Theory of Parsing

Parsing

- A Parser in NLP Uses the grammar rules(formal grammar rules) to verify if the input text is valid or not syntactically.
- The parser helps us to get the meaning of the provided text.
- The parser helps us to Analyze the syntax error in the text, so the parsing process is also known as the syntax analysis and syntactic analysis.
- The syntax analysis uses the formal grammar rules to check the text and if the text follows the rules of grammar then it accepts the text else it rejects the text. For example, if the input text is: Give me hot ice cream.. Then the parser or the syntactic analyzer will reject the text as it does not follow the rules of formal grammar.

Parsing

- We have mainly two types of parsing techniques : **Top Down parsing and Bottom up parsing**.
- In the **Top Down Parsing** Approach the construction of parse tree starts from the root node. And in the **Bottom up Parsing Approach ,**the construction of parse tree starts from the leaf node.
- the parser plays a very vital role as it reports any kind of syntax error in the text. The parser not only tells us about the syntax errors but also helps us to recover from the most commonly occurring errors so that the processing of the rest of the program does not get halted

Function of a Parser

- The parser reports to us any syntax error in the syntax of the text.
- The parser helps us to recover from the most commonly occurring errors so that the processing of the rest of the program does not get halted.
- Parser helps in generating the parse tree.
- Parser also helps in creating the symbol table that is used by various stages of the NLP process and thus it is quite important.
- Parses also help in the production of IR or Intermediate Representations.

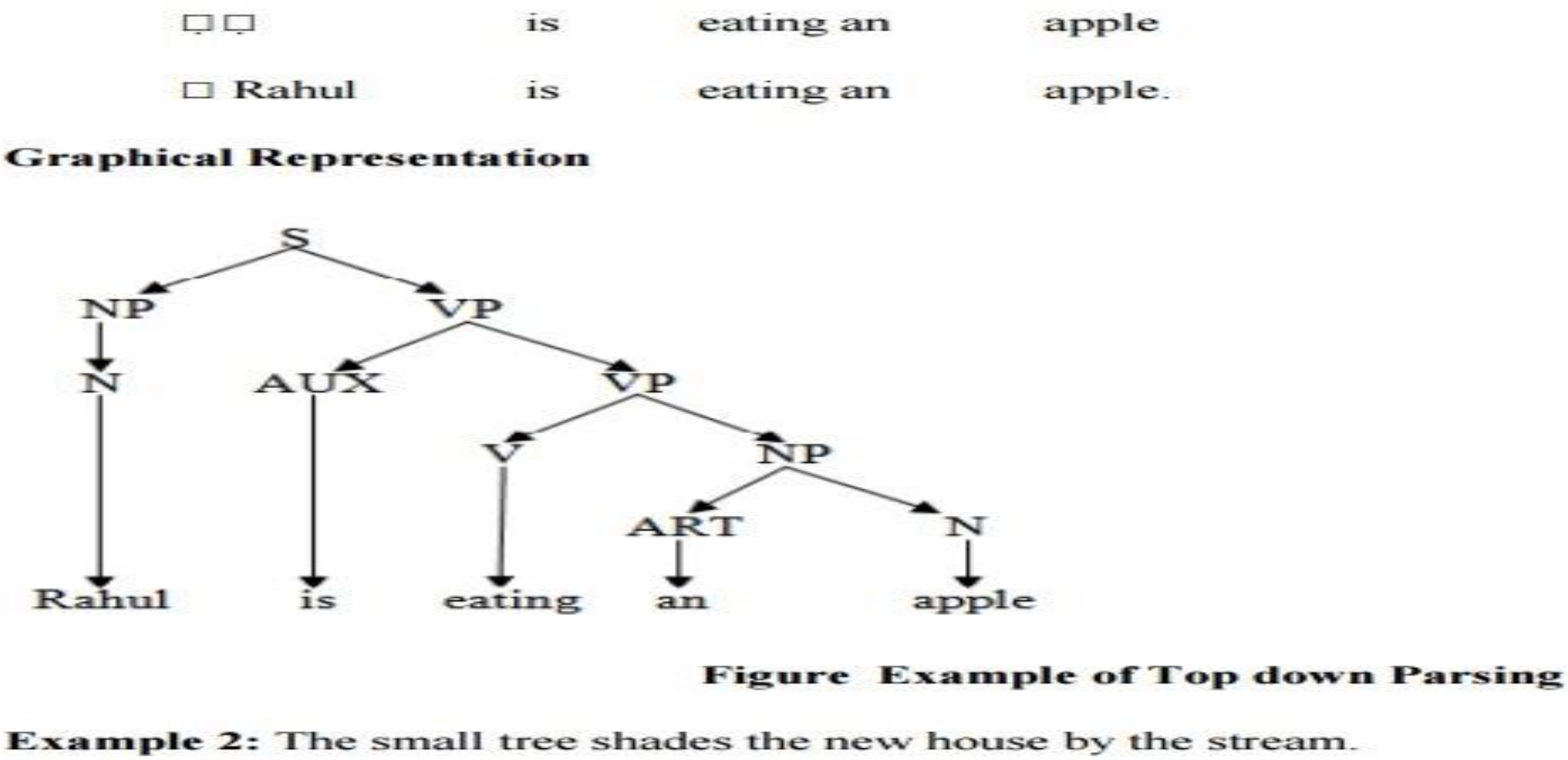
Parsing play a vital Role in NLP Process

