



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

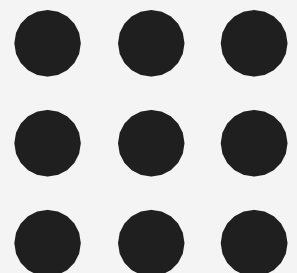
Department of Artificial Intelligence and Data Science

**Course Name – 19AD601 – Natural Language
Processing**

III Year / VI Semester

Unit 3 – SYNTACTIC ANALYSIS

Topic 2- Grammar Rule



Context Free Grammar

Formal Definition of Context-Free Grammar

A context-free grammar G is defined by four parameters: N ; Σ ; R ; S (technically it is a “4-tuple”).

N a set of **non-terminal symbols** (or **variables**)
 Σ a set of **terminal symbols** (disjoint from N)
 R a set of **rules** or productions, each of the form $A \rightarrow \beta$,
where A is a non-terminal,
 β is a string of symbols from the infinite set of strings $(\Sigma \cup N)^*$
 S a designated **start symbol** and a member of N

Capital letters like A , B , and S	Non-terminals
S	The start symbol
Lower-case Greek letters like α , β , and γ	Strings drawn from $(\Sigma \cup N)^*$
Lower-case Roman letters like u , v , and w	Strings of terminals

Context Free Grammar

Derivation is then a generalization of direct derivation:

Let $\alpha_1, \alpha_2, \dots, \alpha_m$ be strings in $(\Sigma \cup N)^*$, $m \geq 1$, such that

$$\alpha_1 \Rightarrow \alpha_2, \alpha_2 \Rightarrow \alpha_3, \dots, \alpha_{m-1} \Rightarrow \alpha_m$$

We say that α_1 **derives** α_m , or $\alpha_1 \xRightarrow{*} \alpha_m$.

Noun \rightarrow *flights* | *flight* | *breeze* | *trip* | *morning*
Verb \rightarrow *is* | *prefer* | *like* | *need* | *want* | *fly* | *do*
Adjective \rightarrow *cheapest* | *non-stop* | *first* | *latest*
| *other* | *direct*
Pronoun \rightarrow *me* | *I* | *you* | *it*
Proper-Noun \rightarrow *Alaska* | *Baltimore* | *Los Angeles*
| *Chicago* | *United* | *American*
Determiner \rightarrow *the* | *a* | *an* | *this* | *these* | *that*
Preposition \rightarrow *from* | *to* | *on* | *near* | *in*
Conjunction \rightarrow *and* | *or* | *but*

Figure 17.2 The lexicon for \mathcal{L}_0 .

Context Free Grammar

Derivation is then a generalization of direct derivation:

Let $\alpha_1, \alpha_2, \dots, \alpha_m$ be strings in $(\Sigma \cup N)^*$, $m \geq 1$, such that

$$\alpha_1 \Rightarrow \alpha_2, \alpha_2 \Rightarrow \alpha_3, \dots, \alpha_{m-1} \Rightarrow \alpha_m$$

We say that α_1 **derives** α_m , or $\alpha_1 \xRightarrow{*} \alpha_m$.

Noun \rightarrow *flights* | *flight* | *breeze* | *trip* | *morning*
Verb \rightarrow *is* | *prefer* | *like* | *need* | *want* | *fly* | *do*
Adjective \rightarrow *cheapest* | *non-stop* | *first* | *latest*
| *other* | *direct*
Pronoun \rightarrow *me* | *I* | *you* | *it*
Proper-Noun \rightarrow *Alaska* | *Baltimore* | *Los Angeles*
| *Chicago* | *United* | *American*
Determiner \rightarrow *the* | *a* | *an* | *this* | *these* | *that*
Preposition \rightarrow *from* | *to* | *on* | *near* | *in*
Conjunction \rightarrow *and* | *or* | *but*

Figure 17.2 The lexicon for \mathcal{L}_0 .

Context Free Grammar

Grammar Rule:

Grammar Rules	Examples
$S \rightarrow NP VP$	I + want a morning flight
$NP \rightarrow$ <i>Pronoun</i> <i>Proper-Noun</i> <i>Det Nominal</i>	I Los Angeles a + flight
$Nominal \rightarrow$ <i>Nominal Noun</i> <i>Noun</i>	morning + flight flights
$VP \rightarrow$ <i>Verb</i> <i>Verb NP</i> <i>Verb NP PP</i> <i>Verb PP</i>	do want + a flight leave + Boston + in the morning leaving + on Thursday
$PP \rightarrow$ <i>Preposition NP</i>	from + Los Angeles

Figure 17.3 The grammar for \mathcal{L}_0 , with example phrases for each rule.

