A Review of Automatic Math Word Problem Solving Techniques

Harshal Kotwal

Research Scholar
harshalrkotwal@gmail.com

Girish Kumar Patnaik

Professor

SSBT's College Of Engineering & Technology, Jalgaon (M.S.)
patnaik.girish@gmail.com

Abstract

Computers, since their creation, have exceeded human beings with respect to speed and accuracy in solving mathematical calculation. However, even after the advancement in computational logic and artificial intelligence, it is still a big challenge today to design algorithms to automatically solve even primary-school-level math word problems. The problem becomes more challenging due to the semantic gap between human understanding and machine logic in formulating the solution to the Math Word Problems. This paper studies significant efforts in solving math word problems. Though researchers claim to have considerable improvement through their work, the application of the proposed solution to large datasets does not produce accurate results. In this paper, various techniques to solve math word problems including rule-based methods and deep neural network based method; and dataset repositories are studied, to motivate and attract more contribution in this area.

Keywords: *Math Word Problem, Natural Language Processing, Semantic Parsing, Deep Neural Network*

1. Introduction

Algebraic Math Word Problems are those problems which involve some kind of rule or logic and require fundamental mathematics to solve it. Math Word Problems are typical taught during elementary and primary school level to test the quantitative aptitude and reasoning ability of the students. The idea of math word problems can be extended to other subject domains including mechanics, physics, geometry and algebra. Such problems are themselves challenging for students not because of their complexity but because of the hidden logic and inability to infer the mathematical equation residing the problem. Attempts to address and solve math word problem is reported to have started during early 1960s. Since then, many researchers with different approaches have contributed and achieved success subject to the constraints they have put during research. Math Word Problem solving domain have attracted focus of research society because developing intelligent system for understanding and solving natural language mathematical problem is very challenging.

Evolution of Math Word Problem Solvers is mainly categorized into three generations based on the technology platform used by these solvers, as proposed by Zhang et al. [1]. During the time period of 1960 to 2010, the technology was based on mainly rule based matching which used hand craft rules to solve math word problems. Early effort during this time period relied on human interventions and could work in very limited scenario. During the second generation of 2011 to 2017, the technology bed was based on sematic parsing, feature engineering and statistical learning. Majority of the contribution is reported during this time period. Standard Dataset repositories were built and made available during this second generation. Contributions made during this generations laid

down a fundamental framework for future work. Significant number of publications with high accuracy and diverse datasets are reported in this generation.

Presently we are into the third generation, in which the technology base has shifted to deep learning and reinforcement learning. Due to the advancement in the area of artificial intelligence, researchers preferred designing learning systems to imitate the human in addressing math word problem solving technique. Advancement in machine learning and natural language processing tools has created new opportunities for young researchers.

2. Types of Math Word Problems

Math Word Problems can be differentiated into various categories based on their feature characteristics. Algebra and Arithmetic Word Problems are those math problems in which there in only one unknown. Based on the mathematical input described in the problem one or more known variable are to be identified and mathematical equation is derived to find the value of unknown variable. Most of the times, these problems are single step solution to the problem, meaning that complexity of these problems is very low. Equation Set Problems may have one or more unknown variable and are more complex. The unknown variable are dependent on each other hence, step by step equation formulation is required to find the value of required unknown variable. The multi-step nature of these problems makes them more complex and more advance techniques are required to solve them. Study of review article by authors Zhang et al. [1], reveals that most of the attempts in solving math word problem is on arithmetic word problems. Equation Set Problems provides a future scope for interested people. Authors Mukherjee et al. [2], have listed other types of math word problems from different subject domains including geometry problems, rate problems, probability problems, matrix problems, physics problems, mechanics problems, chemistry problems and other general numeric problems. Based on the nature of different math word problems, some additional computing tools and techniques may be required along with natural language processing. For example, to solve geometry problems, computer vision and graphic tools are required as these problems include geometry shapes.

3. First Generation Math Word Problems

Efforts in designing automatic math word problem solver (MWPs) started during 1960s [3], [4], and since then have continually attracted intensive research. Ever since, the problem has remained challenging until now, because of the sematic gap between human understanding and computational logic. After emergence of artificial intelligence and natural language processing techniques, there has been a significant improvement in math word problem solvers but still more advance algorithms are required to prove their efficiency for the large datasets [5], [6] and to successful solve MWPs accurately.

In 1964, Bobrow [1], developed the STUDENT problem solving system which was programmed in LISP. The system was designed to accept restricted set of English to express wide variety of algebra story problems. The STUDENT program could find the set of kernel sentences and transformed them into a set of simultaneous equations. These equations are then solved for requested unknowns. Author in this paper confirms STUDENT system to be a first step towards natural language communication with computers.