- **Sentiment Analysis**.
- **Relation Extraction**.
- **Question Answering**.
- **Speech Recognition**.
- **Machine Translation**.
- **Grammar Checking**.

Top Down Parsing

- The top-down parsing uses the leftmost derivation approach to construct the parsing tree of the text.
- The top-down parser in NLP does not support the grammar having common prefixes.
- The top-down parser in NLP also guarantees that the grammar is free from all ambiguity and has followed the left recursion.
- We Can perform Top Down parsing in two ways : 1) Using Backtracking   2) without using backtracking

Top Down parser



□□ is eating an apple

□ Rahul is eating an apple.

Graphical Representation

Figure Example of Top down Parsing

Example 2: The small tree shades the new house by the stream.

Top Down Parsing: **Recursive descent parser**

- The **recursive descent** parser in NLP follows the top-down parsing approach. This parser checks the syntax of the input stream of text by reading it from left to right (hence, it is also known as the Left-Right Parser).
- The parser first reads a character from the input stream and then verifies it or matches it with the grammar's terminals. If the character is verified, it is accepted else it gets rejected.

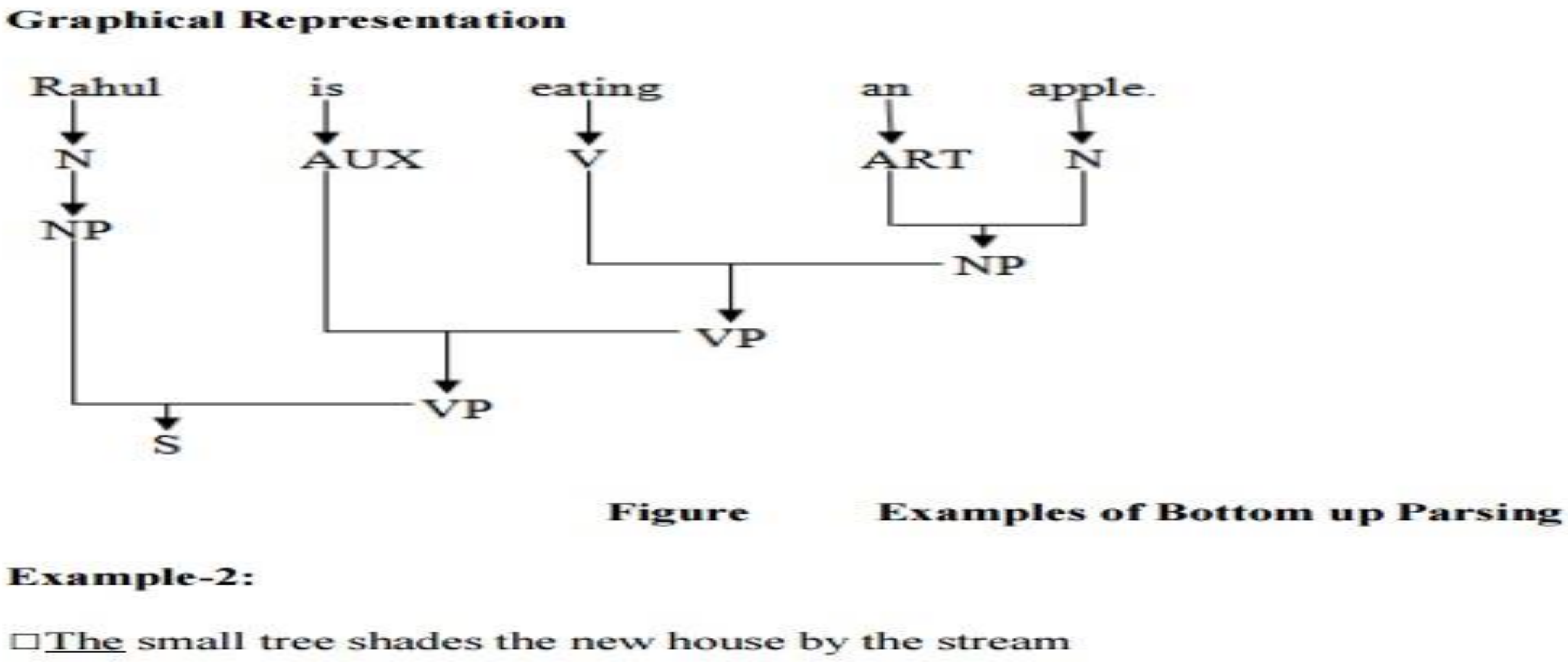
Bottom up parsing

- The bottom-up parsing approach follows the leaves-to-root approach or technique. In the bottom-up parsing approach, the construction of the parse tree starts from the leaf node.
- So, first, the leaf node is formed and then the generation goes subsequently up by generating the parent node, and finally, the root node is generated (following the bottom-to-up approach).
