An analysis of NLP

(Natural Language Processing)

ARYA BRIJITH
IIPP Research Intern
Asia University
Taichung, Taiwan
(e-mail: arya.brijithk@gmail.com).

ABSTRACT
The goal of this in-depth analysis of natural language processing is to shed light on the complexities and promise of language-driven computer intelligence by outlining its basic features, applications in many fields, and future possibilities.

KEYWORDS natural language processing, robots, emotional intelligence, artificial intelligence, programs.

I. INTRODUCTION

A language is a set of finite-length sentences, constructed using a finite alphabet set, or in terms of language syntax, they are constructed using a finite vocabulary of symbols. Since the alphabet set is finite. For our study of languages, we usually consider the languages without any bound, i.e., infinite, but for specific languages, that are of practical use, we limit our study to a finite set. This is because, the machines we would like to use for processing the languages, are finite machines, they have finite memory, and have finite processing power.

Natural Language Processing is a subfield of Artificial Intelligence and linguistics, devoted to making computers understand the statements or words written in human languages. Natural language refers to the ordinary language that is spoken or written by people(humans) for general-purpose communication.

II. THE LANGUAGE APPROACH

Phonology
This level deals with the interpretation of speech sounds within and between words.
there are three types of rules used in phonological analysis:
1) phonetic rules – for sounds in words;

2) phonemic rules - for pronunciation variations in the context of words let's talk and;
3) prosodic rules - change stress and intonation super sentence An NLP system that accepts speech input has sound waves analyzed and encoded into a digitized signal for interpretation using various rules or compared to the specific language model used.

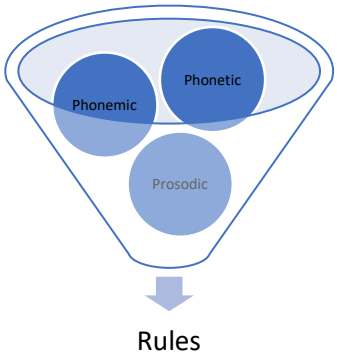

Figure 2: Phonology

Morphology
This level deals with the componential nature of words, which are composed of
morphemes – the smallest units of meaning. For example, the word preregistration can
be morphologically analyzed into three separate morphemes: the prefix pre, the root
registerer, and the suffixation. Since the meaning of each morpheme remains the same
across words, humans can break down an unknown word into its constituent morphemes
to understand its meaning. Similarly, an NLP system can recognize the meaning
conveyed by each morpheme to gain and represent meaning. For example,
adding the suffix –ed to a verb conveys that the action of the verb took place in the past. This is the central meaning, and it often appears only in the text
use of the morpheme -ed.

Lexical
At this level, humans, as well as NLP systems, interpret the meaning of individual words. Several types of processing contribute to word-level comprehension—the first of these
each word is assigned a single part-of-speech identifier. In this reading, the words
that can function as more than one part of speech, the most likely part are determined
by speech recognition based on the context in which they occur.

In addition, at the lexical level, those words that have only one possible meaning or
meaning can be replaced by a semantic representation of that meaning. The character
of the representation varies depending on the semantic theory used in the NLP system. The lexical level may require a special approach to vocabulary and NLP
the system determines whether vocabulary is used and its nature and extent
of information encoded in the dictionary.

Vocabulary can be quite simple, containing only words and their parts of speech, or increasingly complex, containing information about the semantic class of the word, the arguments required for it, and the semantic restrictions of those arguments. definitions of the meaning(s) of the semantic representation used in a given system, and even the semantic field in which each meaning of a polysemous word is used.

Syntactic
This level focuses on analyzing the words in a sentence to reveal grammar
sentence structure This requires both a grammar and a parser. The result of this
level of processing is a sentence representation that reveals
Structural dependencies between words. Different grammars can be exploited which in turn influences the choice of analyzer. Not all NLP
applications require full sentence parsing, so the rest of the challenges are active
the structuring of the prepositional phrase entity and the scope of the conjunction are no longer disturbed in those applications where clause-by-clause dependencies are sufficient. Syntax conveys meaning in most languages because it is affected by order and dependence
importance

Semantics
This is the level at which most people think the meaning is determined as best as possible
See definitions of levels above, all levels contribute to meaning. Semantic processing determines the possible meanings of a sentence by focusing on the sentence
the interaction of word-level meanings of a sentence. This level of processing can
include semantic clarification of ambiguous words; in a similar way
as syntactic specification words that can function as several parts of speech
are achieved at the syntactic level. A semantic specification allows one and only one
meaning of polysemous words to be selected and incorporated into the semantic presentation
sentence If information is needed about another sentence for clarification,
the semantic rather than the lexical level would be unambiguous.