Similar to STUDENT, another program called CARPS (Calculus Rate Problem Solver) was written in LISP and CONVERT by Eugene Charniak [7]. CARPS was restricted to solve rate problems. Structures which were stored internally as trees would hold the information about the objects. The program solved 14 calculus problems taken from standard calculus texts and proved to outperform the STUDENT program in several aspects.

Another program DEDUCOM, written in LISP by James R. Slagle [8], accepted facts about questions to be asked and attempted to answer the questions. DEDUCOM turned out to be a single system answering different kinds of questions rather than the previous approaches which used separate systems for answering.

In 1984, author Charles R. Fletcher [9], described a computer program called "WORDPRO" written in Interlisp-D, simulated the psychological process involved in understanding and solving simple arithmetic word problems. WORDPRO implemented the theory of the comprehension and solution to simple arithmetic word problems.

All of these initial contributions in automatically solving math word problems are based on manual craft rules as per the requirements and problem domain. It also used schemas for pattern matching. The major limitations of these efforts are their heavy dependency on human interventions during problem solving and very limited ability to solve diversified complex math word problems.

4. Data Driven Approaches for MWPs

Authors Wang et al. [10] in their experiment attempted to solve algebra math word problems by generating answer rationales and human readable mathematical expression using natural language sequences. The program was evaluated over a newly created 1oo,ooo sample dataset questions, answers and rationales.

Authors propose to solve math word problems by inducing and designing programs to generate answers and their rationale. This paper has contributed in three different ways. First, they created a new dataset of more than 100,000 algebraic word problems with answers. Secondly, it proposes a sequence to sequence model that generates sequence of instructions to create rationale. Finally, a technique to generate answer from rationale is also proposed in this paper. Experimental results performed by authors show that this method performance better that existing neural models in generational of rationale and ability to solve problems.

Recently published article by authors Benjamin et al. [11], proposes a data driven approach for solving algebra word problems over large scale datasets. Provided a large training data, data driven model can learn to map math word problem text to system of complex equations. Authors proposed that a fine-tuned equation classifier can perform much better than previous math word problem solvers. Experiments were performed on datasets including Chinese language dataset Math23K and English language datasets DRAW and MAWPS. To conclude, authors suggest incorporating sematic and wording knowledge to improve accuracy of algebraic word problem solvers.

Authors Lei Wang et al. [12], have used deep learning reinforcement learning to solve arithmetic word problem. MathDQN program proposed in this article, yield significant improvements on large dataset and with better precision. In the proposed framework, for an input math word problem quantity schema is adopted to identify relevant quantities. Irrelevant quantities are discarded bottom layer of expression tree generated using DQN framework. Positive or negative rewards are returned depending on next state or status of partial tree. Experiments are performed on various datasets including AI2, IL, CC, ArithS; and ArithM, which is combination of AI2, IL and CC itself. Experimental results using MathDQN show that, it achieves the best performance and significant improvement in accuracy in AI2 and CC datasets. Author also claims that, MathDQN is the only method all three datasets.

Another attempt by authors Purvanshi Mehta et al. [13] in solving arithmetic word problem used Deep Neural Network based system called DILTON. Given a math word problem as input, DILTON first predicts the sequence of arithmetic operation to be performed. The four basic arithmetic operators including addition, subtraction, multiplication and division are considered. DILTON uses a deep neural network based

approach to generate the answer. During the process, the input text problem is divided into two parts, wordstate, which covers input data and the query, which contains the question part of the math word problem. Both the parts are processed separately in two different networks and finally merged to generate the answer. The major contributions of this paper include, automated system that learns to solve word problems with two operands, deep neural network based model to predict mathematical operation involved in word problem and identifying relevant quantities through similarity between context of quantity and corresponding question.

5. Dataset Repository for MWPs

In order to test the accuracy of math word problem solvers data repositories are required. The datasets are either manually prepared or collected through online resources. Different types of dataset repositories are made available for researchers. The comparative study of available datasets is as shown below in Table 1. Most of the datasets are available in English language. There is a need for datasets in other popular languages. Dataset of equation set problems and other different types of problems also needs to be made available so as to have a comparative study of different approaches.

