Plural Semantics for Natural Language Understanding
A Computational Proof-Theoretic Approach

Uta Schwertel

Presentation Ph.D. Thesis
Philosophical Faculty, University of Zurich
October 30, 2003
Overview

• Objectives
• Semantic ambiguity
• Views of semantics
• Computational proof-theoretic approach
  – Semantic representation
  – Disambiguation
  – Reasoning
• Applications (DRoPs, Attempto)
• Conclusion
Objectives

- Natural language plurals occur frequently
- Semantics irreducible to singular semantics
  
  *Four men lift a table.*

- Applications require plural component
- No suitable worked-out computational approach
- Practical plural semantics for logic-oriented NLU
  
  ⇔ Computational proof-theoretic approach
    
    Representation – disambiguation – reasoning
Semantic Ambiguity

- Plurals cause semantic ambiguities
  
  *Four men lifted three tables.*
  
  - Collective, distributive, cumulative, neutral readings
  - Scope ambiguities

- Combinatorial explosion of ambiguities

- Ambiguity processing necessary
  
  - Real ambiguity vs. indeterminacy
  - Resolution of real ambiguity
Number of Plural Readings

• Ambiguity bw. collective and distributive reading
  *Two men lifted a table.*

• Indeterminacy of collective reading
  *The students read the newspapers.*
  – Strictly collective, cumulative or other instantiations

• Difficulty
  – Cumulativity in knowledge-bases $\Rightarrow$ inference rules
  – Prevent upward monotonicity $\Rightarrow$ explicit conditions
Views and Purpose of Semantics

Truth-Conditional
• Sentence ⇒ Formula $S_F$
• Real world ⇒ Model $M$
• Truth: $[[ S_F ]]^M = 1$

Proof-Theoretic
• Sentence ⇒ Formula $T_F$
• NL Text ⇒ Formulas $KB$
• Truth: $KB \land Ax \vdash T_F$

Practical Features
• World: static
• Inference: semantic
• Declarative sentences

Practical Features
• World: dynamic
• Inference: syntactic
• Different sentence types
⇒ Suitable for NLU
Overview of Thesis Approach

• Proof-theoretic semantics
• Reflects concepts of existing formal approaches
• Central concepts
  – Discourse tradition: referring vs. quantificational NPs
  – Readings: distributive and collective
  – Ambiguity: global
  – Ontology: lattice-theoretic
  – Representation: flat first-order DRSs
• Disambiguation: tractable algorithm
Representation

Two men lift a table.
Representation

Two men lift a table. (Collective reading)

[A,B,C,D,E]
structure(B,group)-1
quantity(B,cardinality,A,count_unit)-1
value(A,eq,2)-1

[F]
structure(F,atomic)-1
part_of(F,B)-1
=>
[]
object(F,man)-1

structure(D,atomic)-1
quantity(D,cardinality,C,count_unit)-1
value(C,eq,1)-1
object(D,table)-1
structure(E,event)-1
predicate(E,lift,B,D)-1
Two men lift a table.

```
[A, B, C, D, E]
structure(B, group)-1
quantity(B, cardinality, A, count_unit)-1
value(A, eq, 2)-1
[F]
  structure(F, atomic)-1
  part_of(F, B)-1
=>
[]
  object(F, man)-1
structure(D, atomic)-1
quantity(D, cardinality, C, count_unit)-1
value(C, eq, 1)-1
object(D, table)-1
structure(E, event)-1
predicate(E, lift, B, D)-1
```

Only predefined relation symbols
Representation

Two men lift a table.

Only predefined relation symbols

“Predicates” as arguments
Representation

Two men lift a table.

Only predefined relation symbols

“Predicates” as arguments

Typing of discourse variables
Representation

Two men lift a table.

\[
\begin{align*}
\{A, B, C, D, E\} & \quad \text{structure}(B, \text{group}) - 1 \\
\text{quantity}(B, \text{cardinality}, A, \text{count_unit}) - 1 & \quad \text{value}(A, \text{eq}, 2) - 1 \\
\{F\} & \quad \text{structure}(F, \text{atomic}) - 1 \\
\text{part_of}(F, B) - 1 & \quad \Rightarrow \\
\{\} & \quad \text{object}(F, \text{man}) - 1 \\
\text{structure}(D, \text{atomic}) - 1 & \quad \text{quantity}(D, \text{cardinality}, C, \text{count_unit}) - 1 \\
\text{value}(C, \text{eq}, 1) - 1 & \quad \text{object}(D, \text{table}) - 1 \\
\text{structure}(E, \text{event}) - 1 & \quad \text{predicate}(E, \text{lift}, B, D) - 1
\end{align*}
\]

Only predefined relation symbols

“Predicates” as arguments

Typing of discourse variables

Quantity information
Representation

Two men lift a table.

