

Cognitive Technologies

Editor-in-chief

Daniel Sonntag, DFKI, Saarbrücken, Saarland, Germany

The series Cognitive Technologies encompasses artificial intelligence and its subfields and related areas, such as natural-language processing and technologies, high-level computer vision, cognitive robotics, automated reasoning, multiagent systems, symbolic learning theories and practice, knowledge representation and the semantic web, and intelligent tutoring systems and AI and education. Cognitive science, including human and animal cognition and artificial life, is within the scope of this series, as is the integration of symbolic and subsymbolic computation. The series includes textbooks, monographs, coherent thematic state-of-the-art collections, and multiauthor anthologies.

Editorial Board

Franz Baader (Technische Universität Dresden—Knowledge Representation; Automated Deduction)

Jaime G. Carbonell (Carnegie Mellon University—Machine Learning; Language Technologies)

Anthony G. Cohn (University of Leeds—Automated Reasoning)

Luis Fariñas del Cerro (IRIT Toulouse—Logic; Reasoning)

James A. Hendler (Rensselaer Polytechnic Institute—Semantic Web; Knowledge Discovery)

Toru Ishida (Kyoto University—Multiagent Systems; Services Computing; Social Informatics; Design)

Anthony Jameson (DFKI Saarbrücken—Human-Artificial Intelligence)

Martin Kay (Stanford University—Computational Linguistics)

Robert A. Kowalski (Imperial College London—Logic Programming)

Sarit Kraus (Bar-Ilan University—Intelligent Agents; Multiagent Systems)

Luís C. Lamb (Federal University of Rio Grande do Sul (UFRGS)—Cognitive Reasoning)

Maurizio Lenzerini (Università di Roma, Sapienza—Information Integration; Information Systems)

John W. Lloyd (The Australian National University—Agents; Machine Learning)

Mark T. Maybury (MITRE Corp.—Security; Privacy)

Johanna D. Moore (University of Edinburgh—Multimodal Communication)

Bernhard Nebel (Universität Freiburg—Knowledge Representation; Reasoning)

Sharon Oviatt (Incaa Designs—Multimodal Multisensory Interaction)

Stuart J. Russell (University of California, Berkeley—Robust and Beneficial Artificial Intelligence; Machine Learning)

Erik Sandewall (Linköping University—Cognitive Robotics; Human-Robot Interaction)

Oliviero Stock (FBK-irst Trento—Intelligent User Interfaces)

Peter Stone (University of Texas at Austin—Machine Learning; Multiagent Systems; Robotics)

Wolfgang Wahlster (DFKI Saarbrücken—User Modeling; Mobile and Multimodal User Interfaces)

Michael Wooldridge (University of Oxford—Multiagent Systems; Logic, Knowledge Representation; Computational Complexity; Game Theory)

Feiyu Xu (Lenovo, Beijing—Language Technology)

Founding Editors

Dov M. Gabbay (King's College London)

Jörg Siekmann (DFKI, Saarbrücken)

More information about this series at <http://www.springer.com/series/5216>

Yohei Murakami · Donghui Lin
Toru Ishida
Editors

Services Computing for Language Resources

Editors

Yohei Murakami
Kyoto University
Kyoto
Japan

Toru Ishida
Kyoto University
Kyoto
Japan

Donghui Lin
Kyoto University
Kyoto
Japan

ISSN 1611-2482

ISSN 2197-6635 (electronic)

Cognitive Technologies

ISBN 978-981-10-7792-0

ISBN 978-981-10-7793-7 (eBook)

<https://doi.org/10.1007/978-981-10-7793-7>

Library of Congress Control Number: 2017964249

© Springer Nature Singapore Pte Ltd. 2018

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd. part of Springer Nature

The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

We started working on a service-oriented language infrastructure called the Language Grid in 2006. Our vision remains to shift from language resources to language services. The research goal is not to repackaging language resources but to interconnect them flexibly as Web services. Since language is used everywhere in our daily life in different ways, attempting to develop complete general translation packages will not overcome all language barriers: We need to encourage end users to create their own customized multilingual environments to cover the various situations they actually face. Our first Springer book titled “The Language Grid: service-oriented collective intelligence for language resource interoperability” in 2011 focused on our research goal, system architecture, and applications, and the final part was titled Towards Federation of Service Grids. This book summarizes the technologies developed in recent years to actually realize the federation of service grids as well as the designs of applications and analyses of future issues likely to arise in language service networks.

This book includes 13 chapters in four parts. The first part describes two types of language service platforms to interconnect language services across service grids. One platform federates heterogeneous as well as homogeneous language service platforms on the cloud. To share language service registries and compose language services across platforms, a federated grid architecture with upper-level ontology for a service grid has been proposed and illustrated with a case study of a federated operation involving the Language Grid. The other platform connects a language service platform on a user’s device to a public one on the cloud to efficiently construct personalized multilingual applications. This architecture allows users to combine language services regardless of the policy of the service grid operator.

The second part details various language service composition technologies that improve the reusability, efficiency, accuracy, and fault tolerance of composite services. One technology allows users to dynamically bind atomic language services to a language service workflow for reusing it. Another predicts and maximizes the parallel processing efficiency of a composite language service under parallel execution policies of atomic language services. A third optimizes a language service workflow consisting of crowd workers by using various performance models

of crowd workers. The last one analyzes the robustness of composite language services to cascading failure, a frequent cause of failure in dependent language services.

The third part reports research work and activities on creating language resources and services. One aspect addresses language service creation for low-resource languages. To provide an efficient, robust, and accurate bilingual lexicon creation method for low-resource languages, a constraint approach to pivot-based lexicon induction has been introduced, where a new bilingual lexicon of a closely related language pair is induced from two existing dictionaries using a distant language as a pivot. Another aspect is to create complicated language services suitable for real-world demands. To consider how to compose different language services to meet the requirements posed by various users, a user-centered service design approach has been proposed and illustrated with a field study of a multi-language communication service.

