Site Update: DFKI & CoLi, Saarbrücken/Berlin

Yi Zhang

LT-Lab, DFKI, Germany

Computational Linguistics, Saarland University, Germany

DELPH-IN Summit 2012
Sofia, Bulgaria

Overview of Activities

- WSJ treebanking with ERG
- Robust HPSG parsing with approximating PCFGs
- CLIMB: Code Generation for Grammar Development
- Grammar spring-cleaning
- Linguistically-enriched Bulgarian-to-English SMT models (EuroMatrix++)
- Successful completion of TAKE
- New project Deependance
- MCG: Mandarin Chinese grammar
- Documenting GG (Grammar Engineering course project)
- Russian grammar and treebank

- Funded by German Federal Ministry of Education and Research (BMBF)
- Duration: 01.2012 – 12.2014
- Deep dependency oriented analysis with non-discrete constraints
- Aim to improve DLP methods by extending the data-driven dependency parsing schemes with additional elements of linguistic and cognitive sophistication
- German, English
- Applications
    - Relation Extraction
    - Computer-Assisted Language Learning

- Extract linguistic knowledges (lexicon, constructions) from the existing HPSG grammars (GG, ERG) and treebanks
- Extend the TDL formalism to model feature structure disjunctions
- Obtain preferences of soft constraints from corpora
- Transition-based parsing models with preference-controlled constraint satisfaction

Mandarin Chinese Grammar

- Preparing for a new release in July/August
- Reasonable coverage on phenomenon test suite
- Open access: `http://mcg.opendfki.de`
- Emerging multi-site collaboration
- Merge with ManGO on the way