Recognizing Textual Entailment

Models and Applications

Synthesis Lectures on Human Language Technologies

Editor
Graeme Hirst, ***University of Toronto***

Synthesis Lectures on Human Language Technologies is edited by Graeme Hirst of the University of Toronto. The series consists of 50- to 150-page monographs on topics relating to natural language processing, computational linguistics, information retrieval, and spoken language understanding. Emphasis is on important new techniques, on new applications, and on topics that combine two or more HLT subfields.

Recognizing Textual Entailment: Models and Applications
Ido Dagan, Dan Roth, Mark Sammons, and Fabio Massimo Zanzotto
2013

Web Corpus Construction
Roland Schäfer and Felix Bildhauer
2013

Semi-Supervised Learning and Domain Adaptation in Natural Language Processing
Anders Søgaard
2013

Linguistic Fundamentals for Natural Language Processing: 100 Essentials from Morphology and Syntax
Emily M. Bender
2013

Semantic Relations Between Nominals
Vivi Nastase, Preslav Nakov, Diarmuid Ó Séaghdha, and Stan Szpakowicz
2013

Computational Modeling of Narrative
Inderjeet Mani
2012

Natural Language Processing for Historical Texts
Michael Piotrowski
2012

Sentiment Analysis and Opinion Mining
Bing Liu
2012

Discourse Processing
Manfred Stede
2011

Bitext Alignment
Jörg Tiedemann
2011

Linguistic Structure Prediction
Noah A. Smith
2011

Learning to Rank for Information Retrieval and Natural Language Processing
Hang Li
2011

Computational Modeling of Human Language Acquisition
Afra Alishahi
2010

Introduction to Arabic Natural Language Processing
Nizar Y. Habash
2010

Cross-Language Information Retrieval
Jian-Yun Nie
2010

Automated Grammatical Error Detection for Language Learners
Claudia Leacock, Martin Chodorow, Michael Gamon, and Joel Tetreault
2010

Data-Intensive Text Processing with MapReduce
Jimmy Lin and Chris Dyer
2010

Semantic Role Labeling
Martha Palmer, Daniel Gildea, and Nianwen Xue
2010

Spoken Dialogue Systems
Kristiina Jokinen and Michael McTear
2009

Introduction to Chinese Natural Language Processing
Kam-Fai Wong, Wenjie Li, Ruifeng Xu, and Zheng-sheng Zhang
2009

Introduction to Linguistic Annotation and Text Analytics
Graham Wilcock
2009

Dependency Parsing
Sandra Kübler, Ryan McDonald, and Joakim Nivre
2009

Statistical Language Models for Information Retrieval
ChengXiang Zhai
2008

Recognizing Textual Entailment: Models and Applications

Ido Dagan, Dan Roth, Mark Sammons, and Fabio Massimo Zanzotto

ISBN: 978-3-031-01023-1 paperback
ISBN: 978-3-031-02151-0 ebook

DOI 10.1007/978-3-031-02151-0

A Publication in the Springer series
SYNTHESIS LECTURES ON HUMAN LANGUAGE TECHNOLOGIES

Lecture #23
Series Editor: Graeme Hirst, *University of Toronto*
Series ISSN
Synthesis Lectures on Human Language Technologies
Print 1947-4040 Electronic 1947-4059

Recognizing Textual Entailment

Models and Applications

Ido Dagan
Bar-Ilan University, Israel

Dan Roth
University of Illinois, Urbana, IL

Mark Sammons
University of Illinois, Urbana, IL

Fabio Massimo Zanzotto
University of Rome "Tor Vergata," Italy

SYNTHESIS LECTURES ON HUMAN LANGUAGE TECHNOLOGIES #23

ABSTRACT

In the last few years, a number of NLP researchers have developed and participated in the task of Recognizing Textual Entailment (RTE). This task encapsulates Natural Language Understanding capabilities within a very simple interface: recognizing when the meaning of a text snippet is contained in the meaning of a second piece of text. This simple abstraction of an exceedingly complex problem has broad appeal partly because it can be conceived also as a component in other NLP applications, from Machine Translation to Semantic Search to Information Extraction. It also avoids commitment to any specific meaning representation and reasoning framework, broadening its appeal within the research community. This level of abstraction also facilitates evaluation, a crucial component of any technological advancement program. This book explains the RTE task formulation adopted by the NLP research community, and gives a clear overview of research in this area. It draws out commonalities in this research, detailing the intuitions behind dominant approaches and their theoretical underpinnings. This book has been written with a wide audience in mind, but is intended to inform all readers about the state of the art in this fascinating field, to give a clear understanding of the principles underlying RTE research to date, and to highlight the short- and long-term research goals that will advance this technology.

