

Kurt Eberle<sup>1</sup>, Kerstin Eckart<sup>2</sup>, Boris Haselbach<sup>2</sup>, Ulrich Heid<sup>3</sup>, Nadine Siegmund<sup>2</sup>  
<sup>1</sup>Lingenio GmbH, Heidelberg; <sup>2</sup>IMS, Universität Stuttgart; <sup>3</sup>IwiSt, Universität Hildesheim

### **Eine Infrastruktur für Inspektion und Vergleich von automatisch oder manuell erstellten FUDRT-Analysen**

Bei der detaillierten Untersuchung von Phänomenen an der Syntax-Semantik-Schnittstelle besteht Bedarf, semantische Repräsentationen zu notieren, zu editieren, zu inspizieren und zu verwalten. Im Stuttgarter Sonderforschungsbereich SFB 732 *Incremental Specification in Context* arbeiten mehrere Teilprojekte an Fragestellungen, für die Analysen von Wörtern, Sätzen und Texten auf der Basis von Diskursrepräsentationstheorie (DRT, Kamp und Reyle 1993) anfallen; je nach theoretischem Hintergrund und Fokus der Forschung unterscheiden sich die einzelnen Analysen im Detail; sie sind aber im Format von DRT notierbar.

Ziel der hier beschriebenen Arbeiten ist es, eine Infrastruktur bereitzustellen, die die Verwaltung und den Vergleich solcher Analysen unterstützt und als Arbeitsumgebung für die gemeinsame Arbeit dienen kann, ggf. auch über den SFB 732 hinaus. Die Infrastruktur besteht aus vier verschiedenen Bausteinen, die semantische Repräsentationen im Format der flachen Unterspezifizierten DRT (FUDRT, Eberle 2004) verarbeiten; FUDRT liefert eine kompakte getypte Repräsentation, mit flacher Unterspezifikation. Die Komponenten sind:

1. Automatische Analyse von Sätzen: eine FUDRS-Konstruktion, die an das syntaktische Analysewerkzeug *B3Tool* (Eberle et al. 2008) angeschlossen ist. *B3Tool* beruht auf der deutschen Grammatik des MÜ-Systems *translate* von Lingenio (URL: <http://www.lingenio.de/>, 20.08.2013).
2. Editieren/Visualisieren von DRSen: *VisDRS*, ein interaktives Werkzeug, mit dem DRSen schrittweise manuell erstellt oder automatisch erzeugte DRSen visualisiert und bearbeitet werden können.
3. Ausgabe/Visualisierung im Text: Konverter zu LaTeX (Haselbach et al. 2011).
4. Verwaltung: Import- und Exportfunktionen, sowie Repräsentation und Versionierung von und Suche auf den Analyseergebnissen in der Datenbank *B3DB* (Eckart et al. 2010).

Alle Komponenten nehmen FUDRT-Repräsentationen als Input (Datenbank, Konverter nach LaTeX) oder liefern sie als Output (automatische Analyse, interaktives Erstellungs- und Visualisierungstool *VisDRS*). Die Arbeitsumgebung unterstützt den Vergleich von Analysen und zeigt gleichzeitig die Nützlichkeit von FUDRT als Format für die interne Repräsentation und den Austausch semantischer Analysen.

### **Referenzen**

- Eberle, K. (2004): *Flat underspecified representation and its meaning for a fragment of German*, Universität Stuttgart.
- Eberle, K.; Heid, U.; Kountz, M. und Eckart, K. (2008): „A Tool for Corpus Analysis using partial Disambiguation and Bootstrapping of the Lexicon“. in Storrer, A. et al. (Hrsgg.) *Text Resources and Lexical Knowledge*, 145-157, Berlin: Walter de Gruyter.
- Eckart, K.; Eberle, K. und Heid, U. (2010) „An Infrastructure for More Reliable Corpus Analysis“, in *Proceedings of the Workshop on Web Services and Processing Pipelines in HLT: Tool Evaluation, LR Production and Validation (LREC 2010)*, 8-14, Valletta, Malta.
- Haselbach, B.; Eckart, K. und Eberle, K. (2011): „Datenbankrepräsentation und Visualisierung semantischer Repräsentationen in Flat Underspecified DRT“ Poster bei *DGfS-CL Postersession 2011*, Göttingen.
- Kamp, H. und Reyle, U. (1993): *From Discourse to Logic. Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and DRT*, Dordrecht: Kluwer.