Learning to Solve Arithmetic Word Problems Using Sentence Simplification

Vishal Rajpal
Northeastern University
rajpal.vi@husky.neu.edu

Abstract

In order to respond to an arithmetic word problem correctly one needs to understand the question to an extent which allows determining the constraints. These are mostly semantic and are imposed by the question on its answer. Constraints from individual sentences suggest a mathematical operation and when the operators from these sentences are used collectively the answer can be derived. To extract the constraints efficiently, a concept of Syntactic Pattern is introduced which is generated by parsing the sentence using a dependency parser. It encapsulates all the relevant information in a sentence including the subject, verb and object with its quantity. Another important method for this thesis which relies on syntactic patterns is Sentence Simplification. The idea is to have a subject, verb, an object and other necessary parts of speech in the sentence so that it suggests a single operation. Based on the identified patterns a sentence may be simplified to multiple sentences. This would make the classification process easier since the sentences are less complex. To this end, it would be the classifier's job to classify the sentence to a mathematical operator. The identified operators for individual sentences are used to build a mathematical equation for the entire word problem. The results based on 3 datasets are reported and seem promising as compared to the existing systems.

1 Introduction

Answering arithmetic word problems has gained a lot of interest in recent years. The problem is attractive to NLP since the text is concise and relatively straightforward with identifiable semantic constraints. As the arithmetic word problems are directed towards elementary school students, they begin by describing a partial world state, followed by simple quantified updates or elaborations and end with a quantitative question. This information

1

can be mapped to basic operators(addition, subtraction, multiplication and division) and a equation corresponding to the entire problem can be created.

There have been a number of attempts to solve arithmetic word problems through Machine Learning. All the approaches which are not template based (Hosseini et al., 2014), (Roy et al., 2015) and (Roy and Roth, 2015) use different methods to extract similar information. Based on different ways the information is represented, an equation is built for the problem text. The template absed method of (Kushman et al., 2014) implicitly assumes that the solution will be generated from a set of predefined equation templates. Some of these methods only solve addition and subtraction problems (Hosseini et al., 2014), (Roy et al., 2015) while (Roy and Roth, 2015) and (Kushman et al., 2014) can solve problems for all operations.

The approach presented in this thesis can solve a general class of addition and subtraction arithmetic word problems without any predefined equation templates. In particular, it can handle an arithmetic problem as shown in Table 1.

Example 1:
Joan found 70 seashells on the beach . she gave Sam some of her seashells . She has 27 seashell . How many seashells did she give to Sam ?

Table 1: Example Arithmetic Word Problem.

To derive the solution to this problem, the approach needs to understand that after giving some number of seashells from initial 70, Joan was left with 27.

While a solution to these problems requires extracting information and composing numeric expressions, if the sentence is too complex it is hard to extract information accurately.

At the heart of the technical approach, the novel notion of *Sentence Simplification* is involved. Once the sentences in the problem text are simplified, extracting information becomes easier. Each sentence in the problem is simplified to a level where it consists information which is able to be mapped to a single operator. This allows us to decompose the entire problem to a collection of simplified sentences as operator prediction problems, each sentence representing the quantitative information with the mapped operator. These predictions(operators with the quantitative information) can be combined together to form a mathematical equation.

The approach focuses on addition and subtraction problems currently,

but learning to classify operators will allow us to generalize the approach to multiplication and division problems as well. In particular, the system was able to solve Example 1 although it had never seen the problem before and required both addition and subtraction operations.

The approach is evaluated on 3 datasets, achieving competitive performance on all of them. The next section describes the related work in the area of automated arithmetic word problem solving. The theory of sentence simplification is then presented and the information decomposition strategy that is based on it. Section 4 presents the overall computational approach, including the way we learn to classify simplified sentences to operators. Finally, experimental study is discussed followed with conclusion.

References

Hosseini, M. J., Hajishirzi, H., Etzioni, O., and Kushman, N. (2014). Learning to solve arithmetic word problems with verb categorization. In Moschitti, A., Pang, B., and Daelemans, W., editors, *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP 2014, October 25-29, 2014, Doha, Qatar, A meeting of SIGDAT, a Special Interest Group of the ACL*, pages 523–533. ACL.

Kushman, N., Artzi, Y., Zettlemoyer, L., and Barzilay, R. (2014). Learning to automatically solve algebra word problems. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 271–281, Baltimore, Maryland. Association for Computational Linguistics.

Roy, S. and Roth, D. (2015). Solving general arithmetic word problems. In Mrquez, L., Callison-Burch, C., Su, J., Pighin, D., and Marton, Y., editors, *EMNLP*, pages 1743–1752. The Association for Computational Linguistics.

Roy, S., Vieira, T., and Roth, D. (2015). Reasoning about quantities in natural language. *Transactions of the Association for Computational Linguistics*, 3:1–13.