

# 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 these 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 can be mapped

to basic operations(addition, subtraction, multiplication and division) and an equation can be created corresponding to the problem text.

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 based 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:
For Halloween Debby and her sister combined the candy they received. Debby had 32 pieces of candy while her sister had 42. If they ate 35 pieces the first night, how many pieces do they have left?

Table 1: Example Arithmetic Word Problem.

To derive the solution to this problem, the approach needs to understand that *they* refers to *Debby and her sister*. Hence, the number of candies need to be summed up and then *35 candies* need to be subtracted from the total number of candies.

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 represents 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, later the information decomposition strategy that is based on it is discussed. Section 4 presents the overall computational approach, including the way classifier learns to classify simplified sentences to operators. Finally, experimental study is discussed followed with a conclusion.

## 2 Related Work

Most of the previous work in automated arithmetic problem solvers has focused on a restricted subset of problems. The approach described in (Roy et al., 2015) handles problems with all basic operations but makes assumptions about the number of quantities in the problem text and the number of steps required to solve the problem. In contrast, our approach does not make any assumptions about the data in the problem text. Kushman’s approach (Kushman et al., 2014) tries to map numbers from the problem text to predefined equation templates and implicitly assume that similar equation forms have been seen in the training data. In contrast, our system does not rely on pre-defined templates and hence is able to solve questions of type which have never been seen before. The approach described in (Hosseini et al., 2014) might be the most related to ours. It handles addition and subtraction problems and tries to predict an operator for the verb in the problem text. Hence, it requires additional annotated data for verb categories. Our approach uses the information of the verb present in the sentence and handles addition and subtraction problems for now, but there is no requirement of additional annotated data.

Most of the methods mentioned above are able to solve problems with all basic operations, but our approach is easily generalizable to all the operators and also it performs competitively as compared to all other approaches.

### 3 Sentence Simplification and Problem Decomposition

Sentences in an arithmetic word problem are sometimes complex. Hence, its difficult to extract information from such sentences. More than extracting the information it is difficult to predict the impact of the sentence on the result.

Example 2:
Henry had 11 dollars. For his birthday he got 18 more dollars but spent 10 on a new game. How much money does he have now?

Table 2: Example Arithmetic Word Problem.

The second sentence in the above example is complex for a machine learning algorithm. It has addition and subtraction operation in a single sentence. Our idea is to simplify the sentence to multiple simple sentences so that each simplified sentence has a single operation. Below are the simplified sentences for the second sentence in the above example:

<b>For his birthday he got 18 more dollars but spent 10 on a new game.</b>
For his birthday he got 18 more dollars.
He spent 10 dollars on a new game.

Table 3: Simplified Sentences.

We create a mapping for each sentence in the problem text to its simplified sentences by extracting the relational dependencies for each sentence from the stanford dependency parser<sup>1</sup>. Multiple sentences are created based on the dependencies of the actual sentence. Currently, our simplification system simplifies sentences based on conjunctions and commas. There are certain rules when simplifying the sentence.

#### 3.1 Rules for simplifying sentences based on conjunctions.

When a conjunction is encountered, our simplification system attempts to create two simplified sentences from the actual sentence. The first sentence

---

<sup>1</sup><http://stanfordnlp.github.io/CoreNLP/>

is the part before the conjunction while the second sentence is the part after the conjunction.

But, after the split there may be some words which would be required in the second sentence. Consider the example in Table 4:

<b>Example Sentence.</b>
The school cafeteria ordered 42 red apples and 7 green apples for students lunches.

In the sentence above, the split based on conjunction 'and' will result in the first sentence with a Noun while the second will not have a Noun and a Verb. Hence, there are some rules for adding words to the second simplified sentence.

## 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.