|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Finolex Academy of Management and Technology, Ratnagiri | | | |
| **Department of Information Technology** | | | |
| **Subject:** | **Intelligent Systems (BEITL703)** | | | |
| **Class:** | **BE IT / Semester – VII (CBCGS) / Academic year: 2019-20** | | | |
| **Name of Student:** | **Kazi Jawwad A Rahim** | | | |
| **Roll No:** | **29** | | **Date of performance (DOP) :** |  |
| **Assignment/Experiment No:** | | **10** | **Date of checking (DOC) :** |  |
| **Title: Case study on Natural Language Processing using Vlabs** | | | | |
| **Marks:** | |  | **Teacher’s Signature:** |  |

**Aim**: To study Natural Language Processing in artificial intelligence.

**2. Prerequisites**:

1. Learn basics of artificial intelligence.
2. Knowledge and reasoning in artificial intelligence.

**3. Hardware Requirements**:

1. PC with minimum 2GB RAM

**4. Software Requirements:**

1. Windows installed
2. Swi ProDT

**5. Learning Objectives:**

1. To study an expert system.
2. To Analyze the current expert system.

**6. Course Objectives Applicable: CO 6**

**7. Program Outcomes Applicable: PO 2, PO 3, PO 4, PO 5, PO 6, PO 7, PO 8, PO 9, PO 11**

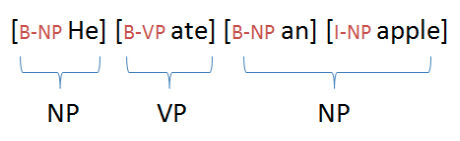
**8. Program Education Objectives Applicable: PEO 2, PEO 3, PEO 4, PEO 5, PEO 6**

**Theory:**

Chunking of text invloves dividing a text into syntactically correlated words. For example, the sentence 'He ate an apple.' can be divided as follows:



Each chunk has an open boundary and close boundary that delimit the word groups as a minimal non-recursive unit. This can be formally expressed by using IOB prefixes.



Chunking of text invloves dividing a text into syntactically correlated words.

Eg: He ate an apple to satiate his hunger.

[NP He ] [VP ate ] [NP an apple] [VP to satiate] [NP his hunger]

**Chunk Types**

The chunk types are based on the syntactic category part. Besides the head a chunk also contains modifiers (like determiners, adjectives, postpositions in NPs).

The basic types of chunks in English are:

Chunk Type Tag Name

1. Noun NP

2. Verb VP

3. Adverb ADVP

4. Adjectivial ADJP

5. Prepositional PP

The basic Chunk Tag Set for Indian Languages

1 . Noun Chunk- NP

2.1 Finite Verb Chunk -VGF

2.2 Non-finite Verb Chunk -VGNF

2.3 Verb Chunk (Gerund) -VGNN

3. Adjectival Chunk- JJP

4. Adverb Chunk -RBP

1. **NP Noun Chunks**

Noun Chunks will be given the tag NP and include non-recursive noun phrases and postposition for Indian languages and preposition for English. Determiners, adjectives and other modifiers will be part of the noun chunk.

Eg:



((in/IN the/DT big/ADJ room/NN))NP

1. **Verb Chunks**

The verb chunks are marked as VP for English, however they would be of several types for Indian languages. A verb group will include the main verb and its auxiliaries, if any.

For English:

I (will/MD be/VB loved/VBD)VP

The types of verb chunks and their tags are described below.

* 1. **VGF Finite Verb Chunk**

The auxiliaries in the verb group mark the finiteness of the verb at the chunk level. Thus, any verb group which is finite will be tagged as VGF. For example,



* 1. **VGNF Non-finite Verb Chunk**

A non-finite verb chunk will be tagged as VGNF.



* 1. **VGNN Gerunds**

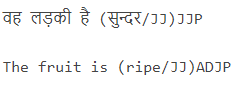
A verb chunk having a gerund will be annotated as VGNN.



1. **JJP/ADJP Adjectival Chunk**

An adjectival chunk will be tagged as ADJP for English and JJP for Indian languages. This chunk will consist of all adjectival chunks including the predicative adjectives.

Eg:



Note: Adjectives appearing before a noun will be grouped together within the noun chunk.

1. **RBP/ADVP Adverb Chunk**

This chunk will include all pure adverbial phrases.

Eg:



He walks (slowly/ADV)/ADVP

PP Prepositional Chunk

This chunk type is present for only English and not for Indian languages. It consists of only the preposition and not the NP argument.

