Dependency Grammar

Linguistics 614

With thanks to Sandra Kübler and Joakim Nivre

Spring 2015

Not a coherent grammatical framework: wide range of different kinds of DG
- just as there are wide ranges of “generative syntax”
- Different core ideas than phrase structure grammar

Dependency grammar is important for those interested in CL:
- Increasing interest in dependency-based approaches to syntactic parsing in recent years

Basic Concepts of Dependency Syntax

The basic idea:
- Syntactic structure consists of lexical items, linked by binary asymmetric relations called dependencies.
- In the (translated) words of Lucien Tesnière [Tesnière(1959)]:
  - The sentence is an organized whole, the constituent elements of which are words. [1.2] Every word that belongs to a sentence ceases by itself to be isolated as in the dictionary. Between the word and its neighbors, the mind perceives connections, the totality of which forms the structure of the sentence. [1.3] The structural connections establish dependency relations between the words. Each connection in principle unites a superior term and an inferior term. [2.1] The superior term receives the name governor. The inferior term receives the name subordinate. Thus, in the sentence Alfred parle [...], parle is the governor and Alfred the subordinate. [2.2]
Overview: constituency

(1) Small birds sing loud songs

What you might be more used to seeing:

```
S
  NP  VP
    Small birds  sing  NP
         loud songs
```

Overview: dependency

A corresponding dependency tree representation [Hudson(2000)]:

```
  OBJ  nmod
     Small birds sing loud songs
     nmod
     OBJ
     nmod
  OBJ
```

What are Dependency Relations?

- DG is based on dependency relations between words:
  - A → B means A governs B or B depends on A ...
  - Dependency relations can refer to syntactic properties, semantic properties, or a combination of the two
    - Some variants of DG separate syntactic and semantic relations by representing different layers of dependency structures
    - These relations are generally syntactic functions: subject, object/complement, adjunct, etc.
    - Subject/Agent: John fished.
    - Object/Patient: Mary hit John
- PSG is based on groupings, or constituents:
  - Grammatical relations are not usually seen as primitives, but as being derived from structure
Simple relation example

For the sentence *John loves Mary*, we have the relations:
- loves →_{subj} John
- loves →_{obj} Mary

Both *John* and *Mary* depend on *loves*, which makes *loves* the head, or root, of the sentence (i.e., there is no word that governs *loves*)

- The structure of a sentence consists of the set of pairwise relations among words.

---

Terminology

Superior Inferior

<table>
<thead>
<tr>
<th>Superior</th>
<th>Inferior</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head</td>
<td>Dependent</td>
</tr>
<tr>
<td>Governor</td>
<td>Modifier</td>
</tr>
<tr>
<td>Regent</td>
<td>Subordinate</td>
</tr>
</tbody>
</table>

Economic news had little effect on financial markets.
Economic news had little effect on financial markets.
Economic news had little effect on financial markets.

- Dependency structures explicitly represent:
  - head-dependent relations (directed arcs),
  - functional categories (arc labels),
  - possibly some structural categories (parts-of-speech).

- Phrase structures explicitly represent:
  - phrases (nonterminal nodes),
  - structural categories (nonterminal labels),
  - possibly some functional categories (grammatical functions).

- Hybrid representations may combine all elements.
**Relation to phrase structure**

What is the relation between DG and PSG?

- If a PS tree has heads marked, then you can derive the dependencies.
- Likewise, a DG tree can be converted into a PS tree by grouping a word with its dependents.
  - This only works for projective trees (no crossing branches).
  - Some constituent distinctions are not possible: e.g., binary-branching vs. flat structures for the same head.
  - No categorization into phrasal levels.

