

**KNOWLEDGE SHARING OF NATURAL LANGUAGE PROCESSING IN
ARTIFICIAL INTELLIGENCE**Aye Mya Sandar ^{*1}Khin Myat Nwe Win ²^{*1}Information Technology, Science and Maintenance Department²Faculty of Computer Science^{*1,2}University of Computer Studies (Mandalay), Myanmarayemyasandarpaing@gmail.comkhinmyatnwewin@gmail.com

ABSTRACT

Natural language processing (NLP), a subfield of Artificial Intelligence (AI), has recently enlarged much attention for representing and analyzing human language computationally and its technology is considered to be important to achieving artificial intelligence that can substitute for humans in certain roles. Natural language processing derived into presence to ease the user's work and to satisfy the wish to communicate with the computer in natural language. Natural Language Processing provides solution in a variety of different fields associated with the social and cultural context of language learning drop. The paper distinguishes knowledge sharing about components of NLP, NLP terminology and level of NLP supporting on Artificial Intelligence.

KEYWORDS:Artificial intelligence, Natural language processing, Knowledge process

INTRODUCTION

Artificial Intelligence is the field of study that describe the capability of machine learning just like humans and the ability to respond to certain behaviors and also a technical field that is difficult to define, natural language processing technology appears to be a major element. This is clearly verified by the use of natural language conversations to measure computer intelligence in the Turing test, which is a well-known technique for determining whether a certain machine is intelligent. In measuring the intelligence of something in which intellectual level cannot be directly observed, it is natural to adopt an approach in which intelligence is assessed based on responses to queries, since linguistic expression is surely a typical means of exhibiting intelligence. [1]

Natural Language processing is a branch of computer science, artificial intelligence and linguistics, that is the science of language which includes Phonology that refers to sound, Morphology word formation, Syntax sentence structure, Semantics syntax and Pragmatics which refers to understanding, concerned with the interactions between computers and human (natural) language. Natural languages are languages spoken by humans. Natural language is any language that humans learn from their environment and use to communicate with each other. Natural language processing is the collection of techniques employed to try and accomplish that goal. The field of natural language processing (NLP) is deep and diverse. Natural language processing (NLP) is a collection of techniques used to extract grammatical structure and meaning from input in order to perform a useful task as a result, natural language generation builds output based on the rules of the target language and the task at hand. NLP is useful in the tutoring systems, duplicate detection, computer supported instruction and database interface fields as it provides a pathway for increased interactivity and productivity. [3]

The field of Natural Language Processing is related with different theories and techniques that deal with the problem of natural language of communicating with the computers. Ambiguity is one of the major problem of natural language which is usually faced in syntactic level which has subtask as lexical and morphology which are concerned with the study of words and word formation. Each of these levels can produce ambiguities that can be solved by the knowledge of the complete sentence. The ambiguity can be solved by various methods such as Minimizing Ambiguity, Preserving Ambiguity, Interactive Disambiguity and Weighting Ambiguity [4]. Some of the methods proposed by researchers to remove ambiguity is preserving ambiguity, e.g. (Shemtov 1997; Emele & Dorna 1998; Knight & Langkilde 2000) [4][5][6] Their objectives are closely in line with the last of these: they cover a wide range of ambiguities and there is a statistical element implicit in their approach.

COMPONENTS OF NATURAL LANGUAGE PROCESSING

A language can be defined as a set of rules or set of symbol. Symbol is combined and used for conveying information or broadcasting the information and are tyrannized by the Rules. Natural Language Processing basically can be classified into two parts i.e.

a) **Natural Language Understanding (NLU)**: this deals with computer reading and comprehension. Understanding involves the following tasks –

- Mapping the given input in natural language into useful representations.
- Analyzing different aspects of the language [11]

b) **Natural Language Generation (NLG)**: this deals with the production of written or spoken discourse. [7] It is the process of producing meaningful phrases and sentences in the form of natural language from some internal representation.

