

Grant R305F100007
Year of Study: 2012

Title: Assessing the use of multiple sources in student essays

Authors: Hastings, P., Houghs, S., Magliano, J.P., Goldman, S.R., & Lawless, K.

Citation: Hastings, P., Houghs, S., Magliano, J.P., Goldman, S.R., & Lawless, K. (2012). Assessing the use of multiple sources in student essays. *Behavioral Research Methods*, 44(3), 622-633.

Strand of Work: Tools

Abstract

Purpose and Questions

Imagine a situation in which a student is asked to write a research paper on the causes of climate change and, in particular, to argue that the primary causes are based on human activities. Presumably, the student would need to identify and integrate information from multiple text sources to write such a paper. Evaluating how well students engage in this task is a serious challenge for teachers, which may lead some to avoid tasks that require integrating information across texts. This puts students at risk for learning the literacy and critical thinking skills necessary to succeed in college and beyond.

One approach to help teachers assess essays that require students to integrate across multiple texts is to develop computer systems that can do some or all of the work for them. The present study explored different approaches for automatically scoring student essays that were written on the basis of multiple texts. Specifically, these approaches were developed to classify whether or not the essays contained important elements of the texts that reflect integration across texts in relation to the essay prompt.

Research Context or Methodology

This study involved analyzing a corpus of 460 essays written by 5th – 8th graders in the Chicago Public School system on the topic of why Chicago became a large city. They read three texts, of which none provided an explicit answer to the question that prompted the essay, but content for these could be used to address the question. We explored three approaches for evaluating the essays automatically. The first was a simple pattern-matching approach called “multi-word” that allowed for flexible matching of words and phrases in the sentences. The second technique was latent semantic analysis (LSA), which was used to compare student sentences to original source sentences using its high-dimensional vector-based representation. Finally, the third was a machine-learning technique, support vector machines, which learned a classification scheme from the corpus.

General statement of findings

The results of the study suggested that the LSA-based system was superior for detecting the presence of explicit content from the texts, but the multi-word pattern-matching approach was better for detecting inferences outside or across texts.

Implications

These results suggest that the best approach for analyzing essays of this nature should draw upon multiple natural language processing approaches. These results also suggest that it is possible to develop computer systems to help support the evaluation of essays and in particular those that involve integrating information across texts. Further development of these system could provide the tools to help encourage teachers to use task that help students develop the critical literacy skills they need to succeed in college and beyond.

Acknowledgements

The work and the preparation of this article were supported by the Institute of Education Sciences, U.S. Department of Education (Grant R305F100007, titled Reading for Understanding Across Grades 6 through 12: Evidence-Based Argumentation for Disciplinary Learning, Susan Goldman, PI). The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.