OPEN ACCESS

International
Journal
Of Modern Engineering Research (IJMER)

Engineering English Grammar: A Rule-Based Approach Using CFG

[1]Prof. Rati Ranjan Sahoo, [2]Prof. Subhadra Biswal, [3]Biswajit Harpal
[1,2,3]*Department of CSE, Einstein Academy of Technology and Management, Khordha, Bhubaneswar*

Abstract:
This study illustrates different types of sentences to develop a deep-rooted understanding of their structure and constituent grammatical elements. Using Context-Free Grammar (CFG), we extract key components such as noun phrases (NP) and verb phrases (VP) and parse input sentences based on custom-defined grammar rules. Several tools, including NLTK and Matplotlib, are utilized, along with the CFG and Tree modules, to generate and visualize parse trees.

I. Introduction

Grammar parsing is essential for understanding the structural composition of sentences, enabling applications in natural language processing [2], machine translation, and grammar-checking tools. Traditional approaches to grammar learning often rely on memorization, whereas an engineering approach provides a systematic method to analyze sentence structures using formal rules. By leveraging rule-based techniques like Context-Free Grammar (CFG), we can break down sentences into their fundamental components, such as noun phrases (NP) and verb phrases (VP), allowing for precise syntactic analysis [1,4]. This structured approach not only enhances comprehension but also enables automation in linguistic processing, making it valuable for both linguistic research and AI-driven applications.

This article aims to analyze English sentence structures using a rule-based approach with Context-Free Grammar (CFG). By defining custom grammar rules, we extract key components such as noun phrases (NP) and verb phrases (VP) to parse sentences systematically. The study employs the Natural Language Toolkit (NLTK) for implementing CFG-based parsing and utilizes Matplotlib to visualize parse trees, providing a clear representation of syntactic structures. This approach demonstrates how computational methods can enhance grammar analysis, making it useful for natural language processing and linguistic applications [1,3,4].

II. Background & Theoretical Foundations

Grammar and parsing are fundamental in computational linguistics, enabling the analysis of sentence structure through defined syntax and rules [3,4] . Context-Free Grammar (CFG) is widely used for this purpose as it provides a structured, rule-based method to break sentences into components like noun phrases (NP) and verb phrases (VP). Unlike statistical or machine learning-based parsing, which relies on large datasets, a rule-based CFG approach ensures precise and interpretable sentence analysis. This makes CFG an ideal choice for studying English grammar systematically, allowing for better control over parsing rules and sentence structure representation.

III. Methodology: Implementing Rule-Based Parsing

Grammar Rules (CFG):

S -> MainClause SubClause | NP VP | NP VP S1 | S Conj S | NP | VP

S1 -> WH-phrase VP

WH-phrase -> WH Prep

WH -> "how"

MainClause -> NP VP

SubClause -> Conj NP VP | Rel VP

NP -> N |N N |AdjP N | N ProperNoun|Quantifier N |Adj NP| ProperNoun | ProperNoun N | Gerund |NP Conj NP | D NP |NP PP| Pronoun |Pronoun NP | Quantifier Adj Adj N | D Adj N

D -> "this" | "each" | "A" | "a" | "an" | "the"|"The" |"An"| "some"

Pronoun -> "I" | "There" | "it" | "its" |"your" | "us" | "you"

ProperNoun -> "Jagarnath" |"Radhe" | "Ram" | "Sita" | "Puja" | "Sipra" | "It"

VP -> VP Adv | VP PP | V NP PP Adv | V | Adv VP | V Adv Adj | V NP | V NP Adv| V PP | Aux NP VP | Aux VP | Aux V InfVP |V Adj InfVP | V Pronoun VP

Aux -> "are" | "is" | "am" |"will" | "be"| "has"

AdjP -> Adj Conj Adj

Adj ->"current" | "past" | "sorry" | "tiring" | "following" | "valuable" | "distinct" | "visual" | "great"

V -> "teaches" | "solve" | "prepare" | "having" | "handling" | "narrated" | "like" |"touch" |"am" | "hear" | "graduated" | "plucking" | "have" | "let" | "is" | "taking" | "going" | "playing" | "singing" | "said" | "be" | "changes"| "acknowledges" | "likes" |"like" | "created" | "is" | "playing" |"singing"

InfVP -> Inf V PP

PP -> PP PP | Prep NP | Prep Num | Prep Gerund Conj Gerund | Prep Gerund | PP NP

Adv ->"also" |"today" | "very" | "separately"