KEYWORDS

natural language processing, textual entailment, textual inference, knowledge acquisition, machine learning

Contents

List of Figures xiii

List of Tables xv

Preface xvii

Acknowledgments xix

1 Textual Entailment **1**

1.1 Motivation and Rationale 1
1.2 The Recognizing Textual Entailment Task 3
1.2.1 The Scope of Textual Entailment 3
1.2.2 The Role of Background Knowledge 6
1.2.3 Textual Entailment versus Linguistic Notion of Entailment 7
1.2.4 Extending Entailment Recognition with Contradiction Detection 8
1.2.5 The Challenge and Opportunity of RTE 9
1.3 Applications of Textual Entailment Solutions 10
1.3.1 Question Answering 10
1.3.2 Relation Extraction 12
1.3.3 Text Summarization 13
1.3.4 Additional Applications 14
1.4 Textual Entailment Evaluation 16
1.4.1 RTE-1 through RTE-5 17
1.4.2 RTE-6 and RTE-7 19
1.4.3 Other Evaluations of Textual Entailment Technology 21
1.4.4 Future Directions for Entailment Evaluation 22

2 Architectures and Approaches **25**

2.1 An Intuitive Model for RTE 25
2.2 Levels of Representation in RTE Systems 27
2.2.1 Lexical-level RTE 28
2.2.2 Structured Representations for RTE 29

2.3 Inference in RTE systems 33
2.3.1 Similarity-based approaches 34
2.3.2 Alignment-focused Approaches 35
2.3.3 "Proof Theoretic" RTE 35
2.3.4 Hybrid Approaches 43
2.4 A conceptual architecture for RTE systems 43
2.4.1 Preprocessing 44
2.4.2 Enrichment 44
2.4.3 Candidate Alignment Generation 46
2.4.4 Alignment Selection 46
2.4.5 Classification 47
2.4.6 Main Decision-making Approaches 47
2.5 Emergent Challenges 47
2.5.1 Knowledge Acquisition Bottleneck: Acquiring Rules 48
2.5.2 Noise-tolerant RTE architectures 48

3 Alignment, Classification, and Learning 51

3.1 An Abstract Scheme for Textual Entailment Decisions 52
3.2 Generating Canditates and Selecting Alignments 54
3.2.1 Anchors: Linking Texts and Hypotheses 54
3.2.2 Formalizing Candidate Alignment Generation and Alignment 56
3.3 Classifiers, Feature Spaces, and Machine Learning 58
3.4 Similarity feature spaces 59
3.4.1 Token-level Similarity Features 59
3.4.2 Structured Similarity Features 61
3.4.3 Entailment Trigger Feature Spaces 67
3.4.4 Rewrite Rule Feature Spaces 69
3.4.5 Discussion 75
3.5 Learning Alignment Functions 76
3.5.1 Learning Alignment from Gold-standard Data 76
3.5.2 Learning Entailment with a Latent Alignment 78

4 Case Studies 81

4.1 Edit Distance-based RTE 81
4.1.1 Open Source Tree Edit-based RTE System 82
4.1.2 Tree Edit Distance with Expanded Edit Types 82
4.2 Logical Representation and Inference 83

4.2.1 Representation . . . 84
4.2.2 Logical Inference with Abduction . . . 86
4.2.3 Logical Inference with Shallow Backoff System . . . 88
4.3 Transformation-based Approaches . . . 90
4.3.1 Transformation-based Approach with Integer Linear Programming . . . 90
4.3.2 Syntactic Transformation with linguistically motivated rules . . . 92
4.3.3 Syntactic Transformation with a Probabilistic Calculus . . . 94
4.3.4 Syntactic Transformation with Learned Operation Costs . . . 96
4.3.5 Natural Logic . . . 96
4.4 Alignment-focused Approaches . . . 100
4.4.1 Learning Alignment Selection Independently of Entailment . . . 100
4.4.2 Hand-Coded Alignment Function . . . 104
4.4.3 Leveraging Multiple Alignments for RTE . . . 105
4.4.4 Aligning Discourse Commitments . . . 107
4.4.5 Latent Alignment Inference for RTE . . . 110
4.5 Paired Similarity Approaches . . . 112
4.6 Ensemble Systems . . . 115
4.6.1 Weighted Expert Approach . . . 115
4.6.2 Selective Expert Approach . . . 117
4.7 Discussion . . . 119

5 Knowledge Acquisition for Textual Entailment . . . 123

5.1 Scope of Target Knowledge . . . 123
5.2 Acquisition from Manually Constructed Knowledge Resources . . . 125
5.2.1 Mining computation-oriented knowledge resources . . . 125
5.2.2 Mining human-oriented knowledge resources . . . 129
5.3 Corpus-based Knowledge Acquisition . . . 130
5.3.1 Distributional Similarity Methods . . . 131
5.3.2 Co-occurrence-based Methods . . . 139
5.3.3 Acquisition from Parallel and Comparable Corpora . . . 142
5.4 Integrating Multiple Sources of Evidence . . . 144
5.4.1 Integrating Multiple Information Sources . . . 144
5.4.2 Simultaneous Global Learning of Multiple Rules . . . 146
5.5 Context Sensitivity of Entailment Rules . . . 148
5.6 Concluding Remarks and Future Directions . . . 151

