Automated Grammatical Error Detection for Language Learners, Second Edition

Claudia Leacock[1], Martin Chodorow[2], Michael Gamon[3], and Joel Tetreault[4]
([1]CTB McGraw-Hill; [2]Hunter College and the Graduate Center, City University of New York; [3]Microsoft Research; [4]Yahoo! Labs)

Morgan & Claypool (Synthesis Lectures on Human Language Technologies, edited by Graeme Hirst, volume 25), 2014, xv+154pp; paperbound, ISBN 978-1-62705-013-5

Reviewed by
Xiaofei Lu
Pennsylvania State University

Research in automated grammatical error detection and correction has gained considerable momentum in the past few years. Although much progress has been made in this area, numerous challenges and opportunities exist for further research. For NLP researchers and students who are considering following or pursuing work in this area and for English language teaching practitioners and researchers who are interested in utilizing systems resulting from research in this area, this book by Leacock, Chodorow, Gamon, and Tetreault provides a timely, updated review of the state of the art in automatic learner error detection and correction.

The introductory chapter first highlights the growth in prominence of the field of grammatical error detection and correction since the publication of the first edition of the volume. It then orients the reader to the book by detailing the changes from the first edition, providing a working definition of grammatical error, justifying the book's focus on English language learners, clarifying the use of the term "English language learner" to refer to learners of English as either a second or foreign language, putting forth a few claims about the relationship between NLP and Computer-Assisted Language Learning (CALL), specifying the intended audience of the book, and outlining the structure of the book.

Chapter 2 provides the historical context for automated grammatical error detection and correction. The chapter first discusses how varying degrees of tolerance of grammatical errors are achieved in computational grammars used in grammar-checking and proofreading tools. It then offers a quick overview of data-driven and hybrid approaches to error detection.

Chapter 3 summarizes the types of errors made by English language learners as reported in previous empirical and corpus-based learner error studies, touches upon the influence of L1 on learner English, and then delves into a detailed discussion of three specific problem areas for English language learners: prepositions, articles, and collocations.

Chapter 4 focuses on the evaluation of error detection systems. The chapter first introduces traditional evaluation measures (e.g., precision, recall, F-score, accuracy, and kappa) and the evaluation measures adopted in recent shared tasks on grammatical error correction. It then discusses evaluation using a corpus of correct usage, questioning the value of this approach for indicating how well the system may perform on actual learner data. The advantages and challenges of evaluating system performance on

doi:10.1162/COLI_r_00211

© 2015 Association for Computational Linguistics

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/COLI_r_00211 by guest on 29 March 2024

learner writing are then examined. The authors advocate the use of multiple annotators and crowdsourcing as a means to improve the reliability of manual evaluation. The chapter concludes with a discussion of how statistical significance testing of differences in system performance may be performed and provides a checklist for consistent reporting of system results.

Chapter 5 zooms in on data-driven approaches to detecting and correcting article and preposition errors. The chapter first looks at four types of information used by different systems, including lexical, syntactic, semantic, and source information. It then describes three prevailing types of data used to train statistical models of grammatical error correction, namely, well-formed text, artificial errors, and error-annotated learner corpora. Next, it discusses classification methods and *n*-gram statistics-related methods for error detection and correction. Finally, the chapter presents two end-to-end systems, Criterion and MSR ESL Assistant.

Chapter 6 focuses on collocation errors. Following a brief discussion of the properties of collocations, the chapter introduces a range of metrics commonly used to measure the association strength between pairs of words and reviews a number of recent collocation error detection and correction systems.

Chapter 7 moves beyond article, preposition, and collocation errors and examines rule-based and statistical methods for detecting and correcting verb-form, spelling, and punctuation errors and for identifying ungrammatical sentences. A small number of error detection systems for learners of languages other than English are also discussed.

Chapter 8 covers issues with learner error annotation, describes examples of comprehensive and targeted annotation schemes, and proposes three methods for improving the efficiency of large-scale annotation: sampling, crowdsourcing, and mining online revision logs.

Chapter 9 highlights several exciting emerging directions in the field of automated grammatical error correction. These include three recent shared tasks on grammatical error correction, the use of machine translation techniques for grammatical error correction, real-time crowdsourcing of grammatical error correction, and longitudinal studies on the efficacy of automated error correction systems for improving the writing of users of such systems.

Chapter 10 concludes the book with several suggestions for future research in the field, including annotation for evaluation, error detection for underrepresented languages, research on understudied error types, L1-specific error detection modules, collaboration with second language learning and education groups, and applications of grammatical error correction. The book also includes an Appendix that contains a list of textual learner corpora with at least some publicly accessible URLs or references.

I very strongly recommend this book to all NLP students and researchers who are interested in learning about, following, or pursuing research in automated grammatical error detection and correction. Not only does it offer a comprehensive systematic review of research in this field, but it also highlights real challenges and opportunities for fruitful future research. Practitioners and researchers in the language teaching and CALL community can also use this book to obtain a realistic understanding of the state of the art of the field. They will also most certainly welcome the volume's repeated emphasis on the importance for the NLP community to join efforts with the language teaching and CALL community to assess the effect of automated grammatical correction systems on improving the quality of language learners' writing.

I have just two minor quibbles with this book. First, it seems that the title may better reflect the content of the book and the goals of the research field in question with the words "and correction" added after "detection." Second, in both Chapter 3

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/COLI_r_00211 by guest on 29 March 2024

and Chapter 6, native speakers' preference of *powerful computer* over *strong computer* is cited as an example of the arbitrariness of collocations. I do not intend to delve into the debate over whether collocations are arbitrary or linguistically motivated, but I note that the preference in this particular case is not arbitrary but semantically motivated (consider the semantic difference between *powerful man* and *strong man*). In fact, this preference merely reflects the fact that we are generally more concerned about the functional power of computers than their physical build, and it would be a false positive if a system marks *strong computer* as an error in a sentence that talks about a computer that is well-built and not easily breakable.

This book review was edited by Pierre Isabelle.

Xiaofei Lu is Gil Watz Early Career Professor in Language and Linguistics, and Associate Professor of Applied Linguistics and Asian Studies at the Pennsylvania State University. His research interests are primarily in computational linguistics, corpus linguistics, and intelligent computer-assisted language learning. Lu's e-mail address is xx113@psu.edu.

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/COLI_r_00211 by guest on 29 March 2024

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/COLI_r_00211 by guest on 29 March 2024