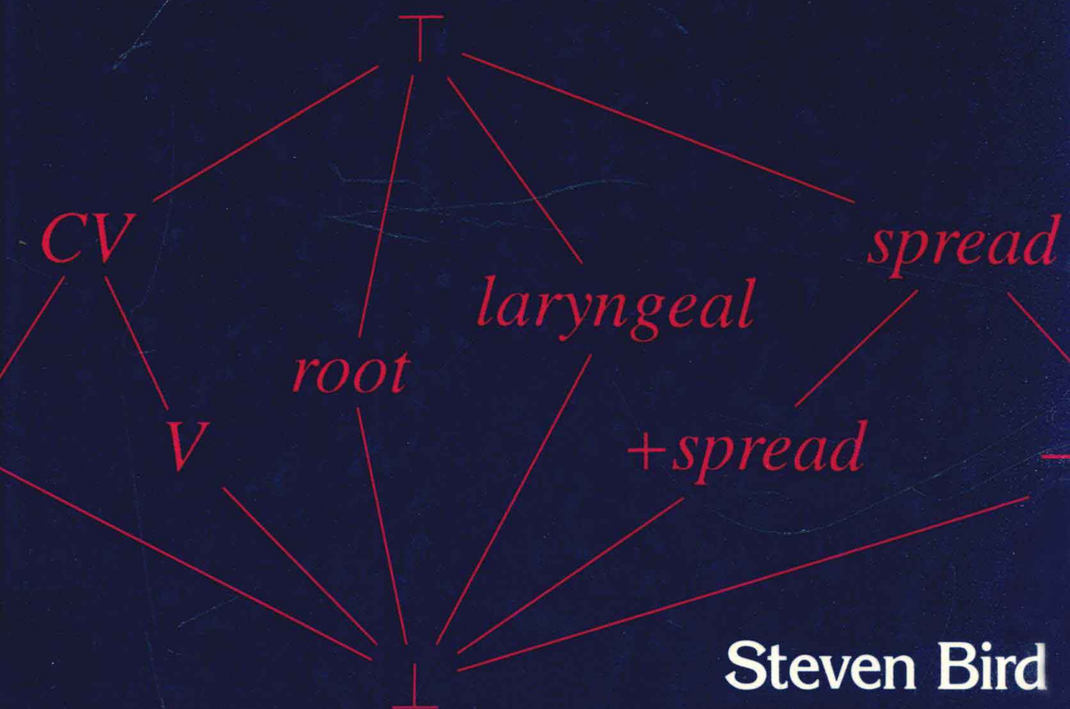
Studies in
NATURAL
LANGUAGE
PROCESSING

COMPUTATIONAL PHONOLOGY

A constraint-based approach



Steven Bird

Computational phonology

A constraint-based approach

by

STEVEN BIRD

University of Edinburgh

CAMBRIDGE
UNIVERSITY PRESS

Published by the Press Syndicate of the University of Cambridge
The Pitt Building, Trumpington Street, Cambridge CB2 1RP
40 West 20th Street, New York, NY 10011-4211, USA
10 Stamford Road, Oakleigh, Victoria 3166, Australia

First published 1995

Printed in Great Britain at the University Press, Cambridge

A catalogue record for this book is available from the British Library

Library of Congress cataloguing in publication data

Bird, Steven.
Computational phonology: a constraint-based approach / Steven Bird
p. cm. – (Studies in natural language processing)
Includes bibliographical references and indexes.
ISBN 0 521 47496 5 (hardback)
1. Grammar, Comparative and general – Phonology – Data processing.
2. Computational linguistics. 3. Constraints (Artificial intelligence)
4. Language and logic. I. Title. II. Series.
P217.3.B57 1995
414'.0285–dc20 94-30865 CIP

ISBN 0 521 47496 5 hardback

SGB

Computational phonology is one of the newest areas of computational linguistics, and is experiencing rapid growth as its practitioners apply the wealth of theories, technologies and methodologies of computational linguistics to phonology. This book is the first to survey these developments, and it does so in a way that is accessible to computational linguists, phonologists and computer scientists alike. The interests of these diverse groups overlap in the subject area of constraints. The goal of this book is to explore the use of constraints in modern non-linear phonology and then – drawing on insights from constraint-based grammar and constraint logic programming – to formalise and implement a constraint-based phonology.

Studies in Natural Language Processing

Computational Phonology

Studies in Natural Language Processing

Also in the series

Semantic Processing for Finite Domains, M.S. Palmer
Reference and Computation, A. Kronfeld
Text Generation, K.R. McKeown
Planning English Sentences, D.E. Appelt
Relational Models of the Lexicon, M.W. Evens
Systemic Text Generation as Problem Solving, T. Patten
Memory and Context for Language Interpretation, H. Alshawi
Machine Translation, edited by S. Nirenburg
Semantic Interpretation and the Resolution of Ambiguity, G. Hirst
The Linguistic Basis of Text Generation, L. Danlos
Computational Linguistics, R. Grishman
Studies in Natural Language Parsing, edited by D. Dowty, L. Karttunen and A. Zwicky
Language and Spatial Cognition, Annette Herskovits
Machine Translation Systems, edited by Jonathan Slocum
Challenges in Natural Language Processing, edited by M. Bates and R.M. Weischedel
Computational Linguistics and Formal Semantics, M. Rosner and R. Johnson
Inheritance, Defaults and the Lexicon, edited by E.J. Briscoe, A. Copestake and V. de Paiva

For Andrew and Alison

Preface

This book is an expanded and reorganised version of the author's PhD thesis 'Constraint-Based Phonology' (Edinburgh University, 1990). It has been substantially revised to take account of recent developments in computational phonology.