6 Research Directions in RTE **157**

6.1 Development of better/more flexible preprocessing tool chain 158

6.2 Knowledge Acquisition and Specification 159

6.3 Open Source Platform for Textual Entailment 160

6.4 Task Elaboration and Phenomenon-specific RTE Resources 161

6.5 Learning and Inference: efficient, scalable algorithms 162

6.6 Conclusion 163

A Entailment Phenomena **165**

Bibliography **171**

Authors' Biographies **199**

List of Figures

1.1 Representative RTE examples, including contradiction. 8

2.1 Sample entailment pair with notional meaning representation. 26

2.2 Sample non-entailing entailment pair with notional meaning representation. . . 27

2.3 Sample entailing entailment pair with pure lexical meaning representation. . . . 28

2.4 Sample non-entailing entailment pair with pure lexical meaning representation. 29

2.5 Sample entailing entailment pair with lexical alignment. 30

2.6 Sample entailing entailment pair with syntactic parse annotation. 31

2.7 Sample entailing entailment pair with semantic role annotation. 32

2.8 Entailing entailment pair with rephrased Hypothesis, with syntactic parse annotation. 33

2.9 Example of the proof process using Modus Ponens. 37

2.10 Example of the proof process using Refutation. 38

2.11 Block diagram for an RTE system based around a (refutation resolution-based) Theorem Prover. 39

2.12 Block diagram for generic RTE framework, with Machine Learning applied to different aspects of the entailment process. 45

3.1 An alignment of an entailment pair, where edges represent "similar" or "compatible" terms. 54

3.2 Candidate Alignment Generation: linking similar edges for an entailment pair. 55

3.3 A syntactic dependency analysis of an entailment pair with anchors. 63

3.4 A constituency syntactic analysis of a sentence pair. 65

3.5 A sample hand-annotated lexical alignment for an entailment pair. 77

4.1 EDITS workflow. 84

4.2 COGEX syntax-based logical form; Raina et al. uses a similar representation. . 86

4.3 A schematic view of Bos and Markert's Discourse Representation Structure representation for Textual Entailment pairs. 86

4.4 BLUE's logical form. 87

4.5 BLUE system architecture. 88

4.6 Representation of text used by Braz et al. 93

4.7 An example of a syntactic transformation in Bar-Haim et al.'s system. 94

4.8 The entailment relations of Natural Logic. 98

4.9 The behavior of semantic functions in Natural Logic. 98

4.10 Interactions between negation and NatLog's entailment relations. 100

4.11 An example of projection of polarity resulting from an edit to a semantic composition tree. 100

4.12 Composition calculus for Natural Logic entailment relations. 101

4.13 Block diagram representation of Iftene and Moruz's textual entailment system. 104

4.14 Example of discourse commitments. 108

4.15 Hickl and Bensley's textual entailment framework. 109

4.16 Tree Skeleton representation of an entailment pair. 115

4.17 COGEX syntax-based logical form augmented with semantic role and temporal predicates, as used in RTE-2 and RTE-3. 117

4.18 Wang and Neumann's divide-and-conquer system architecture. 118

List of Tables

1.1 Examples of Text-Hypothesis pairs from from the RTE-2 development set, with Task and Judgment for each. 5

4.1 Characteristics of selected edit distance-based RTE systems. 85
4.2 Characteristics of selected logic-based RTE systems with abduction. 89
4.3 Characteristics of selected hybrid systems with logic-based components. 91
4.4 Characteristics of selected transformation-based systems. 95
4.5 Characteristics of selected transformation-based systems. 97
4.6 Characteristics of the NatLog system. 101
4.7 Characteristics of selected alignment-focused systems. 103
4.8 Characteristics of selected alignment-focused systems. 106
4.9 Characteristics of selected alignment-focused systems. 112
4.10 Characteristics of selected pairwise similarity systems. 114
4.11 Characteristics of selected pairwise similarity systems. 116
4.12 Characteristics of selected ensemble systems. 119

5.1 The five pattern groups suggested by Chklovsky and Pantel and an example pattern for each group 141

A.1 Comparison phenomena 166
A.2 Ellipsis phenomena 167
A.3 Interpretation phenomena part 1 168
A.4 Interpretation phenomena part 2 169

Preface

In the last few years, a number of NLP researchers have developed and participated in the task of Recognizing Textual Entailment (RTE). This task encapsulates Natural Language Understanding capabilities within a very simple interface: recognizing when the meaning of a text snippet is contained in the meaning of a second piece of text. This simple abstraction of an exceedingly complex problem has broad appeal partly because it can be conceived also as a component in other NLP applications, from Machine Translation to Semantic Search to Information Extraction. It also avoids commitment to any specific meaning representation and reasoning framework, broadening its appeal within the research community. This level of abstraction also facilitates evaluation, a crucial component of any technological advancement program. This book explains the RTE task formulation adopted by the NLP research community and gives an overview of research in this area. It draws out commonalities in this research, detailing the intuitions behind dominant approaches and their theoretical underpinnings.