Only predefined relation symbols

“Predicates” as arguments

Typing of discourse variables

Quantity information

Avoid special operators
Representation

Two men lift a table.

Only predefined relation symbols

“Predicates” as arguments

Typing of discourse variables

Quantity information

Avoid special operators

Index for tracking within reasoning
Distributive Reading

*Two men each lift a table.*

**Two men**

[A,B]
structure(B,group)-1
quantity(B,cardinality,A,count_unit)-1
value(A,eq,2)-1

[C]
structure(C,atomic)-1
part_of(C,B)-1
=>
[]
object(C,man)-1

**each**

[D]
structure(D,atomic)-1
part_of(D,B)-1
=>

**lift a table**

[E,F,G]
structure(F,atomic)-1
quantity(F,cardinality,E,count_unit)-1
value(E,eq,1)-1
object(F,table)-1
structure(G,event)-1
predicate(G,lift,D,F)-1
Evaluation of Representation

www.ifi.unizh.ch/cgi-bin/schwertel/demo/plural/plural_new.html

• First-order solution for difficult phenomena
  – Non-upward monotonicity (*at most n, exactly n, all*)
  – Vague and context-dependent determiners (*few, many*)
  – Proportional quantifiers (*most*), partitives (*two of the*)
  – Part-structure modifiers (*as a whole, simultaneously*)
  – Measurements (*two ounces of gold*), coordination

• Advantages, problems and further research
  + General, rich, flat, first-order ⇒ automated deduction
  – Extend coverage, efficiency, maximality conditions
Disambiguation of Plurals

• Humans often use context or world-knowledge

  \textit{Five men lift a piano.} \quad \textit{Five men lift a chair.}

• Context computationally not manageable

• Compromise: Parameter based disambiguation
  – Accessible lexical and structural disambiguation factors
  – Disambiguation constraints and preferences
  – Hierarchy of plausible readings
  – Automatic selection of best reading
Disambiguation Information

Two men tell a story to several children. ⇒ 19 different possible readings

• Explicit triggers (floated quantifiers)

• Lexical information
  – Determiner (distributivity type: ‘c>d’)
  – Verb (arity: dv, collectivity type: m)

• Structural information
  – Grammatical function (subj, do, io, po)
  – Syntactic structure (n, np)
  – Linear order

• Assign numeric values to parameters
Disambiguation Information

Two men *tell a story to several children.*

- Explicit triggers (floated quantifiers)

- Lexical information
  - Determiner (distributivity type: ‘c>d’)
  - Verb (arity: dv, collectivity type: m)

- Structural information
  - Grammatical function (subj, do, io, po)
  - Syntactic structure (n, np)
  - Linear order

- Assign numeric values to parameters

- Add info to underspecified store
  
  ```
  [[Matrix_DRS], [np(F_1), np(F_2), np(F_3)]]
  ```

  ```
  [[drs([[A],[structure(A,event),
predicate(A,tell_to,B,C,D)]),
[np(F_1 & syn:index:B &
drs:out:drs([[B],A6],[structure(B,group),
  quantity(B,cardinality,A6,count_unit),
  value(A6,eq,2),drs([[B6],
  [structure(B6,atomic),part_of(B6,B)])
  =>drs([],object(B6,man))]) &
  sem:quant:quant_type:card &
  quantity:value:2 & num_rel:eq &
disambig:default:dist_type:'c>d' &
  reading:result:coll &
  stability:pref &
  gra_fct:subj &
  subcat:dv &
  syn_struct:noun &
  scop_lex:2
  scope_value:local:10&
  monotone:up &
  string:[two,men] &
  v_info:default:dist_type:m &
  log_rel:man ), np(F_2), np(F_3)]]]
  ```
Scoping Preferences

Two men *tell a story to several children.*

6 possible scopings

DRoPs Scoping Algorithm

Weighted list of two plausible scopings

[[18]-[np(F₁), np(F₂), np(F₃)],
[16]-[np(F₁), np(F₃), np(F₂)]]
Plural Preferences

*Two men tell a story to several children.*

- Scoping algorithm ⇒ 2 plausible scopings
  
  - Collective/distributive readings
    - Constraints (C1-9)
    - Preferences (P1-12)
  
