LSA.343
Precision Grammar Implementation for Linguistic Hypothesis Testing

Emily M. Bender
University of Washington
ebender@u.washington.edu

Dan Flickinger
CSLI Stanford
danf@csli.stanford.edu

Stephan Oepen
Universitet i Oslo & CSLI Stanford
oe@csli.stanford.edu
So, What is Computational Linguistics?

... the scientific study of human language—specifically of the system of rules and the ways in which they are used in communication—using mathematical models and formal procedures that can be realized and validated using computers; a cross-over of many disciplines. (Stanford Professor)
So, What is Computational Linguistics?

... the scientific study of human language—specifically of the system of rules and the ways in which they are used in communication—using mathematical models and formal procedures that can be realized and validated using computers; a cross-over of many disciplines. (Stanford Professor)

... a cornerstone of our pioneering .NET initiative and the operating systems of the future; innovative technology that will change our world. (President of US-Based Software Company)
So, What is Computational Linguistics?

... the scientific study of human language—specifically of the system of rules and the ways in which they are used in communication—using mathematical models and formal procedures that can be realized and validated using computers; a cross-over of many disciplines. (Stanford Professor)

... a cornerstone of our pioneering .NET initiative and the operating systems of the future; innovative technology that will change our world. (President of US-Based Software Company)

... a sub-discipline of our Artificial Intelligence programs. (CMU Professor)
What About (Computational) Grammar?

Wellformedness

- *Kim was happy because ____ passed the exam.*
- *Kim was happy because ____ final grade was an A.*
- *Kim was happy seeing ____ appear on television.*
What About (Computational) Grammar?

Wellformedness

- Kim was happy because ____ passed the exam.
- Kim was happy because ____ final grade was an A.
- Kim was happy seeing ____ appear on television.

Meaning

- Kim gave Sandy a book.
- Kim gave a book to Sandy.
- Sandy was given a book by Kim.
What About (Computational) Grammar?

**Wellformedness**
- Kim was happy because ____ passed the exam.
- Kim was happy because ____ final grade was an A.
- Kim was happy when she saw ____ on television.

**Meaning**
- Kim gave Sandy a book.
- Kim gave a book to Sandy.
- Sandy was given a book by Kim.

**Ambiguity**
- I saw the astronomer with the telescope.
- Have her report on my desk immediately!
What We Are About to Do

Course Outline

- Construct testable linguistic hypotheses
- Employ software validation of these hypotheses using both constructed and naturally occurring data
- Explore implications for linguistic theory:
  - What requirements are imposed?
  - How much testing is needed for validation?
- Understand issues in test suite design
- Learn to use open-source software and knowledge base resources
Why Computational Grammars?

- **research** formalize linguistic theories with complex interactions of language phenomena; identify cross-language generalizations;
- **education** teach frameworks or analyses in formal morphology, syntax, and semantics; support student experimentation;
- **applications** embed grammar-based natural language analysis or generation in research prototypes and commercial applications.
Student Experimentation — Immediate Gratification

Precision Grammar Implementation (6)
Some Applications of Computational Grammars

**Machine Translation**
- Traditional: analyse source to some degree, transfer, generate target

**Text ‘Understanding’**
- Email auto-response: interpret customer requests
- Semantic Web: annotate WWW with structured, conceptual data

**(Spoken) Dialogue Systems**

**Grammar & Controlled Language Checking**

**Summarization & Text Simplification**
### Some Areas of Descriptive Grammar

<table>
<thead>
<tr>
<th>Area</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phonetics</td>
<td><em>The study of speech sounds</em></td>
</tr>
<tr>
<td>Phonology</td>
<td><em>The study of sound systems</em></td>
</tr>
<tr>
<td>Morphology</td>
<td><em>The study of word structure</em></td>
</tr>
<tr>
<td>Syntax</td>
<td><em>The study of sentence structure</em></td>
</tr>
<tr>
<td>Semantics</td>
<td><em>The study of language meaning</em></td>
</tr>
<tr>
<td>Pragmatics</td>
<td><em>The study of language use</em></td>
</tr>
</tbody>
</table>
Some Areas of Descriptive Grammar

- **Phonetics**: The study of speech sounds
- **Phonology**: The study of sound systems
- **Morphology**: The study of word structure
- **Syntax**: The study of sentence structure
- **Semantics**: The study of language meaning
- **Pragmatics**: The study of language use
Grammar Engineering from a CS Perspective

Implementation Goals

- Translate linguistic constraints into specific formalism → formal model
- Provide mapping between form and meaning: computational grammar
- Assign correct analyses to grammatical inputs; reject ungrammatical
- Apply mapping in either direction: parsing and generation algorithms

Analogy to (Object-Oriented) Programming

- Computational system with observable behavior: immediately testable
- typed feature structures as a specialized (OO) programming language
- make sure that all the pieces fit together; revise – test – revise – test ...
Course Organization
DELPH-IN Software Resources

- LKB: Linguistic Knowledge Builder
- [incr tsdb()]: competence and performance profiling platform
- Grammar Matrix: starter kit for new grammars
- Redwoods: dynamic treebanking environment
- ERG: English Resource Grammar

www.delph-in.net
Comments on Background Literature

Formal Syntax


The Linguistic Knowledge Builder