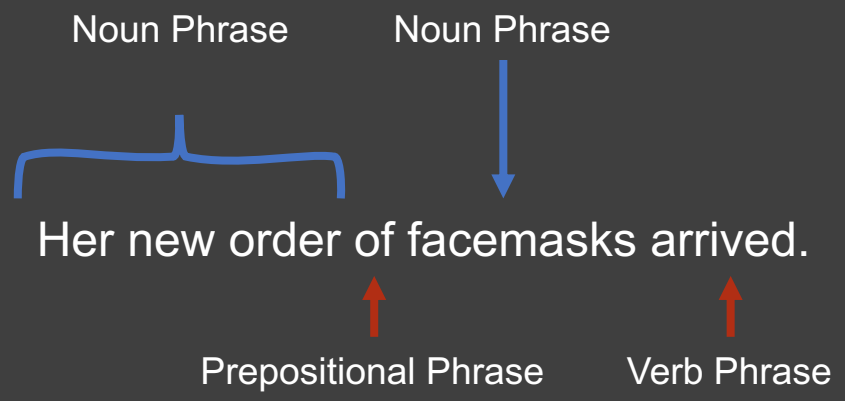
Partial Parsing

Natalie Parde
UIC CS 421

What if we don’t need a full parse tree?

- Full parse trees can be complex and time-consuming to build
- Many NLP tasks don’t require full hierarchical parses

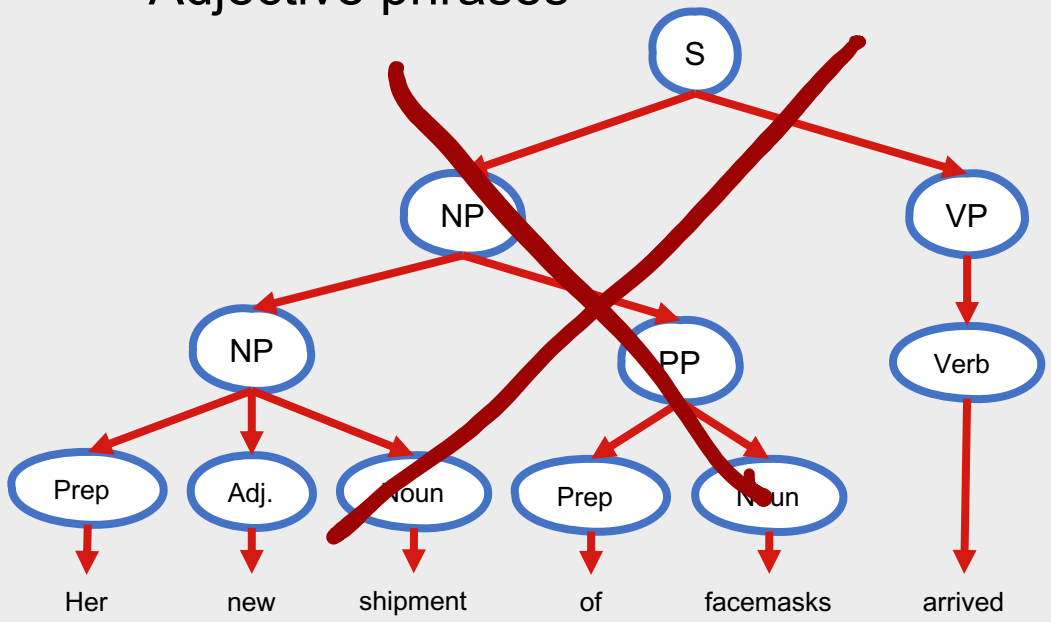


Easier solution?

