured parts so that the computer can easily understand and process it. In order for the parsing algorithm to construct this parse tree, a set of rewrite rules, which describes what tree structures are legal, must be available. These rules say that a certain symbol may be expanded in the tree by a sequence of other symbols. The rewrite rules of this example is as follows:

S -> NP VP
NP -> DET N | DET ADJ N
VP -> V NP
DET -> the
ADJ -> big | fat
N -> cat | cats | rice
V -> eat | eats | ate

Context-Free Grammars are simply grammars consisting entirely of rules with a single symbol on the left-hand side of the rewrite rules. The obvious advantage of CFG is that it is simple to define. Many of the grammars used for NLP systems are CFG, as such they have been widely studied and understood and hence highly efficient parsing mechanisms have been developed to apply them to their input.

However, CFG also have some severe disadvantages. Consider the above rewrite rules, since V can be replaced by both "eat" or "eats", sentences like "The cat eat the rice" would be allowed. Therefore, additional sets of grammar would have to be implemented for both singular and plural sentences. Moreover, completely different sets of rules would also be needed for passive sentences, e.g. "The rice was eaten by the cat". This means that an extremely large set of rules would have to be created which makes it difficult to handle. Many different grammar formalisms like the unification grammar and the categorial grammar have been developed to capture the rules of syntax more concisely, but we won't go into them.

Note that sentences which are syntactically sound but makes no sense are acceptable by the parser. For example, "The rice eats the cat" is totally nonsensical yet acceptable by the parser. Thus the next stage (Semantic Analysis) would be needed to filter out these nonsensical parses.

Top-Down Parser. A simple parsing algorithm used is the Top-Down Parsing Algorithm. Here, the parser starts with the S symbol and attempts to rewrite it into a sequence of terminal symbols that matches the classes of the words in the input sentence until it consists entirely of terminal symbols. These are then checked with the input sentence to see if it matched. If not, the process is started over again with a different set of rules. This is repeated until a specific rule is found which describes the structure of the sentence. An improvement to this algorithm is to use a depth first search with backtracking strategy. Here, when the first terminal symbol in the grammar is reached, we immediately check whether the first word of the sentence belongs to the same category as the terminal symbol. If it is, then the process is continued for the rest of the sentence. If not, the process is backtracked and an alternative rule is tried.

The main problem with this type of parsers is inefficiency, as the search process has to be repeated if an error occurs. Other kinds of more efficient parsers used in NLP includes the transition network parsers and the chart parsers.

Semantic Analysis. The process of semantic analysis is a complex one and will not be tackled in this article. The main function and importance of the process is discussed here. Basically in semantic analysis, we try to obtain an initial representation of the meaning of the sentence from the possible parses of the syntactic analysis stage. This initial representation of the meaning of the sentence is known as the logical form.

Natural language is ambiguous by nature. A simple sentence can be interpreted in many different ways. In order for the computer to process the sentence, it needs to know the exact meaning of the sentence. Thus, the logical form is needed as an intermediate unambiguous representation of the meaning of the sentence.

At this stage of the analysis, only the context-independent meaning of the sentence is derived. By this, we mean that the sentence is treated as a stand-alone line by itself, disregarding any knowledge of previous sentences or the world. Because of this, certain sentences which are dependent on knowledge of previous sentences will not be fully interpreted. That is when the last stage comes in. Anyway, once the semantic relationships are determined, certain words senses which are found to be impossible are then eliminated from consideration and the remaining choices are passed to the next stage.

Contextual Interpretation. Contextual Interpretation is the last stage of the analysis. In this stage, the meaning is elaborated based on the contextual and world knowledge. Here, the logical form is being mapped to the final knowledge representation (KR) language. The main function at this stage is to work out referents of expressions. For example, the pronoun "she" in the sentence "She left for school" refers to a particular girl, but we are unable to establish her identity from the sentence. However, we can usually find out more about her identity from previous sentences. Another task done at this stage of analysis is to try and determine the goals underlying utterances. By knowing the goal of the utterance, a more appropriate and meaningful reply could be given. For example, when someone asks about the score of a football match, it can be assumed that they will also be interested in the scorer, and thus the scores as well as the scorers could be returned.

References:

- <http://www.computerhope.com/jargon/v/voicereco.htm>
- http://www.webopedia.com/TERM/V/voice_recognition.html
- http://www.mind.ilstu.edu/curriculum/protothinker/natural_language_processing.php
- http://www.doc.ic.ac.uk/~nd/surprise_97/journal/vol1/ctt/

Куанышбай Д., Орынбетов З., Қарибоз Д.

Магистранты университета имени Сулеймана Демиреля, Казахстан, Каскелен

ГОЛОСОВАЯ ПОМОЩЬ ДЛЯ МОБИЛЬНЫХ СИСТЕМ ПРИ РАСПОЗНАВАНИИ ЕСТЕСТВЕННОГО ЯЗЫКА

Аннотация: В работе рассматриваются вопросы создания мобильного приложения для ускоренного введения слов или предложений на естественном языке (казахском языке) в мобильные телефоны или компьютеры на основе анализа нескольких вариантов ввода. Приложение может быть использовано в системах с речевыми технологиями. Таким образом, статья посвящена обработке естественного языка и анализу предложений.

В данной статье авторы постарались изложить проблемы, существующие в создании мобильных приложений и в использовании различных технологий при распознавании естественного языка. Предлагаемые авторами методы довольно просты, но, в то же время, фундаментальны, применив которые можно убедиться в их эффективности.

Тем не менее, несмотря на их новизну и востребованность, они могут решить лишь определенную часть задач в ограниченных областях. Работа в данном направлении ведется уже около 40 лет, но и сегодня эти исследования очень актуальны и требуют еще очень много усилий для достижения совершенства.

Ключевые слова: распознавание звука, искусственный интеллект, обработка естественного языка.

Куанышбай Д., Орынбетов З., Қарибоз Д.

Сулейман Демирель атындағы университеттің магистранттары, Қазақстан, Қаскелен

ТАБИҒИ ТІЛДІ ӨНДЕУ ЖӘНЕ ТАЛДАУДАҒЫ ДАУЫСТЫҚ КӨМЕК

Аңдатпа: Мақалада қазақ тіліндегі сөздерді немесе сөйлемді енгізудің бірнеше нұсқасын салыстыру негізінде мобильдік телефондарға немесе компьютерлерге жылдам енгізу мәселелері қарастырылады. Қосымша әртүрлі сөйлеу технологияларымен жабдықталған жүйелерде қолданылады. Сонымен, мақала табиғи тілді өндеу және талдау мәселелеріне арналған.

Бұл мақалада, авторлар мобильді қосымшаны құру кезіндегі табиғи тілді әртүрлі технологияларда қолданудың қиындықтары туралы мәлімет берген.

Авторлармен берілген әдістер анағұрлым оңай, бірақ фундаменталды, яғни оларды қолдану кезінде тиімді екеніне көз жеткізуге болады. Алайда, олардың жаңалығы мен қажеттілігіне қарамастан, олар шектеулі облыстарындағы шешімдердің тек белгілі бір бөлігін шеше алады. Бұл бағыттағы жұмыстар 40 жыл бойы жасалып келе жатыр, бірақ бұл зерттеулер бүгінгі күнге әлі де өте өзекті және жетілдіру үшін көп күш-жігерін талап етеді.

Кілт сөздер: дауыстану, зерделі жүйе, табиғи тілді өндеу, дамыту, тәжірибе.