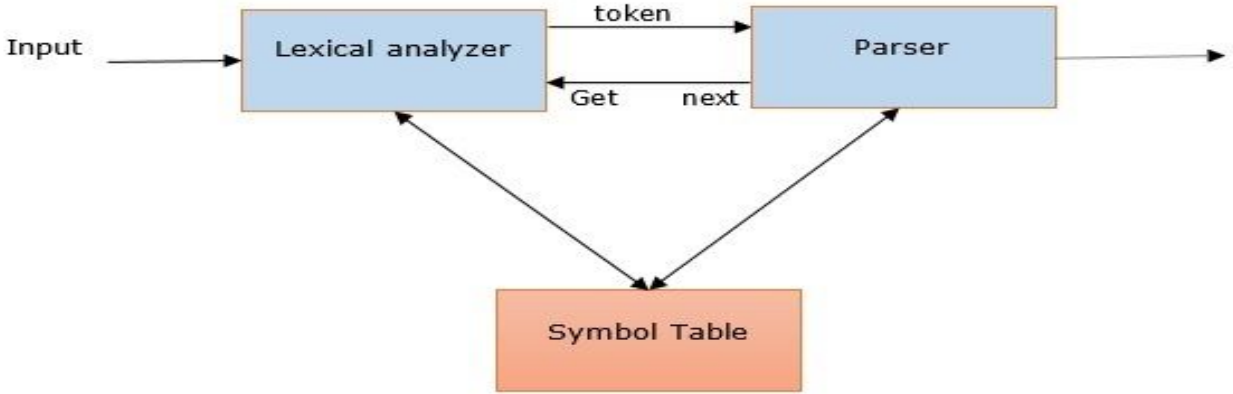
Parsing

- Gouthami Sirpurkar
11450270

Definition

The word ‘Parsing’ whose origin is from Latin word ‘pars’ (which means ‘part’), is used to draw exact meaning or dictionary meaning from the text.



Relevance of parsing in NLP

Parser is used to report any syntax error.

Parse tree is created with the help of a parser.

Parser is used to create symbol table, which plays an important role in NLP.

Parser is also used to produce intermediate representations (IR).

Context Free Grammar

Context-free grammars are the backbone of many formal models of the syntax of natural language.

they play a role in many computational applications, including grammar checking, semantic interpretation, dialogue understanding, and machine translation.

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 to- lexicon gther, and a lexicon of words and symbols

Interpreting Language is Hard

"Parsing" resolves structural ambiguity in a formal way

"I saw a girl with a telescope"

Dependency Parsing

In dependency formalisms, phrasal constituents and phrase-structure rules do not play a direct role.

Instead, the syntactic structure of a sentence is described solely in terms of directed binary grammatical relations between the words.

Relations among the words are illustrated above the sentence with directed, labeled arcs from heads to dependents.

A root node explicitly marks the root of the tree, the head of the entire structure

