MRSs in translation: Interlingua vs. semantic transfer

Fefor discussion June 16th 2006

Lars Ahrenberg
NLPLab
Department of Computer and Information Science, Linköpings universitet

1

Bond et al. (2005)

- Future work
  - How much of the semantic representation can be shared between languages (and thus require little or no transfer)?

2

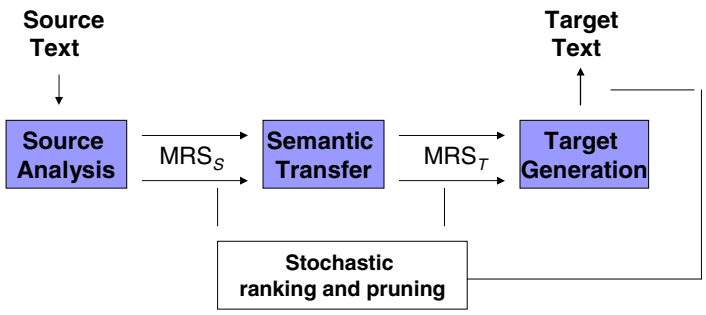
Issues for discussion

- Semantic types and features
  - Language-specific or shared
- Granularity of semantic representations
  - unanalysed or decomposed
- Structure of the semantic space
  - none, simple hierarchy, network structure, …
- Architecture
  - Pipelined or embedded components

3

MRS in translation

- The *DELPH-IN* MT architecture
  - LOGON and Bond et al. (2005)



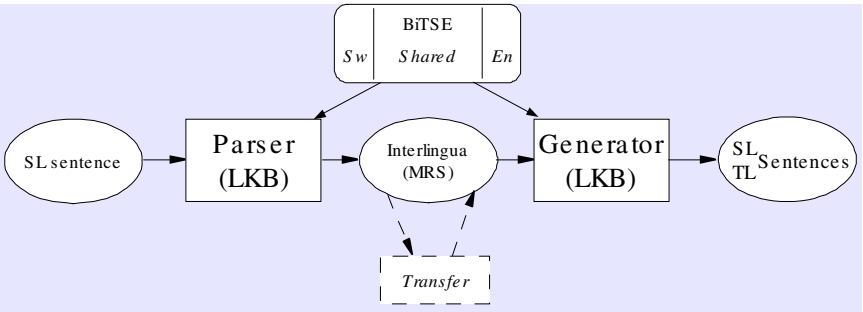
4

MRS in translation

- Characteristics of the *DELPH-IN* approach
  - Grammars, parsers and generators of different frameworks can be combined freely as long as they use MRS as semantic representations (as in LOGON)
  - MRSs are language-specific
  - ”… source and target grammars do most of the work”
  - Transfer must use all EPs of a source MRS and often results in more than one target MRS
  - Stochastic approach to selection/disambiguation

5

BiTSE architecture



Design of the MT-system

6

MRS in translation

- Characteristics of the BiTSE approach
  - One grammar using language features to separate languages
  - Either language can be taken as source
  - MRSs are shared across languages
  - The grammar does all of the work
  - Translation amounts to selection from (sets of) sets of sentences sharing a MRS (paraphrase sets)
  - Stochastic ranking may be used to support selection (but is currently not)
  - Transfer may be used when a paraphrase set is mono-lingual

7

Paraphrase sets

{ the dog barks, hunden skäller }

<h1,e2,
 {h3:**def_q**(x4,h5,h6),
  h7:**dog**(x4),
  h9:**bark**(e2,x4),
  h1:**prop_m**(h10),
 {h5 qeq h7, h10 qeq h9}}>

8

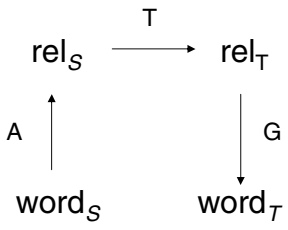
Paraphrase sets for ambiguous sentences

{ We saw her smile, Vi såg henne le, … }
{ We saw her smile, Vi såg hennes leende, … }

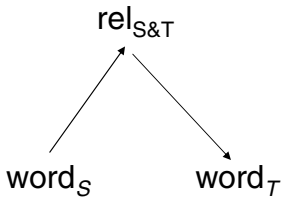
9

Linking words to relations

DELPH-IN



BiTSE



10

Pros and cons

- Flexibility, application development
  - Transfer rules model translational relations much more flexibly than a common MRS, but …
- Theoretical insights
  - Transfer rules, whether called semantic or not, do not (usually) model semantic relations.

11

Transfer and semantics

- Equivalence

  dog(x_i) → **hund**(x_i)
- Narrowing

  wall(x_i) → **vägg**(x_i)

  wall(x_i) → **mur**(x_i)
- Broadening

  ceiling(x_i) → **tak**(x_i)

12

Transfer and semantics

- Words vs. constructions

  give$(e_i,x_j,x_k,x_l) \wedge$ **answer**$(x_l) \rightarrow$ **svara**(e_i,x_j,x_k)

  embarass$(e_i,x_j,x_k) \rightarrow$ **göra**$(e_i,x_j,e_m) \wedge$ **generad**(e_m,x_k)

  Note: the problem is the same for text understanding, question-answering with a single language

13

Pros and cons

- The relation of translation to semantics
  - there is definitely more to translation than semantic equivalence, but a translation must not misrepresent the content of the source. It should give “more or less” the same information.
- Grammar harmonisation
  - Multilingual grammars encourage you to consider several languages in the same framework, and evaluate proposed solutions for more than one language at a time.

14

Pros and cons

- Translatability and the nature of languages
  - all meanings are not shared, though in the global village a large part of the world *is* shared.
    - a dog is a dog whether pet or food?'
  - TL may require grammatical distinctions that are not made in the SL
    - Strict common IL is then out of the question
    - Both approaches require extra machinery

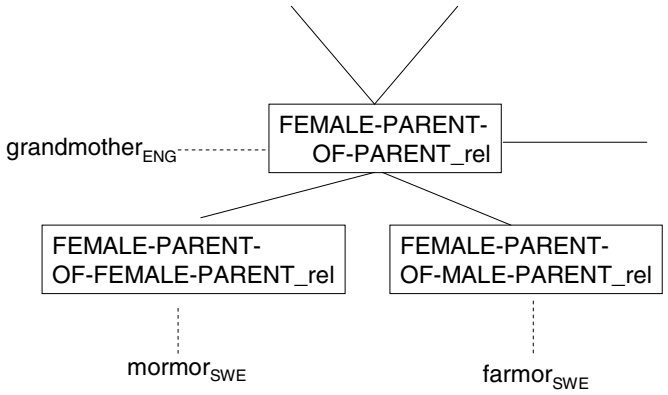
15

Structure of a semantic space

- Semantic relations
  - homonymy, partonomy, antonomy, …
  - sense distinctions
  - decomposition
- Semantic proximity and semantic neigborhoods

16

Matching on proximity



17

Issues for discussion

- Semantic types and features
  - Language-specific or shared
- Granularity of semantic representations
  - unanalysed or decomposed
- Structure of the semantic space
  - none, simple hierarchy, network structure, …
  - measures of proximity
- Architecture
  - Pipelined or embedded components

18