



K. J. Somaiya College of Engineering, Mumbai-77

Batch: A2

Roll No.: 1911027

Experiment / assignment / tutorial No. 10

Grade: AA / AB / BB / BC / CC / CD /DD

Signature of the Staff In-charge with date

Title: Self Learning Experiment using Vlab in NLP

Objective: To study NLP to introduce concepts like word analysis, word generation and building POS Tagger.

Expected Outcome of Experiment:

Course Outcome	After successful completion of the course students should be able to
CO4	Analyse applications of AI and understand planning & learning processes in advanced AI applications

Books/ Journals/ Websites referred:

1. <http://nlp-iiith.vlabs.ac.in/List%20of%20experiments.html?domain=ComputerScience>

Pre Lab/ Prior Concepts:

Knowledge Engineering

Historical Profile:

What is Natural Language Processing?

Natural Language Processing, usually shortened as NLP, is a branch of artificial intelligence that deals with the interaction between computers and humans using the



K. J. Somaiya College of Engineering, Mumbai-77

natural language. The ultimate objective of NLP is to read, decipher, understand, and make sense of the human languages in a manner that is valuable. Most NLP techniques rely on machine learning to derive meaning from human languages.

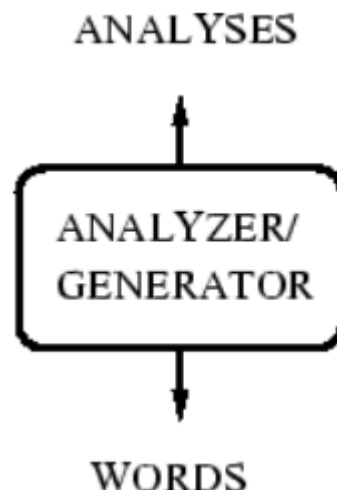
New Concepts learnt: (point wise)

- Understood morphological analysis of a word and explored its types such as inflectional and derivational morphology.
- Natural language processing and how it can be used to enable machines understand human language.
- Explored components of natural language processing such as natural language generation and natural language understanding.
- Understood how to derive various morphological features of a word such as its root word, tense, category, case, etc.
- Explored how chat bots, automatic essay generators, automatic text summarizers work internally

Chosen Experiment (Title with reference):

Title: Word Analysis

A word can be simple or complex. For example, the word 'cat' is simple because one cannot further decompose the word into smaller part. On the other hand, the word 'cats' is complex, because the word is made up of two parts: root 'cat' and plural suffix '-s'.



Analysis of a word into root and affix(es) is called as Morphological analysis of a word. It is mandatory to identify root of a word for any natural language processing task. A root word can have various forms. For example, the word 'play' in English has the following forms: 'play', 'plays', 'played' and 'playing'. Hindi shows more number of



K. J. Somaiya College of Engineering, Mumbai-77

forms for the word 'खेल' (khela) which is equivalent to 'play'. The forms of 'खेल'(khela) are the following:

खेल(khela), खेला(khelaa), खेली(khelii), खेलूंगा(kheluungaa), खेलूंगी(kheluungii), खेलेगा(khelegaa), खेलेगी(khelegii), खेलते(khelate), खेलती(khelatii), खेलने(khelane), खेलकर(khelakar)

For Telugu root ఆడడం (Adadam), the forms are the following:

Adutaanu, AdutunnAnu, Adenu, Ademu, AdevA, AdutAru, Adutunnaru, AdadAniki, Adesariki, AdanA, Adinxi, Adutunxi, AdinxA, AdeserA, Adestunnaru, ...

Thus we understand that the morphological richness of one language might vary from one language to another. Indian languages are generally morphologically rich languages and therefore morphological analysis of words becomes a very significant task for Indian languages.

Types of Morphology:

Morphology is of two types,

1) Inflectional morphology:

Deals with word forms of a root, where there is no change in lexical category. For example, 'played' is an inflection of the root word 'play'. Here, both 'played' and 'play' are verbs.

2) Derivational morphology:

Deals with word forms of a root, where there is a change in the lexical category. For example, the word form 'happiness' is a derivation of the word 'happy'. Here, 'happiness' is a derived noun form of the adjective 'happy'.

Morphological Features:

All words will have their lexical category attested during morphological analysis. A noun and pronoun can take suffixes of the following features: gender, number, person, case For example, morphological analysis of a few words is given below:

Language	input:word	output:analysis
Hindi	लडके (ladake)	rt=लड़का(ladakaa), cat=n, gen=m, num=sg, case=obl
Hindi	लडके (ladake)	rt=लड़का(ladakaa), cat=n, gen=m, num=pl, case=dir
Hindi	लड़कों (ladakoM)	rt=लड़का(ladakaa), cat=n, gen=m, num=pl, case=obl
English	boy	rt=boy, cat=n, gen=m, num=sg
English	boys	rt=boy, cat=n, gen=m, num=pl



K. J. Somaiya College of Engineering, Mumbai-77

A verb can take suffixes of the following features: tense, aspect, modality, gender, number, person 'rt' stands for root. 'cat' stands for lexical category.

