

ITMO UNIVERSITY

NLP – Dependency Parsing

Quick Introduction

Mar 12th, 2020

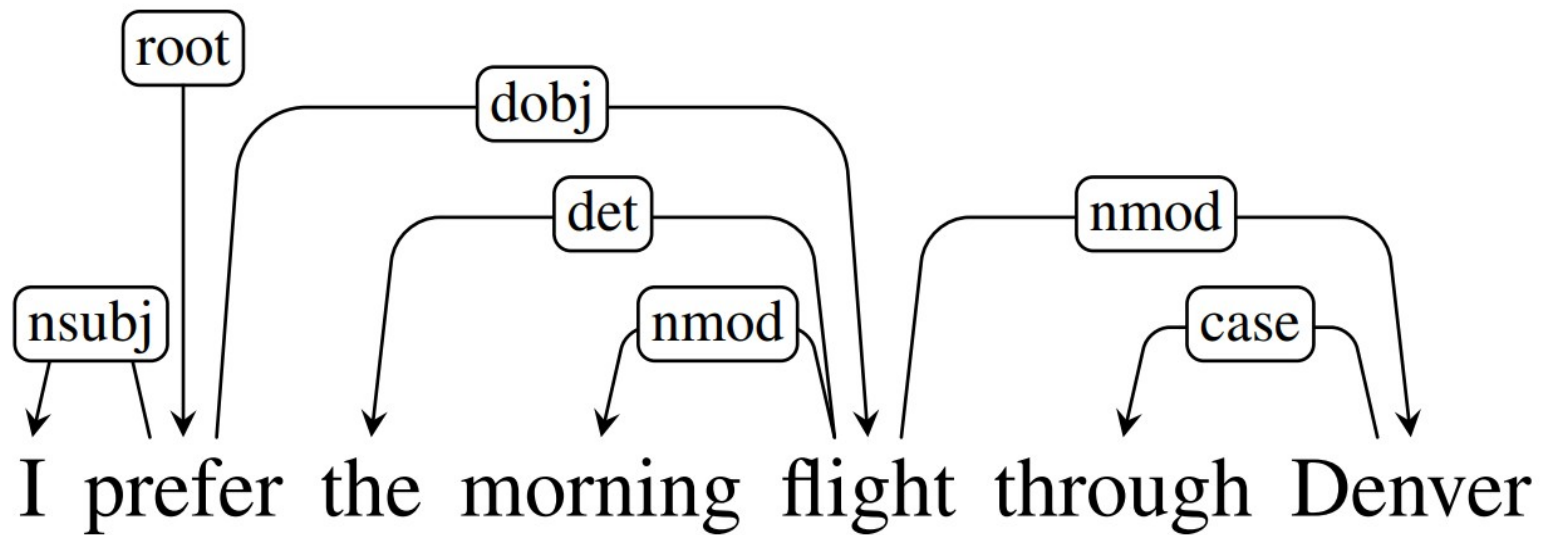
References

- ✓ <https://web.stanford.edu/~jurafsky/slp3/15.pdf>
- ✓ State-of-the-art methods:
http://nlpprogress.com/english/dependency_parsing.htm

Dependency Parsing

- ✓ **Goal:** we want to understand the **grammatical structure of a sentence**, similar to context-free constituent parsing
- ✓ In contrast to constituent parsing, **not based on rules** for the structure of the sentence, but **more flexible**, well-suited for example for languages like Russian, which have a flexible word order.
DPs abstracts over word order ...

Example



(15.1)

Basics

- ✓ We look at binary relationships between words.
- ✓ The relations are always between a **head** and a **dependent**
- ✓ The relationships are **directed** and **labeled**
- ✓ The label show the type of grammatical relationships
- ✓ There is always **one root node** (the predicate of the sentence)

Motivation – Why DP?

- ✓ **Question answering:** Who invented the light bulb? → If we understand the structure of the sentence, we can more easily answer the query. When was the light bulb invented → modifier.
- ✓ Information extraction
- ✓ Co-reference resolution

- ✓ → this information is very helpful / necessary for many NLP problems

Relation types

- ✓ There is a fixed number of relation types

Clausal Argument Relations	Description
NSUBJ	Nominal subject
DOBJ	Direct object
IOBJ	Indirect object
CCOMP	Clausal complement
XCOMP	Open clausal complement
Nominal Modifier Relations	Description
NMOD	Nominal modifier
AMOD	Adjectival modifier
NUMMOD	Numeric modifier
APPOS	Appositional modifier
DET	Determiner
CASE	Prepositions, postpositions and other case markers
Other Notable Relations	Description
CONJ	Conjunct
CC	Coordinating conjunction

Figure 15.2 Selected dependency relations from the Universal Dependency set. (de Marneffe et al., 2014)

Universal dependency set (UD) / Treebanks

- ✓ **Standardized set of labels**
- ✓ linguistically motivated, computationally useful, cross-linguistically applicable
- ✓ See example on previous slide
- ✓ **Treebanks** are basically **datasets** with manually or semi-automatically labelled examples
- ✓ **They can be used for training and evaluation of algorithms**

Another Example