It involves –

- **Text planning** – It includes retrieving the relevant content from knowledge base.
- **Sentence planning** – It includes choosing required words, forming meaningful phrases, setting tone of the sentence.
- **Text Realization** – It is mapping sentence plan into sentence structure [11]

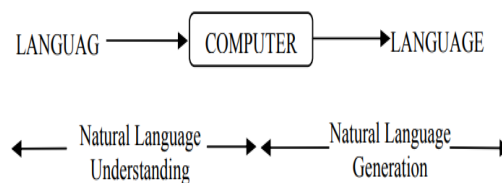


Fig. 1 Natural Language Processing

NATURAL LANGUAGE PROCESSING STEPS

There are general five steps –

- **Lexical Analysis** – It involves identifying and analyzing the structure of words. Lexicon of a language means the collection of words and phrases in a language. Lexical analysis is dividing the whole chunk of txt into paragraphs, sentences, and words.
- **Syntactic Analysis (Parsing)** – It involves analysis of words in the sentence for grammar and arranging words in a manner that shows the relationship among the words. The sentence such as “The school goes to boy” is rejected by English syntactic analyzer.
- **Semantic Analysis** – It draws the exact meaning or the dictionary meaning from the text. The text is checked for meaningfulness. It is done by mapping syntactic structures and objects in the task domain. The semantic analyzer disregards sentence such as “hot ice-cream”.
- **Discourse Integration** – The meaning of any sentence depends upon the meaning of the sentence just before it. In addition, it also brings about the meaning of immediately succeeding sentence.
- **Pragmatic Analysis** – During this, what was said is re-interpreted on what it actually meant. It involves deriving those aspects of language which require real world knowledge.[11]

TERMINOLOGY OF NATURAL LANGUAGE PROCESSING

According to paper [2], terminology of language is one of the most explanatory method for representing the Natural Language processing which helps to generate the NLP text by realizing Content Planning, Sentence Planning and Surface Realization phases. Linguistic is the science which involves meaning of language, language context and various forms of the language. The various important terminologies of Natural Language Processing are: -

A. Phonology

Phonology is the part of Linguistics which refers to the systematic arrangement of sound. It is study of organizing sound systematically. [11] The term phonology comes from Ancient Greek and the term phono- which means voice or sound, and the suffix logy refers to word or speech. It could be explained as, phonology proper is concerned with the function, behaviour and organization of sounds as linguistic items. Phonology includes semantic use of sound to encode meaning of any Human language. (Clark et al.,2007) [8].

B. Morphology

It is a study of construction of words from primitive meaningful units. [11] The different parts of the word represent the smallest units of meaning known as Morphemes. Morphology which comprise of Nature of words, are initiated by morphemes, that is primitive unit of meaning in a language. [11] An example of Morpheme could be, the word pre-cancellation can be morphologically scrutinized into three separate morphemes: the prefix pre, the root cancella, and the suffixation. The interpretation of morpheme stays same across all the words, just to understand the meaning humans can break any unknown word into morphemes. [2]

C. Lexical

In Lexical, humans, as well as NLP systems, interpret the meaning of individual words. Sundry types of processing bestow to word-level understanding – the first of these being a part-of-speech tag to each word. In this processing, words that can act as more than one part-of-speech are assigned the most probable part-of speech tag based on the context in which they occur. At the lexical level, Semantic representations can be replaced by the words that have one meaning. In NLP system, the nature of the representation varies according to the semantic theory deployed. [2]

D. Syntactic

It refers to arranging words to make a sentence. It also involves determining the structural role of words in the sentence and in phrases. [11] This level emphasis to scrutinize the words in a sentence so as to uncover the grammatical structure of the sentence. Both grammar and parser are required in this level. The output of this level of processing is representation of the sentence that divulge the structural dependency relationships between the words. There are various grammars that can be impeded, and which in twirl, whack the option of a parser. Not all NLP applications require a full parse of sentences, therefore the abide challenges in parsing of prepositional phrase attachment and conjunction audit no longer impede that plea for which phrasal and clausal dependencies are adequate [9]. Syntax conveys meaning in most languages because order and dependency contribute to connotation. [2]

E. Semantic

It is concerned with the meaning of words and how to combine words into meaningful phrases and sentences. [11] In semantic most people think that meaning is determined, however, this is not it is all the levels that bestow to meaning. Semantic processing determines the possible meanings of a sentence by pivoting on the interactions among word-level meanings in the sentence. This level of processing can incorporate the semantic disambiguation of words with multiple senses; in a cognate way to how syntactic disambiguation of words that can errand as multiple parts-of-speech is adroit at the syntactic level. The semantic level scrutinizes words for their dictionary elucidation, but also for the elucidation they derive from the milieu of the sentence. Semantics milieu that most words have more than one elucidation but that we can spot the appropriate one by looking at the rest of the sentence. [10]

