Overview

- Multilingual grammar engineering
- Assumptions
- Data: Armenian, Farsi, Swedish
- MRS primer and the problem
- Proposed solution
- Conclusions

Multilingual Grammar Engineering

- Monolingual grammar engineering tests linguistic analyses
- for consistency
- against corpora
- Multilingual grammar engineering tests cross-linguistic hypotheses
- LinGO Grammar Matrix: precision grammar starter-kit

Cross-linguistic Hypotheses

- HPSG notions of headedness and valence
- Semantic compositionality (so-called “Frege’s Principle”) (Pelletier 2001)
- Minimal Recursion Semantics for typed feature structure based semantic composition (Copestake et al 2003, Flickinger and Bender 2003)

Assumption 1

- A monotonic system for compositional semantics is desirable.
  - bidirectionality
  - partial interpretation from partial parses

Assumption 2

- In an adequate semantic representation, all nominal indices are bound by quantifiers.
Assumption 3

- Scopal adjectives exist.
  - Type 1: fake, alleged, former
  - Type 2: probable, likely
  - By contrast, most adjectives are intersective: good, red, tired

Assumption 4

- Quantifiers can take scope between scopal adjectives and the nouns they modify.
  - The most probable winner of every medal was disqualified.

Assumption 4 (cont)

- every > the > probable: Different probable winners for each medal, all disqualified.
- the > every > probable: One super athlete is favored in each competition individually and disqualified.
- the > probable > every: No one person dominated all events, but if anyone were to win all the medals, it would X, who was disqualified.

Assumption 5

- Determiners combine with nominal expressions to produce quantifier-expressing constituents.

Armenian

- Indefinite article is a separate word, while the definite article appears to be an affix:
  a. դղայ մեն ե բոյ  INDEF be.3sg
     ‘He is a boy.’ (Bardakjian and Thomson 1977:18)
  b. դուն-ե մեճ ե հուս-DEF big  be.3sg
     ‘The house is big.’ (Andonian 1966:22)
Armenian

- There are also what appear to be scopal adjectives:

  - hawanagan k'ayleri masin
  - probable step.pl.gen about
  - ‘... concerning probable steps'

Farsi

- ro/-o appears on definite direct objects
- or, in combination with -i, on specific indefinite direct objects

  - mi xahæ nd xane-i-ro bexæ ræ nd
  - preswant.3pl house-indef-acc/def buy
  - ‘They want to buy a house. (A certain house)’ (Mace 2003)

French

- -i marks NPs as indefinite and non-generic
- -i marks NPs as indefinite and non-generic

  - ketab: the book
  - ketab-i: a book, books in general

Swedish

- There are also what appear to be scopal adjectives:

  - mest troligaorsakentillhjärtafarkt
  - the most probable-cause-def of heart attack

  - IGLO: http://www.hum.uib.no/cvenonius/lingua/
The Problem

Syntax suggests Compositional semantics requires

(a) NP   (b) NP
  A     A
     |     N
   NP    N
   probable probable
     winner-def winner-def

The Problem

MRS Primer

- MRS representations are underspecified for scope
- Fixed scopal relations are represented as constraints on possible complete scopings
- Quantifiers are free to ‘float’ into any space

Constraints on Scope

top
every(y)
  medal(y)  def_q(z)
    probable  disqualify(z,x)
      winner(x,y)

Constraints on Scope

def_q(z)
top  every(y)
    medal(y)  probable
      disqualify(z,x)
        winner(x,y)

Constraints on Scope

def_q(z)
probable  disqualify(z,x)
    medal(y)
      winner(x,y)

Constraints on Scope

def_q(z)
top
    every(y)
      medal(y)  winner(x,y)
The Problem Again

\[
\text{def}_q(x) = \ldots \text{probable}\ldots \text{winner}(x, y)
\]

Solutions

- Loosen the tight coupling between syntax and semantics
- Reject the assumptions that require treating the (in)definiteness markers as contributing semantic quantifiers

More Flexible Interfaces

- CLLS (Constraint Language for Lambda Semantics) uses more flexible dominance relations plus semantic types to reduce ambiguity (Egg et al. 2001)
- GLUE Semantics formalizes 'weak compositionality', semantic representations projected off complete syntactic representations (Asudeh et al. 2002)

(In)definiteness Markers as Non-Quantifiers

- Definiteness affixes contribute other information about NP semantics (givenness, specificity) (cf. Borthe and Hauge to appear)
- Actual quantifier built by non-branching construction
(In)definiteness Markers as Non-Quantifiers

\[
\begin{align*}
\text{NP} & \quad \quad = \\
N' & \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \queue
Conclusions

- *MRS makes interesting cross-linguisitic predictions*
- *Grammar engineering supports linguistic hypothesis testing*
- *Computational linguistic resources should be designed to be cross-linguistically applicable*

Acknowledgments

- *Thanks to Markus Egg, Dan Flickinger, Ann Copestake for helpful discussion*
- *Thanks to Arman Maghbouleh, Tom Johnson, and Oskar Gjertsson for assistance with data*