Context Free Grammar

Parse Tree

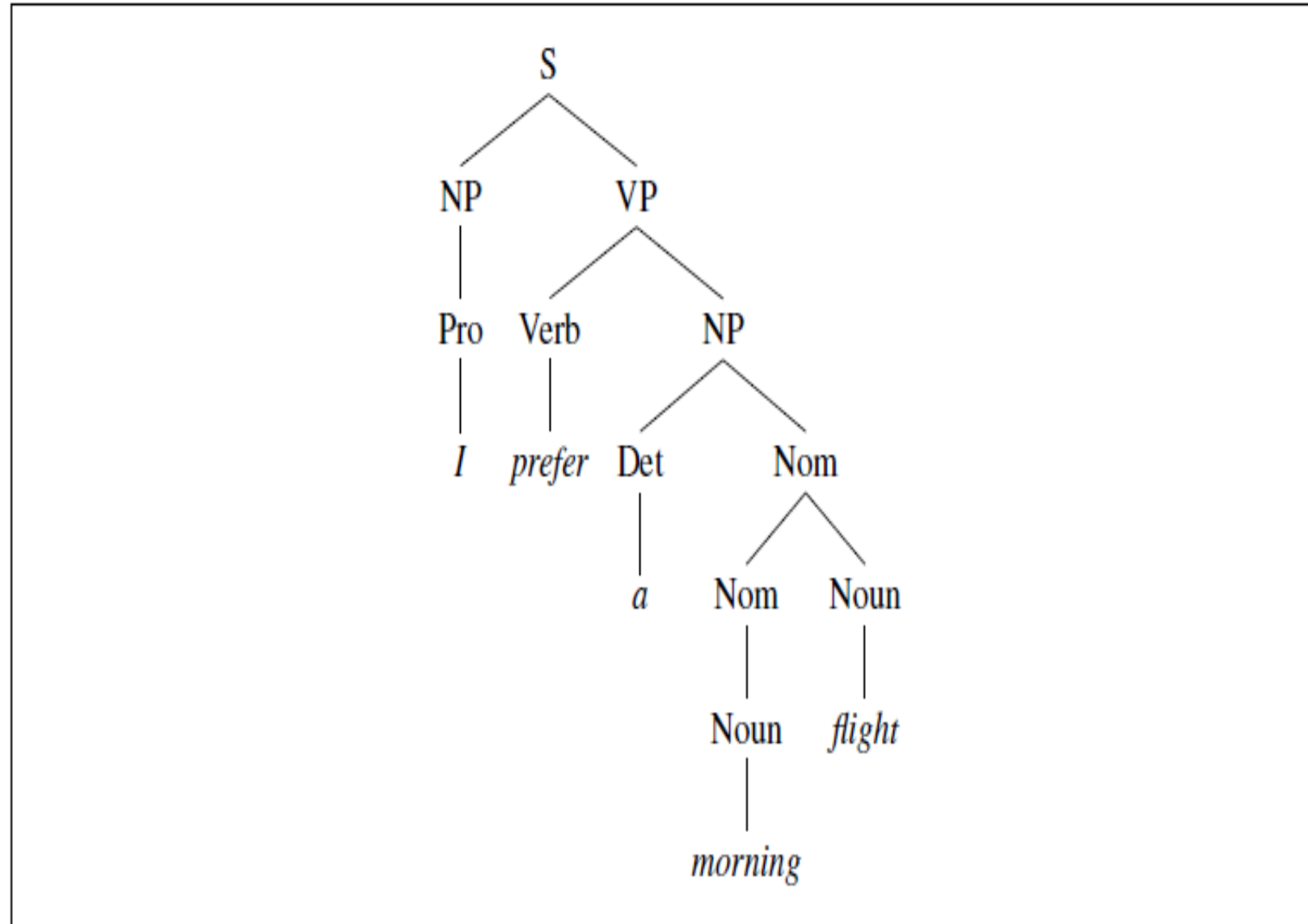


Figure 17.4 The parse tree for “I prefer a morning flight” according to grammar \mathcal{L}_0 .



Context Free Grammar



Treebanks

A corpus in which every sentence is annotated with a parse tree is called a treebank.

Treebanks play an important role in parsing as well as in linguistic investigations of syntactic phenomena.

Treebanks are generally made by parsing each sentence with a parse that is then hand-corrected by human linguists.



THANK YOU