This book has been composed with three audiences in mind: computational linguists, phonologists and computer scientists. The interests of these diverse groups overlap in the subject area of CONSTRAINTS, a central theme of this work. In computational linguistics there is a well-established research tradition known as CONSTRAINT-BASED GRAMMAR. Phonologists are actively studying the role of constraints in the lexicon, in derivations, on rules and on surface forms. The logic programming community in computer science is concerned with developing languages for CONSTRAINT PROGRAMMING and with efficient algorithms for constraint resolution.

For all three groups, a CONSTRAINT expresses a generalisation which should be true of all candidate solutions. Constraints INTERACT in interesting and potentially complex ways, mutually constraining the solution space. Furthermore, the notion of ORDERING is largely absent: the sequence in which constraints are applied does not affect the end result.

The goal of this book is to clarify the role that constraints play in phonology and then – drawing on insights from constraint-based grammar and constraint-programming – to formalise and implement a CONSTRAINT-BASED PHONOLOGY. For practical reasons it has been necessary to restrict this exercise to the established core of practice in contemporary phonology (e.g. feature geometry, autosegmental association, prosodic hierarchy, licensing) and avoid some of the more controversial and less well-understood devices. On the computational side, it has been necessary to avoid the use of defaults in the interests of producing an implementation in the style of a constraint solver.

There are several motivations for this work. First, the practising phonologist faces severe limitations when it comes to developing and testing a sizeable theory; computerisation promises to alleviate this burden. Second, work in the field of natural language processing is limited to languages without complex phonological processes represented in the orthography; an implemented contemporary phonological model may help to overcome such a restriction and enable natural language processing technology to be applied to a much wider range of languages. Third, it is sometimes claimed that rule-based speech recognition systems exhibit poor performance as a result of employing the 1960's SPE rule system; contemporary non-linear phonological models bear a closer resemblance to the speech stream and they promise to provide a fresh source of symbolic techniques to guide speech recognition systems. Finally, it is possible to view phonology as a bridge between the speech technology community and the natural language processing community; attention to developments in phonology may help to achieve the long-term prospect of having integrated speech and language systems.

It is a pleasure to acknowledge my debt to several people who have been instrumental in helping me transform a collection of ideas into the form of a monograph: Patrick Blackburn, Jo Calder, John Coleman, Robin Cooper, Mark Ellison, Dafydd Gibbon, Mark Johnson, András Kornai, Marcus Kracht, Bob Ladd, Chris Mellish, Dick Oehrle, Geoff Pullum, Mike Reape, Jim Scobbie, Henry Thompson, Richard Weise, Pete Whitelock, and many others. The material presented here has also been used in courses at the *Second and Fifth European Summer Schools in Logic, Language and Information* (Leuven 1990, Lisbon 1993), the *32nd Annual Meeting of the ACL* (Las Cruces, 1994) and the *2nd Australian Linguistics Institute* (Melbourne, 1994), and I would like to thank the participants for helping to test the material and for providing valuable feedback. In particular I would like to thank Ewan Klein, for his careful guidance during the PhD on which this work is based, and for his clarity of thought and practical wisdom about presentation, which were nothing short of inspirational. I am indebted to my other teachers Roland Sussex, John Upton and Jean-Louis Lassez who fostered my interests in linguistics, mathematics and constraint-logic programming while I was a student at Melbourne University.

It has been a pleasure to collaborate with Bran Boguraev (the series editor) and the staff at Cambridge University Press in the production of this book. I am also grateful to the following organisations for financial support: the Overseas Research Studentship Awards Scheme, Edinburgh University, the Victoria League, the Linguistic Society of America and the Science and Engineering Research Council.

Greater debts are nearer to home. My christian friends in Scotland have been an extended family to me while I have lived in exile from Australia. My parents, by their example, taught me perseverance and dedication to the task, for which I am eternally grateful. Finally, I thank Kay for being such a rich blessing on my life throughout this work.

Contents

1 Introduction

The last two decades have witnessed a vigorous growth of new descriptive notational devices in phonology. These devices have had enormous heuristic value in helping practitioners to see and intuitively understand complex phenomena. However, linguistic notations should be 'perfectly explicit' and ought not 'rely on the intelligence of the understanding reader' (Chomsky, 1965, 4). It is not clear that modern non-linear phonology, to any great extent, meets these fundamental requirements of generative grammar. If current work in computational phonology and speech technology is focused on the outdated SPE model (Chomsky and Halle, 1968) it is because nothing more recent has surpassed SPE's formal explicitness. Therefore, it is time for these new phonological frameworks to be placed on a formal footing.

