SNS COLLEGE OF ENGINEERING

Department of Artificial Intelligence and Data Science

Course Name – 19AD601 – Natural Language Processing

III Year / VI Semester

Unit 3 – SYNTACTIC ANALYSIS

Topic 5- Syntactic Parsing, Ambiguity

Graph-Based Dependency Parsing

- Graph-based parsers are more accurate than transition-based parsers, especially on long sentences.

- Graph-based parsers can produce non-projective trees. Graph-based dependency parsers search through the space of possible trees for a given sentence for a tree (or trees) that maximize some score.

- These methods encode the search space as directed graphs and employ methods drawn from graph theory to search the space for optimal solutions.

- More formally, given a sentence S we're looking for the best dependency tree in Gs, the space of all possible trees for that sentence, that maximizes some score.

$$\hat{T}(S) = \underset{t \in \mathcal{G}_S}{\mathrm{argmax}} \, \mathrm{Score}(t, S)$$

We'll make the simplifying assumption that this score can be edge-factored, meaning that the overall score for a tree is the sum of the scores of each of the scores of the edges that comprise the tree.

$$Score(t,S) = \sum_{e \in t} Score(e)$$

Graph-based algorithms have to solve two problems: (1) assigning a score to each edge, and (2) finding the best parse tree given the scores of all potential edges.

Ambiguity

Ambiguity is the most serious problem faced by syntactic parsers. Structural ambiguity occurs when the grammar can assign more than one parse to a sentence.

Groucho Marx's well-known line as Captain Spaulding in Animal Crackers is ambiguous because the phrase in my pajamas can be part of the NP headed by elephant or a part of the verb phrase headed by shot.

Structural ambiguity, appropriately enough, comes in many forms. Two common kinds of ambiguity are
- attachment ambiguity and
- coordination ambiguity
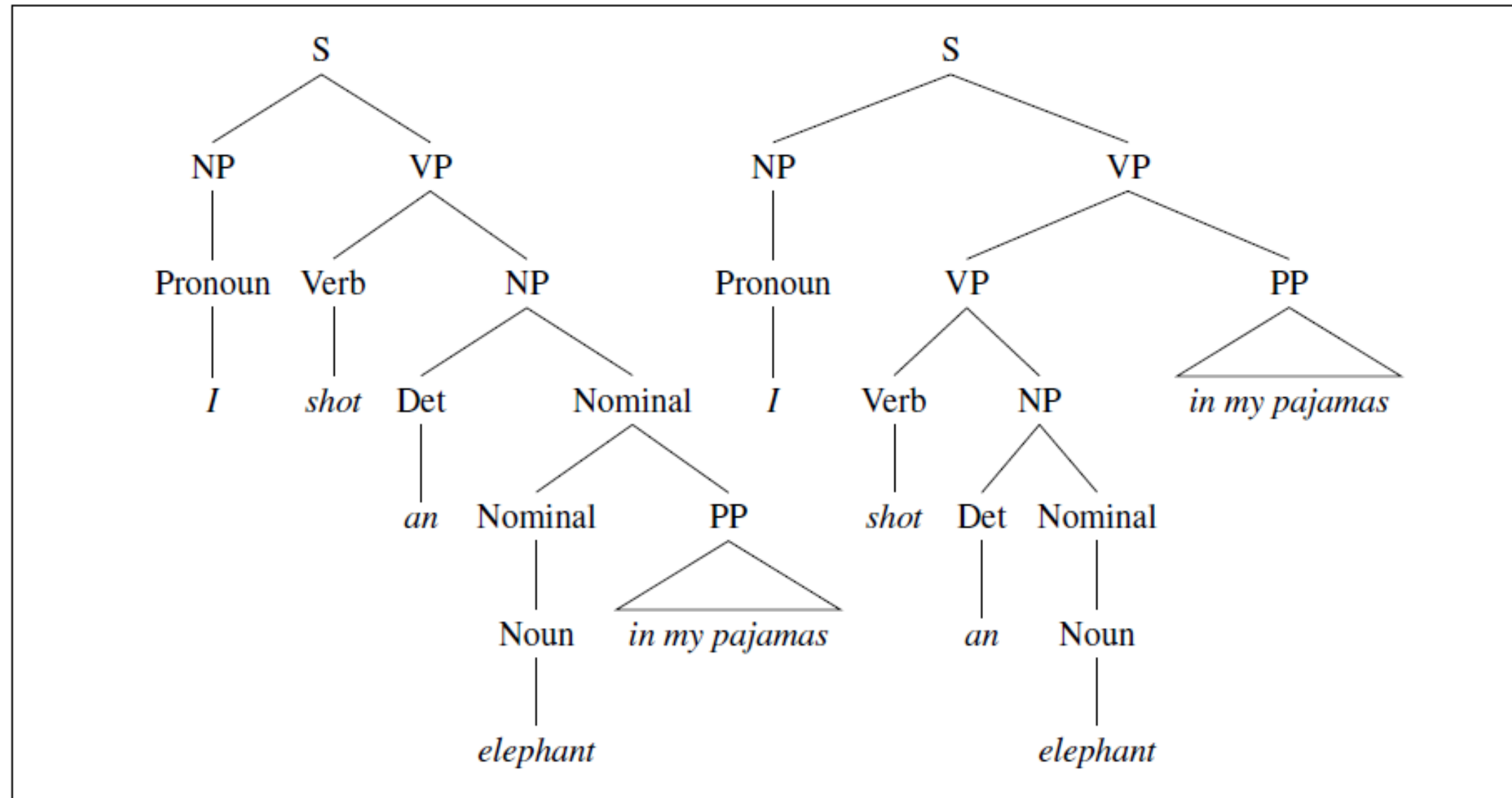
Syntactic Parsing, Ambiguity

Figure 17.9 Two parse trees for an ambiguous sentence. The parse on the left corresponds to the humorous reading in which the elephant is in the pajamas, the parse on the right corresponds to the reading in which Captain Spaulding did the shooting in his pajamas.

Syntactic Parsing, Ambiguity

- A sentence has an attachment ambiguity if a particular constituent can be attached to the parse tree at more than one place.

- The Groucho Marx sentence is an example of PP-attachment ambiguity: the preposition phrase can be attached either as part of the NP or as part of the VP.

- In coordination ambiguity phrases can be conjoined by a conjunction like and.

- For example, the phrase old men and women can be bracketed as [old [men and women]], referring to old men and old women, or as [old men] and [women], in which case it is only the men who are old.

- The fact that there are many grammatically correct but semantically unreasonable parses for naturally occurring sentences is an irksome problem that affects all parsers.

- Fortunately, the CKY algorithm below is designed to efficiently handle structural ambiguities.

THANK YOU