

# Methodology To Make Natural Language as Computer Programming Language

<sup>1</sup>Nalina H J, <sup>2</sup> Mr.Vinay T R

Department of Computer Science and Engineering, VTU, Bengaluru-562157

[nalina0309@gmail.com](mailto:nalina0309@gmail.com), [tr.vinay@gmail.com](mailto:tr.vinay@gmail.com)

**ABSTRACT**-Natural Language Processing(NLP) holds great promise for making computer interaction easier with the naïve users or non-computer programmers. We propose a methodology which converts natural language statements into a computer programming language. The non-computer programmers to express their ideas in computer language (programming language) is difficult. So, here we propose a methodology where the naïve users can easily express their ideas in their natural language on computers. We can either build a natural language compiler but, this takes time and effort for developing and learning. So our methodology helps the naïve users to simply express their idea in their own natural way and our process converts it into a computer program and generates the output.

**Keywords:** Natural Language, Computer Language, Naïve users, Computer Programmers, Natural language programming.

## 1. Introduction

Natural language processing (NLP) is a field of [computer science](#), [artificial intelligence](#), and [computational linguistics](#) concerned with the interactions between [computers](#) and [human \(natural\) languages](#). As such, NLP is related to the area of [human-computer interaction](#)[1]. Many challenges in NLP involve: [natural language understanding](#), enabling computers to derive meaning from human or natural language input; and others involve [natural language generation](#).

Modern NLP algorithms are based on [machine learning](#), especially [statistical machine learning](#). The paradigm of machine learning is different from that of most prior attempts at language processing. Prior implementations of language-processing tasks typically involved the direct hand coding of large sets of rules. The machine-learning paradigm calls instead for using general learning algorithms often, although not always, grounded in [statistical inference](#) to automatically learn such rules through the analysis of large [corpora](#) of typical real-world examples. A corpus is a set of documents that have been hand-annotated with the correct values to be learned[2].

Many different classes of machine learning algorithms have been applied to NLP tasks. These algorithms take as input a large set of "features" that are generated from the input data. Some of the earliest-used algorithms, such as [decision trees](#), produced systems of hard if-then rules similar to the systems of hand-written rules that were then common. Increasingly, however, research has focused on [statistical models](#), which make soft, [probabilistic](#) decisions based on attaching [real-valued](#) weights to each input feature[3]. Such models have the advantage that they can express the relative certainty of many different possible answers rather than only one, producing more reliable results when such a model is included as a component of a larger system.

The main difference between the Natural Language and the Computer Language should be known in order to understand the Natural Language Processing.

Natural Language Processing (NLP) and Programming Languages are both established areas in the field of Computer

Science, each of them with a long research tradition. Natural Language Processing involves a wide range of techniques that enable the automated parsing and processing of natural language[4]. This paper tries to address this gap by proposing a methodology that attempts to convert natural language text into computer programs. This methodology is particularly useful for non-computer programmers to express their ideas on computer[5].

Natural Language	Computer Language
1)It is generally peoples uses to share their feeling or for communication purpose.	1)It is a derived from Natural Language which is used to instruct machines.
2)It is understood and developed naturally by humans.	2) It is developed by humans in the form of programming language to coomunicate with machines.
3)Used to communicate between humans.	3) Used to communicate with the machines and human interactions.
4)Very Strict syntax is not followed.	4)Very Strict syntax has to follow.
5) The Lifespan is endless.	5)The Lifespan is limited.

Table 1: Natural language v/s Computer language

NLP is a huge field in computer science about language based interactions between computers and humans. For any person to express the ideas,he has to learn computer programming language. All most all compilers today are built using English like language. So all the prospective programmers should know English language[6]. For non-English persons, he has to learn English language then learn a computer programming language. So to help non-English prospective programmers we have to build a natural language compiler for their language or use a conversion tool, which can convert natural language statements into programming statements which is executable[7].

## 2. Existing system

Since, its inception in 1950s, NLP research has been focusing on tasks such as machine translation, information retrieval,

text summarization, question answering, information extraction, topic modeling, and more recently, opinion mining. Most NLP research carried out in the early days focused on syntax analysis.

Although the semantic problems and needs of NLP were clear from the very beginning, the strategy adopted by the research community was to tackle syntax first, for the more direct applicability of machine learning techniques. Another popular model for the description of natural language is production rule. A production rule system keeps a working memory of on-going memory assertions[8].

There are no tools available in the market to convert the natural language into a programming language but tools are available which convert natural language words into its equivalent English words. Nowadays only the computer programmers are able to develop or implement the ideas. The naïve users are just using the computer systems but they are not able to develop or express their ideas on computer. As they don't know the programming languages[9]. Also the naïve users must depend on the computer programmers to implement their ideas.

### 3. Proposed System

The proposed system explains to converting methodology for moving from naturally speaking language to machine understandable language. To develop a natural language compiler it takes time and effort. Also we have to train the people as how to write computer programs using natural language compiler. Here we do proposing a methodology which will convert a natural language statements into a computer language statements[10]. And this methodology can be further applied to any natural language.