Language	input:word	output:analysis
Hindi	हँसी(hansii)	rt=हँस(hans), cat=v, gen=fem, num=sg/pl, per=1/2/3 tense=past, aspect=pft
English	toys	rt=toy, cat=n, num=pl, per=3

The value of lexical category can be noun, verb, adjective, pronoun, adverb, preposition. 'gen' stands for gender. The value of gender can be masculine or feminine. 'num' stands for number. The value of number can be singular (sg) or plural (pl). 'per' stands for person. The value of person can be 1, 2 or 3. The value of tense can be present, past or future. This feature is applicable for verbs. The value of aspect can be perfect (pft), continuous (cont) or habitual (hab). This feature is not applicable for nouns. 'case' can be direct or oblique. This feature is applicable for nouns. A case is an oblique case when a postposition occurs after noun. If no postposition can occur after noun, then the case is a direct case. This is applicable for hindi but not english as it doesn't have any postpositions. Some of the postpositions in hindi are: का(kaa), की(kii), के(ke), को(ko), में(meM).

Snapshots of Execution (Min 3) :

Select a Language which you know better

English

Select a word from the below dropbox and do a morphological analysis on that word

studying

Select the Correct morphological analysis for the above word using dropboxes (NOTE : na = not applicable)

WORD	studying	
ROOT	study	✓
CATEGORY	verb	✓
GENDER	male	✓
NUMBER	singular	✓
PERSON	third	✓
CASE	na	✓
TENSE	present-perfect-continuous	✓
Check	Right answer!!!	



K. J. Somaiya College of Engineering, Mumbai-77

Select a Language which you know better

English

Select a word from the below dropbox and do a morphological analysis on that word

walks

Select the Correct morphological analysis for the above word using dropboxes (NOTE : na = not applicable)

WORD	walks	
ROOT	walk	✓
CATEGORY	verb	✓
GENDER	male	✓
NUMBER	singular	✓
PERSON	third	✓
CASE	na	✓
TENSE	simple-present	✓
Check	Right answer!!!	

Select a Language which you know better

English

Select a word from the below dropbox and do a morphological analysis on that word

played

Select the Correct morphological analysis for the above word using dropboxes (NOTE : na = not applicable)

WORD	played	
ROOT	play	✓
CATEGORY	verb	✓
GENDER	male	✓
NUMBER	singular	✓
PERSON	third	✓
CASE	na	✓
TENSE	past-perfect	✓
Check	Right answer!!!	

Select a Language which you know better

English

Select a word from the below dropbox and do a morphological analysis on that word

chaired

Select the Correct morphological analysis for the above word using dropboxes (NOTE : na = not applicable)

WORD	chaired	
ROOT	chair	✓
CATEGORY	verb	✓
GENDER	male	✓
NUMBER	singular	✓
PERSON	second	✓
CASE	na	✓
TENSE	simple-past	✓
Check	Right answer!!!	



K. J. Somaiya College of Engineering, Mumbai-77

Select a Language which you know better

English

Select a word from the below dropbox and do a morphological analysis on that word

sofa

Select the Correct morphological analysis for the above word using dropboxes (NOTE : na = not applicable)

WORD	sofa	
ROOT	sofa	✓
CATEGORY	noun	✓
GENDER	na	✓
NUMBER	singular	✓
PERSON	na	✓
CASE	na	✓
TENSE	na	✓
Check	Right answer!!!	

Select a Language which you know better

Hindi

Select a word from the below dropbox and do a morphological analysis on that word

चलती

Select the Correct morphological analysis for the above word using dropboxes (NOTE : na = not applicable)

WORD	चलती	
ROOT	चल	✓
CATEGORY	verb	✓
GENDER	female	✓
NUMBER	singular	✓
PERSON	third	✓
CASE	na	✓
TENSE	simple-present	✓
Check	Right answer!!!	

Select a Language which you know better

Hindi

Select a word from the below dropbox and do a morphological analysis on that word

कमाएंगे

Select the Correct morphological analysis for the above word using dropboxes (NOTE : na = not applicable)

WORD	कमाएंगे	
ROOT	कमा	✓
CATEGORY	verb	✓
GENDER	male	✓
NUMBER	plural	✓
PERSON	third	✓
CASE	na	✓
TENSE	simple-future	✓
Check	Right answer!!!	



K. J. Somaiya College of Engineering, Mumbai-77

Select a Language which you know better

Hindi

Select a word from the below dropdown and do a morphological analysis on that word

लड़को

Select the Correct morphological analysis for the above word using dropdowns (NOTE : na = not applicable)

WORD	लड़को	
ROOT	लड़का	✓
CATEGORY	noun	✓
GENDER	male	✓
NUMBER	plural	✓
PERSON	na	✓
CASE	oblique	✓
TENSE	na	✓
<div>Check</div>		Right answer!!!

Select a Language which you know better

Hindi

Select a word from the below dropdown and do a morphological analysis on that word

देखूँगा

Select the Correct morphological analysis for the above word using dropdowns (NOTE : na = not applicable)

WORD	देखूँगा	
ROOT	देख	✓
CATEGORY	verb	✓
GENDER	male	✓
NUMBER	singular	✓
PERSON	first	✓
CASE	na	✓
TENSE	simple-future	✓
<div>Check</div>		Right answer!!!



K. J. Somaiya College of Engineering, Mumbai-77

Select a Language which you know better

Hindi

Select a word from the below dropbox and do a morphological analysis on that word

बालको

Select the Correct morphological analysis for the above word using dropboxes (NOTE : na = not applicable)

WORD	बालको	
ROOT	बालक	✓
CATEGORY	noun	✓
GENDER	male	✓
NUMBER	plural	✓
PERSON	na	✓
CASE	oblique	✓
TENSE	na	✓
<div>Check</div>		Right answer!!!

Findings/Analysis/Conclusion:- Understood the concept of morphological analysis and natural language processing. Explored how to derive morphological features of a given word. Using virtual labs derived morphological features from a given word. Explored how natural language processing can be used in building chatbots.