



Impressions from the Hankø WeSearch MRS meeting

Emily M. Bender, Stephan Oepen
DELPH-IN Summit
2 July 2012
Sofia, Bulgaria

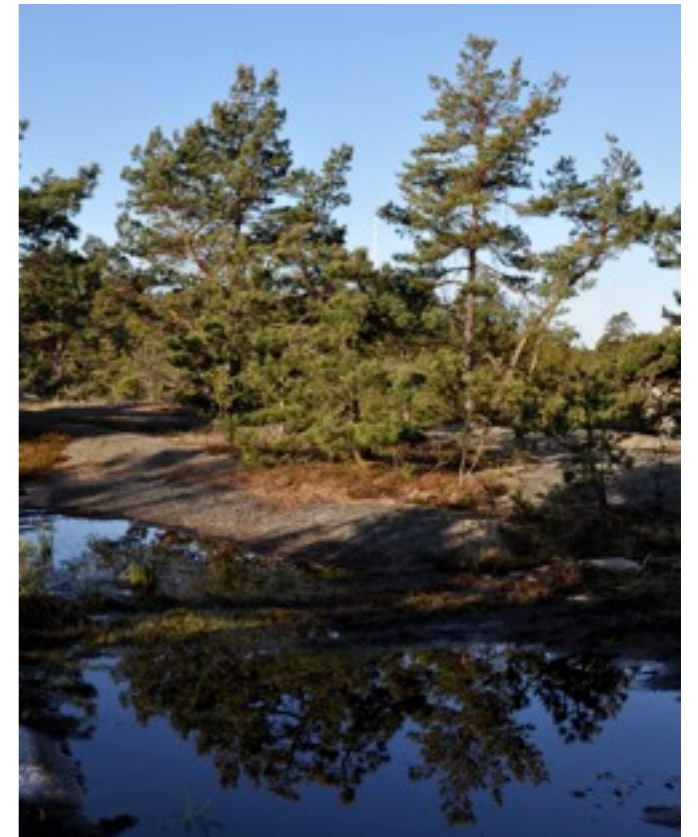
Workshop goals



WeSearch: Towards a stable semantic interface

MRS Design Principles identified

- The interface representation should include all information that is constrained by the grammar
- The interface representation should be highly normalized, abstracting away from details of surface syntax
- Close paraphrases should lead to comparable or identical structures in the interface representation
- Minimize ambiguity
 - Corollary: Differentiate lexical predicates only if that distinction corresponds to morphosyntactic differences
- Decomposition is desirable, but not a goal in itself.



Points of detailed discussion

- Interaction of coordination (RNR) and label identification, for both scopal and non-scopal arguments
- Predicates of underspecified arity
- Representation of comparatives
- Simplified semantics of complex proper names, inspired by treatment of complex number names (fewer quantifiers!)
- ICONS: anaphora links (possible, likely, prohibited); information structure; variable property identification for generator input; underspecification of PP attachment; appositives, non-restrictive relatives, parentheticals...
- Underspecification of modifier attachment (PPs, NNNN compounds, names)



Notes: <http://moin.delph-in.net/WeSearch/HankøSchedule>

ICONS (brief preview)

- Constraints involving two referential entities (e or x)
- Additional list parallel to RELS, HCONS
- Need different types, if they are to be applied in all those different ways
 - Coreference (anaphora resolution)
 - Not coreference (Principles A, B)
 - Coreferent and share variable properties (*to bite one's tongue*, etc)
 - Trigger search for set of variables + also label sharing (underspecified PP attachment)
 - Topic, focus, contrast (information structure)
- Expect some to come from the grammar, others to come from post-processing (e.g., anaphoric links)



Underspecification of modifier attachment: Possible approach



- Attach PP-type modifiers only low; force one type of branching on NNN compounds
- Provide a special link between modifier's ARG and head's INDEX (ICONS) that triggers search at next stage for possible other values
- Search or other constraints capture interactions between attachment decisions: *Kim saw the astronomer on the hill with the telescope* (high attachment of *on* blocks attachment of *with* to *astronomer*)
 - Possibly through reasoning about (non-)projectivity in DMRS
- Treebanking extended to include hand-resolution of such dependencies
 - Possibly allowing only partial resolution in cases where context doesn't fully constrain (*airline reservation counter; be careful when using fire in the forest*)

Underspecification of modifier attachment: Motivation

- Smaller parse forests
- Compatible with the goals of saying as much about the semantics as we can and no more
- Create an interesting data set for studying attachment ambiguity
- Possibly beneficial for ML over (distributional) semantics for parse selection
- Possibly a better account of *In Paris on Tuesday Kim wants to read a book.*



Underspecification of modifier attachment: Challenges

- Disambiguation could be more difficult as a two-stage process (though possibly artificially easy now)
- Compositionality: how to force low attachment in the grammar
- Fundamental change to the processing machinery
- Second stage: how to (partially) resolve the underspecification, both in manual annotations and automated processing
- Challenge to mono-stratal roots?