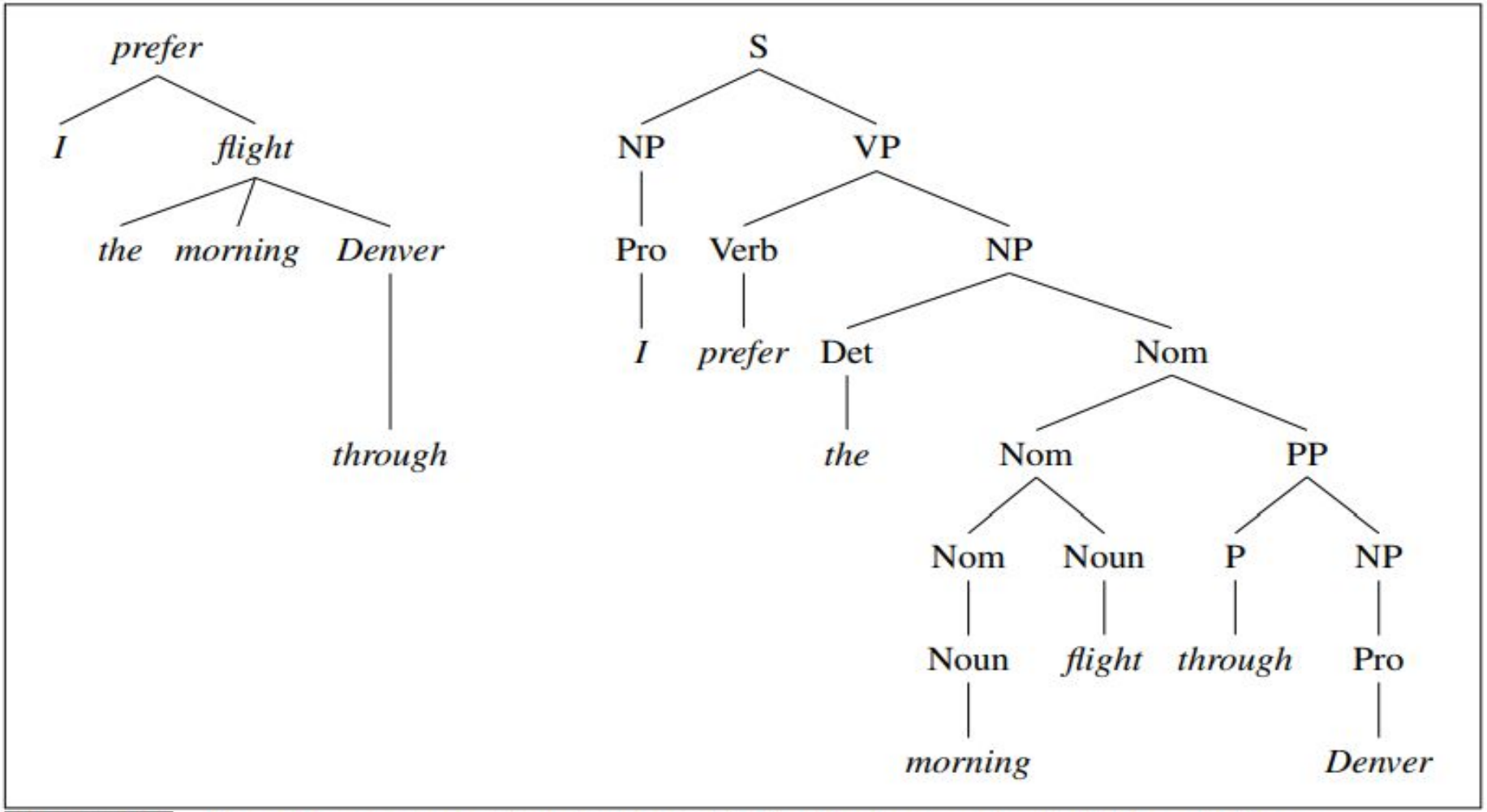


Figure 14.1 Dependency and constituent analyses for *I prefer the morning flight through Denver.*

Advantages:

A major advantage of dependency grammars is their ability to deal with lanfree word order guages that are morphologically rich and have a relatively free word order

Reusability of labour

Broad coverage, not just few intuitions

A way to evaluate systems

Dependency structure

Binary asymmetric relationships between lexical items(words)

The arrows represent dependencies

Heads and dependent

Root

Dependency Tree

A dependency structure can be represented as a directed graph G = (V,A), consisting of a set of vertices V, and a set of ordered pairs of vertices A, which we'll call arcs.

a dependency tree is a directed graph that satisfies the following constraints:

1. There is a single designated root node that has no incoming arcs.

2. With the exception of the root node, each vertex has exactly one incoming arc.

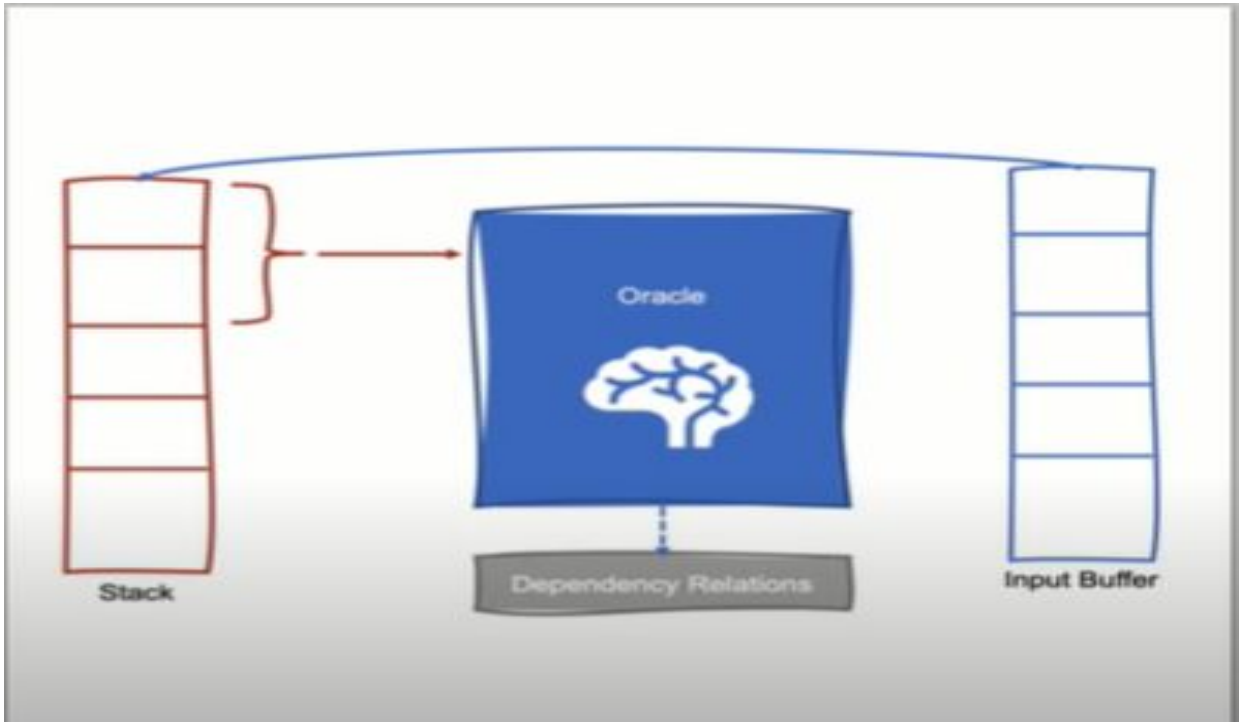
3. There is a unique path from the root node to each vertex in V.