AI2 has 395 single-step or multi-step arithmetic word problems. It consists of math problems that can be solved using addition and subtraction operation. IL is another dataset collected from k5learning.com and dadsworksheets.com. It has 562 single-step word problems with only one operator. CC dataset consists of 600 multi-step math word problems. The problems are collected from commoncoresheets.com which includes various combinations of fundamental mathematical operations. SingleEQ dataset consists of both single-step and multi-step math word problems. AllArith dataset is created using resources of AI2, IL, CC and SingleEQ with total 831 numbers of problems. DolphinS which is a subset of Dolphin18K contains 115 single operator and 6955 multi operator problems.

Dataset repositories for equation set problem includes Dolhin18K containing 18,460 complex math word problems. Another popular dataset in this category is ALG514 which is collected from Algebra,com. It consists of total 514 math word problems in English. DRAW1K comes up with 1000 linear equation problems. Dolphin1878 has collected problems from Algebra.com and answer.yahoo.com with 1878 different math word problems. Math23K is presently the largest dataset in Chinese with over 23,000 math word problems. At present there is a need for a very large diversified dataset of math word problems to verify the testability and accuracy of math word problems. Neural Network based approaches require systems to be trained and thus large datasets will definitely prove helpful to improve efficiency of intelligent systems.

Table 1. Data Repositories for MWPs

Dataset Name	Problem Type	No. of Problems	Language
AI2	Arithmetic	395	English
IL	Arithmetic	562	English
CC	Arithmetic	600	English
SingleEQ	Arithmetic	508	English
AllArith	Arithmetic	831	English
Dolphin-18K	Arithmetic	18,460	English
Dolphin-S	Arithmetic	115	English
Math23K	Arithmetic	23,161	Chinese
ALG514	Equation Set	514	English
Dolphin1878	Equation Set	1878	English

DRAW1K	Equation Set	1000	English

6. Conclusion

The sematic gap between human understanding of math word problems and machine logic make it difficult to design efficient math word problem solvers. Early efforts in this respect were constraint based and worked in very restricted environment. Contributions made during second generation are very significant. They paved future direction and paved path for upcoming researchers. More advance algorithms incorporating Artificial Intelligence and Neural Network can help is solving complex multi-step equation set problems. The journey of math solvers has been very challenging but still there is a lot of scope for improvement in this domain.

References

[1] D. Zhang, L. Wang, N. Xu, B. T. Dai and H. T. Shen, "The Gap of Semantic Parsing: A Survey on Automatic Math Word Problem Solvers", arXiv preprint arXiv:1808.07290, August, 2018.

[2] Anirban Mukherjee and Utpal Garain, "A review of methods for automatic understanding of natural language mathematical problems," Artif. Intell. Rev., vol. 29, no. 2, pp. 93–122, Apr. 2008.

[3] D. Bobrow, "Natural language input for a computer problem solving system," in Semantic information processing, M. Minsky, Ed. MIT Press, 1964, pp. 146–226.

[4] E. A. Feigenbaum and J. Feldman, Computers and Thought. New York, USA: McGraw-Hill, Inc., 1963.

[5] P. Clark, "Elementary school science and math tests as a driver for AI: take the aristo challenge!" in AAAI, 2015, pp. 4019–4021.

[6] P. Clark and O. Etzioni, "My computer is an honor student - but how intelligent is it? standardized tests as a measure of AI," AI Magazine, vol. 37, no. 1, pp. 5–12, 2016.

[7] E. Charniak, "Computer solution of calculus word problems," in Proceedings of the 1st International Joint Conference on Artificial Intelligence, ser. IJCAI'69, 1969, pp. 303–316.

[8] J. R. Slagle, "Experiments with a deductive question-answering program," Commun. ACM, vol. 8, no. 12, pp. 792–798, Dec. 1965.

[9] C. R. Fletcher, "Understanding and solving arithmetic word problems: A computer simulation," Behavior Research Methods, Instruments, & Computers, vol. 17, no. 5, pp. 565–571, Sep 1985.

[10] W. Ling, D. Yogatama, C. Dyer and P. Blunsom,"Program induction by rationale generation: Learning to solve and explain algebraic word problems," In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, pages 158–167, 2017.

[11] Benjamin R., R. Koncel-Kedziorski, Hannaneh H.,"Data-driven methods for solving algebra word problems," CoRR, abs/1804.10718, 2018.

[12] L. Wang, D. Zhang, L. Gao, J. Song, L. Guo, H. T. Shen,"Mathdqn: Solving arithmetic word problems via deep reinforcement learning," In Thrity-Second AAAI Conference on Artificial Intelligence, 2018

[13] P. Mehta, P. Mishra, V. S. Athavale, M. Shrivastava, D. M. Sharma, "Deep Neural Network based system for solving Arithmetic Word problems," 8th International Joint Conference on Natural Language Processing, 2017.