Integrating grammatically relevant lexicalized meaning into morphological analyzers
Katina Bontcheva, Nicolas Kimm, Natalia Mamerow
Heinrich-Heine-University Düsseldorf

Conference „Concept Types and Frames“ (CTF 2012), Düsseldorf, August 22nd-24th 2012
Outline

1. Motivation
2. Theoretical Background
   1. Realizational morphology
   2. Vendler’s four time schemata
   3. Scalarity and manner/result complementarity
3. Examples
4. Future work
5. References
Motivation

- **Aim:** Construction of a morphological analyzer that includes semantic properties (e.g. (non-)scalarity, valency).
- **Valency information** is necessary for syntactic parsing and has been used in Constraint Grammar shallow parsers and in dependency parsers.

**Advantages of our approach:**
- The valency-pattern tags are added to classes of verbs rather than to individual lexical entries.
- It is possible to provide alternative outputs for the integrated aspectually relevant semantic information.
Outline

1. Motivation

2. Theoretical Background
   1. Realizational morphology
   2. Vendler’s four time schemata
   3. Scalarity and manner/result complementarity

3. Examples

4. Future work

5. References
Theoretical background

- Realizational morphology

- Vendler classes

- Scularity and manner/result complementarity
Vendler classes: pros and contras

- Vendler’s four time schemata:
  - States
  - Activities
  - Accomplishments
  - Achievements
- Pro: Vendler’s classification was widely accepted and is used in most current studies on aspect.
- Contras:
  - Vendler does not classify verbs but VPs.
  - Part of the features used to differentiate between the classes are not lexicalized by the verb but can be determined at the VP level.
  - This classification allows multiple class membership even for the same word sense. Thus *run* can be activity and accomplishment, e.g., *running/running a mile.*
Rappaport Hovav & Levin (RH & L) classes

- Dynamic verbs either lexicalize scales (scalar verbs) or do not (non-scalar verbs).
- Non-scalar verbs lexicalize manner.
- Scalar verbs lexicalize result.
- Scalar verbs lexicalize two major types of scales – multi-point scales and two-point scales.
- The chosen aspectually relevant properties are complementary.
- All lexical distinctions described here have grammatical consequences which are relevant to aspectual composition.
Pros of RH & L classes

- The verbs fall into disjunctive classes. There is no multiple class membership (for the same word sense).
- The aspectual properties are lexicalized exclusively by the verb and are not computed at the VP level.
- The lexicalized aspectual properties constrain the syntactical behavior of the verb.
- Manner verbs in English show a uniform argument-realization pattern: they can appear with unspecified and non-subcategorized objects.
- Result verbs are more constrained and less uniform in their argument realization patterns. Transitivity (in contrast to the manner verbs) is an issue.
Vendler classes and scalability (Van Valin (to appear), following RH 2008)

Vendler Classes

- change
- states

+ change

- scalar
- activities

+ scalar

two-point scale
multi-point scale

achievements
accomplishments
Intersecting inflectional and semantic classes

- Inflectional classes (regular, irregular and suppletive verbs) are disjunctive by default.
- RH & L aspectual classes (manner, multi-point scale and two-point scale verbs) are also disjunctive (in contrast, e.g., to Vendler classes).
- The intersections of inflectional classes with aspectual classes produce sets of bases that share the same inflectional class and semantic class.
Outline

1. Motivation
2. Theoretical Background
   1. Realizational morphology
   2. Vendler’s four time schemata
   3. Scalerity and manner/result complementarity
3. Examples
4. Future work
5. References
Example

*Killer wants to go to school* (Paul Simon, slightly adapted)

1 killer  killer  +N+Nom+Sg
2 wants  want  +N+Nom+Pl
          +V+Pres+3P+Sg
3 to    to    +Prep
          +InfMark
4 go    go    +V+Pres+Non3PSg  +V+Inf
          +N+Nom+Sg
5 to    to    +Prep
          +InfMark
6 school  school  +V+Pres+Non3PSg  +V+Inf
          +N+Nom+Sg
7 .    .    +Punct
Example

A shallow syntactic parser typically uses mainly the POS tags:

1 killer    killer    +N+Nom+Sg
2 wants     want      +N+Nom+Pl
           +V+Pres+3P+Sg
3 to       to        +Prep
           +InfMark
4 go       go        +V+Pres+Non3PSg  +V+Inf
           +N+Nom+Sg
5 to       to        +Prep
           +InfMark
6 school   school    +V+Pres+Non3PSg  +V+Inf
           +N+Nom+Sg
7 .        .          +Punct
Example

Accordingly, the tokens can receive the following phrase-syntax tags:

1. killer   NPHead   PCompl   Premod
2. wants   NPHead   PCompl   MainV+F
3. to      PPHead   Premarker
4. go      NPHead   PCompl   MainV+F   MainV-F
5. to      PPHead   Premarker
6. school  NPHead   PCompl   MainV+F   MainV-F
Example

