Open Session on DELPH-IN Tools and Deep Learning

Ann Copestake (chairing)

July 16, 2019

DELPH-IN: Deep Linguistic Processing using HPSG

- LinGO (Linguistic Grammars Online): started in early 1990s at CSLI, Stanford to build computational grammars based on the HPSG linguistic framework.
- DELPH-IN collaboration on tools and grammars started in 2001, 18 groups including LinGO: see `http://www.delph-in.net/`
- Some large grammars: ERG for English.
- Shared technology for parsing, realization etc, common semantic framework (MRS and variants).
- Grammar Matrix: framework/starter kit for the development of grammars from all language families: over 500.
- All grammars and technology are Open Source

An ERG example

Very few of the Chinese construction companies consulted were even remotely interested in entering into such an arrangement with a local partner.

An ERG example

Very few of the Chinese construction companies consulted were even remotely interested in entering into such an arrangement with a local partner.

modified quantifier

An ERG example

Very few of the Chinese construction companies consulted were even remotely interested in entering into such an arrangement with a local partner.

partitive

An ERG example

Very few of the Chinese construction companies consulted were even remotely interested in entering into such an arrangement with a local partner.

compound nominal

An ERG example

Very few of the Chinese construction companies consulted were even remotely interested in entering into such an arrangement with a local partner.

reduced relative

An ERG example

Very few of the Chinese construction companies consulted were even remotely interested in entering into such an arrangement with a local partner.

modified modifier

An ERG example

Very few of the Chinese construction companies consulted were even remotely interested in entering into such an arrangement with a local partner.

predeterminer

DELPH-IN grammars

- Broad-coverage grammars are combinations of many thousands of constraints for a language.
- Used for parsing and realization of language.
- Constructed manually but:
  - ▶ A core grammar matrix has been developed which can be shared across languages.
  - ▶ Some tools for semi-automatic grammar construction.
  - ▶ Large grammars use automatic acquisition for straightforward parts of the lexicon.
  - ▶ Machine learning for ranking of parses and realizations.
- Compositional semantic construction producing MRS, but very little lexical semantics in the grammars.
- Treebanked data (various genres) and large quantities of automatically parsed data, especially for English.

Using DELPH-IN technology with NNs

- Buys and Blunsom (2017): Robust Incremental Neural Semantic Graph Parsing.
- Yufei Chen, Weiwei Sun, Xiaojun Wan (2018): Accurate SHRG-based semantic parsing.
- Today:
    1. Goals of computational semantics: Guy Emerson
    2. Computational support for the constructivist approach: Weiwei Sun.
    3. Neural text generation from rich semantic representations: Michael Goodman (Hajdik et al, NAACL paper, 2019).
    4. Deep learning evaluation using ShapeWorld: Alex Kuhnle.
    5. Captioning in ShapeWorld: Huiyuan Xie.