- The main aim of bottom-up parsing is to reduce the input string of text to get the start symbol by using the rightmost derivation technique

Bottom up parsing

- The bottom-up parsing approach is used by various programming language compilers such as C++, Perl, etc.
- This technique can be implemented very efficiently.
- The bottom-up parsing technique detects the syntactic error faster than other parsing techniques.

Bottom up Parsing



Graphical Representation

Rahul
is
eating
an
apple.
N
AUX
V
ART
N
NP
NP
VP
VP
S

Figure Examples of Bottom up Parsing

Example-2:

□The small tree shades the new house by the stream

Bottom up parsing

The bottom-up parsers or LR parsers are of four types:

- **LR(0)**
- **SLR(1) or Simple LR**
- **LALR or Look Ahead LR**
- **CLR or Canonical LR**

Finite State Parsing Method

- The finite-state parsing methods have evolved in the past two decades and provided us with very interesting studies on Natural Language Processing.
- Morphological Parsing and Dependency parsing are some of the most important finite-state parsing methods.
- In the morphological parsing, we try to find out the morphemes of the provided input word.
- The morphological parsing helps us to yield important information that is used in NLP applications.

Finite State Parsing Method

- In the dependency parsing method, our main aim is to find the linguistic units or words that are related to each other via direct link.
- This direct link is depicting the dependencies of the words..
- The NLTK Python package provides us with two ways of performing dependency parsing namely- Probabilistic, projective dependency parsing, and Stanford parsing