International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)
Web Site: www.ijettcs.org Email: editor@ijettcs.org
Volume 4, Issue 2, March-April 2015 **ISSN 2278-6856**

NATURAL LANGUAGE PROCESSING USING ARTIFICIAL INTELLIGENCE

Unnati Dhavare [1], Prof. Umesh Kulkarni [2]

[1]A.R.M.I.E.T A.S.Rao Nagar, Sapgaon, Shahapur, Thane, Maharashtra 400708

[2]V.I.T. Vidyalankar campus, Vidyalankar College Marg, Wadala (East), Mumbai, Maharashtra 400037

Abstract

Artificial Intelligence (AI) is the study of how to make computers (machines) do things which, at the moment, people do better. There are many applications of the artificial intelligence. NATURAL LANGUAGE PROCESSING (NLP) is one of the upcoming applications of AI.The goal of the Natural Language Processing is to design and build software that will analyze, understand, and generate languages that humans use naturally, so that eventually you will be able to address your computer as though you were addressing another person. Natural Language Processing is the use of computers to process written and spoken language for some practical, useful, purpose: to translate languages, to get information from the web on text data banks so as to answer questions, to carry on conversations with machines, and to get advice about, say, pensions and so on. Practical applications of natural language processing are machine translation, database access, information retrieval, text categorization, extracting data from text etc.Still no such completely working system has developed yet but research is going on. And it may be done soon. Some basic systems are already developed like ELIZA, INTELLISHRINK, and AMALGAM etc.

KEYWORDS: Artificial Intelligence, Natural Language Processing.

1. INTRODUCTION

Artificial Intelligence is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable. As a theory in the philosophy of mind, artificial intelligence (or AI) is the view that human cognitive mental states can be duplicated in computing machinery. Accordingly, an intelligent system is nothing but an information processing system. NATURAL LANGUAGE PROCESSING (NLP) is one of the upcoming applications of AI. The goal of the Natural Language Processing (NLP) is to design and build software that will analyze, understand, and generate languages that humans use naturally, so that eventually you. will be able to address your computer as though you were addressing another person. This goal is not easy to reach. "Understanding" language means, among other things, knowing what concepts a word or phrase stands for and knowing how to link those concepts together in a meaningful way. It's ironic that natural language, the symbol system that is easiest for humans to learn and use, is hardest for a computer to master. Long after machines have proven capable of inverting large matrices with speed and grace, they still fail to master the basics of our spoken and written languages.

2. LITERATURE SURVEY

A linguistic ontology of space for natural language processing John A. Bateman, , Joana Hois, Robert Ross, Thora Tenbrink 2010[1]. Natural Language Processing (Almost) from Scratch published in the journal of machine learning research in 2011 Ronan Collobert, Jason Weston [2]. Natural Language Processing For Indian Languages: A Literature Survey Paperback – August 23, 2012 by P. J. Antony (Author), K. P. Soman [3]. Seth Grimes on BI, text/content analytics, sentiment analysis, and more March 4, 2013.

3. COMPONENTS OF NLP SYSTEM

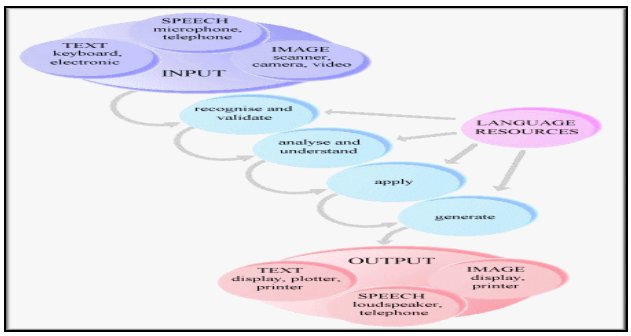

Figure 1: Model of a Language Enabled System

Fig.1 shows a model of a Language Enabled System. Within this general model there are, of course, many different configurations. Depending on the application of the technology, not all these components are needed. The basic processes of Natural Language Processing are shown in the diagram above. These are broadly concerned with:

- Entering material into the computer, using speech, printed text or handwriting, or text either keyed in or introduced electronically.
- Recognizing the language of the material, distinguishing separate words, for example, recording it in symbolic form and validating it.
- Building an understanding of the meaning of the material, to the appropriate level for the particular application.

