

# AUTOMATIC ANSWER GENERATION FOR MATH WORD PROBLEMS

Kulakshi Fernando

178090C

Thesis/Dissertation submitted in partial fulfillment of the requirements for the  
degree Master of Science in Computer Science and Engineering

Department of Computer Science & Engineering

University of Moratuwa

Sri Lanka

May 2019

## 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:

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

Name of the Supervisor: Dr. Surangika Ranathunga

Signature of the Supervisor:

Date:

Name of the Supervisor: Prof. Gihan Dias

Signature of the Supervisor:

Date:

## ACKNOWLEDGEMENTS

I am sincerely grateful for the advice and guidance of my supervisors, Dr. Surangika Ranathunga and Prof. Gihan Dias. Without their support, mentoring, and encouragement this project would not have been completed. I would like to thank them for taking time out of their busy schedule to be available anytime that was needed with help and advice.

I would also like to thank my progress review committee, Dr. Dulani Mee-deniyaya and Dr. Supunmali Ahangama. Their valuable insights and guidance helped me to enhance the quality of my research work.

I would like to thank the entire staff of the Department of Computer Science and Engineering, both academic and non-academic for all their help during the course of this work and for providing me with any resources necessary to conduct my research.

This work was funded by a Senate Research Committee (SRC) Grant of the University of Moratuwa.

Finally, I would like to express my gratitude to my friends, specially Mr. Mokanarangan Thayaparan and Ms. Ishadi Jayasinghe who were always supporting me at my research work and to my family for their immense support.

# ABSTRACT

## Automatic Answer Generation for Math Word Problems

A math word problem (MWP) is a mathematical problem expressed using natural language. In this research, elementary level set-related word problems in which information is given in set notation are considered. As per our knowledge, this is the first research addressing set theory related word problems.

This research introduces an abstract representation to interpret mathematical semantics of set expressions and relations between sets. Two methods to extract given set related expressions were implemented: rule based method and a statistical method. Results show that statistical method is more robust to typing errors and unexpected expression formats. A parser based on a context free grammar is introduced to validate set related expressions and give feedback to the user when there are incorrect expressions. Along with these functionalities, we present a complete set problem solver system that understand and solve a given set word problem.

In addition to the solver, we experiment in extracting mathematical expressions from unstructured plain text using sequential classifiers. Several sequential classification models including conditional random-fields (CRF) and Long-Short Term Memory (LSTM) networks were compared with word and character level features. The results show that using character level features significantly increase the performance of mathematical expression extraction.

**Keywords:** Set theory. Answer generation. Math word problems.

## LIST OF FIGURES

Figure 3.1	Overview of the system	21
Figure 3.2	Venn diagram of set A, the set of all positive integers between 1 and 9	23
Figure 3.3	The Venn diagram denoting information in question Q1. Numbers denote the cardinalities of sets and bordered labels denote the set that represent by each region	25
Figure 3.4	The graphical representation of subset-superset relations between all the possible sets in Q1. The binary representation of the sets are written with set names. Edges of the graph point subsets to supersets.	29
Figure 3.5	Extracting information from Q2 using rule based approach	31
Figure 3.6	Network architecture of the W-LSTM model	34
Figure 3.7	Network architecture of the W-Bi-LSTM model	34
Figure 3.8	Network architecture of the W-CH-Bi-LSTM model	36
Figure 3.9	Finding the universal set in Q2	39
Figure 4.1	CRF performance against cumulatively added feature sets from the set A to set G listed in Table 3.4	49
Figure 4.2	Performance of RNN models with respect to number of epochs	50

## LIST OF TABLES

Table 3.1	Problem categories based on information part of the problem	24
Table 3.2	Binary representation of sets given in Q1 and cardinality information about each set	26
Table 3.3	Regular expressions used to extract set information	30
Table 3.4	Features used for CRF divided into categories. The left most column contains a label for each of the set of features	35
Table 3.5	Examples for set expressions that can be validated by the CFG parser	36
Table 3.6	Generated equations for sets in Q1	42
Table 4.1	Datasets used to evaluate solver system	46
Table 4.2	Statistics of problems in the dataset, DExpr	47
Table 4.3	Accuracy, Recall, Precision and F1-score of the best performance of all models	50
Table 4.4	Performance of statistical and rule-based expression extractors	51
Table 4.5	Statistics of the data used to evaluate the parser	51
Table 4.6	Evaluation results of the parser	52
Table 4.7	Statistics of the data used to evaluate the solver	52
Table 4.8	Evaluation results of the solver	52
Table 4.9	Evaluation results of the complete system	53

## LIST OF ABBREVIATIONS

MWP	Math Word Problem
GCE	General Certificate of Education
IGCSE	International General Certificate of Secondary Education
GCSE	General Certificate of Secondary Education
JCE	Junior Certificate Examination
NCERT	National Council of Educational Research and Training
NLP	natural Language Processing
CRF	Conditional Random Field
LSTM	Long-Short Term Memory
Bi-LSTM	Bi-directional Long-Short Term Memory
CFG	Context Free Grammar

# TABLE OF CONTENTS

Declaration of the Candidate & Supervisor	i
Acknowledgement	ii
Abstract	iii
List of Figures	iv
List of Tables	v
List of Abbreviations	vi
Table of Contents	vii
1 Introduction	1
1.1 Background	1
1.2 Research Problem	2
1.3 Research Objectives	3
1.4 Contributions	3
1.5 Scope and Limitations	4
1.6 Publications	4
1.7 Organization	4
2 Literature Survey	6
2.1 Math Word Problem Solving Systems	6
2.1.1 Rule-based approaches	7
2.1.2 Statistical approaches	8
2.1.3 Hybrid approaches	9
2.1.4 MWP domains addressed in previous research	15
2.1.5 Discussion on problem solving systems	15
2.2 Expression Extraction Methods	16
2.3 Information Extraction Using Sequential Classifiers	17
2.4 Summary	19
3 Methodology	20
3.1 System Overview	20
3.2 Set Problem Categorization	20



3.2.1	Presentation formats of sets	20
3.2.2	Properties of sets	22
3.2.3	Problem categorization	23
3.3	Abstract Representation of Sets	23
3.3.1	Representing any number of Sets	26
3.3.2	Representing main sets	26
3.3.3	Representing derivable sets	27
3.3.4	Finding subsets and supersets	28
3.4	Expression Extraction	28
3.4.1	Rule based expression extraction	29
3.4.2	Limitations of regular expressions	30
3.4.3	Expression extraction using statistical approach	32
3.5	Expressions Parsing	33
3.6	Mapping to Data Representation	38
3.7	Question Validation	39
3.8	Answer Generation	41
3.8.1	Generating equations and calculating cardinalities	41
4	Evaluation	45
4.1	Data Sets	45
4.2	Experimental Setup	46
4.3	Evaluations	48
4.3.1	Evaluation of math expressions extraction based on sequential classifiers	48
4.3.2	Evaluation of set expressions extraction	49
4.3.3	Parsing sets expressions	51
4.3.4	Solving a set problem	52
4.3.5	End-to-end performance	52
5	Conclusion	54
5.1	Future Work	55
	References	56