Lecture Notes in Computer Science

Edited by G. Goos and J. Hartmanis

91

Derick Wood

Grammar and L Forms: An Introduction

Springer-Verlag
Berlin Heidelberg New York 1980

Editorial Board

W. Brauer P. Brinch Hansen D. Gries C. Moler G. Seegmüller
J. Stoer N. Wirth

Author

Derick Wood
McMaster University, Unit for Computer Science
1280 Main Street West
Hamilton, Ontario
Canada L8S 4K1

AMS Subject Classifications (1979): 68 A 05, 68 A 30
CR Subject Classifications (1974): 5.22, 5.23

ISBN 3-540-10233-7 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-10233-7 Springer-Verlag New York Heidelberg Berlin

Library of Congress Cataloging in Publication Data. Wood, Derick, 1940- Grammar and L forms. (Lecture notes in computer science; 91) Bibliography: p. Includes index.
1. Formal languages. I. Title. II. Series.
QA267.3.W66. 511.3. 80-20426

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically those of translation, reprinting, re-use of illustrations, broadcasting, reproduction by photocopying machine or similar means, and storage in data banks. Under § 54 of the German Copyright Law where copies are made for other than private use, a fee is payable to the publisher, the amount of the fee to be determined by agreement with the publisher.

© by Springer-Verlag Berlin Heidelberg 1980
Printed in Germany

Printing and binding: Beltz Offsetdruck, Hemsbach/Bergstr.
2145/543210-3140

To Deborah Jane, Doreen, Ethel, Henry and Sharon Clare

PREFACE

The study of form theory as a new discipline within formal language theory is less than ten years old. It was initiated by Armin Cremers and Seymour Ginsburg around 1974/75 and since this time it has truly blossomed.

The primary goal of these lecture notes is to provide an introductory and unified account of "context-free" form theory as it stands today. Because of its introductory rather than exhaustive nature many results are only mentioned in passing rather than being given the space they deserve. However I trust the reader will be convinced by the material in these notes that the study of grammar and L forms has indeed contributed towards (a) the understanding of generative devices, (b) the deeper understanding of the context-free and EOL languages and (c) the development of a theory of grammatical similarity. Moreover many interesting and open problems remain to be solved.

A word of warning about proofs is in order. A number of the proofs are given in an incomplete fashion, that is a construction and/or a proof sketch is all that is usually given. The reason for this is that many proofs for forms reduce to "standard" proofs for grammars. Therefore it is assumed that the reader has a good working knowledge of formal language theory, particularly of context-free EOL and ETOL grammars and languages. With this warning in mind the material should be accessible to advanced undergraduate students, graduate students and researchers in formal language theory.

These lecture notes consist of five chapters and an extensive bibliography. Chapter I reviews the basic terminology, notation and conventions needed in the following chapters and also carries out a rapid survey of various approaches to grammatical similarity. One important aspect of form theory is its contribution to this topic. Chapters II and III form the core material of these notes; context-free grammar forms are studied in Chapter II and EOL and ETOL forms in Chapter III. In contradistinction, Chapter IV discusses a number of results for "non-context-free" forms. Finally Chapter V presents the historical background of the material presented herein, discusses some omitted results and finally presents some open problems in addition to those mentioned throughout the text.

Two possible study sequences are to first read: Chapter I, Section I.1 then

either Chapter II, Sections II.1, II.2, II.3.3, II.3.4, II.4, II.6.1 and II.6.2

or Chapter III, Sections III.1, III.2, III.3.1, III.4.1, III.4.2, III.5.1, III.6.1 and III.6.4

followed by Chapter V, depending on whether your interest lies in context-free grammar forms or EOL forms.

These lecture notes would not have been assembled without the assistance, encouragement and support of my two longstanding collaborators Hermann A. Maurer and Arto Salomaa. We have had many enjoyable times together proving theorems and, most important, taking löyly. Second I wish to thank Aristid Lindenmayer, the originator of L systems for his continuing willingness to discuss them. Also there are those people who either through collaboration and/or by reading a draft version have contributed to this manuscript. These are: W. Ainhirn, B. von Braunmühl, K. Culik II, S. Ginsburg, J. Hagauer, E. Hotzel, R. Leipälä, Th. Ottmann, M. Penttonen, G. Rozenberg and E. Welzl. However the many errors and inconsistencies that surely remain are completely attributable to myself.

