The Handbook of Computational Linguistics and Natural Language Processing

Edited by

Alexander Clark, Chris Fox and Shalom Lappin

WILEY-BLACKWELL

Contents

List of Figures

List of Tables

Notes on Contributors

Preface

Introduction

Part I Formal Foundations

1 Formal Language Theory

1 Introduction
2 Basic Notions
3 Language Classes and Linguistic Formalisms
4 Regular Languages
5 Context-Free Languages
6 The Chomsky Hierarchy
7 Mildly Context-Sensitive Languages
8 Further Reading

2 Computational Complexity in Natural Language

1 A Brief Review of Complexity Theory
2 Parsing and Recognition

3 Complexity and Semantics
4 Determining Logical Relationships between Sentences

3 Statistical Language Modeling
1 Introduction to Statistical Language Modeling
2 Structured LanguageModel
3 Speech Recognition Lattice Rescoring Using the Structured Language Model
4 Richer Syntactic Dependencies
5 Comparison with Other Approaches
6 Conclusion

4 Theory of Parsing
1 Introduction
2 Context-Free Grammars and Recognition
3 Context-Free Parsing
4 Probabilistic Parsing
5 Lexicalized Context-Free Grammars
6 Dependency Grammars
7 Tree Adjoining Grammars
8 Translation
9 Further Reading

Part II Current Methods

5 Maximum Entropy Models
1 Introduction

2 Maximum Entropy and Exponential Distributions
3 Parameter Estimation
4 Regularization
5 Model Applications
6 Prospects

6 Memory-Based Learning
1 Introduction
2 Memory-Based Language Processing
3 NLP Applications
4 Exemplar-Based Computational Psycholinguistics
5 Generalization and Abstraction
6 Generalizing Examples
7 Further Reading

7 Decision Trees
1 NLP and Classification
2 Induction of Decision Trees
3 NLP Applications
4 Advantages and Disadvantages of Decision Trees
5 Further Reading

8 Unsupervised Learning and Grammar Induction
1 Overview
2 Computational Learning Theory

3 Empirical Learning
4 Unsupervised Grammar Induction and Human Language Acquisition
5 Conclusion

9 Artificial Neural Networks
1 Introduction
2 Background
3 Contemporary Research
4 Further Reading

10 Linguistic Annotation
1 Introduction
2 Review of Selected Annotation Schemes
3 The Annotation Process
4 Conclusion

11 Evaluation of NLP Systems
1 Introduction
2 Fundamental Concepts
3 Evaluation Paradigms in Common Evaluation Settings
4 Case Study: Evaluation ofWord-Sense Disambiguation
5 Case Study: Evaluation of Question Answering Systems
6 Summary

Part III Domains of Application

12 Speech Recognition

- 1 Introduction
- 2 Acoustic Modeling
- 3 Search
- 4 Case Study: The AMI System
- 5 Current Topics
- 6 Conclusions

13 Statistical Parsing

- 1 Introduction
- 2 History
- 3 Generative Parsing Models
- 4 Discriminative Parsing Models
- 5 Transition-Based Approaches
- 6 Statistical Parsing with CCG
- 7 OtherWork
- 8 Conclusion

14 Segmentation and Morphology

- 1 Introduction
- 2 Unsupervised Learning ofWords
- 3 Unsupervised Learning of Morphology
- 4 Implementing ComputationalMorphologies
- 5 Conclusions

15 Computational Semantics

- 1 Introduction
- 2 Background
- 3 State of the Art

4 Research Issues
5 Corpus-Based andMachine Learning Methods
6 Concluding Remarks

16 Computational Models of Dialogue
1 Introduction
2 The Challenges of Dialogue
3 Approaches to Dialogue System Design
4 Interaction and Meaning
5 Extensions
6 Conclusions

17 Computational Psycholinguistics
1 Introduction
2 ComputationalModels of Human Language Processing
3 Symbolic Models
4 Probabilistic Models
5 Connectionist Models of Sentence Processing
6 HybridModels
7 Concluding Remarks

