**NLP – WORKSHEET 4**

**(Solutions)**

1.

A) It consists of a set of production rules

B) The production rules are of the following form:

A-> BC where A is non terminal while B, C can be either terminal or non-terminal.

C) These grammars are free of context in which they are used, they will remain same regardless of the context in which they are used.

2.

A) All the production rules in PCFG has probability associated with them while in CFG we do not have Probability of a production rule.

B) With PCFG we can find the most probable parse tree of a sentence which we cannot find CFG.

3.

D) All of the above

4.

A) Free word parsing

5.

A) It establish dependencies between words of a sentence

B) The dependencies are established in terms of subject-object-verb and other dependencies.

6.

A) Chunking

C) unigram chunker

D) bigram chunker

7.

A) It uses the POS tag of a word and find the most probable IOB label for that POS tag

8.

C) It assigns that IOB label which has the maximum probability based on the POS tag

9.

A) It is a sequential modeling process for assigning POS tags to the word

B) It uses a dictionary of IOB labels to assign a IOB entity label

C) It uses the POS tag of the word and its previous word to assign the most probable IOB label

10.

B) Bigram chunker

11.

A) Word Tokenization

B) Lemmatization

C) Dependency Parsing

D) POS tagging

12.

D) All of the above

13.

A) It starts with start symbol S

B) we use the CFG production rule to generate the sentence from the S start symbol

14.

A) pattern = “#\w\*”

15.

C) pattern = “@\w\*”