Prep -> "about" | "to" | "with" | "by" | "of" | "for" |"at" | "with" | "to"

Inf -> "to"

N -> "book" | "problems" | "math" | "timetable" | "branches" | "branch" | "faculty member" | "subjects" | "stories" | "college" | "achievements" |"Alumnus"|"dog" | "flowers" |"flower" | "fun" | "sir" | "mathematics" | "class" | "football" | "temple" | "reference"|"object" | "position" | "respect" | "frame" | "motion" | "reference" | "manuscript" | "book" | "academics" | "contribution"| "council" | "song" | "football" | "flower" | "boy"|"girl" | "systems" | "rods" | "cones" | "workout"

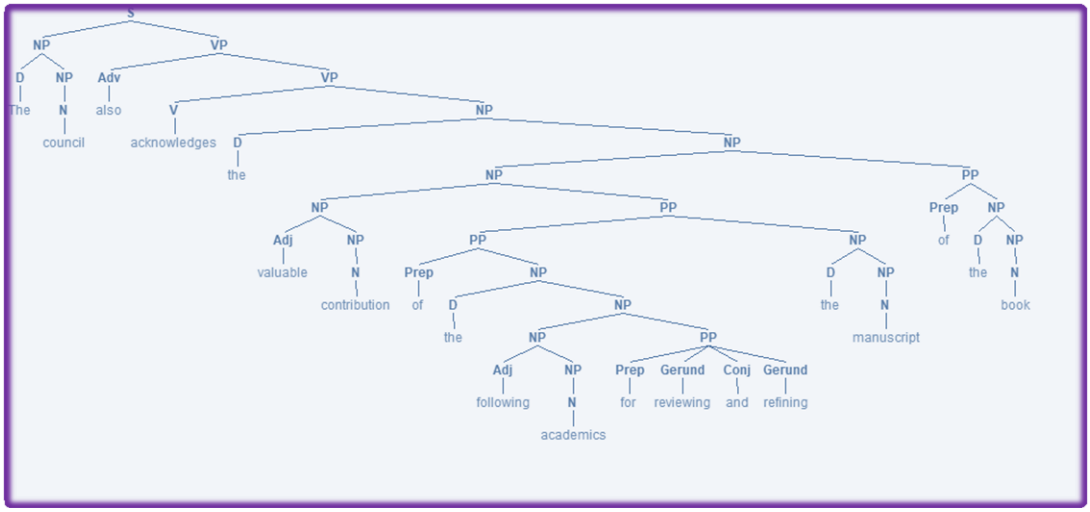
Gerund -> "swimming" |"reviewing" | "refining" | "tiring"

Conj -> "and" | "when" | ";"

Parsing Sentences Using NLTK[5]:

sentence = "The council also acknowledges the valuable contribution of the following academics for reviewing and refining the manuscript of the book".split()

IV. Experiment & Results



V. Conclusion

In conclusion, this study demonstrates how a rule-based approach using Context-Free Grammar (CFG) can effectively analyze English sentence structures. By leveraging NLTK for parsing and Matplotlib for visualization, we illustrate the systematic breakdown of sentences into their grammatical components. This method provides a precise and interpretable way to understand syntax, making it valuable for linguistic research and natural language processing applications. Future work can explore extending CFG with probabilistic methods to handle ambiguity and complex sentence structures more efficiently.

References

[1]. Yogish, D., Manjunath, T. N., & Hegadi, R. S. (2019). Review on Natural Language Processing Trends and Techniques Using NLTK. In *Communications in Computer and Information Science* (Vol. 1037). Springer Singapore. https://doi.org/10.1007/978-981-13-9187-3_53

[2]. Lauriola, I., Lavelli, A., & Aiolli, F. (2022). An introduction to Deep Learning in Natural Language Processing: Models, techniques, and tools. Neurocomputing, 470(xxxx), 443–456. https://doi.org/10.1016/j.neucom.2021.05.103

[3]. Gao, C. (2011). Theory and Practice in Language Studies. *Theory and Practice in Language Studies*, *1*(5).

[4]. Cambria, E. (2014). *jumping-nlp-curves*. *May*.

[5]. Navlani, A. (2019). Text ANALYTICS for Beginners using NLTK. *Datacamp*, 1–14. https://www.datacamp.com/community/tutorials/text-analytics-beginners-nltk