Discourse
While syntax and semantics work with units of sentence length, NLP at the discourse level
works with texts longer than a sentence. Thus, he does not interpret multi-sentence texts only as chain sentences, each of which can be interpreted separately. Rather, discourse focuses on those features of the text as a whole that convey meaning and create relationships between component statements. Multiple chat types can occur at this level, the two most common of which are anaphora resolution and recognizing discourse/text structure. Anaphoric resolution is the replacement of such words as semantically empty pronouns with the appropriate entity for which they are signaled

Pragmatic
This level is about appropriate language use in situations and ways to use it to understand the context of the content of the text The goal is to explain as additional meaning is read into texts without actually encoding them. This requires a lot of world knowledge, including understanding and understanding intentions, plans goals Some NLP applications may use databases and inference modules.

III. APPLICATIONS

Natural language processing provides both theory and applications for many different languages. Some of these are as follows:

- Searching for information - Since there is a lot of text in this application, it is surprising that so few applications use NLP. Recently, statistical approaches to NLP have been used more.

- Information Extraction (IE) - focuses on recognition, labeling, and decoding into a structured representation, specific key data elements such as people, companies, places, organizations, and large text collections. These extracts can then be used for various applications including query answering, visualization, and data mining.

- Answer questions – unlike an information search which provides a list of possibly relevant documents in response to a user's request, answering questions offers the user either only the text of the answer or the text that offers the answer paragraphs

- Summarization - The higher levels of NLP, especially the discourse level, can an accelerator that turns larger text into shorter but richer text formed an abridged narrative of the original document.

- Machine Translation - Perhaps the oldest of all NLP applications at various levels NLP has been used in MT systems since the "word-based" approach applications involving advanced analytics.

- Dialog systems - provided by major end-user application manufacturers. Dialogue systems which usually focus on a strictly defined application (such as a refrigerator or home audio system), currently use the phonetic and lexical levels of the language.

- Morphological segmentation -Separate words into individual morphemes and identify the class of the morphemes. The difficulty of this task depends greatly on the complexity of the morphology i.e. the structure of words of the language being considered.

- Named entity recognition (NER)- It describes a stream of text to determine which lexicon in the text relates properly to sentence

IV. FUTURE PROSPECT

NLP is a relatively recent area of research and application, as compared to other
information technology approaches and the main goal is to make machines able to interpret language at a Human level, or human-readable natural language it is going to be the perfect problem for artificial intelligence, this provides the answer to AI solving and making Computers that are just as smart as humans when it comes to problem thinking like people and perform functions that humans cannot. As natural language understanding or readability improves, computers or machines or devices can learn information online and practice what you learn in the real world Along with natural language generation and becoming more human-like.

V. CONCLUSIONS

NLP is a relatively recent area of research and application, and being able to use natural language for query specification and retrieval baggs over the keyword, and keyphrase approaches makes me believe that the restricted use of natural language in captions for multimedia data abstraction is a less cumbersome task and built upon not only for abstracting images but also the form so multimedia.

VI REFERENCES

1. Liddy, E. D. (2001). Natural language processing.
   Church, K. W., & Rau, L. F. (1995). Commercial applications of natural language processing. *Communications of the ACM*, *38*(11), 71-79.

2. Nadkarni, P. M., Ohno-Machado, L., & Chapman, W. W. (2011). Natural language processing: an introduction. *Journal of the American Medical Informatics Association*, *18*(5), 544-551.

3. Chowdhary, K., & Chowdhary, K. R. (2020). Natural language processing. *Fundamentals of artificial intelligence*, 603-649.

4. Joshi, A. K. (1991). Natural language processing. *Science*, *253*(5025), 1242-1249.

5. Chopra, A., Prashar, A., & Sain, C. (2013). Natural language processing. *International journal of technology enhancements and emerging engineering research*, *1*(4), 131-134.

6. Alsmirat, M. A., Jararweh, Y., Al-Ayyoub, M., Shehab, M. A., & Gupta, B. B. (2017). Accelerating compute intensive medical imaging segmentation algorithms using hybrid CPU-GPU implementations. Multimedia Tools and Applications, 76, 3537-3555.
7. Tripathi, S., Gupta, B., Almomani, A., Mishra, A., & Veluru, S. (2013). Hadoop based defense solution to handle distributed denial of service (ddos) attacks.
8. Almomani, A., Gupta, B. B., Wan, T. C., Altaher, A., & Manickam, S. (2013). Phishing dynamic evolving neural fuzzy framework for online detection zero-day phishing email. arXiv preprint arXiv:1302.0629.
9. Gupta, B. B., Joshi, R. C., & Misra, M. (2012). ANN based scheme to predict

number of zombies in a DDoS attack. Int. J. Netw. Secur., 14(2), 61-70. http://ijns.jalaxy.com.tw/contents/ijns-v14-n2/ijns-2012-v14-n2-p61-70.pdf