Part IV Applications

18 Information Extraction
1 Introduction
2 Historical Background
3 Name Extraction
4 Entity Extraction

5 Relation Extraction
6 Event Extraction
7 Concluding Remarks

19 Machine Translation
1 Introduction
2 The State of the Art: Phrase-Based Statistical MT
3 Other Approaches to MT
4 MT Applications
5 Machine Translation at DCU
6 Concluding Remarks and Future Directions
7 Further Reading

20 Natural Language Generation
1 High-Level Perspective: Making Choices about Language
2 Two NLG Systems: SumTime and SkillSum
3 NLG Choices and Tasks
4 NLG Evaluation
5 Some NLG Research Topics
6 NLG Resources

21 Discourse Processing
1 Discourse: Basic Notions and Terminology
2 Discourse Structure
3 Discourse Coherence
4 Anaphora Resolution
5 Applications

6 Further Reading

22 Question Answering

1 What is Question Answering?
2 Current State of the Art in Open Domain QA
3 Current Directions
4 Further Reading

References

Author Index

Subject Index

Praise for The Handbook of Computational Linguistics and Natural Language Processing

“All in all, this is very well compiled book, which effectively balances the width and depth of theories and applications in two very diverse yet closely related fields of language research.”

Machine Translation

“This Handbook is exceptionally broad and exceptionally deep in its coverage. The contributions, by noted experts, cover all aspects of the field, from fundamental theory to concrete applications. Clark, Fox and Lappin have performed a great service by compiling this volume.”

Richard Sproat, Oregon Health & Science University

Blackwell Handbooks in Linguistics

This outstanding multi-volume series covers all the major subdisciplines within linguistics today and, when complete, will offer a comprehensive survey of linguistics as a whole.

Already published:

The Handbook of Child Language
Edited by Paul Fletcher and Brian MacWhinney

The Handbook of Phonological Theory, Second Edition
Edited by John A. Goldsmith, Jason Riggle, and Alan C. L. Yu

The Handbook of Contemporary Semantic Theory
Edited by Shalom Lappin

The Handbook of Sociolinguistics
Edited by Florian Coulmas

The Handbook of Phonetic Sciences, Second Edition
Edited by William J. Hardcastle and John Laver

The Handbook of Morphology
Edited by Andrew Spencer and Arnold Zwicky

The Handbook of Japanese Linguistics
Edited by Natsuko Tsujimura

The Handbook of Linguistics
Edited by Mark Aronoff and Janie Rees-Miller

The Handbook of Contemporary Syntactic Theory
Edited by Mark Baltin and Chris Collins

The Handbook of Discourse Analysis

Edited by Deborah Schiffrin, Deborah Tannen, and Heidi E. Hamilton

The Handbook of Language Variation and Change
Edited by J. K. Chambers, Peter Trudgill, and Natalie Schilling-Estes

The Handbook of Historical Linguistics
Edited by Brian D. Joseph and Richard D. Janda

The Handbook of Language and Gender
Edited by Janet Holmes and Miriam Meyerhoff

The Handbook of Second Language Acquisition
Edited by Catherine J. Doughty and Michael H. Long

The Handbook of Bilingualism and Multilingualism,Second Edition
Edited by Tej K. Bhatia and William C. Ritchie

The Handbook of Pragmatics
Edited by Laurence R. Horn and Gregory Ward

The Handbook of Applied Linguistics
Edited by Alan Davies and Catherine Elder

The Handbook of Speech Perception
Edited by David B. Pisoni and Robert E. Remez

The Handbook of the History of English
Edited by Ans van Kemenade and Bettelou Los

The Handbook of English Linguistics
Edited by Bas Aarts and April McMahon

The Handbook of World Englishes

Edited by Braj B. Kachru; Yamuna Kachru, and Cecil L. Nelson

The Handbook of Educational Linguistics
Edited by Bernard Spolsky and Francis M. Hult

The Handbook of Clinical Linguistics
Edited by Martin J. Ball, Michael R. Perkins, Nicole Müller, and Sara Howard

The Handbook of Pidgin and Creole Studies
Edited by Silvia Kouwenberg and John Victor Singler

The Handbook of Language Teaching
Edited by Michael H. Long and Catherine J. Doughty