- **Partial parsing**, or **shallow parsing**
- How to generate a partial parse?
  - Cascades of finite state transducers
  - **Chunking**

Chunking

- Process of finding the non-overlapping, non-recursive constituents in an input text
  - Noun phrases
  - Verb phrases
  - Prepositional phrases
  - Adjective phrases



$[\text{Her new shipment}]_{NP}\ [\text{of}]_{PP}\ [\text{facemasks}]_{NP}\ [\text{arrived}]_{VP}$

Chunking: Fundamental Tasks

1

Segmentation: Identify the non-overlapping, fundamental phrases

[Her new order] [of] [facemasks] [arrived]

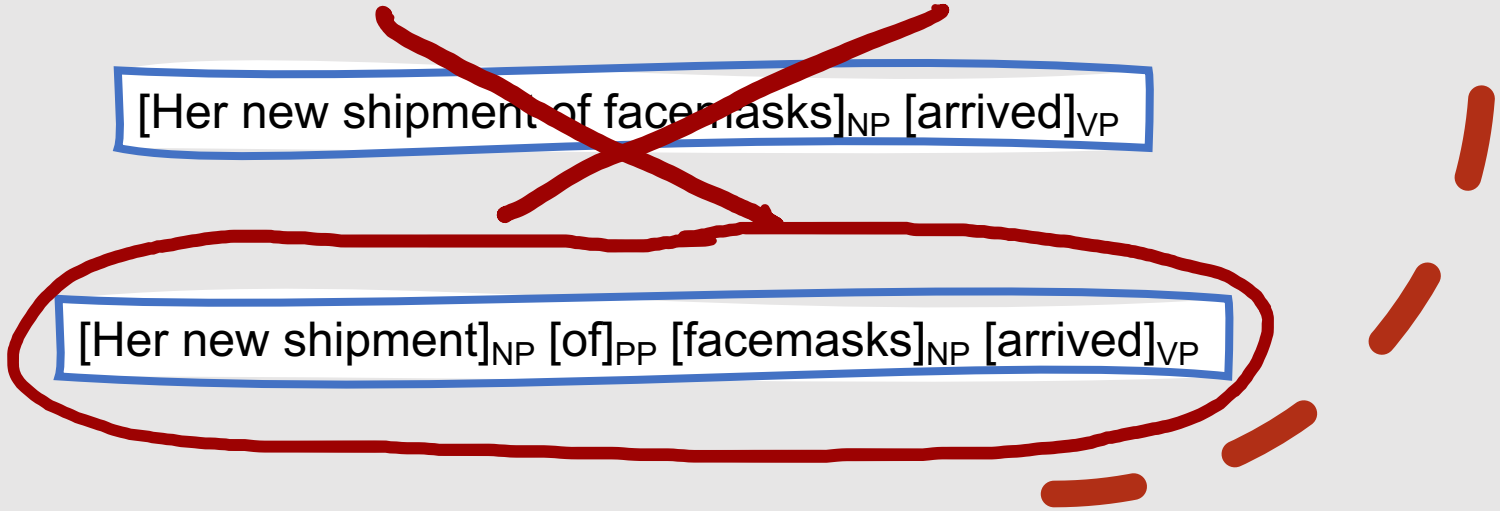
2

Labeling: Assign labels to those phrases

$[\text{Her new order}]_{NP}$ $[\text{of}]_{PP}$ $[\text{facemasks}]_{NP}$ $[\text{arrived}]_{VP}$

What is, and is not, a chunk?