In looking over my correspondence I see that my proposal for this book was first made to Springer-Verlag on September 13, 1977. I wish to thank them for their encouragement and patience over the intervening two and a half years, when I failed to meet deadline after deadline. However I feel that these notes are the better for the delay since it has enabled much new material to be included.

I wish to express my heartfelt gratitude to Suzanna den Bleker who labored long to produce a readable first draft from my standard scribble, to Robin Robertson who provided last minute assistance with her usual cheerful manner and to Margaret Belec who has done an excellent job in producing the final copy.

Finally, I want to ascribe credit to the nineteenth century English novelist Anthony Trollope who taught me the importance of a regular writing schedule and to the Lord of the Universe, Jesus Christ who helped me to stick to it!

McMaster University
2nd June 1980

CONTENTS

			Page
PREFACE			
CHAPTER I:	BEGINNINGS		1
	I.1	Basic Terminology and Notation	2
	I.1.1	Context-Free Grammars and Languages	7
	I.1.2	EOL and ETOL Grammars and Languages	11
	I.2	Notions of Grammatical Similarity	16
	I.2.1	Weak and Structural Equivalence	16
	I.2.2	Covers	18
	I.2.3	Translations and Szilard Languages	19
	I.2.4	Grammar Morphisms	21
	I.2.5	Topological Similarity	22
	I.2.6	Grammar Collections	25
CHAPTER II:	CONTEXT-FREE GRAMMAR FORMS		26
	II.1	The Basics of Context-Free Grammar Forms	27
	II.1.1	Two Kinds of Interpretation	27
	II.1.2	Grammar and Grammatical Families	31
	II.2	Isolation and Simulation	35
	II.2.1	Isolation	37
	II.2.2	Simulation	39
	II.3	Collections of Grammar Families	47
	II.3.1	Pre-Orders and Closure Operators	47
	II.3.2	Composition of Grammatical Closure Operators	48
	II.3.4	Strong Form Equivalence and Lattices	62
	II.3.4	Minimal Grammar Forms	66
	II.4	Normal Form, Closure and Characterization Results	71
	II.4.1	Reduction Results	71
	II.4.2	Closure Properties	86
	II.4.3	Characterization Theorems	90

	II.5	Syntax Analysis	100
	II.5.1	Syntax Analysis of s-grammatical Families	100
	II.5.2	Precedence Relations	106
	II.5.3	Pushdown Acceptor Forms	113
	II.6	Dense Collections of Grammatical Families	121
	II.6.1	Preliminary Notions	121
	II.6.2	Denseness and s-grammatical Families	124
	II.6.3	MSW Spaces	130
	II.6.4	Two-symbol Grammatical Families	136
CHAPTER III:	EOL AND ETOL FORMS		146
	III.1	The Basics of EOL and ETOL Forms	147
	III.1.1	Derivation Relationships	152
	III.1.2	Isolation and Simulation	153
	III.2	Reduction Theorems	159
	III.2.1	Some Preliminary Reduction and Non-Reduction Results	159
	III.2.2	Propagating and Two-table Normal Forms	164
	III.3	Completeness	170
	III.3.1	Two-symbol EPOL Forms	170
	III.3.2	Short {S,A,a}-synchro-EPOL Forms	186
	III.3.3	ETOL Forms and Completeness	191
	III.3.4	Regular-completeness	196
	III.4	Goodness and Very-completness	199
	III.4.1	Preliminaries	199
	III.4.2	The EOL Situation	201
	III.4.3	The ETOL Situation	205
	III.4.4	Relative and Mutual Goodness	209
	III.5	Decidability Questions	212
	III.5.1	Strong Form Equivalence	212
	III.5.2	Uniform Interpretations of EOL Forms	215
	III.5.3	OL Forms and Form Equivalence	221

	III.6	Generative Capacity of EOL Forms	227
	III.6.1	Closure and Non-closure Properties	227
	III.6.2	Sub-Regular Grammatical Families are EOL Form Families	229
	III.6.3	Regular-completeness, -boundedness and -sufficiency	233
	III.6.4	Generators	239
CHAPTER IV:	OTHER TOPICS		248
	IV.1	Phrase Structure Grammar Forms	249
	IV.2	EIL Forms	258
	IV.3	Context-free Matrix Forms	266
	IV.4	Controlled Grammar Forms	275
CHAPTER V:	PAST, PRESENT AND FUTURE		284
	V.1	Historical Background	285
	V.2	A Discussion of Some Omitted Results	289
	V.3	New Directions and Open Problems	297
BIBLIOGRAPHY			302