ASSESSMENT AND ERROR IDENTIFICATION OF ANSWERS TO MATHEMATICAL WORD PROBLEMS

Jayanath Chamindu Sandanuwan Kadupitiya
(168007X)

Thesis submitted in partial fulfillment of the requirements for the Degree of Master of Science

Department of Computer Science & Engineering

University of Moratuwa
Sri Lanka

April 2017

DECLARATION

I declare that this is my own work and this dissertation does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief, it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature: Date:

Name: J.C.S Kadupitiya

The above candidate has carried out research for the Masters Dissertation under my supervision.

Signature of the supervisor: Date:

Name of the supervisor: Dr. Surangika Ranathunga

Signature of the co-supervisor: Date:

Name of the supervisor: Prof. Gihan Dias

Abstract

In Mathematics, the term “word problem” is often used to refer to any mathematical exercise where significant background information on the problem is presented as text rather than in mathematical notation. This research focuses on word problems that have simple numerical and/or algebraic answers. These types of word problems can be further categorized according to the domain, such as interest calculation questions, percentages, shares and mensuration. These word problems can be found in many international examinations. Existing research has produced solutions that focus on questions only for some of the aforementioned categories. Moreover, they have not focused on assessment based on a marking rubric.

This thesis presents a system that is capable of assessing answers to both numerical and algebraic type word problems using a (teacher-provided) marking rubric. We automatically identify the exact errors (if any) made by students by using the marking rubric. This system is modularized and can be extended to support different types of word problems. If the answer contains a short sentence phrase along with the numerical or algebraic expression, it is also evaluated in order to check whether the student has actually understood the question. Our main focus is the questions from the GCE Ordinary Level (O/L) Mathematics syllabus in Sri Lanka. Many students take this examination in Sinhala (an official language in Sri Lanka). Therefore short sentence evaluation had to be done for Sinhala. This requirement led us to conduct the first research on short sentence similarity measurement for Sinhala. The unsupervised similarity measurement technique we used showed comparable results to that of English.

The system was thoroughly evaluated with student answers to questions from GCE O/L examination. It was further tested for answers to some word problems from the Cambridge Ordinary Level and the Australian year 10 international examinations, which demonstrated that the system is able to deal with variations in questions in different examinations.

ACKNOWLEDGEMENT

I would first like to thank my supervisors, Dr. Surangika Ranathunga and Prof. Gihan Dias, for the continuous support given for the success of this research. They consistently allowed this research to be my own work, but steered me in the right direction whenever they thought I needed it.

This research was supported by the LK Domain Registry, the University of Moratuwa Senate Research Grant, and the DL4D (Digital Learning for Development) Grant. I would like to covey my sincere gratitude to colleagues of the Maths Tutor project, the typists and annotators who helped me in this research.

I would like to thank all the staff from the Department of Computer Science and Engineering, University of Moratuwa, for ther kindness they expressed on all occasions.

Finally, I must express my very profound gratitude to my parents for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them.

Thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ALGORITHMS

LIST OF ABBREVIATIONS

CAT -	Common Admission Test
GATE -	Graduate Aptitude Test
GRE -	Graduate Record Examination
GCE -	General Certificate of Education
O/L -	Ordinary Level
A/L -	Advanced Level
ALE -	Adaptive Learning Environments
ITS -	Intelligent Tutoring System
MOOC -	Massive Open Online Course
CAA -	Computer Aided Assessment
CAS -	Computer Algebra System
LSA -	Latent Semantic Analysis
ESA -	Explicit Semantic Analysis
STS -	Semantic Text Similarity
CDSM -	Compositional Distributional Semantic Model
RTM -	Referential Translation Machine
MAML -	Mathematics Assessment Markup Language