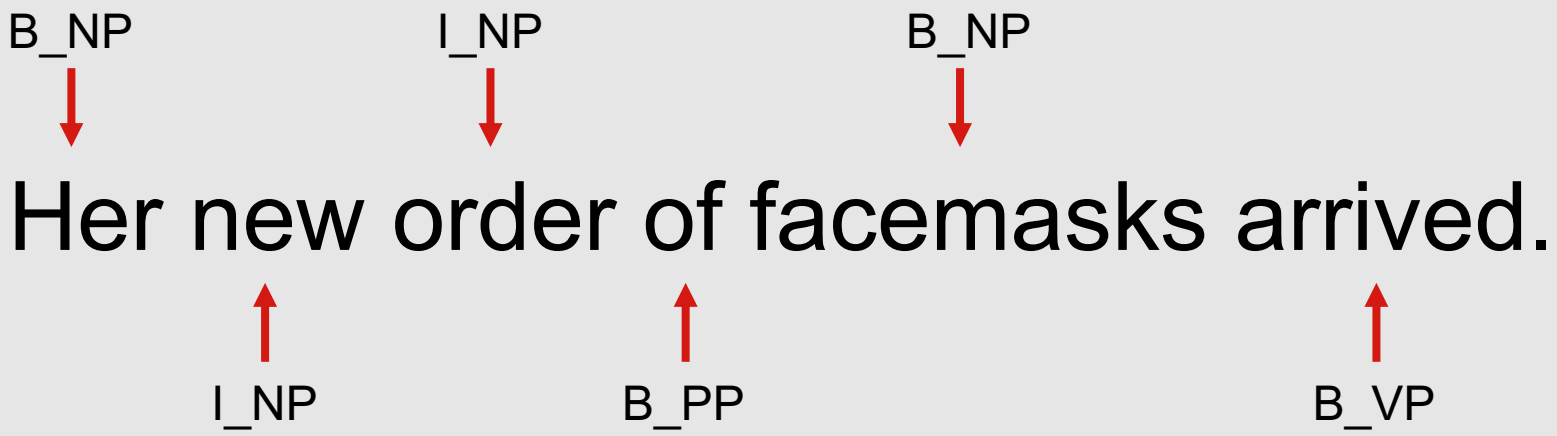
- Depends on the task!
- General guidelines:
  - Non-recursive
  - When chunking phrases that would otherwise be parsed recursively:
    - Keep head word
    - Keep all material belonging to constituent that occurs before the head word



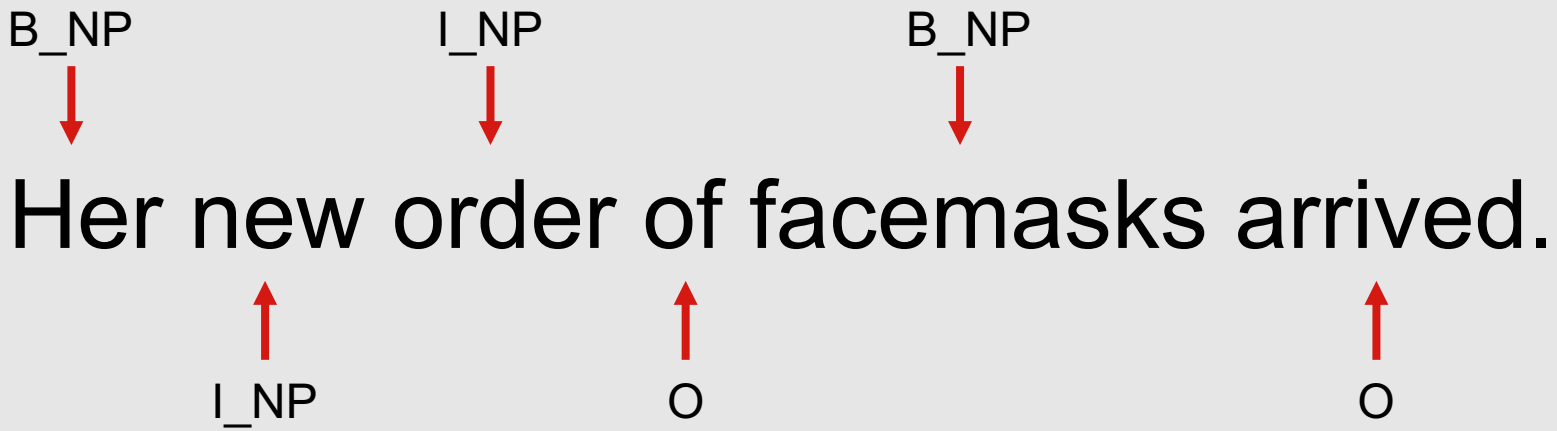
How do we segment text into chunks?

- **IOB tagging**
  - **I:** Tokens **inside** a chunk
  - **O:** Tokens **outside** any chunk
  - **B:** Tokens **beginning** a chunk
- Generally framed as a **sequence labeling** task

Task: IOB Tagging (All Constituent Types)



Task: IOB Tagging (Noun Phrases)



How do we evaluate chunking systems?

- Standard text classification metrics, comparing predictions with a gold standard
  - Precision
  - Recall
  - F-measure