**Theoretical Frameworks**

- **Word Grammar (WG)** [Hudson(1984), Hudson(1990)]
- **Functional Generative Description (FGD)** [Šgall et al.(1986), Šgall, Hajicová and Panevová]
- **Dependency Unification Grammar (DUG)** [Hellwig(1986), Hellwig(2003)]
- **Meaning-Text Theory (MTT)** [Mel’čuk(1988)]
- **Functional Dependency Grammar (FDG)** [Tapanainen and Järvinen(1997), Järvinen and Tapanainen(1998)]
- **Topological/Extensible Dependency Grammar ([T/X]DG)** [Duchier and Debusmann(2001), Debusmann et al.(2004)]

**Theoretical Issues**

- Dependency structure sufficient as well as necessary?
- Mono-stratal or multi-stratal syntactic representations?
- What is the nature of lexical elements (nodes)?
  - Morphemes?
  - Word forms?
  - Multi-word units?
- What is the nature of dependency types (arc labels)?
  - Grammatical functions?
  - Semantic roles?
- What are the criteria for identifying heads and dependents?
- What are the formal properties of dependency structures?
Criteria for Heads and Dependents

- Criteria for a syntactic relation between a head $H$ and a dependent $D$ in a construction $C$ [Zwicky(1985), Hudson(1990)]:
  1. $H$ determines the syntactic category of $C$; $H$ can replace $C$.
  2. $H$ determines the semantic category of $C$; $D$ specifies $H$.
  3. $H$ is obligatory, $D$ may be optional.
  4. $H$ selects $D$ and determines whether $D$ is obligatory.
  5. The form of $D$ depends on $H$ (agreement or government).
  6. The linear position of $D$ is specified with reference to $H$.

- Issues:
  1. Syntactic (and morphological) versus semantic criteria
  2. Exocentric versus endocentric constructions

Some Clear Cases

<table>
<thead>
<tr>
<th>Construction</th>
<th>Head</th>
<th>Dependent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exocentric</td>
<td>Verb</td>
<td>Subject (sbj)</td>
</tr>
<tr>
<td></td>
<td>Verb</td>
<td>Object (obj)</td>
</tr>
<tr>
<td>Endocentric</td>
<td>Verb</td>
<td>Adverbial (vmod)</td>
</tr>
<tr>
<td></td>
<td>Noun</td>
<td>Attribute (nmod)</td>
</tr>
</tbody>
</table>

Economic news suddenly affected financial markets.

Some Tricky Cases

- Complex verb groups (auxiliary ↔ main verb)
- Subordinate clauses (complementizer ↔ verb)
- Coordination (coordinator ↔ conjuncts)
- Prepositional phrases (preposition ↔ nominal)
- Punctuation

I can see that they rely on this and that.
Some Tricky Cases

- Complex verb groups (auxiliary ↔ main verb)
- Subordinate clauses (complementizer ↔ verb)
- Coordination (coordinator ↔ conjuncts)
- Prepositional phrases (preposition ↔ nominal)
- Punctuation

I can see that they rely on this and that.
Some Tricky Cases

- Complex verb groups (auxiliary ↔ main verb)
- Subordinate clauses (complementizer ↔ verb)
- Coordination (coordinator ↔ conjuncts)
- Prepositional phrases (preposition ↔ nominal)
- Punctuation

I can see that they rely on this and that.
Some Tricky Cases

- Complex verb groups (auxiliary ↔ main verb)
- Subordinate clauses (complementizer ↔ verb)
- Coordination (coordinator ↔ conjuncts)
- Prepositional phrases (preposition ↔ nominal)
- Punctuation

I can see that they rely on this and that.
A dependency structure can be defined as a directed graph $G$, consisting of
- a set $V$ of nodes,
- a set $E$ of arcs (edges),
- a linear precedence order $<$ on $V$
  (not in every theory)

Labeled graphs:
- Nodes in $V$ are labeled with word forms (and annotation).
- Arcs in $E$ are labeled with dependency types.

Notational conventions $(i, j \in V)$:
- $i \rightarrow j \equiv (i, j) \in E$
- $i \rightarrow^* j \equiv i = j \lor \exists k : i \rightarrow k, k \rightarrow^* j$

**Intuitions:**
- Syntactic structure is complete (Connectedness).
- Syntactic structure is hierarchical (Cyclicity).
- Every word has at most one syntactic head (Single-Head).