- Using this understanding in an application such as transformation (e.g. speech to text), information retrieval, or human language translation.
- Generating the medium for presenting the results of the application finally, presenting the results to human users via a display of some kind: a printer or a plotter; a loud speaker or the telephone.

4. ARCHITECTURE OF NLP SYSTEM

By far the largest part of human linguistic communication occurs as speech. Written language is a fairly recent invention and still plays a less central role than speech in most activities. But processing written language is easier, in some ways, than processing speech. For example to build a program that understands spoken language, we need all the facilities of a written language understand as well as enough additional knowledge to handle all noise and ambiguities of the audio signal. Thus it is useful to divide the entire language processing problem into two tasks:

1) Processing written text, using lexical, syntactic, and semantic knowledge of the language as well as the required real world information.

2) Processing spoken language, using all the information needed above plus additional knowledge about phonology as well as enough added information to handle the further ambiguities that arise in speech

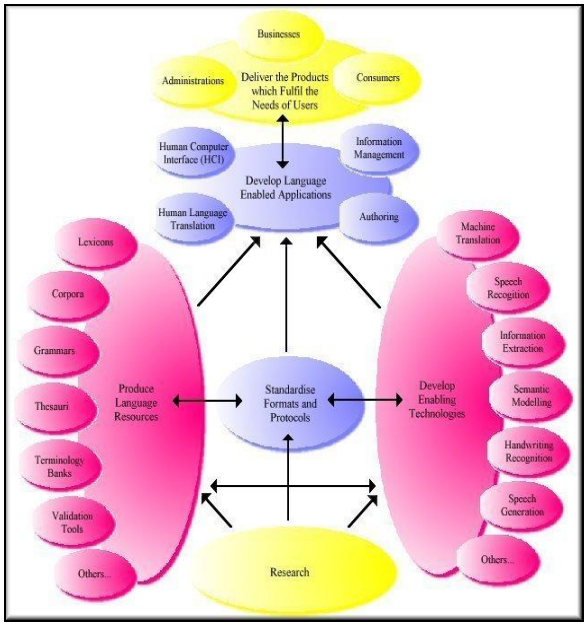

Figure 2: Architecture of NLP System

The diagram above depicts the chain of activities which are involved in Language Engineering, from research to the delivery of language-enabled and language enhanced products and services to end-users. The process of research and development leads to the development of techniques, the production of resources, and the development of standards. These are the basic building blocks. Language Engineering is applied at two levels. At **the first level there are a number of generic classes of application, such as:**

- language translation
- information management (multi-lingual)
- authoring (multi-lingual)
- human/machine interface (multi-lingual voice and text)

At the second level, these enabling applications are applied to real world problems across the social and economic spectrum. So, for example:

- Information management can be used in an information service, as the basis for analyzing requests for information and matching the request against a database of text or images, to select the information accurately.
- Authoring tools are typically used in word processing systems but can also be used to generate text, such as business letters in foreign languages, as well as in conjunction with information management, to provide document management facilities
- Human language translation is currently used to provide translator workbenches and automatic translation in limited domains
- Most applications can usefully be provided with natural language user interfaces, including speech, to improve their usability

4.1 MAIN STEPS IN THE PROCESS

1. Morphological Analysis**:** Individual words are analysed into their components, and non word tokens, such as punctuation, are separated from the words.
2. Syntactic analysis: Linear sequences of words are transformed into structures that show how the words relate to each other. Some word sequences may be rejected if they violate the language's rules for how words may be combined. For example, an English syntactic analyzer would reject the sentence "Boy the go the store."
3. Semantic Analysis: The structures created by the syntactic analyzer are assigned meaning. In other words, a mapping is made between the syntactic structures and the objects in the task domain. Structures for which no such mapping is possible may be rejected. For example, in most universes, the sentence "Colorless green ideas sleep furiously" would be rejected as semantically anomalous.
4. Discourse Integration: The meaning of an individual sentence may depend on the sentences that precede it and may influence the meaning of the sentences that follow it. For example, the word "it" in the sentence, “John wanted it”, depends on the prior discourse context, while the word "John" may influence the meaning of later sentences.
5. Pragmatics Analysis: The structure representing what was said is reinterpreted to determine what was actually meant. For example, the sentence "Do you know what time it is?" should be interpreted as a request to be told the time

