

Using Wikipedia and Conceptual Graph Structures to Generate Questions for Academic Writing Support

Ming Liu, Rafael A. Calvo, *Senior Member, IEEE*, Anindito Aditomo, and Luiz Augusto Pizzato

Abstract—In this paper, we present a novel approach for semiautomatic question generation to support academic writing. Our system first extracts key phrases from students' literature review papers. Each key phrase is matched with a Wikipedia article and classified into one of five abstract concept categories: Research Field, Technology, System, Term, and Other. Using the content of the matched Wikipedia article, the system then constructs a conceptual graph structure representation for each key phrase and the questions are then generated based the structure. To evaluate the quality of the computer generated questions, we conducted a version of the Bystander Turing test, which involved 20 research students who had written literature reviews for an IT methods course. The pedagogical values of generated questions were evaluated using a semiautomated process. The results indicate that the students had difficulty distinguishing between computer-generated and supervisor-generated questions. Computer-generated questions were also rated as being as pedagogically useful as supervisor-generated questions, and more useful than generic questions. The findings also suggest that the computer-generated questions were more useful for the first-year students than for second or third-year students.

Index Terms—Automatic question generation, writing support, natural language processing

1 INTRODUCTION

REVIEWING the literature to summarize and build upon current knowledge about a topic is a key part of academic writing [1]. According to Steward [2] a good review should present a critical discussion of relevant materials, with the goal of developing new ideas from available evidence and knowledge. Unfortunately, determining the relevance of materials, engaging in critical discussion, and synthesizing evidence are precisely what students often find difficult in writing literature reviews [3]. This paper discusses the provision of reflective questions as a way to help students overcome some of these problems.

Generic trigger questions have been widely used for literature review support [4]. These are questions such as "Have you critically analyzed the literature you use?" and "Have you discussed how your project will contribute to that discipline or field?" While generic questions may be useful, students may benefit more from questions which are specific to the content of their document. However, creating such specific questions is typically difficult and time consuming. The goal of our research is to develop a fully automated method to generate specific questions to support academic writing.

Automatic Question Generation (AQG) is a challenging task which involves natural language understanding and generation [5]. Three major aspects of AQG have been addressed in the literature: selection of the target content (what to ask about), selection of the question types (e.g., Who, Why, Yes/No), and construction of the actual questions. An increasing number of automatic question generation techniques have been explored [6], [7], [8], [9], [10]. Most of this research has focused on generating questions for supporting reading comprehension or vocabulary assessment using factual or nonfactual questions. These questions can be asked to any student in the cohort and assessed by anyone qualified in the topic. In contrast, our work focuses on generating trigger questions to support academic writing. In this context the questions are based on the individual students' work and therefore can only be assessed by them.

In a previous study [11], we described and evaluated a QG approach that focuses on citation sentences in a literature review. In that approach, we classified citation sentences into these categories: Opinion, Aim, Result, Method, System, and Other. For example, if a student describes an opinion found in a citation: "Cannon (1927) challenged this view mentioning that physiological changes were not sufficient to discriminate emotions," the system would generate trigger questions about, for instance, the evidence regarding the opinion: "Why did Cannon challenge this view mentioning that physiological changes were not sufficient to discriminate emotions? (What evidence is provided by Cannon to prove the opinion?) Does any other scholar agree or disagree with Cannon?"

Our evaluation showed that the automatically generated questions outperformed generic questions; furthermore, they were judged as being equally useful compared to questions generated by human supervisors. In another

• M. Liu and R.A. Calvo are with the School of Electrical and Information Engineering, The University of Sydney, NSW 2006, Australia. E-mail: {Ming.Liu, Rafael.Calvo}@sydney.edu.au.

• A. Aditomo is with the University of Sydney, Room 237, A35 Education Building, NSW 2006, Australia. E-mail: aadi4954@uni.sydney.edu.au.

• L.A. Pizzato is with the University of Sydney, Room 328, J12-School of Information Technologies Building, NSW 2006, Australia. E-mail: luiz.pizzato@sydney.edu.au.

Manuscript received 5 Sept. 2011; revised 28 Nov. 2011; accepted 27 Feb. 2012; published online 5 Mar. 2012.

For information on obtaining reprints of this article, please send e-mail to: lt@computer.org, and reference IEEECS Log Number TLT-2011-09-0095. Digital Object Identifier no. 10.1109/TLT.2012.5.