SNS COLLEGE OF ENGINEERING

Kurumbapalayam(Po), Coimbatore – 641 97
Accredited by NAAC-UGC with 'A' Grade
Approved by AICTE, Recognized by UGC & Affiliated to Anna University, Chennai

Department of Artificial Intelligence and Data Science

Course Name – 19AD601 – Natural Language Processing

III Year / VI Semester

Unit 3 – SYNTACTIC ANALYSIS

Topic 1- Context Free Grammar

Context Free Grammar

- A widely used formal system for modeling constituent structure in natural language is the context-free grammar, or CFG.
- Context-free grammars are also called phrase-structure grammars, and the formalism is equivalent to Backus-Naur form, or BNF.
- A context-free grammar consists of a set of rules or productions, each of which expresses the ways that symbols of the language can be grouped and ordered together, and a lexicon of words and symbols.
- For example, the following productions express that an NP (or noun phrase) can be composed of either a ProperNoun or a determiner (Det) followed by a Nominal; a Nominal in turn can consist of one or more Nouns.

$$
\begin{aligned}
NP &\rightarrow Det\ Nominal \\
NP &\rightarrow ProperNoun \\
Nominal &\rightarrow Noun \mid Nominal\ Noun
\end{aligned}
$$

Context Free Grammar

The symbols that are used in a CFG are divided into two classes

- The symbols that correspond to words in the language (“the”, “nightclub”) are called terminal symbols; the lexicon is the set of rules that introduce these terminal symbols.

- The symbols that express abstractions over these terminals are called non-terminals.

- In each context-free rule, the item to the right of the arrow (□) is an ordered list of one or more terminals and non-terminals;

- To the left of the arrow is a single non-terminal symbol expressing some cluster or generalization.

A CFG can be thought of in two ways:

- as a device for generating sentences and
- as a device for assigning a structure to a given sentence

Context Free Grammar

The symbols that are used in a CFG are divided into two classes

- The symbols that correspond to words in the language (“the”, “nightclub”) are called terminal symbols; the lexicon is the set of rules that introduce these terminal symbols.

- The symbols that express abstractions over these terminals are called non-terminals.

- In each context-free rule, the item to the right of the arrow (□) is an ordered list of one or more terminals and non-terminals;

- To the left of the arrow is a single non-terminal symbol expressing some cluster or generalization.

A CFG can be thought of in two ways:

- as a device for generating sentences and
- as a device for assigning a structure to a given sentence

Context Free Grammar

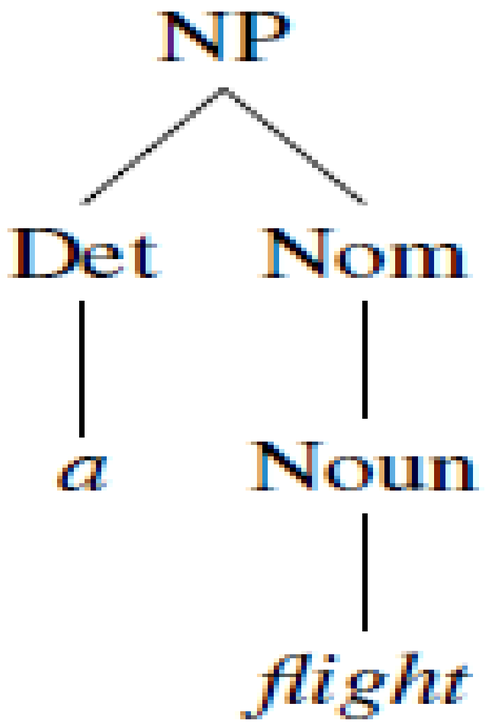
We say the string a flight can be derived from the non-terminal NP. Thus, a CFG can be used to generate a set of strings.

This sequence of rule expansions is called a derivation of the string of words. It is common to represent a derivation by a parse tree.

So starting from the symbol: NP
we can use our first rule to rewrite NP as: Det Nominal
and then rewrite Nominal as: Noun
and finally rewrite these parts-of-speech as: a flight

We say the string a flight can be derived from the non-terminal NP. Thus, a CFG can be used to generate a set of strings. This sequence of rule expansions is called a derivation of the string of words. It is common to represent a derivation by a parse tree.

Context Free Grammar



Context Free Grammar

Each grammar must have one designated start symbol, which is often called S.

Let’s add a few additional rules to our inventory. The following rule expresses the fact that a sentence can consist of a noun phrase followed by a verb phrase:

S → NP VP I prefer a morning flight

verb phrase in English consists of a verb followed by assorted other things; for example, one kind of verb phrase consists of a verb followed by a noun phrase:

VP → Verb NP prefer a morning flight

Or the verb may be followed by a noun phrase and a prepositional phrase:

VP → Verb NP PP leave Boston in the morning

Or the verb phrase may have a verb followed by a prepositional phrase alone:

VP → Verb PP leaving on Thursday

A prepositional phrase generally has a preposition followed by a noun phrase

PP → Preposition NP from Los Angeles

THANK YOU