Formal Methods for Testing Grammars

Inari Listenmaa

The defence of this thesis for the degree of Doctor of Philosophy
will be held in room **EC**, Hörsalsvägen 11,
Chalmers University of Technology,
on **Friday 8th March, 2019** at **10:00**.

Faculty opponent: **Fred Karlsson**
University of Helsinki

Abstract

Grammar engineering has a lot in common with software engineering. Analogous to a program specification, we use descriptive grammar books; in place of unit tests, we have gold standard corpora and test cases for manual inspection. And just like any software, our grammars still contain bugs: grammatical sentences that are rejected, ungrammatical sentences that are parsed, or grammatical sentences that get the wrong parse.

This thesis presents two contributions to the analysis and quality control of computational grammars of natural languages. Firstly, we present a method for finding *contradictory grammar rules* in Constraint Grammar, a robust and low-level formalism for part-of-speech tagging and shallow parsing. Secondly, we generate *minimal and representative test suites* of example sentences that cover all grammatical constructions in Grammatical Framework, a multilingual grammar formalism based on deep structural analysis.