

Linux Command Tutor

Vidhate Shailendra, Ghumare Karishma, Jangam Priyanka, Jivani Reshma

Professor, Dept. of Computer Engineering, MET BKC, Maharashtra, Nashik

Student, Dept. of Computer Engineering, MET BKC, Maharashtra, Nashik

Student, Dept. of Computer Engineering, MET BKC, Maharashtra, Nashik

Student, Dept. of Computer Engineering, MET BKC, Maharashtra, Nashik

Abstract – The aim of Natural Language Processing is to interpret the English sentence and perform appropriate actions. This concept of NLP is used in our project which will help beginners to use our application for the ease of open source. Natural Language Process in Open Source is a desktop application for all the users who are willing to work in an open source environment. The intended purpose of this project is that the users who are not familiar with the open source commands will be now able to use it very efficiently without any need of learning the command. Thus, our system will provide user friendly platform just like that of proprietary operating system. It is mostly useful for fresher's who are unaware of the commands and scripts of open source OS which interrupts them to work with the system comfortably.

Key Words: Natural Language processing, Stoplist, Information Retrieval, Command mapping, Auto-suggestion

1. INTRODUCTION

The goal of the Natural Language Processing (NLP) is to design and build software that will examine in detail, understand, and generate languages that humans use naturally, so that eventually you will be able to address your computer as though you were speaking to another person. This goal is difficult to reach. Understanding language means, knowing what concepts a word or phrase stands for and knowing how to link those concepts together in a meaningful way. Natural language communication with computers has long been a major goal of artificial intelligence. There are several applications of NLP developed over the years[1]. They can be mainly divided into two parts as follows .

i. **Text-based applications:** This involves applications such as searching for a certain topic or a keyword in a data base, extracting.

Information from a large document, translating one language to another or summarizing text for different purposes.

ii. **Dialogue based applications:** Some of the typical examples of this are answering systems that can answer questions, services that can be provided over a telephone without an operator, teaching systems, voice controlled machines (that take instructions by speech) and general problem solving systems[2].

2. Literature Survey

Following are the existing systems:

• **Automatic summarization:** Automatic summarization is used to produce curtail version of text which contains information that is important or relevant to the user.

E.g.: abstraction of any document, article report etc. Also simplifying text by compressing sentences.

Radev et al. (2002) define a summary as “a text that is produced from one or more texts, that conveys important information in the original text(s), and that is no longer than half of the original text(s) and usually significantly less than that”. This simple definition captures three important aspects that characterize research on automatic summarization:

1. Summaries may be produced from a single document or multiple documents,
2. Summaries should preserve important information,
3. Summaries should be short [3].

• **Machine translation:** Machine translation is a computational linguistic which uses a software to translate text or speech from one natural language to another. Many MT systems across the globe have already been developed for the most commonly used natural languages such as English, Russian, Japanese, Chinese, Spanish, Hindi and other Indian languages etc. Followings are the existing machine translation systems and various approaches used in developing these systems:

1. Direct
2. Transfer-based

3. Statistical
4. Example-based
5. Hybrid[4].

• **Natural language generation:** NLG or Natural Language Generation is the process of constructing natural language outputs from nonlinguistic inputs.[5].

This means that NLG is automatic generation of text for user which can be used in website, weather report, graphics, newspaper, document etc.

E.g.: summarization of weather on daily basis, prediction of weather, summary of past events.

• **Natural language understanding:** NLG is the inverse of NLU (Natural Language Understanding) or NLI (Natural Language).

Interpretation. NLU maps from text to meaning. NLG is easier than NLU because a NLU system cannot control the complexity of language structure it receives as input [6].

• **Information retrieval (IR):** Information extraction is a complex procedure and process because of the collection of information based online as well as heterogeneous content available such as text, image, audio, video etc.

Information retrieval (IR) is finding material (usually documents)of an unstructured nature (usuallytext) that satisfies an information need from within large collections (usually stored on computers). Information retrieval is fast becoming the dominant form of information access, overtaking traditional database style searching [7].

• **Information extraction (IE):**

Its major goal is to transform the data from unstructured form into structured representation. The information extraction (IE) task can be expressed as to process the collection of texts which belong to a particular field and derive from each of them a previously defined set of name types, relations between them and events in which they participate. There are two approaches to the information extraction system design, namely knowledge engineering and automatic training approaches. Both of them have their own benefits and drawbacks and are applied depending on the resources available to the system’s designer [8].

