Book Reviews

Automated Grammatical Error Detection for Language Learners

Claudia Leacock, Martin Chodorow, Michael Gamon, and Joel Tetreault
(Butler Hill Group, Hunter College, Microsoft Research, Educational Testing Service)

Reviewed by
Stephen Pulman
University of Oxford

This book is a useful survey of the current state of the art in automated grammatical error detection in (mostly) non-native English, by some of the leading researchers in the field, aimed at audiences in computational linguistics and computer-aided language learning. The book begins with a working definition of grammatical error, distinguishing these from typos and some kinds of spelling and punctuation errors, and points out that whereas (only) 400 million people have some kind of English as their first language, another billion speak it as their second (or are trying to learn it as such).

Chapter 2 presents a brief historical overview of the field. Until the advent of statistical methods in the 1990s, most systems used some kind of rule-based parsing supplemented with mechanisms for dealing with ill-formed input. These mechanisms included pattern matching, the addition of special "mal-rules" for spotting frequently occurring forms of error, or different types of "parse fitting" and "relaxation" techiques (e.g., in the case of unification grammars, allowing for some types of unification failure). Probably the most successful of these, both in commercial terms and in terms of performance, was the Epistle system originating from IBM and subsequently included in Microsoft Word, although its primary purpose was first-language rather than second-language correction.

In Chapter 3 the authors explore the problems faced by second-language learners. On the basis of evidence from the Cambridge University Press corpus of learner English, it seems that the most frequent error type is simply an incorrect choice of content word, followed in turn by incorrect preposition, and incorrect determiner choice. Like many collocation errors, it is difficult to use the techniques described in the book to address incorrect content word choice, because this usually does not result in ungrammaticality, and the state of the art is not such that semantic anomaly can be reliably detected. Many systems therefore have concentrated on preposition and determiner errors.

Chapter 4 surveys available data resources: corpora of learner language, in several different languages. The field of automated error detection suffers from a lack of large-scale annotated corpora: They exist, at least for English, but for commercial reasons are not easily or cheaply available. There is a real need in this area for good corpus resources, and some shared benchmark testing data. At present, it is literally impossible to compare the performance of different systems, because they are all tested on different data sets, often using different metrics; these issues are explored further in Chapter 5. Chapter 6 focuses in more detail on article and preposition errors, which are those where the techniques discussed—typically various types of classifier, often trained on grammatically "correct" data like the British National Corpus (BNC) or

equivalents—are able to achieve reasonable performance on error detection (80–90%) although error correction is typically less accurate. Two particular complete systems, ETS's Criterion and Microsoft's ESL Assistant, are described here.

In Chapter 7, collocation errors are discussed. As noted earlier, these are particularly challenging. Current methods for measuring strength of association for detecting collocations are described. There are few systems which aim to detect collocation errors, but one described here, aimed specifically at Chinese speakers learning English, achieves good performance: Noting that many of their collocation errors are the result of inappropriate direct translation from Chinese, the system first checks whether verb+noun combinations occur in the BNC, and, if not, using bilingual dictionaries and an aligned English-Chinese parallel corpus, suggests alternative translations for one of the words, and then checks whether the resulting combination occurs in the BNC.

Chapter 8 discusses techniques aimed at particular types of error, such as verb forms or agreement, and Chapter 9 describes annotation schemes for error corpora and discusses the various issues that arise in the context of trying to develop such resources. In Chapter 10 some alternative error detection techniques using Google n-grams, on-line translation systems, and Web counts are briefly discussed, and Chapter 11 summarizes and concludes.

This volume serves the purpose of the series: It is a valuable and clear overview of the current state of the art, and would be a good starting place for those new to the field. Of course, grammatical errors and typos are not unique to second-language learners. One might expect the same technology to improve the quality of some apparently hurriedly written text: There is an amusingly large collection of grammatical errors and typos in this book: "they are likely do some additional verification", "... errors in the use of auxiliary verbs, gerunds and infinitives in." "grammatical error typess", etc., etc. Quis custodiet ipsos custodes?

This book review was edited by Pierre Isabelle.

Stephen Pulman is Professor of Computational Linguistics at the Department of Computer Science (formerly the Computing Laboratory) at Oxford University. He has worked in computational morphology, syntax and parsing, and dialogue, as well as practical applications including automated assessment, error detection in L2 English, and sentiment analysis. His research currently focuses on semantics and inference. Pulman's address is: Department of Computer Science, Oxford University, Wolfson Building, Parks Road, Oxford, OX13QD; e-mail: stephen.pulman@comlab.ox.ac.uk.