Chapter 1 provides the context for textual entailment research. It describes the motivation and rationale for the entailment recognition task and specifies its scope and positioning. The chapter further reviews the utility of entailment recognition in various NLP applications and describes in detail the evaluation methodologies used for RTE. Chapter 2 focuses on the intuitive model underlying RTE systems developed to date, and describes a generic architecture to which those systems conform. It describes limitations of the naive approach and motivates the more detailed discussion that follows. Chapter 3 describes Machine Learning techniques applied to the RTE task, and fleshes out the theoretical basis of the various models developed for RTE. Chapter 4 surveys some specific RTE systems, describing them in terms of the framework developed in the preceding chapters in order to facilitate comparison. Chapter 5 addresses the problem of the "knowledge acquisition bottleneck," i.e., the problem of acquiring the background knowledge needed for broad textual inference, surveying research in this area. Chapter 6 concludes the book with a short exploration of open research questions in this area.

This book has been written with a wide audience in mind, but is intended to inform all readers about the state of the art in this fascinating field, to give a clear understanding of the principles underlying RTE research to date, and to highlight the short- and long-term research goals that will advance this technology.

Ido Dagan, Dan Roth, Mark Sammons, and Fabio Massimo Zanzotto
May 2013

Acknowledgments

The authors are indebted to the publisher, Mike Morgan, and editor, Graeme Hirst, for their hard work in coordinating this project, and for their patience. We thank the three anonymous reviewers for their thorough reviews and insightful suggestions that helped us significantly improve the book. We thank Jonathan Berant and Eyal Shnarch for contributing input to the chapter on Knowledge Acquisition. We are grateful to the colleagues who were intensively involved in establishing and organizing the Recognizing Textual Entailment (RTE) challenges, including Oren Glickman, Bernardo Magnini, Luisa Bentivogli, Danilo Giampiccolo, Hoa Trang Dang, Lucy Vanderwende, Bill Dolan, and Peter Clark, as well as the PASCAL Network of Excellence, NIST, and CELCT for funding the challenges and Michele Sebag from PASCAL for her continuous support. We also thank students and colleagues whose joint work contributed ideas and perspectives that found their way into this book: Lorenzo Dell'Arciprete, Maayan Zhitomirsky-Geffet, Idan Szpektor, Roy Bar-Haim, Shachar Mirkin, Jonathan Berant, Eyal Shnarch, Asher Stern, Lili Kotlerman, Jacob Goldberger, Sebastian Pado, Marco Pennacchiotti, Alessandro Moschitti, V.G. Vinod Vydiswaran, Vasin Punyakanok, Rodrigo de Salvo Braz, Vivek Srikumar, Ming-Wei Chang, Dan Goldwasser, Roxana Girju, and other members of the Cognitive Computation Group at Illinois that helped this effort directly or indirectly. We thank the substantial support of the EXCITEMENT project to RTE research and infrastructure, under the European Community's Seventh Framework Programme (FP7/2007-2013), grant agreement no. 287923, and to other research sponsors that have funded our RTE-related work: Boeing; the U.S. government via ARDA, DHS, the DARPA Machine Reading program, and the DARPA Deep Exploration and Filtering of Text program; the Israel Science Foundation; the Israeli Ministry of Science, Technology and Space; and the collaboration project of FBK-irst, University of Haifa, and Bar-Ilan University.

Ido thanks Taly, his wife, for endless patience and support along the years, without which his Textual Entailment endeavor would not have been possible, as well as Yuval, Ori, and Inbal for accommodating the long hours spent on this book. He also thanks his students at the Natural Language Processing Lab for their dedication, and the great atmosphere they created along the joint effort.

Dan thanks his wife, Michal, for her love and support, and Noam, Edo, and Ella for their love and for reminding him what really matters. He also thanks his students and colleagues in the Cognitive Computation Group for their hard work over the years and for making it all possible.

Mark thanks Lauren, Alex, and Teddy for their enduring support and understanding, and his colleagues and research partners in the Cognitive Computation Group for their hard work on numerous RTE-related projects.

Fabio thanks Emanuela Bizzarri for her supportive help during these years used to develop, collect, and write the ideas in this book.

Ido Dagan, Dan Roth, Mark Sammons, and Fabio Massimo Zanzotto
May 2013