Towards a (Logical) Extraction of Semantic Relations from Syntactic Annotations

Jens Michaelis (Universität Osnabrück) & Uwe Mönnich (Universität Tübingen)

Motivation
- Relevant (linguistic) annotation models essentially provide tree-like structures (cf. figure).

Tree-like Graph Grammars (Goals)
- Formal properties of (logically-)algebraically definable families of tree-like graph grammars.
- Transductions between tree-like graph structures.
- Query systems over tree-like graph structures.
  - Potential (logically) query language provided by Monadic Second Order Logic (MSO).
    - First order logic + quantifications over (monadic) sets.

Advantages:
- Expressiveness
- Efficient querying in linear time depending on the number of nodes of a tree-like graph.
  - Guaranteed by bounded tree-width.
- Automata models at hand.

Transductions (Goals)
- Annotation graphs (Bird et al. 2000) as a general model of (linguistic) multilayer annotations.

Example
- The concept of tree-decomposition and (bounded) treewidth (Robertson and Seymour 1986) as an indicator for the tree-likeness of a graph.
- Specification of logically definable transductions on tree-like structures.
  - Preferably, in plain MSO-terms, “at least” in terms of MS$_2$O (MSO with node set and edge set quantifications).

- Restriction to a finite set of annotation graphs guarantees a class of tree-like graphs of bounded treewidth.
- Extraction via MSO query/transduction on a class of graphs of bounded treewidth can be done in linear time depending on the number of vertices (Courcelle 1990).
- If a set of graphs has a decidable MS$_2$O theory then this set is the subset of (the homomorphic image under an MSO transduction) of a recognizable set of trees, i.e., the corresponding set of graphs is interpretable in MS$_2$O-terms as a subset of a recognizable set of trees (Seese 1991).
  - Replacing graphs by, e.g., planar graphs, MS$_2$O can be replaced by MSO.

Prospective envoi
- Semantic extraction ideally results in information consistent with a (graph) representation of a given ontology.
- In case the information is “new” it can be (algorithmically) integrated as a “subgraph” into the ontology yielding an adaptation.
- Annotation graphs provide a general concept for an amalgamation of hybrid specifications.