

THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Learning Language (with) Grammars

From Teaching Latin to Learning Domain-Specific Grammars

HERBERT LANGE

Presentation:

September 16th, 2020, 10:00

Room EDIT 8103, EDIT Building

Hörsalsvägen 11 (Rännvägen 6),

Chalmers University of Technology, Campus Johanneberg

Online: <http://www.cse.chalmers.se/~langeh/defense.html>

Faculty Opponent:

Johanna Björklund

Associate professor at Department of Computing Science

Umeå University, Sweden



UNIVERSITY OF GOTHENBURG

The thesis is available at:
Department of Computer Science & Engineering
University of Gothenburg
Gothenburg, Sweden, 2020

Phone: +46 31 772 1000

Abstract

This thesis describes work in three areas: grammar engineering, computer-assisted language learning and grammar learning. These three parts are connected by the concept of a grammar-based language learning application. Two types of grammars are of concern. The first we call resource grammars, extensive descriptions a natural languages. Part I focuses on this kind of grammars. The other are domain-specific or application-specific grammars. These grammars only describe a fragment of natural language that is determined by the domain of a certain application. Domain-specific grammars are relevant for Part II and Part III. Another important distinction is between humans learning a new natural language using computational grammars (Part II) and computers learning grammars from example sentences (Part III).

Part I of this thesis focuses on grammar engineering and grammar testing. It describes the development and evaluation of a computational resource grammar for Latin. Latin is known for its rich morphology and free word order, both have to be handled in a computationally efficient way. A special focus is on methods how computational grammars can be evaluated using corpus data. Such an evaluation is presented for the Latin resource grammar.

Part II, the central part, describes a computer-assisted language learning application based on domain-specific grammars. The language learning application demonstrates how computational grammars can be used to guide the user input and how language learning exercises can be modeled as grammars. This allows us to put computational grammars in the center of the design of language learning exercises used to help humans learn new languages.

Part III, the final part, is dedicated to a method to learn domain- or application-specific grammars based on a wide-coverage grammar and small sets of example sentences. Here a computer is learning a grammar for a fragment of a natural language from example sentences, potentially without any additional human intervention. These learned grammars can be based e.g. on the Latin resource grammar described in Part II and used as domain-specific lesson grammars in the language learning application described Part II.

Keywords: Latin, Latin syntax, Latin morphology, Grammar engineering, Grammar testing, Corpus-based evaluation, Computer-Assisted Language Learning, Grammar learning, Constraint-satisfaction, Constraint-optimization

ISBN: 978-91-7833-986-0 (Print) 978-91-7833-987-7 (PDF)

URL: <http://hdl.handle.net/2077/65453>