Connectedness can be enforced by adding a special root node.

- $G$ is (weakly) connected:
  - For every node $i$ there is a node $j$ such that $i \rightarrow j$ or $j \rightarrow i$.
- $G$ is acyclic:
  - If $i \rightarrow j$ then not $j \rightarrow^* i$.
- $G$ obeys the single-head constraint:
  - If $i \rightarrow j$, then not $k \rightarrow j$, for any $k \neq i$.
- $G$ is projective:
  - If $i \rightarrow j$ then $i \rightarrow^* k$, for any $k$ such that $i < k < j$ or $j < k < i$. 
Projectivity

A dependency is projective provided that every word between the head and the dependent is a subordinate of A.

- subordinate (base case): dependent of A
- subordinate (recursive case): dependent of a subordinate of A

Most theoretical frameworks do not assume projectivity. Non-projective structures are needed to account for:

- long-distance dependencies,
- free word order.

Non-projectivity in principle could allow any word order. This would clearly overgenerate for most languages.

Some DGs use projectivity constraints [Hudson(1990)]:

(2) with great difficulty
(3) *great with difficulty

*great with difficulty is ruled out because branches would have to cross in that case.

In general, this is too strong of a constraint.
Obtaining word order constraints

More often, DGs specify word order separately from dependencies
- e.g., \( \text{Seq} (\text{det}(w), \text{adj}(w), \text{n}(w)) \):
  - determiner precedes adjective which precedes noun

Grammaticality

Important concept in DG is valency: ability of a word to take arguments

Possible lexicon fragment

[Duchić et al. (2003)]:

```
<table>
<thead>
<tr>
<th>Slot_1</th>
<th>Slot_2</th>
<th>Slot_3</th>
</tr>
</thead>
<tbody>
<tr>
<td>sink_1</td>
<td>ACT(nom)</td>
<td>PAT(acc)</td>
</tr>
<tr>
<td>sink_2</td>
<td>PAT(nom)</td>
<td></td>
</tr>
<tr>
<td>give</td>
<td>ACT(nom)</td>
<td>PAT(acc)</td>
</tr>
</tbody>
</table>
```

To determine grammaticality:
- Words have valency requirements that must be satisfied
- Constraints apply to the valencies to see if a sentence is valid

Lexicon

Example lexical entry from a different framework [Duchier (1999)]:

```
<table>
<thead>
<tr>
<th>string</th>
<th>loves</th>
<th>cat</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>agr</td>
<td>&lt;sing 3 nom&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>comp</td>
<td>{subject, object}</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```
Adjuncts & Complements

Two main kinds of dependencies for A (head) → B (dependent):
- Head-Complement: if A has slot for B, B is a complement
- Head-Adjunct: if B has slot for A, B is an adjunct

Adjunct/complement distinction may also be captured in the type (label) of the dependency relation

Grammaticality

Using layers of dependencies

Different frameworks allow for differing layers of dependencies
- e.g., FGD distinguishes tectogrammatical & analytical layers

Example from MTT [Mel'čuk(1988)]:

Mutual dependence: verb selects subject (& other arguments), but verb form depends on the subject:

(4) a. The child is playing.
   b. The children are playing.

One solution:
- Dependence of child/children on the verb is syntactic
- Dependence of the verb(form) on the subject is morphological

Double dependencies

Likewise, clean could depend both on the verb wash & on the noun dish:

(5) Wash the dish clean.

One solution:
- Dependence of clean on wash is syntactic (cf. case)
- Dependence of clean on dish is semantic (cf. gender)

(6) My našli
    We found the hall masc.pust-ym masc.sg inst
Advantages and Disadvantages of DG

Advantages:
- Close connection to semantic representation
- More flexible structure for, e.g., non-constituent coordination
- Easier to capture some typological regularities
- Large body of computational work on dependency parsing

Disadvantages:
- No constituents makes analyzing coordination difficult
- No distinction between modifying a constituent vs. an individual word
- Harder to capture things like, e.g., subject-object asymmetries

References


REFERENCES


REFERENCES