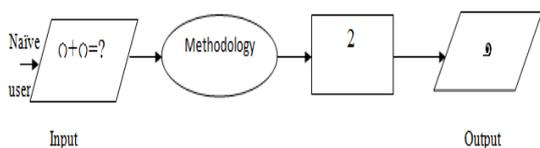


Fig 1 : High level design of Proposed Methodology

The above figure 1 describes the processing steps towards from natural to programming language. The user or the naïve user who will give the input in the form of natural language statements as 0 + 0 =? which will be given to our process, where the computer program will be executed in order to make the input understandable by computer. After the process is done inside the methodology the output will be generated in the English language, and again that will be converted into the natural language output.

#### 1. Algorithm

These steps give a processing operation from input to executable output.

- [1] Given statements is broken into a meaningful phrases (Lexical Analysis).

- [2] Phrases is compared with the stop words.

- [3] Phrases are analysed to parts of speech.

- [4] Operation to be performed with respect to input statements will be analysed.

- [5] Computer program is executed.

- [6] Output will be generated.

- [7] Later the generated output will be converted back to the natural language.

**Case Study:** The following statements is taken for case study.

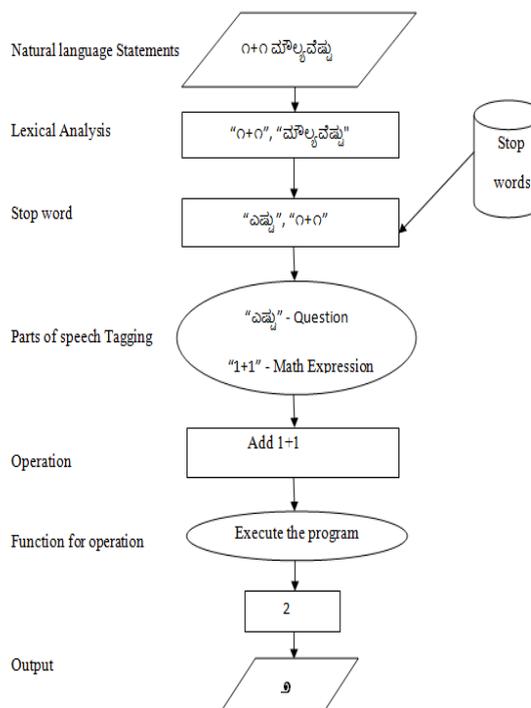


Fig 2 : A Simple Case Study

Fig 2 shows an example program that asks for a addition of 2 numbers in Kannada. Here the first statement is 0+0

|| which is a Natural Language statement. This statement is later broken into meaning small pieces called tokens which we call Lexical Analysis of the given statement 0,+,0,||, -. Next the word - || is reduced to - || using stop words. Now - || is analysed as questioning and 0 is converted into 1. So 0+0 is converted into mathematical expression 1+1. The operation will be performed to generate the output through the computer program as -2||. The last step is to convert back the above generated output(2) into natural language as (೨).

Starting with the natural language text as an input we would ideally like to generate a computer program. While this is still a long term goal, in this section we show how we can automatically generate computer program skeletons that can

be used as a starting point for creating procedural computer programs. Specifically, we focus on the description of a system for natural language procedural programming: Mathematical expression.

#### **4. Conclusion**

This methodology is developed in order to make a naïve users to express their ideas on computer. This may lead to bust in the number of programmers and many new novel ideas may be developed. This will help all the naive users to program and develop their ideas. Hence due to this the world of technology becomes easier. As there is no necessary to learn the programming language to program the ideas. So the common people can also work in the technology field by quickly adopting.

#### **REFERENCE**

[1] NLP (Natural Language Processing) for NLP (Natural Language Programming) by Rada Mihalcea, Hugo Liu, and Henry Lieberman. Springer-Verlag Berlin Heidelberg 2010.  
[2] From NLP (Natural Language Processing) to MLP (Machine Language Processing) by Peter Teu, Udo Payer and Guenter Lackner. Volume 3685 of Lecture notes in computer science. Springer (2012).

[3] Jumping NLP Curves: A Review of Natural Language Processing Research by Erik Cambria, Bebo White. IEEE Computational intelligence magazine | may 2014.

[4] Natural Language Processing Laboratory Plan by Mohammad Azadnia, Sina Rezagholizadeh and Alireza Yari. December 2009.

[5] TechWare: Spoken Language Understanding Resources by Gokhan Tur, Ye-Yi Wang and Dilek Hakkani-Tür. IEEE SIGNAL PROCESSING MAGAZINE May 2013.

[6] TAN Ying-wei, MO Qian. Design and Realization of Supervise Adaptive Topic Tracking System. Journal of Zhengzhou University (Natural Science Edition). 2014.

[7] Pang B, Lee L, Vaithyanathan S. Thumbs up? Sentiment classification using machine learning techniques. In: Isabelle P, ed. Proc. of the EMNLP 2002. Morristown: ACL, 2010.

[8] Ye Shiren, T-S Chua, M-Y Kan, and L. Qiu, -Document concept lattice for text understanding and summarization, || Inf. Process. Manage., vol. 43, 2014.

[9] S. Pado and M. Lapata, -Dependency-based construction of semantic space models, || Assoc. Comput. Linguist., vol. 33, no. 2, 2014.

[10] R. Keefer and N. Bourbakis, -Interactions with a mobile reading device for the visually impaired, || in Proc. 21st Int. Conf. Tools with Artif. Intel., Nov. 2012.