The fourth part provides various applications and tools for understanding and designing language services that well support intercultural collaboration. Researchers proposed a series of mechanisms and guidelines to address the issue of content inconsistency in knowledge sharing among multi-language communities. In parallel, controlled experiments have been conducted to understand how various language services can effectively support the listening comprehension of non-native speakers, which is useful for designing real-time adaptive systems with greater listening comprehension. To support the users in multilingual communication, the translation agent metaphor is proposed as a novel interactive way to promote the efficiency of communication and reduce the communication breaks caused by translation errors. Moreover, a multiagent-based gaming simulation environment was developed to understand the strategies of various stakeholders in the language service infrastructure. Furthermore, an intercultural communication environment was developed by an international NPO to support agricultural knowledge transfer in rural areas in Vietnam.

The Language Grid has been contributing to the research into language resources for ten years as a concrete example of an effective language service infrastructure and its applications. In June 2017, we launched the Language Grid Association, a nonprofit organization, as a legal entity to continue to support the research and development of language services. We are grateful to the many researchers, students, and field-workers who have collaborated with, and supported our project.

Kyoto, Japan
August 2017

Yohei Murakami
Donghui Lin
Toru Ishida

Contents

Part I Language Service Platform

| | |
|--|----|
| Federated Grid Architecture for Language Services | 3 |
| Yohei Murakami, Takao Nakaguchi, Donghui Lin and Toru Ishida | |
| Language Mashup: Personalized Language Service Platform | 21 |
| Masayuki Otani, Nguyen Cao Hong Ngoc, Takao Nakaguchi and Donghui Lin | |

Part II Language Service Composition

| | |
|---|----|
| Language Service Composition Based on Higher Order Functions | 41 |
| Takao Nakaguchi, Yohei Murakami, Donghui Lin and Toru Ishida | |
| Policy-Aware Language Service Composition | 57 |
| Trang Mai Xuan, Yohei Murakami and Toru Ishida | |
| Optimizing Crowdsourcing Workflow for Language Services | 75 |
| Shinsuke Goto, Toru Ishida and Donghui Lin | |
| Cascading-Failure Tolerance for Language Service Networks | 91 |
| Kemas M. Lhaksana, Toru Ishida and Yohei Murakami | |

Part III Language Resources and Services Creation

| | |
|--|-----|
| A Constraint Approach to Lexicon Induction for Low-Resource Languages | 109 |
| Mairidan Wushouer, Donghui Lin, Toru Ishida and Yohei Murakami | |
| Language Service Design Based on User-Centered QoS | 125 |
| Donghui Lin, Toru Ishida and Yohei Murakami | |

Part IV Understanding and Designing Language Services

| | |
|---|-----|
| Consistency Analysis in Multi-language Knowledge Sharing System | 141 |
| Amit Pariyar, Yohei Murakami, Donghui Lin and Toru Ishida | |
| Supporting Non-native Speakers' Listening Comprehension with Automated Transcripts | 157 |
| Xun Cao, Naomi Yamashita and Toru Ishida | |
| Translation Agent | 175 |
| Chunqi Shi, Toru Ishida and Donghui Lin | |
| Gaming for Language Services | 193 |
| Yuu Nakajima, Ryutaro Otsuka, Reiko Hishiyama, Takao Nakaguchi and Naoyuki Oda | |
| Youth Mediated Communication: Knowledge Transfer as Intercultural Communication | 209 |
| Toshiyuki Takasaki, Yumiko Mori, Toru Ishida and Masayuki Otani | |
| Author Index | 225 |

Contributors

Xun Cao Department of Social Informatics, Kyoto University, Kyoto, Japan

Shinsuke Goto Department of Social Informatics, Kyoto University, Kyoto, Japan

Reiko Hishiyama Faculty of Science and Engineering, Waseda University, Tokyo, Japan

Toru Ishida Department of Social Informatics, Kyoto University, Kyoto, Japan

Kemas M. Lhaksmana School of Computing, Telkom University, Bandung, Indonesia

Donghui Lin Department of Social Informatics, Kyoto University, Kyoto, Japan

Trang Mai Xuan Department of Social Informatics, Kyoto University, Kyoto, Japan

Yumiko Mori NPO Pangaea, Kyoto, Japan

Yohei Murakami Unit of Design, Kyoto University, Kyoto, Japan

Takao Nakaguchi Department of Web Business Technology, School of Applied Information Technology, The Kyoto College of Graduate Studies for Informatics, Kyoto, Japan; Department of Social Informatics, Kyoto University, Kyoto, Japan

Yuu Nakajima Faculty of Science, Toho University, Chiba, Japan

Nguyen Cao Hong Ngoc Department of Social Informatics, Kyoto University, Kyoto, Japan

Naoyuki Oda Faculty of Science, Toho University, Chiba, Japan

Masayuki Otani Department of Informatics, Faculty of Science and Engineering, Kindai University, Osaka, Japan

Ryutaro Otsuka Faculty of Science and Engineering, Waseda University, Tokyo, Japan

Amit Pariyar Institute of Social Informatics and Technological Innovations, University Malaysia Sarawak, Sarawak, Malaysia

Chunqi Shi Department of Social Informatics, Kyoto University, Kyoto, Japan

Toshiyuki Takasaki NPO Pangaea, Kyoto, Japan

Mairidan Wushouer Department of Computer Science, Xinjiang University, Urumqi, Xinjiang Uyghur Autonomous Region, People's Republic of China

Naomi Yamashita NTT Communication Science Labs, Kyoto, Japan