Eg:

(with/IN)PP a pen IOB prefixes

Each chunk has an open boundary and close boundary that delimit the word groups as a minimal non-recursive unit. This can be formally expressed by using IOB prefixes: B-CHUNK for the first word of the chunk and I-CHUNK for each other word in the chunk. Here is an example of the file format:

**Tokens POS Chunk-Tags**

He PRP B-NP

ate VBD B-VP

an DT B-NP

apple NN I-NP

to TO B-VP

satiate VB I-VP

his PRP$ B-NP

hunger NN I-NP

**Objective**

The objective of this experiment is to understand the concept of chunking and get familiar with the basic chunk tagset.

**Procedure**

STEP1: Select a language

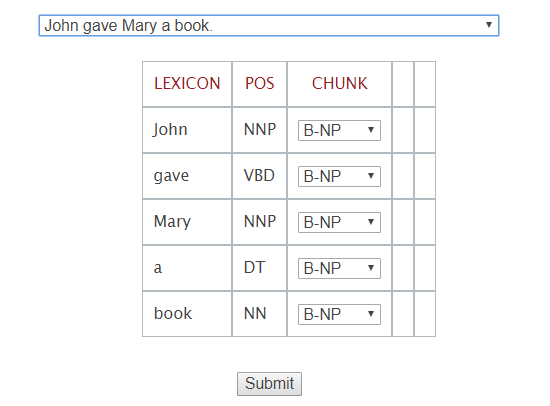
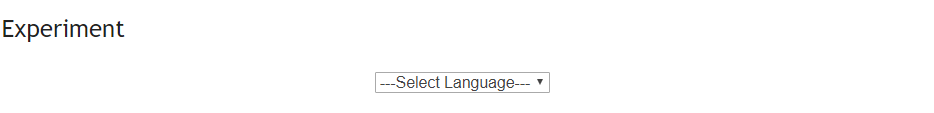
STEP2: Select a sentence

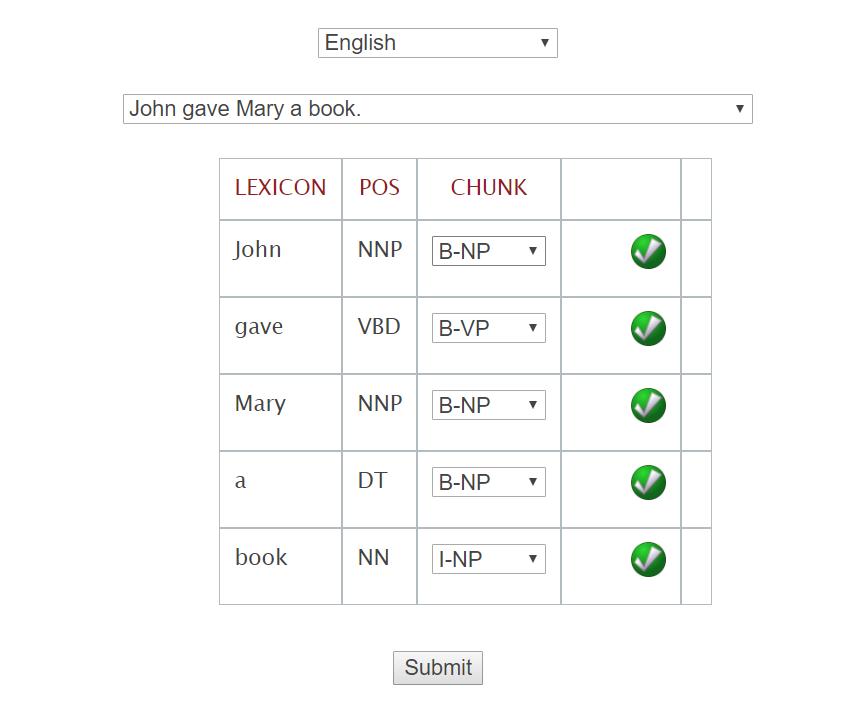
STEP3: Select the corresponding chunk-tag for each word in the sentence and click the **Submit** button.

OUTPUT1: The submitted answer will be checked.

Click on **Get Answer** button for the correct answer.

**Result:**





**10. Learning Outcomes Achieved**

1. Understanding the natural language processing using artificial intelligence.
2. Studying how chunking of words are done.

**11. Conclusion:**

Natural Language Processing focuses on interaction between computer and human .

**References**:

[1] Artificial Intelligence: A modern approach, Stuart Russel and Peter Norvig, Pearson.

[2] Artificial Intelligence, Elaine Rich and Kevin Knight, Tata McGraw.

[3] Principles of Artificial Intelligence, Nils J. Nilson, Narosa Publications.

**Viva Questions**

1. What is NLP?
2. What are the components of NLP?
3. What is programming analysis in NLP?
4. What are the functional areas of NLP?