The Handbook of Language Contact
Edited by Raymond Hickey

The Handbook of Language and Speech Disorders
Edited by Jack S. Damico, Nicole Müller, Martin J. Ball

The Handbook of Computational Linguistics and Natural Language Processing
Edited by Alexander Clark, Chris Fox, and Shalom Lappin

The Handbook of Language and Globalization
Edited by Nikolas Coupland

The Handbook of Hispanic Linguistics
Edited by Manuel Díaz-Campos

The Handbook of Language Socialization
Edited by Alessandro Duranti, Elinor Ochs, and Bambi B. Schieffelin

The Handbook of Intercultural Discourse and Communication
Edited by Christina Bratt Paulston, Scott F. Kiesling, and Elizabeth S. Rangel

The Handbook of Historical Sociolinguistics
Edited by Juan Manuel Hernández-Campoy and Juan Camilo Conde-Silvestre

The Handbook of Hispanic Linguistics
Edited by José Ignacio Hualde, Antxon Olarrea, and Erin O'Rourke

The Handbook of Conversation Analysis
Edited by Jack Sidnell and Tanya Stivers

The Handbook of English for Specific Purposes
Edited by Brian Paltridge and Sue Starfield

The Handbook of Computational Linguistics and Natural Language Processing

Edited by

Alexander Clark, Chris Fox, and Shalom Lappin

WILEY-BLACKWELL
A John Wiley & Sons, Ltd., Publication

This paperback edition first published 2013
© 2013 Blackwell Publishing Ltd except for editorial material and organization
© 2013 Alexander Clark, Chris Fox, and Shalom Lappin

Edition History: Blackwell Publishing Ltd (hardback, 2010)

Blackwell Publishing was acquired by John Wiley & Sons in February 2007. Blackwell's publishing program has been merged with Wiley's global Scientific, Technical, and Medical business to form Wiley-Blackwell.

Registered Office
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

Editorial Offices
350 Main Street, Malden, MA 02148-5020, USA
9600 Garsington Road, Oxford, OX4 2DQ, UK
The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

For details of our global editorial offices, for customer services, and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com/wiley-blackwell.

The right of Alexander Clark, Chris Fox, and Shalom Lappin to be identified as the authors of the editorial material in this work has been asserted in accordance with the UK Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by the UK Copyright, Designs and Patents Act 1988, without the prior permission of the publisher.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be

available in electronic books.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The publisher is not associated with any product or vendor mentioned in this book. This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

Library of Congress Cataloging-in-Publication Data

The handbook of computational linguistics and natural language processing/edited by Alexander Clark, Chris Fox, and Shalom Lappin.
p. cm. - (Blackwell handbooks in linguistics)
Includes bibliographical references and index.

ISBN 978-1-4051-5581-6 (hardcover: alk. paper) ISBN 978-1-118-34718-8 (paperback: alk. paper)
1. Computational linguistics. 2. Natural language processing (Computer science).
I. Clark, Alexander (Alexander Simon) II. Fox, Chris, 1965- III. Lappin, Shalom.
P98.H346 2010
410_.285-dc22

2010003116

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

Cover image: Theo van Doesburg, Composition IX, opus 18, 1917. Haags Gemeentemuseum,The Hague, Netherlands /

The Bridgeman Art Library.
Cover design byWorkhaus.

For Camilla

לאחיי דוד ודניאל, ולאחותי נעמי באהבה ובהומור

List of Figures

1.1 Chomsky's hierarchy of languages.
2.1 Architecture of a multi-tape Turing machine.
2.2 A derivation in the Lambek calculus.
2.3 Productions of a DCG recognizing the language $\{a^n b^n c^n d^n e^n \mid n \geq 0\}$.
2.4 Derivation of the string aabbccddee in the DCG of Figure 2.3.
2.5 Semantically annotated CFG generating the language of the syllogistic.
2.6 Meaning derivation in a semantically annotated CFG.
2.7 Productions for extending the syllogistic with transitive verbs.
3.1 Recursive linear interpolation.
3.2 ARPA format for language model representation.
3.3 Partial parse.
3.4 A word-and-parse k-prefix.
3.5 Complete parse.
3.6 Before an adjoin operation.
3.7 Result of adjoin-left under NTlabel.
3.8 Result of adjoin-right under NTlabel.
3.9 Language model operation as a finite state machine.
3.10 SLM operation.
3.11 One search extension cycle.
3.12 Binarization schemes.
3.13 Structured language model maximum depth distribution.
3.14 Comparison of PPL, WER, labeled recall/precision error.
4.1 The CKY recognition algorithm.
4.2 Table ***T*** obtained by the CKY algorithm.
4.3 The CKY recognition algorithm, expressed as a deduction system.
4.4 The Earley recognition algorithm.
4.5 Deduction system for Earley's algorithm.
4.6 Table ***T*** obtained by Earley's algorithm.
4.7 Parse forest associated with table ***T*** from Figure 4.2.
4.8 Knuth's generalization of Dijkstra's algorithm, applied to finding the most probable parse in a probabilistic context-free grammar ***G***.
4.9 The probabilistic CKY algorithm.

4.10 A parse of ‘our company is training workers,’ assuming a bilexical context-free grammar.
4.11 Deduction system for recognition with a 2-LCFG. We assume $w = a_1 \cdots a_n$, $a_{n+1} = \$$.
4.12 Illustration of the use of inference rules (f), (c), and (g) of bilexical recognition.
4.13 A projective dependency tree.
4.14 A non-projective dependency tree.
4.15 Deduction system for recognition with PDGs. We assume $w = a_1 \cdots a_n$, and disregard the recognition of $a_{n+1} = \$$.
4.16 Substitution (a) and adjunction (b) in a tree adjoining grammar.
4.17 The TAG bottom-up recognition algorithm, expressed as a deduction system.
4.18 A pair of trees associated with a derivation in a SCFG.
4.19 An algorithm for the left composition of a sentence w and a SCFG $\boldsymbol{G}$.
6.1 An example 2D space with six examples labeled white or black.
6.2 Two examples of the generation of a new hyper-rectangle in NGE.
6.3 An example of an induced rule in RISE, displayed on the right, with the set of examples that it covers (and from which it was generated) on the left.
6.4 An example of a family in a two-dimensional example space and ranked in the order of distance.
6.5 An example of family creation in Fambl.
6.6 Pseudo-code of the family extraction procedure in Fambl.
6.7 Generalization accuracies (in terms of percentage of correctly classified test instances) and F-scores, where appropriate, of MBL with increasing k parameter, and Fambl with $k = 1$ and increasing K parameter.
6.8 Compression rates (percentages) of families as opposed to the original number of examples, produced by Fambl at different maximal family sizes (represented by the x-axis, displayed at a log scale).
7.1 A simple decision tree for period disambiguation.
7.2 State of the decision tree after the expansion of the root node.
7.3 Decision tree learned from the example data.
7.4 Partitions of the two-dimensional feature subspace spanned by the features *‘color’* and *‘shape.’*
7.5 Data with overlapping classes and the class boundaries found by a decision tree.
7.6 Decision tree induced from the data in Figure 7.5 before and after pruning.
7.7 Decision tree with node numbers and information gain scores.

7.8 Decision tree with classification error counts.
7.9 Probabilistic decision tree induced from the data in Figure 7.5. 190
7.10 Part of a probabilistic decision tree for the nominative case of nouns. 194
9.1 A multi-layered perceptron.
9.2 Category probabilities estimated by an MLP
9.3 A recurrent MLP, specifically a simple recurrent network.
9.4 A recurrent MLP unfolded over the sequence.
9.5 The SSN architecture, unfolded over a derivation sequence, with derivation decisions *Dt* and hidden layers *St*.
9.6 An SSN unfolded over a constituency structure.
10.1 An example PTB tree.
10.2 A labeled dependency structure.
10.3 OntoNotes: a model for multi-layer annotation.
12.1 Waveform (top) and spectrogram (bottom) of conversational utterance *'no right I didn't mean to imply that.'*
12.2 HMM-based hierarchical modeling of speech.
12.3 Representation of an HMM as a parameterized stochastic finite state automaton (left) and in terms of probabilistic dependences between variables (right).
12.4 Forward recursion to estimate $\alpha_t(q_j) = p(\boldsymbol{x}_1,\ldots, \boldsymbol{x}_t, q_t = q_j \mid \text{A.})$.
12.5 Hidden Markov models for phonemes can be concatenated to form models for words.
12.6 Connected word recognition with a bigram language model.
12.7 Block processing diagram showing the AMI 2006 system for meeting transcription (Hain et al., 2006)
12.8 Word error rates (%) results in the NIST RT'06 evaluations of the AMI 2006 system on the evaluation test set, for the four decoding passes.
13.1 Example lexicalized parse-tree.
13.2 Example tree with complements distinguished from adjuncts.
13.3 Example tree containing a trace and the gap feature.
13.4 Example unlabeled dependency tree.
13.5 Generic algorithm for online learning taken from McDonald et al. (2005b).
13.6 The perceptron update.
13.7 Example derivation using forward and backward application.
13.8 Example derivation using type-raising and forward composition.
13.9 Example CCG derivation for the sentence *Under new features, participants can transfer money from the new funds.*
14.1 The two problems of word segmentation.

14.2 Word discovery from an MDL point of view.
14.3 A signature for two verbs in English.
14.4 Morphology discoveryaslocal descent.
14.5 BuildinganFST from two FSAs. 15.1 Derivation of semantic representation with storage.
16.1 Basic componentsofaspoken dialogue system.
16.2 Finite state machine for a simple ticket booking application.
16.3 Asimple frame.
16.4 Goal-oriented action schema.
16.5 A single utterance gives rise to distinct updates of the DGB for distinct participants.
17.1 Relative clause attachment ambiguity.
17.2 An example for the parse-trees generated by a probabilistic-context free grammar (PCFG) (adapted from Crocker & Keller 2006).
17.3 The architecture of the SynSem-Integration model, from Pado et al. (2009).
17.4 Asimple recurrent network.
17.5 CIANet: a network featuring scene–language interaction with a basic attentional gating mechanism to select relevant events in a scene with respecttoanunfolding utterance.
17.6 The competitive integration model (Spivey-Knowlton & Sedivy 1995).
18.1 Example dependency tree.
19.1 Asentence-aligned corpus.
19.2 Anon-exact alignment.
19.3 In the word-based translation on the left we see that the noun–adjective reordering into English is missed. On the right, the noun and adjective are translated as a single phrase and the correct ordering is modeled in the phrase-based translation.
19.4 Merging source-to-target and target-to-source alignments (from Koehn 2010).
19.5 All possible source segmentations with all possible target translations (from Koehn 2004).
19.6 Hypothesis expansion via stack decoding (from Koehn 2004).
19.7 An aligned tree pair in DOT for the sentence pair: *he chose the ink cartridge*, *il a choisi la cartouche d'encre*.
19.8 Compositionintree-DOT.
20.1 Human and corpus wind descriptions for September 19, 2000.
20.2 An example literacy screener question (SkillSum input).
20.3 Example text producedbySkillSum.

20.4 Example SumTime document plan.

20.5 Example SumTime deep syntactic structure. 21.1 Exampleofthe RST relation evidence.

22.1 BasicQAsystem architecture.

22.2 An ARDA scenario (from Small & Strzalkowski 2009).

22.3 An answer model for the question: *Where is Glasgow?* (Dalmas & Webber 2007), showing both Scotland and Britain as possible answers.

22.4 Example interaction taken from a live demonstration to the ARDA AQUAINT communityin2005.

22.5 Goal frame for the question: *What is the status of the Social Security system?*

22.6 Two cluster seed passages and their corresponding frames relative to the retirement clarification question.

22.7 Two cluster passages and their corresponding frames relative to the private accounts clarification question.

List of Tables

3.1 Headword percolation rules

3.2 Binarization rules

3.3 Parameter re-estimation results

3.4 Interpolation with trigram results

3.5 Maximum depth evolution during training

6.1 Examples generated for the letter–phoneme conversion task, from the word–phonemization pair *booking–[bukIN],* aligned as [b-ukI-N]

6.2 Number of extracted families at a maximum family size of 100, the average number of family members, and the raw memory compression, for four tasks

6.3 Two example families (represented by their members) extracted from the PP and CHUNK data sets respectively

7.1 Training data consisting of seven objects which are characterized by the features *'size,' 'color,'* and *'shape.'* The first four items belong to class '+,' the others to class '—'

8.1 Comparison of different tag sets on IPSM data

8.2 Cross-linguistic evaluation: 64 clusters, left all words, right $f < 5$

11.1 Structure of a typical summary of evaluation results

11.2 Contingency table for a document retrieval task

16.1 NSUs in a subcorpus of the BNC

16.2 Comparison of dialogue management approaches

17.1 Conditional probability of a verb frame given a particular verb, as estimated using the Penn Treebank

19.1 Number of fragments for English-to-French and French-to-English HomeCentre experiments

20.1 Numerical wind forecast for September 19, 2000

Notes on Contributors

Ciprian Chelba is a Research Scientist with Google. Between 2000 and 2006 he worked as a Researcher in the Speech Technology Group at Microsoft Research.

He received his Diploma Engineer degree in 1993 from the Faculty of Electronics and Telecommunications at "Politehnica" University, Bucuresti, Romania, M.S. in 1996 and PhD in 2000 from the Electrical and Computer Engineering Department at the Johns Hopkins University.

His research interests are in statistical modeling of natural language and speech, as well as related areas such as machine learning and information theory as applied to natural language problems.

Recent projects include language modeling for large-vocabulary speech recognition (discriminative model estimation, compact storage for large models), search in spoken document collections (spoken content indexing, ranking and snipeting), as well as speech and text classification.

Alexander Clark is an Honorary Research Fellow in the Department of Computer Science at Royal Holloway, University of London. His first degree was in Mathematics from the University of Cambridge, and his PhD is from the University of Sussex. He did postdoctoral research at the University of Geneva. In 2007 he was a *Professeur invité* at the University of Marseille. He is on the editorial board of the journal *Research on Language and Computation*, and a member of the steering committee of the International Colloquium on Grammatical Inference. His research is on unsupervised learning in computational linguistics, and in grammatical inference; he has won several prizes and competitions for his research. He has co-authored with

Shalom Lappin a book entitled *Linguistic Nativism and the Poverty of the Stimulus*, which is being published by Wiley-Blackwell in 2010.

Stephen Clark is a Senior Lecturer at the University of Cambridge Computer Laboratory where he is a member of the Natural Language and Information Processing Research Group. From 2004 to 2008 he was a University Lecturer at the Oxford University Computing Laboratory, and before that spent four years as a postdoctoral researcher at the University of Edinburgh's School of Informatics, working with Prof. Mark Steedman. He has a PhD in Artificial Intelligence from the University of Sussex and a first degree in Philosophy from the University of Cambridge. His main research interest is statistical parsing, with a focus on the grammar formalism combinatory categorial grammar. In 2009 he led a team at the Johns Hopkins University Summer Workshop working on "Large Scale Syntactic Processing: Parsing the Web." He is on the editorial boards of *Computational Linguistics* and the *Journal of Natural Language Engineering*, and is a Program Co-Chair for the 2010 Annual Meeting of the Association for Computational Linguistics.

Matthew W. Crocker obtained his PhD in Artificial Intelligence from the University of Edinburgh in 1992, where he subsequently held appointments as Lecturer in Artificial Intelligence and Cognitive Science and as an ESRC Research Fellow. In January 2000, Dr Crocker was appointed to a newly established Chair in Psycholinguistics, in the Department of Computational Linguistics at Saarland University, Germany. His current research brings together the experimental investigation of real-time human language processing and situated cognition in the development of computational cognitive models.

Matthew Crocker co-founded the annual conference on Architectures and Mechanisms for Language Processing

(AMLaP) in 1995. He is currently an associate editor for *Cognition*, on the editorial board of Springer's *Studies in Theoretical Psycholinguistics*, and has been a member of the editorial board for *Computational Linguistics*.

Walter Daelemans(MA, University of Leuven, Belgium, 1982; PhD, Computational Linguistics, University of Leuven, 1987) held research and teaching positions at the Radboud University Nijmegen, the AI-LAB at the University of Brussels, and Tilburg University, where he founded the ILK (Induction of Linguistic Knowledge) research group, and where he remained part-time Full Professor until 2006. Since 1999, he has been a Full Professor at the University of Antwerp (UA), teaching Computational Linguistics and Artificial Intelligence courses and co-directing the CLiPS research center. His current research interests are in machine learning of natural language, computational psycholinguistics, and text mining. He was elected fellow of ECCAI in 2003 and graduated 11 PhD students as supervisor.

Raquel Fernándezis a Postdoctoral Researcher at the Institute for Logic, Language and Computation, University of Amsterdam. She holds a PhD in Computer Science from King's College London for work on formal and computational modeling of dialogue and has published numerous peer-review articles on dialogue research. She has worked as Research Fellow in the Center for the Study of Language and Information (CSLI) at Stanford University and in the Linguistics Department at the University of Potsdam.

Dr **Chris Fox**is a Reader in the School of Computer Science and Electronic Engineering at the University of Essex. He started his research career as a Senior Research Officer in the Department of Language and Linguistics at the University of Essex. He subsequently worked in the Computer Science Department where he obtained his PhD in 1993. After that he spent a brief period as a Visiting

Researcher at Saarbruecken before becoming a Lecturer at Goldsmiths College, University of London, and then King's College London. He returned to Essex in 2003. At the time of writing, he is serving as Deputy Mayor of Wivenhoe.

Much of his research is in the area of logic and formal semantics, with a particular emphasis on issues of formal expressiveness, and proof-theoretic approaches to characterizing intuitions about natural language semantic phenomena.

Jonathan Ginzburg is a Senior Lecturer in the Department of Computer Science at King's College London. He has previously held posts in Edinburgh and Jerusalem. He is one of the managing editors of the journal *Dialogue and Discourse*. He has published widely on formal semantics and dialogue. His monograph *The Interactive Stance: Meaning for Conversation* was published in 2009.

John A. Goldsmith is Edward Carson Waller Distinguished Service Professor in the Departments of Linguistics and Computer Science at the University of Chicago, where he has been since 1984. He received his PhD in Linguistics in 1976 from MIT, and taught from 1976 to 1984 at Indiana University. His primary interests are computational learning of natural language, phonological theory, and the history of linguistics.

Ralph Grishman is Professor of Computer Science at New York University. He has been involved in research in natural language processing since 1969, and since 1985 has directed the Proteus Project, with funding from DARPA, NSF, and other government agencies. The Proteus Project has conducted research in natural language text analysis, with a focus on information extraction, and has been involved in the creation of a number of major lexical and syntactic resources, including Comlex, Nomlex, and NomBank. He is a past President of the Association for Computational

Linguistics and the author of the text *Computational Linguistics: An Introduction*.

Thomas Hain holds the degree Dipl.-Ing. with honors from the University of Technology, Vienna and a PhD from Cambridge University. In 1994 he joined Philips Speech Processing, which he left as Senior Technologist in 1997. He took up a position as Research Associate at the Speech, Vision and Robotics Group and Machine Intelligence Lab at the Cambridge University Engineering Department where he also received an appointment as Lecturer in 2001. In 2004 he joined the Department of Computer Science at the University of Sheffield where he is now a Senior Lecturer. Thomas Hain has a well established track record in automatic speech recognition, in particular involvement in best-performing ASR systems for participation in NIST evaluations. His main research interests are in speech recognition, speech and audio processing, machine learning, optimisation of large-scale statistical systems, and modeling of machine/machine interfaces. He is a member of the IEEE Speech and Language Technical Committee.

James B. Henderson is an MER (Research Professor) in the Department of Computer Science of the University of Geneva, where he is co-head of the interdisciplinary research group Computational Learning and Computational Linguistics. His research bridges the topics of machine learning methods for structure-prediction tasks and the modeling and exploitation of such tasks in NLP, particularly syntactic and semantic parsing. In machine learning his current interests focus on latent variable models inspired by neural networks. Previously, Dr Henderson was a Research Fellow in ICCS at the University of Edinburgh, and a Lecturer in CS at the University of Exeter, UK. Dr Henderson received his PhD and MSc from the University of Pennsylvania, and his BSc from the Massachusetts Institute of Technology, USA.