2. SYSTEM ANALYSIS AND DESIGN

2.1 System Analysis

This section contain data flow i.e. how user gives input in natural language and our system gives output as appropriate command.

Data Flow

In this data flow diagram, data is the query which will be given

by user in sentence form. This sentence will then go through various modules and the end result will be the execution of appropriate Linux command which the user gave input in English sentence. Following is the data flow diagram.

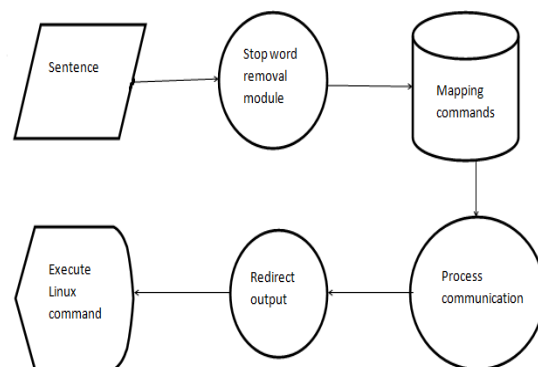


Fig -1: Data Flow Diagram

System Architecture

• **Stop word removal:** User input is given to stop word removal module. These are the words which are ignored at search time as they do not have any meaning which support for command mapping. Some of the common words are “the”, “is”, “at”, “which” and so on.

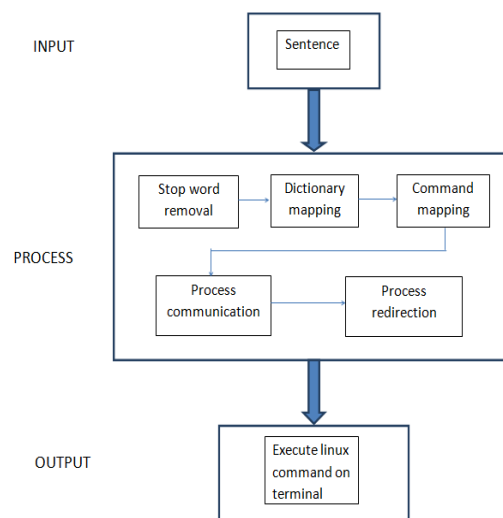


Fig -2: Block Diagram

•**Mapping words in dictionary:** The words left after removal of stop words are mentioned in our dictionary of words. When the match is found, the command associated with it is considered.

•**Process communication:** Our process communicates with terminal processes for command execution process.

•**Process redirection:** Then the output is redirected to our process.

•**Execute Command:** The result is then copied on our GUI

3. ADVANTAGES

- GUI based command execution.
- Enter the query in users (natural) language

4. ADVANTAGES

- Requires huge dictionary of commands.

6. CONCLUSIONS

Our system is a desktop application for all the users who are willing to work in an open source environment, i.e. they use open source operating system. The intended purpose of this project is that the users who are not familiar with the open source commands will be now able to use it very efficiently without any need of remembering or learning any type of command.

REFERENCES

- [1] <https://www.microsoft.com/en-us/research/group/natural-language-processing>
- [2] Preeti and Brahmaleen Kaur Sidhu, "NATURAL LANGUAGE PROCESSING" University College of Engineering, Punjabi University, Patiala, Punjab, India
- [3] Dipanjan Das André F.T. Martins (2007). "A Survey on Automatic Text Summarization", Language Technologies Institute Carnegie Mellon University.
- [4] G V Garje and G K Kharate "SURVEY OF MACHINE TRANSLATION SYSTEMS IN INDIA". Department of Computer Engineering and Information Technology PVG's College of Engineering and Technology, Pune, India.
- [5] Ehud Reiter and Robert Dale, "Building Natural Language Generation Systems", Cambridge University Press, 1999.
- [6] Daniel Jurafsky and James H. Martin, "Speech and Language Processing", 2nd ed., Pearson Prentice Hall, 2008.

[7] Christopher D. Manning Prabhakar Raghavan Hinrich Schütze. "An Introduction to Information Retrieval", Cambridge University Press Cambridge, England.

[8] Varsha pande, Dr. A. S. Khandelwal. "Information Extraction Technique: A Review" Nagpur, India