Approaches

1: Transition Based Dependency Parsing

2: Graph Based Dependency Parsing

Transition Based Dependency Parsing



Graph Based dependency parsing

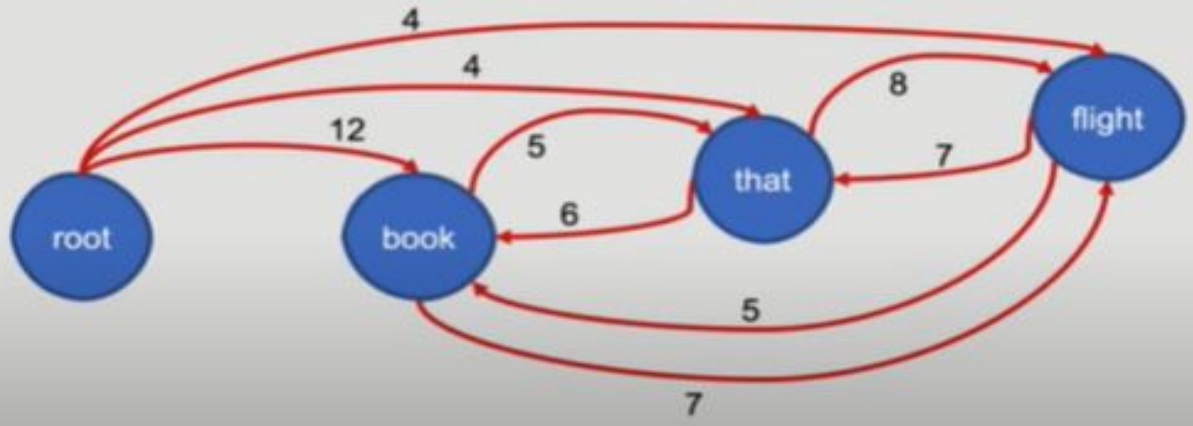
Edge Factored Approach:

the overall score for a tree is the sum of the scores of each of the scores of the edges that comprise the tree.

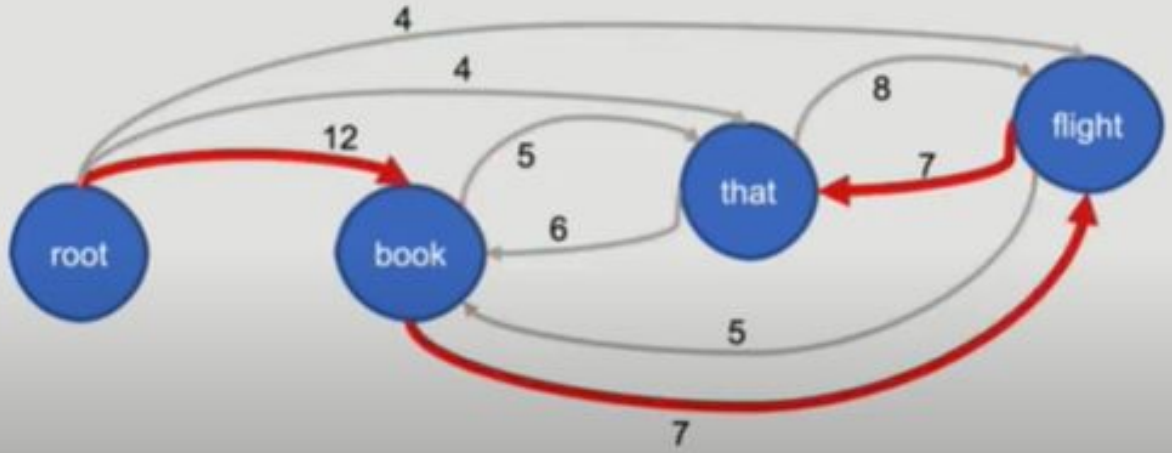
Maximum Spanning Tree:

A maximum spanning tree is the spanning tree with the highest score. Thus a maximum spanning tree of G emanating from the ROOT is the optimal dependency parse for the sentence.

Maximum Spanning Tree: Example



Maximum Spanning Tree: Example



Thank
you !!!
...