If no semantic constrains apply, the following analyses are possible:

<table>
<thead>
<tr>
<th>word 1</th>
<th>word 2</th>
<th>word 3</th>
<th>word 4</th>
<th>word 5</th>
<th>word 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>NPHead</td>
<td>MainV+F</td>
<td>PHead</td>
<td>PCompl</td>
<td>PPHead</td>
<td>PCompl</td>
</tr>
<tr>
<td>Mary</td>
<td>moved</td>
<td>with</td>
<td>John</td>
<td>to</td>
<td>Canada</td>
</tr>
<tr>
<td>NPHead</td>
<td>MainV+F</td>
<td>to</td>
<td>PCompl</td>
<td>Switzerland</td>
<td>to</td>
</tr>
<tr>
<td>Pam</td>
<td>went</td>
<td>to</td>
<td>MainV-F</td>
<td>recover</td>
<td></td>
</tr>
<tr>
<td>NPHead</td>
<td>MainV+F</td>
<td>to</td>
<td>MainV-F</td>
<td>college</td>
<td></td>
</tr>
<tr>
<td>Killer</td>
<td>wants</td>
<td>to</td>
<td>MainV-F</td>
<td>sleep</td>
<td></td>
</tr>
<tr>
<td>NPHead</td>
<td>promised</td>
<td>to</td>
<td>PCompl</td>
<td></td>
<td></td>
</tr>
<tr>
<td>John</td>
<td>whales</td>
<td>to</td>
<td>PCompl</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Premod</td>
<td>NPHead</td>
<td>to</td>
<td>PCompl</td>
<td>Antarctica</td>
<td></td>
</tr>
<tr>
<td>Killer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Example

1 killer killer +N+Nom+Sg
2 wants want +N+Nom+Pl
   +V+Pres+3P+Sg
3 to to +Prep
   +InfMark
4 go go +V+Pres+Non3PSg +V+Inf <multi-point>
   +N+Nom+Sg
5 to to +Prep
   +InfMark
6 school school +V+Pres+Non3PSg +V+Inf
   +N+Nom+Sg
7 . . +Punct

24.08.2012 Integrating Aspectually Relevant Properties of Verbs into a Morphological Analyzer for English
Example

Multi-point scale verbs are used with PPs that mark scale:

<table>
<thead>
<tr>
<th>verb</th>
<th>head</th>
<th>complement</th>
<th>modifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>killer</td>
<td>NPHead</td>
<td>PCompl</td>
<td>Premod</td>
</tr>
<tr>
<td>killer</td>
<td>NPHead</td>
<td></td>
<td></td>
</tr>
<tr>
<td>wants</td>
<td>NPHead</td>
<td>PCompl</td>
<td>MainV+F</td>
</tr>
<tr>
<td>wants</td>
<td>MainV+F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>to</td>
<td>PPHead</td>
<td>Premarker</td>
<td></td>
</tr>
<tr>
<td>to</td>
<td>Premark</td>
<td></td>
<td></td>
</tr>
<tr>
<td>go</td>
<td>NPHead</td>
<td>PCompl</td>
<td>MainV+F</td>
</tr>
<tr>
<td>go</td>
<td>MainV-F</td>
<td></td>
<td>MainV-F</td>
</tr>
<tr>
<td>to</td>
<td>PPHead</td>
<td>Premarker</td>
<td></td>
</tr>
<tr>
<td>to</td>
<td>PPHead</td>
<td></td>
<td></td>
</tr>
<tr>
<td>school</td>
<td>NPHead</td>
<td>PCompl</td>
<td>MainV+F</td>
</tr>
<tr>
<td>school</td>
<td>PCompl</td>
<td></td>
<td>MainV-F</td>
</tr>
</tbody>
</table>
Example

Multi-point scale verbs are used with PPs that mark scale:

killer  NPHead  PCompl  Premod
killer  NPHead
wants  NPHead  PCompl  MainV+F
wants  MainV+F
to     PPHead  Premarker
to     Premark
go     NPHead  PCompl  MainV+F  MainV-F
go     MainV-F
to     PPHead  Premarker
to     PPHead
school NPHead  PCompl  MainV+F  MainV-F
school PCompl
Advantages of the Approach

- Alternative outputs for the integrated aspectually relevant semantic information can be provided.
- Redundancy is avoided by identifying abstract lexicalized properties that are shared across large groups of verbs.
- The computational efficiency of the morphological analyzer is not jeopardized.
- The integration of semantic properties improves the performance of syntax parsers that use the output of the morphological analyzer.
Outline

1. Motivation

2. Theoretical Background
   1. Realizational morphology
   2. Vendler’s four time schemata
   3. Scalarity and manner/result complementarity

3. Examples

4. Future Work

5. References
Future work

- Implementation of state verbs
- Work on noun semantic classes
- Cross-linguistic examination of the approach
Thank you!
References