Those who are suspicious of formalism may cry foul at this point. After all, many a good linguistic insight has been buried under a barrage of definitions and theorems, and a preoccupation with technical hygiene may blinker one's vision of what is really going on. However, the solution is not to retreat to a position where *formulating* a description is synonymous with *formalising* a description. Rather, we need to recognise that formalisation has considerable heuristic value of its own. After all, linguistic theorising in this century has been characterised – possibly even driven – by a tension between attempts at rigorous theories of linguistic structure and attempts to formulate intuitively sensible descriptions of linguistic phenomena.

Beyond this, formalisation is fundamental to the empirical basis of the field. The widespread practice of testing an empirical generalisation on isolated examples leads to unstable theories which are restricted to small fragments of a language. If, as noted with regret by Anderson (1989, 803), outside observers do not always take phonology seriously, then an important reason is different notions of what a scientific theory is and does. As we shall see below, a phonological theory which can be implemented on a computer can meet the dual requirements of rigour and non-trivial empirical content which much current work has unsuccessfully striven to achieve.

Underlying these concerns is the goal of constructing grammars which do not favour generation at the expense of recognition, or vice versa. This connects with the familiar debate about the metatheoretical undesirability of extrinsically ordered rules (Koutsoudas, Sanders and Noll, 1974; Hooper, 1976), and with earlier complaints that the derivational stance of generative phonology was inherently process-oriented. Despite claims to the contrary, many current phonological theories remain performance models. They enumerate the steps which must be taken in moving from a lexical form to a surface form, borrowing heavily on the now dated flowchart model of computation. Crucially, there is no guarantee that such rule systems work in reverse.[1] If we accept that linguistics is the study of that knowledge which is independent of processing tasks, then the statements of a linguistic theory ought to have a declarative semantics: an interpretation which is expressed solely in terms of the utterances which are licensed by theory. Of course, if a theory is going to be useful its statements should also have one or more procedural interpretations, but these ought not to be mistaken for the linguistic theory itself.

Readers with a background in computational syntax and semantics will be wondering how this computational phonology could fit into an overarching computational grammar framework. Here, our starting point is provided by the work of Deirdre Wheeler and Emmon Bach,[2] who showed how the principles of Montague grammar can be applied to phonology. However, the aim here is not to perform this integration of phonology and grammar, but rather to do phonology in such a way that this integration is possible. Therefore, a monostratal approach[3] to phonological description

---

[1] This non-reversibility is a general result, which Bear (1990) has demonstrated for Klamath (Halle and Clements, 1983, 113).

[2] Wheeler (1981, 1988); Bach and Wheeler (1981); Bach (1983).

[3] See §1.4.5 for an explanation of the term 'monostratal'.

has been adopted since this is a requirement for a phonological framework to be integrated into existing constraint-based grammar frameworks.[4]

In this connection it is necessary to introduce two distinctions. The first is the DESCRIPTION/OBJECT distinction: an expression of a linguistic theory DENOTES the class of utterance tokens which SATISFY that expression. These expressions are combined using familiar logical connectives. While there is a fundamental difference in kind between descriptions and objects, and so one might imagine that this configuration is actually polystratal, there remains only a single level of linguistic description. This state of affairs contrasts with the procedural model of traditional generative phonology in which there is no principled upper bound on the number of intermediate levels of description. Frameworks which build in this distinction are sometimes called CONSTRAINT-BASED because their linguistic descriptions act in concert to mutually constrain the solution space.

A second distinction is that of FRAMEWORK versus THEORY. A linguistic framework is essentially a formal notation in which linguistic theories can be stated. As such, a framework makes no empirical claims of its own, though a good framework should facilitate the expression and evaluation of such claims. Just as two theories which make contradictory claims can be expressed in the same framework, a given theory can potentially be encoded in a variety of different frameworks. A computational benefit of frameworks is that once a framework is implemented, a whole family of theories can be easily expressed within it, and it is not necessary to write whole implementations from scratch for each new theory.

These are the essential ingredients of what I shall call CONSTRAINT-BASED PHONOLOGY, a term derived from the established fields of constraint-based grammar and constraint logic programming. It is hoped that the eventual payoff of work in this vein will be the construction of rigorous and empirical phonological theories along with the construction of integrated systems for speech and language processing. However, in the light of this aspiration the immediate goals are more humble. After providing the necessary background material in chapter 1, a logical foundation for phonology cast in a language of classical first-order predicate logic is presented along with

[4]For example, Generalised Phrase Structure Grammar (Gazdar et al., 1985), Categorial Unification Grammar (Uszkoreit, 1986), Head-Driven Phrase Structure Grammar (Pollard and Sag, 1987), and Unification Categorial Grammar (Calder, Klein and Zeevat, 1988). Some evidence of initial progress in this direction can be found in Bird (1992); Bird and Klein (1994).