F. Discourse

It deals with how the immediately preceding sentence can affect the interpretation of the next sentence. [11] While syntax and semantics travail with sentence-length units, the discourse level of NLP travail with units of text longer than a sentence i.e, it does not interpret multi sentence texts as just sequence sentences, apiece of which can be elucidated singly. Rather, discourse focuses on the properties of the text as a whole that convey meaning by making connections between component sentences (Elizabeth D. Liddy,2001) [9]. The two of the most common levels are Anaphora Resolution - Anaphora resolution is the replacing of words such as pronouns, which are semantically stranded, with the pertinent entity to which they refer. Discourse/Text Structure

Recognition - Discourse/text structure recognition sways the functions of sentences in the text, which, in turn, adds to the meaningful representation of the text. [2]

G. *Pragmatic:*

It deals with using and understanding sentences in different situations and how the interpretation of the sentence is affected. [11] Pragmatic is concerned with the firm use of language in situations and utilizes nub over and above the nub of the text for understanding the goal and to explain how extra meaning is read into texts without literally being encoded in them. This requisite much world knowledge, including the understanding of intentions, plans, and goals. For example, the following two sentences need aspiration of the anaphoric term 'they', but this aspiration requires pragmatic or world knowledge (Elizabeth D. Liddy, 2001) [9]

CONCLUSION

Today NLP is growing to the huge improvements in the access to data and the increase in computational power, which is allowing practitioners to achieve meaningful results in areas like healthcare, media, finance and human resources, among others. While NLP is a relatively recent area of research and application, as compared to other information technology approaches, there have been sufficient successes to date that suggest that NLP-based information access technologies will continue to be a major area of research and development in information systems now and far into the future. AI nowadays is being implemented in almost every field of study through several models such as Support Vector Machine (SVM) and Artificial Neural Network (ANN). AI revelation age and therefore we should accept into this change and welcome it too by embracing AI and moving toward a better society.

REFERENCES

- [1]. Yoshihiro Matsuo, Ryuichiro Higashinaka, Hisako Asano, and Toshiro Makino, (2014) "Natural Language Processing Supporting Artificial Intelligence". Feature Articles: Artificial Intelligence Research Activities in NTT Group.
- [2]. Diksha Khurana, Aditya Koli , Kiran Khatter and Sukhdev Singh. "Natural Language Processing: State of The Art, Current Trends and Challenges".
- [3]. Prof. Alpa Reshamwala, Prof. Dhirendra Mishra, Prajakta Pawar.(2013). "REVIEW ON NATURAL LANGUAGE PROCESSING" IRACST – Engineering Science and Technology: An International Journal (ESTIJ), ISSN: 2250-3498, Vol.3, No.1,
- [4]. Shemtov, H. (1997). Ambiguity management in natural language generation. Stanford University.
- [5]. Emele, M. C., & Dorna, M. (1998, August). Ambiguity preserving machine translation using packed representations. In Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics-Volume 1 (pp. 365-371). Association for Computational Linguistics.
- [6]. Knight, K., & Langkilde, I. (2000, July). Preserving ambiguities in generation via automata intersection. In AAAI/IAAI (pp. 697-702)
- [7]. Sudhir K Mishra,(2018)" Artificial Intelligence and Natural Language Processing". Cambridge Scholars Publishing Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK
- [8]. Nation, K., Snowling, M. J., & Clarke, P. (2007). Dissecting the relationship between language skills and learning to read: Semantic and phonological contributions to new vocabulary learning in children with poor reading comprehension. *Advances in Speech Language Pathology*, 9(2), 131-139.
- [9]. Liddy, E. D. (2001). Natural language processing
- [10]. Feldman, S. (1999). NLP Meets the Jabberwocky: Natural Language Processing in Information Retrieval. *ONLINE-WESTON THEN WILTON-*, 23, 62-73.
- [11]. https://www.tutorialspoint.com/artificial_intelligence/artificial_intelligence_natural_language_processing