5. COMPARISION OF NLP SYSTEM

Table 1: Comparison among already developed NLIDB System

S. N.	System Name	Domain	Language	Approaches	Year
1	GINLIDB	Natural	Natural	Lexical Analysis	2009
2	PNLIDB	Agriculture	Punjabi-Sql-Punjabi	Shallow parser	2010
3	HNLIDB	Employee	Hindi-Sql-Hindi	Shallow parser	2011
4	Portable NLIDB	Natural	CPG framework	linguistic semantics	2012

6. ADVANTAGES OF NLP SYSTEM

The benefits to be gained from successful Natural Language Processing are immense. They include:

1. Improved service from our public administration and public service agencies.
2. Wide accessibility of information through easier use of computer systems and Information Services.
3. Enhanced ability to compete in global markets.
4. Saving time by using intelligent computer systems as our agents.
5. Improvements in the quality of information recorded in information systems.
6. Better filtering of information when we need it.
7. More effective international co-operation.
8. Improved safety through 'hands-free' operation of equipment.
9. Greater security through voice verification techniques.

7. CONCLUSION

The complete process of the natural language processing system, at many places had made the work easier. Just by our natural language (any language) we can direct the robot, which can do conversation with computer and there is no need of a person to work as a translator for the conversation between two persons who don't know any common language. Still current program have not reached this level but they may do so very soon. Language technologies can be applied to a wide range of problems in business and administration to produce better, more effective solutions. They can also be used in education, to help the disabled, and to bring new services both to organizations and to consumers. There are a number of areas where the impact is significant such as competing in a global market, offering services directly through tele-business, supporting electronic commerce, enhancing entertainment, leisure and creativity.

ACKNOWLEDGMENT

I would like to express my sincere gratitude towards my guide Prof.Umesh Kulkarni for the help, guidance and encouragement in the development of this methodology. They supported me with scientific guidance, advice and encouragement, and were always helpful and enthusiastic and this inspired me in my work. I have benefitted from numerous discussions with guide and other colleagues.

REFERENCES

[1]. Natural Language Processing (Special Issues of Artificial Intelligence) Paperback – Import, 11 May 1994 by Fernando C N Pereira (Author).

[2]. Eugene Charniak and Drew McDermott, Introduction to Artificial Intelligence, Pearson, 1998, Chapter4.

[3]. Enhanced Text Retrieval Using Natural Language Processing Elizabeth D. Liddy President1, 2 Article first published online: 31 JAN 2005 DOI: 10.1002/bult.91

[4]. Ambient intelligence—the next step for artificial intelligence C Ramos, JC Augusto, D Shapiro - Intelligent Systems, IEEE, 2008ieeexplore.ieee.org References 1. IST Advisory Group, Scenarios for Ambient Intelligence in 2010.

[5]. Natural language processing: an introduction PrakashMNadkarni,
1. LucilaOhnoMachado,2 and Wendy W Chapman2 in volume 18 on page 540.J Am Med Inform Assoc. 2011 Sep-Oct; 18(5): 544–551.

[6]. K.R. Chowdhary Professor & Head CSE Dept. M.B.M. Engineering College, Jodhpur, India.April29, 2012.NaturalLanguageProcessing.

[7]. Rada Mihalce, Hugo Liu, and Henry Lieberman Computer Science Department, University of North Texas rada@cs.unt.edu Media Arts and Sciences, Massachusetts Institute of Technology. NLP (Natural Language Processing) for NLP (Natural Language Programming).