- Possible NP (re-)interpretations for each scoping (2×4 = 8 possibilities)

- Reinterpretation
  - Costs (P24-25)
  - Threshold (P26)

- 4 plausible readings

\[\begin{align*}
[[18]-[np(F_1), np(F_2), np(F_3)],
[16]-[np(F_1), np(F_3), np(F_2)]
\end{align*}\]

\[\begin{align*}
np(F_1 &>disambig:default:dist_type:‘c>d’ & 
reading:result:coll & 
stability:pref)
\end{align*}\]

\[\begin{align*}
np(F_1 &>reading:result:distr & 
stability:unpref)
\end{align*}\]

\[\begin{align*}
[[18,0]-[np(F_1>&coll), np(F_2>&coll), np(F_3>&coll)],
[18,-2]-[np(F_1>&distr), np(F_2>&coll), np(F_3>&coll)],
[16,-3]-[np(F_1>&coll), np(F_3>&distr), np(F_2>&coll)],
[16,-5]-[np(F_1>&distr), np(F_3>&distr), np(F_2>&coll)]
\end{align*}\]
Two men tell a story to several children.

1. Reading ([18,0]): 1. [ two men ]-coll, 2. [ a story ]-coll, 3. [ several children ]-coll
2. Reading ([18,-2]): 1. [ two men ]-distr, 2. [ a story ]-coll, 3. [ several children ]-coll
3. Reading ([16,-3]): 1. [ two men ]-coll, 2. [ several children ]-distr, 3. [ a story ]-coll
4. Reading ([16,-5]): 1. [ two men ]-distr, 2. [ several children ]-distr, 3. [ a story ]-coll

Fully specified store of selected reading

DRS construction \(\Rightarrow\) Unambiguous DRS
Evaluation of Disambiguation

• Advantages
  – Structurally available disambiguation information
  – Example reduction: 19 possible ⇒ 4 plausible ⇒ 1 best
  – Underspecified Logical Forms ⇒ Disambiguated LFs
  – DRS construction after disambiguation (one grammar)

• Further research
  – Parameters: more experiments, fine-tuning, theory
  – Extensions: plural anaphora, vertical scoping, coverage
  – Underspecification vs. full disambiguation
Automated Reasoning for NLU

• Logic-oriented text understanding systems
  – Detect entailments, contradictions, answer questions
  – Realized as logical reasoning

• Automated first-order logical reasoning
  – Tools: existing theorem provers Otter and Satchmo
  – Requirement: Input/output on level of natural language
  – Extensions: tracking in natural language, all solutions

• Reasoning with plurals requires auxiliary axioms
  – lattices, equality, mathematics, evaluables, lexicon
Every company that buys a machine gets a discount.

Six Swiss companies each buy a machine.

A company gets a discount.

∃A(structure(A,group) ∧ …)

∀X(structure(X,group) → ∃Y(structure(Y,atomic) ∧ part_of(Y,X))

∀B(structure(B,atomic) ∧ …)
Entailment Example

Text

Every company that buys a machine gets a discount.
Six Swiss companies each buy a machine.

Query

A company gets a discount.

The reasoner proved that the sentence(s)
A company gets a discount.
can be deduced from the sentence(s)
Every company that buys a machine gets a discount.
Six Swiss companies each buy a machine.
using the auxiliary axiom(s)
(Ax. 9): Definition of proper_part_of.
(Ax. 10-1): Every group consists of atomic parts.
(Ax. 22-1): Number Axiom.

www.ifi.unizh.ch/cgi-bin/schwertel/demo/race/race.html
## Question Answering Example

<table>
<thead>
<tr>
<th>Text</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Six Swiss companies each buy a machine.</em></td>
<td><em>Who buys machines?</em></td>
</tr>
<tr>
<td><em>A German company buys a special machine.</em></td>
<td></td>
</tr>
</tbody>
</table>

1. The reasoner proved that the query (-ies)
   
   *Who buys machines?*
   
   can be answered on the basis of the sentence(s)
   
   *Six Swiss companies each buy a machine.*
   
   using the FOL axiom(s)

   (Ax. 10-2): Groups have atomic parts.
   (Ax. 2): atomic => dom
   (Ax. 9): Definition of proper_part_of.
   (Ax. 11): Atoms have no proper parts.
   (Ax. 15-1): Identity axiom for objects.
   (Ax. 22-1): Number Axiom.

