Underspecified Modelling of Complex Discourse Constraints

Markus Egg, University of Groningen
egg@let.rug.nl
Michaela Regneri, Saarland University
regneri@coli.uni-sb.de

Underspecified discourse structures

- [c1] I try to read a novel  [c2] if I feel bored  [c3] or I am unhappy.
- (read) IF ( (bored) OR (unhappy) )

- represent genuine ambiguity
- gain robustness in discourse parsing
- represent global discourse constraints

Regular Tree Grammars as underspecification formalism

- [c1] I try to read a novel  [c2] if I feel bored
- [c3] because the TV programs disappoint me  [c4] but I can't concentrate.

- integrate constraints using standard algorithms
- RTGs are expressively complete, the resulting representation may exceed the expressive power of graph-based formalisms

Corpus studies reveal interdependencies between discourse relations and tree structure

Representing soft constraints with weighted Regular Tree Grammars

- maximum weight: 9
- medium weight: 3
- minimum weight: 1

Integrating complex constraints by intersection with filter grammars

- cause is never a right child

References: