Lecture Notes in Computer Science 5070

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison
Lancaster University, UK

Takeo Kanade
Carnegie Mellon University, Pittsburgh, PA, USA

Josef Kittler
University of Surrey, Guildford, UK

Jon M. Kleinberg
Cornell University, Ithaca, NY, USA

Alfred Kobsa
University of California, Irvine, CA, USA

Friedemann Mattern
ETH Zurich, Switzerland

John C. Mitchell
Stanford University, CA, USA

Moni Naor
Weizmann Institute of Science, Rehovot, Israel

Oscar Nierstrasz
University of Bern, Switzerland

C. Pandu Rangan
Indian Institute of Technology, Madras, India

Bernhard Steffen
University of Dortmund, Germany

Madhu Sudan
Microsoft Research, Cambridge, MA, USA

Demetri Terzopoulos
University of California, Los Angeles, CA, USA

Doug Tygar
University of California, Berkeley, CA, USA

Gerhard Weikum
Max-Planck Institute of Computer Science, Saarbruecken, Germany

Małgorzata Marciniak
Agnieszka Mykowiecka (Eds.)

Aspects of Natural Language Processing

Essays Dedicated to Leonard Bolc on the Occasion of His 75th Birthday

Springer

Volume Editors

Małgorzata Marciniak
Agnieszka Mykowiecka
Institute of Computer Science
Polish Academy of Sciences
ul. J.K. Ordona 21, 01-237 Warsaw, Poland
E-mail: {Malgorzata.Marciniak, Agnieszka.Mykowiecka}@ipipan.waw.pl

The cover illustration, showing the Rosetta Stone, was retrieved from
`http://de.wikipedia.org/wiki/Datei:Rosetta_Stone_BW.jpeg`

Library of Congress Control Number: 2009935055

CR Subject Classification (1998): I.2, H.3-5, J.5, F.4

LNCS Sublibrary: SL 3 – Information Systems and Application, incl. Internet/Web and HCI

ISSN	0302-9743
ISBN-10	3-642-04734-3 Springer Berlin Heidelberg New York
ISBN-13	978-3-642-04734-3 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

springer.com

© Springer-Verlag Berlin Heidelberg 2009
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Markus Richter, Heidelberg
Printed on acid-free paper SPIN: 12766055 06/3180 5 4 3 2 1 0

Leonard Bolc

Preface

This book is dedicated to Professor Leonard Bolc on the occasion of his 75th birthday, and contains essays written by his friends, former students and colleagues to celebrate his scientific career. For many years Leonard Bolc has played an important role in the Polish computer science community. He was especially known for his clear vision in the development of artificial intelligence, inspiring research, organizational and editorial achievements.

The papers included in this volume, present research in the areas which Leonard Bolc and his colleagues investigated during his long scientific career, i.e., logic, automatic reasoning, natural language processing, and computer applications of natural language or human-like reasoning.

Part I of the book is devoted to logic—the domain which was one of the most explored by Leonard Bolc himself. The first paper in this section describes a new approach to making correct judgments keeping in mind real-world constraints. The second paper addresses the problem of qualitative interpretation of fuzzy-like paraconsistent reasoning with natural language words.

Part II contains papers focusing on different aspects of computational linguistics. This part begins with papers devoted to morphology. Two of them describe morphological aspects of the Polish and Uzbek languages, and the third paper discusses morphological properties of Polish multi-word proper names. The subsequent paper presents a new definition of Polish nominal phrases defined within the metamorphosis grammar formalism, which was first described in one of the numerous publications edited by Professor Bolc. The following four papers are devoted to automatic acquisition of linguistic knowledge: extraction of semantic relations, valence dictionaries, and semantic restrictions on verb arguments.

Part III comprises papers describing different applications in which natural language processing or automatic reasoning plays an important role. The first paper is devoted to speech technology and describes the issues of voice portal implementation. Speech analysis was a subject of interest to Professor Bolc in the mid-1970s, which is commemorated in the following paper. The next two papers are devoted to different aspects of information extraction. The first one presents real-time event extraction, whereas the next one describes the creation of a domain model for information extraction applications in the medical domain. The next topics addressed are methods of searching for exact and approximate patterns in texts, and decision making in the context of a support system that is capable of dynamically adjusting its prediction model. The last paper presents an environment that allows a group of professional analysts to work together on complex problems and exchange their results in active and tacit ways.

By preparing this volume we would like to express our gratitude to our mentor Professor Leonard Bolc, for the many years of his generous and friendly leadership. While preparing this volume we had the unique opportunity to experience the type of work to which Leonard Bolc devoted so much time and effort, and in

which he continues to be very active. We would like to thank all the authors for their positive response to the idea of creating the volume, and for their help in reviewing the papers. We would also like to thank Maciej Korzeniowski for help in proofreading the volume.

We would like to express our thanks for the scientific encouragement Professor Bolc gave us, on behalf of all members of the Linguistic Engineering Group in the Institute of Computer Science, which he organized and led for many years, his former students who were inspired by him to devote their life to scientific research, and also on behalf of those who decided to work in industry. Finally, our colleagues, Piotr Rychlik, Paweł Stefański, Anna Ostaszewska, Agata Wrzos-Kamińska, Jacek Wrzos-Kamiński, Marek Wójcik, Agata Sawicka-Wójcik, and Łukasz Dębowski, who worked in the Group in the space of nearly 25 years, join the authors of the papers included in this volume in wishing Leonard Bolc all the best!

August 2009

Małgorzata Marciniak
Agnieszka Mykowiecka

A Scientific Portrait of Professor Leonard Bolc

Leonard Bolc started his scientific career studying at the Philological Department of the Adam Mickiewicz University in Poznań. From 1961 to 1987 Leonard Bolc worked at the Warsaw University (UW) in the Institute of Computer Science from the position of assistant to associate professor. He received his PhD in applied linguistics from the Adam Mickiewicz University in 1964. In the same year he began his postdoctoral studies at the Institute of Mathematics of the University of Münster. He spent four years there, expanding his knowledge of mathematics and computer science under the supervision of Professor H. Hermes, whose mentor was one of the founders of modern logic, G. Frege. During his stay at Münster, Leonard Bolc attended the lectures of Noam Chomsky, who sparked his interest in computational linguistics, especially in formal grammars and natural language parsers. In 1969, he received his habilitation at the Adam Mickiewicz University in using a context-free grammar formalism for natural language text analysis. The next year was spent in Dresden working on digital processing of speech signals. When he returned to Warsaw, he was nominated the Head of Department of Information Analysis and Synthesis in the Institute of Computer Science UW where he worked on artificial intelligence, focusing on various aspects of natural language processing [2]. From 1987 to 2007 he worked at the Institute of Computer Science, Polish Academy of Sciences, where he organized the Man-Machine Communication Group, that later changed its name to the Linguistics Engineering Group.

Professor Bolc focused his research on logic and computational linguistics, with the aim of creating a computer system with a natural language user interface. To achieve this goal, it is necessary to solve various problems from many fields—speech technologies, language engineering, knowledge representation and reasoning. Professor Leonard Bolc himself explored some of them and inspired his students and colleagues to work on many others. He worked in the fields of formal grammar [8], semantic interpretation of texts [1], searching in large data sets [7], information retrieval [3], [5], non-classical logic [10, 11], uncertain and incomplete knowledge processing [6] and automatic deduction [4].

The wide range of Leonard Bolc's research was the result of his conviction that only by taking into account all aspects of natural language communication—speech understanding, syntactic, semantic and pragmatic analysis of utterances, and automatic deduction based on human reasoning—can we implement a system that processes data in natural language.

Professor Leonard Bolc was very active in the dissemination of many important scientific results. He has been on the editorial board of many journals published by Academic Press, Springer-Verlag, Kluwer Academic Publishing, and Blackwell Scientific Publications. Many important monographs concerning new achievements in various fields of computer science have been edited by him. He has been a co-founder of the series "Symbolic Computation" published by

Springer-Verlag. He has edited more than 40 volumes in world-famous publishing houses, and he continues to be a very active editor. Leonard Bolc is also a very well-known editor in Poland. More than 120 volumes have been published in various book series edited by him. On the basis of many of these publications, Polish scientists received their habilitation or were nominated professors. Over the last few years, he has been the editor-in-chief for many academic handbooks of the Polish-Japanese Institute of Information Technology Press.

During his long scientific career, first at the Warsaw University, then at the Institute of Computer Science PAS, he was a scientific mentor to many students and young scientists. He encouraged many young people to devote their life to research in various aspects of artificial intelligence. All papers in this volume are co-authored by scientists who worked with Professor Bolc at various stages of their scientific careers; eight authors worked in the Group he created.

Selected Publications

1. Bolc, L., Strzalkowski, T.: Transformation Of Natural Language Into Logical Formulas. COLING 1982, 29–36 (1982)
2. Bolc, L., Cichy, M., Różańska, L.: Przetwarzanie języka naturalnego. WNT Warszawa (1982)
3. Bolc, L., Kochut, K., Lesniewski, A., Strzalkowski, T.: Natural Language Information Retrieval System Dialog. EACL 1983, 196–203 (1983)
4. Bolc, L., Rychlik, P.: The Use of Modal Default Reasoning in a Medical Diagnostic System with Natural Language Interface, Proc. of the 6th ECAI, Pisa (1984)
5. Bolc, L., Kowalski, A., Kozlowska, M., Strzalkowski, T.: A Natural Language Information Retrieval System with Extensions Towards Fuzzy Reasoning. International Journal of Man-Machine Studies 23(4), 335–367 (1985)
6. Bolc, L., Borodziewicz, W., Wójcik, M.: Podstawy przetwarzania informacji niepewnej i niepełnej. PWN, Warszawa (1991)
7. Bolc, L., Cytowski, J.: Search Methods for Artificial Intelligence. Academic Press, London (1992)
8. Bolc, L., Mykowiecka, A.: Podstawy przetwarzania języka naturalnego. Wybrane metody formalnego zapisu składni, Akademicka Oficyna Wydawnicza (1992)
9. Bolc, L., Zaremba, J.: Wprowadzenie do uczenia się maszyn. PWN, Warszawa (1993)
10. Bolc, L., Borowik, P.: Many-Valued Logics. Vol. 1: Theoretical Foundations. Springer, Heidelberg (2000)
11. Bolc, L., Borowik, P.: Many-Valued Logics 2: Automated Reasoning and Practical Applications. Springer, Heidelberg (2003)

Springer Acknowledgements

It was a great pleasure for Springer to receive the proposal to publish a book to be dedicated to Professor Leonard Bolc on the occasion of his 75th birthday – after a quick round of evaluation of the proposal in discussion with the series editors of the Lecture Notes in Artificial Intelligence series, there was unanimous agreement to publish this book as part of the LNCS/LNAI Festschrift subline.

The extraordinary scientific achievements and merits of Professor Bolc during his scientific career, now spanning almost half a century, are dignified in an impressive way by all those contributing to this Festschrift. However, beyond being an outstanding scientist and organizer of the computer science community, in Poland as well as internationally, Professor Bolc was also a very prolific author and was successfully involved in editing several books and entire book series and reviewing for various journals. Starting in the 1970s, a substantial part of Professor Bolc's editorial activities resulted in publications by Springer-Verlag and we are very proud to enjoy his cooperation up to the present day.

Starting out in 1978 when acting as volume editor of LNCS 63, *Natural Language Communication with Computers*, Professor Bolc developed close contacts and excellent cooperation with several generations of staff of the Heidelberg-based computer science editorial team in Springer-Verlag. From Springer's point of view, the book series *Artificial Intelligence and Symbolic Computation* initiated by Professor Bolc was a very successful venue for the publication of monographs and edited collections of chapters on timely topics in artificial intelligence; also his own twin-monographs on Many-Valued Logics, coauthored by Piotr Borowik, are landmark publications for the logics community. Overall, in his editorial activities Professor Bolc patiently and considerately encouraged many scientists around the world to write down their consolidated results for monograph publications. Beyond publications in AI&SC and LNCS/LNAI, Professor Bolc brought to Springer quite a few remarkable standalone publications, including the bestseller *Genetic Algorithms + Data Structures = Evolution Programs* by Zbigniew Michalewicz.

As head of the Springer Heidelberg computer science editorial team, it is my privilege to be in close contact and cooperation with Professor Bolc up to the present – at occasional personal meetings and during periodic phone conversations he generously shares his scientific knowledge and community connections and many of the leads he has offered in these exchanges have been turned into successful projects.

On behalf of Springer, in particular on behalf of the present and former members of the computer science editorial team, I would like to express my sincere gratitude to Professor Bolc for the excellent cooperation he has offered over decades and which we hope he will offer for many more years. We wish Professor Bolc all the best for the years to come!

August 2009 Alfred Hofmann

Table of Contents

I Logic

Wisdom Technology: A Rough-Granular Approach 3
Andrzej Jankowski and Andrzej Skowron

Paraconsistent Reasoning with Words 43
Alicja S. Szalas and Andrzej Szałas

II Language

On the Root-Based Lexicon for Polish 61
Joanna Rabiega-Wiśniewska

Representation of Uzbek Morphology in Prolog 83
Gayrat Matlatipov and Zygmunt Vetulani

Inflection of Polish Multi-Word Proper Names with Morfeusz and Multiflex 111
Agata Savary, Joanna Rabiega-Wiśniewska, and Marcin Woliński

A New Formal Definition of Polish Nominal Phrases 143
Marek Świdziński and Marcin Woliński

Morphosyntactic Constraints in the Acquisition of Linguistic Knowledge for Polish 163
Maciej Piasecki and Adam Radziszewski

Towards the Automatic Acquisition of a Valence Dictionary for Polish ... 191
Adam Przepiórkowski

Semantic Annotation of Verb Arguments in Shallow Parsed Polish Sentences by Means of the EM Selection Algorithm 211
Elżbieta Hajnicz

Adjectives: Constructions vs. Valence 241
Anna Kupść

III Applications

User-Centered Design for a Voice Portal 273
Krzysztof Marasek, Łukasz Brocki, Danijel Koržinek, Krzysztof Szklanny, and Ryszard Gubrynowicz

Speech Understanding System SUSY—A New Version of the Speech Synthesis Program 295
Jerzy Cytowski

Exploring Curvature-Based Topic Development Analysis for Detecting Event Reporting Boundaries 311
Jakub Piskorski

Domain Model for Medical Information Extraction—The LightMedOnt Ontology 333
Agnieszka Mykowiecka and Malgorzata Marciniak

A Survey of Text Processing Tools for the Automatic Analysis of Molecular Sequences 359
Andrzej Polański, Rafał Pokrzywa, and Marek Kimmel

Intelligent Decision Support: A Fuzzy Stock Ranking System 379
Adam Ghandar, Zbigniew Michalewicz, and Ralf Zurbruegg

COLLANE: An Experiment in Computer-Mediated Tacit Collaboration .. 411
Tomek Strzalkowski, Sarah Taylor, Samira Shaikh, Ben-Ami Lipetz, Hilda Hardy, Nick Webb, Tony Cresswell, Ting Liu, Min Wu, Yu Zhan, and Song Chen

Author Index 451