2. The reasoner proved that the query (-ies)
   
   *Who buys machines?*
   
   can be answered on the basis of the sentence(s)
   
   *A German company buys a special machine.*
   
   using the FOL axiom(s)

   (Ax. 2): atomic => dom
   (Ax. 11): Atoms have no proper parts.
   (Ax. 15-1): Identity axiom for objects.
   (Ax. 22-1): Number Axiom.
Consistency Example

The reasoner proved that the sentence(s)

Many mountaineers summitted.
Few mountaineers summitted.

and the FOL axiom(s)

(Ax. 24): not(many & few).

are inconsistent.

The reasoner proved that the sentence(s)

Many mountaineers summitted.
Few mountaineers summitted quickly.

are consistent.
Evaluation of Reasoning

• Advantages
  – *Automated* reasoning with natural language plurals possible by first-order proof-theoretic approach
  – Implementation of domain-independent plural axioms
  – Logical analysis of texts, natural language interfaces

• Further research
  – Efficiency, robustness, scalability – Less deep analysis?
  – Reasoning with disambiguated vs. underspecified LFs
  – Reasoning to disambiguate, other reasoning methods
Application: Plurals in Attempto

• Attempto Controlled English (ACE)
  – Controlled subset of English ⇒ clear FOL translation
  – Combines natural language with formal methods
  – Software specifications, interface language, teaching

• Constructive ambiguity processing

• Plurals in ACE
  – Practical application of proof-theoretic plural approach
  – Use DRoPs techniques (readings, DRS, axioms)
  – Constructive disambiguation ⇒ parameter resettings
Plural Disambiguation in ACE

• Constructive strategy $\Rightarrow$ simplification
  – Individual denoting plural NPs get collective reading
    \textit{At least two companies order a machine.}
  – Distributive readings expressed buy \textit{each of}
    \textit{Each of at least two companies orders a machine.}

• Trade-off
  – Goal: reproducible and intelligible disambiguation
  – Rules: simple, systematic, general, easy to learn
    $\Rightarrow$ Sometimes dispreferred interpretation, stilted English
Conclusion: Achievements

• Comprehensive, computational, proof-theoretic plural semantics for logic-oriented NLU
• Computationally suitable first-order representation
• Tractable disambiguation algorithm
• Automated reasoning with plurals using axioms
• Implementation and applications
  – DRoPs (Disambiguating and Reasoning with Plurals)
  – Plurals in Attempto Controlled English
Conclusion: Further Research

• Representation – Coverage and efficiency
  – More phenomena, event semantics, compact notation
  – Relate to other first-order KR languages

• Disambiguation – Partial analysis
  – Empirical evaluation methods, test suites
  – Role of underspecified representations

• Reasoning – Broader view of inference
  – Lexical inferences, ontologies, temporal reasoning
  – Reasoning with underspecified LFs, other techniques
Appendix
Combinatorial Explosion

Four men lift three tables.

 Scoped Readings

<table>
<thead>
<tr>
<th>Scope</th>
<th>R1</th>
<th>R2</th>
<th>R3</th>
<th>R4</th>
</tr>
</thead>
<tbody>
<tr>
<td>wide</td>
<td>men</td>
<td>narrow</td>
<td>tables</td>
<td>=</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cumulative Reading</th>
</tr>
</thead>
<tbody>
<tr>
<td>R9</td>
</tr>
</tbody>
</table>

No scope interaction

= man = table

= man = table
Ambiguity vs. Indeterminacy

• Ambiguity
  – Truth-value in a situation dependent on reading
  – Different logically independent sem. representations
  – Different logical entailments

• Indeterminacy
  – Truth-value in a situation fixed
  – One semantic representation
  – Different elaborations (instantiations) possible
  – Elaborations entail original representation
Existing Formal Approaches

- **Important distinctions**
  - Tradition, ambiguity, readings, ontology, representation
- **Generalized Quantifier Theory (Scha 1981)**
  - Expressive power, flexibility
  - “Dowty-sentences”, anaphora, higher-order, ambiguity
- **Discourse Approach (Link 1983)**
  - Referring vs. quantificational noun phrases
  - Anaphora, lattices, mass domain, first-order
  - Non-monotone quantifiers, maximality, ambiguity
DRoPs Scoping Algorithm

Two men tell a story to several children.

6 possible scopings
DRoPs Scoping Algorithm

Two men tell a story to several children.

• Local numeric scope value of each NP
(P21-23) local = gra_fct + dist_type + syn_struct

• Relative scope value for each NP pair
(P19-20) global = linear order + local

• (P18) Check ambiguity
for each NP pair

• Weighted list of two plausible scopings