MANJARA CHARITABLE TRUST

RAJIV GANDHI INSTITUTE OF TECHNOLOGY, MUMBAI

**DEPARTMENT OF COMPUTER ENGINEERING**

**Academic Year: 2022-2023** **Semester: VII**

Student Name: _____ Roll No:_____

Subject: **Natural Language Processing**

Class/Div: **BE/A**

**Experiment No.:__4__**

**Title of Experiment: Study of Morphological Analysis**

**_____**

**Experiment Evaluation:**

| Experiment No. | Performance/ Coding (04) | Punctuality (03) | Question-Answer (03) | **Total (10)** |
|---|---|---|---|---|
| 1 | | | | |

---

Staff Signature

**Aim:** Study of Morphological analysis.

## THEORY:

Morphology in NLP is defined as the study of the structure of words and how the words are formed. It identifies the root of the word and the prefix and suffix which are attached to the root of the word. For example, take a word "unhappiness", here if we see the formation of the word then we will come to know that prefix is "un", the root is "happy", a suffix is "ness". This study of word formation and identification of the structure of a word is known as Morphology.

Morphological analysis (MA) is a method for identifying, structuring and investigating the total set of possible relationships contained in a given multidimensional problem complex. MA allows small groups of subject specialists to define, link, and internally evaluate the parameters of complex problem spaces, creating a solution space and a flexible inference model.

Problems that involve many governing factors, where most of them cannot be expressed numerically can be well suited for MA.

The conventional approach is to break a complex system into parts, isolate the parts (dropping the 'trivial' elements) whose contributions are critical to the output and solve the simplified system for desired scenarios. The disadvantage of this method is that many real-world phenomena do not have obviously trivial elements and cannot be simplified.

Morphological analysis works backwards from the output towards the system internals without a simplification step.[4] The system's interactions are fully accounted for in the analysis.

Example:

1. Your brother is older than mine.

The first thing we must do is separate it by words to be able to analyse each of them independently. Next in each one of them we will indicate its grammatical category. Let’s see the resulting example:

- You : possessive determiner
- Brother: noun, common, masculine, singular.
- More : adverb of quantity
- Major : adjective, neuter, singular
- What : conjunction
- Mine : possessive pronoun

2. The brown dog runs across the meadow.

- Dog : noun, common, masculine, singular
- Brown : adjective, singular neuter
- Run : verb run, third person singular present indicative
- By : preposition
- A : determinant, article, feminine, singular
- Meadow : noun, common, feminine, singular

**CODE:**

```
import enchant
d= enchant.Dict("en_US")
tokens=[]
def tokenize(st):
    if not st:return
for i in range(len(st),-1,-1):
    if d.check(st[0:i]):
        tokens.append(st[0:i])
        st=st[i:]
        tokenize(st)
        break
tokenize("Rockandroll")
tokenize("hanana")
tokenize("processings")
tokenize("unhappyness")
print(tokens)
```

**CONCLUSION:** Hence, we have imbibed the concepts of Morphological analysis in the linguistic approach and how stem words are separated from prefixes and suffixes.