VisualMath: An Automated Visualization System for Understanding Math Word-Problems

Utkarsh Dwivedi
IBM Research India
New Delhi, India
utkdwive@in.ibm.com

Prasenjit Dey
IBM Research India
New Delhi, India
prasenjit.dey@in.ibm.com

Nitendra Rajput
IBM Research India
New Delhi, India
rnitendra@in.ibm.com

Blessin Varkey
Autism Research Centre, Tamana
New Delhi, India
blessinvarkey@gmail.com

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.
Copyright is held by the owner/author(s).
IUI'17 Companion, March 13-16, 2017, Limassol, Cyprus
ACM 978-1-4503-4893-5/17/03.
http://dx.doi.org/10.1145/3030024.3040989

Abstract
Math word problems are difficult for students to start with since they involve understanding the problem's context and abstracting out its underlying mathematical operations. A visual understanding of the problem at hand can be very useful for the comprehension of the problem. We present a system VisualMath that uses machine learning tools and crafted visual logic to automatically generate appropriate visualizations from the text of the word-problems and solve it. We demonstrate the improvements in the understanding of math word-problems by conducting a user study and learning of meaning of relevant new words by students.

Author Keywords
Natural Language Processing, Automated Visualization.

ACM Classification Keywords
Information Interfaces and Presentation (e.g. HCI), Miscellaneous.

Introduction
Math word-problems are an important pedagogical tool to learn the use of mathematical operations in various real-life scenarios. However, students often find word-problems difficult to start with because of various

Figure 1: VisualMath experiment setup

CODE	EXPLANATION
R1	correct answer, correct interpretation
R2	Keyword based, but correct
R3	Guessing and correct
R4	needed help but answered correctly
W1	Correct interpretation, wrong math
W2	correct interpretation, no answer
W3	keyword based, but wrong answer
W4	wrong interpretation, wrong answer
W5	did not solve

Table 1: Coding scheme for recording the understanding and scoring of the student.

reasons such as, 1) the language is sometimes incongruent to their current level of understanding and vocabulary; 2) they cannot connect to the context of numbers and their operations; 4) they cannot solve the resulting equation or interpret the answer in terms of the problem [2, 4]. For some topics, well-designed word problems are easier to solve than the plain, equivalent equations for problems that are contextually relevant and appropriately represented [2, 3].
These efforts address the problem using verbs or words that the child is familiar yet un-exposed to, it has been found that in general diagrams or multiple representations of the problem can be very useful when shown alongside the text [3].
All these efforts however have been targeted at students who are in grades eight and above, where computer skills and language proficiency are developed to quite an extent. They have not specifically looked at the issues faced by young math learners who have been introduced to math word-problems for the first time, and do not have a well-developed vocabulary or have English as a second language.
We propose a system VisualMath that uses advancements in automated word problem solving and designed animation logic to generate appropriate visualizations from the text of the word-problems to aid in understanding the solution. The generated animation and visualizations are contextual to the text of the word-problem and uses multimodal animated representations to embody the verbs and the mathematical operations more effectively [5]. We conduct a user study for students in age group 8-10 years and report how student's understanding and ability to solve word-problems change with this system. Our work is novel in addressing this young-learner age group who have limited vocabulary and computer skills and have faced math word-problems for the first time.

Solution approach

We used ARIS [2] to solve the word problem which has a 77.7% accuracy across datasets. ARIS consists of four steps (1) training a model to classify verb categories for each sentence (2) grouping the problem into entities and containers, and (3) tracking and updating, using verb categories, the world states (4) forming equations to be solved and solving them.

Using ARIS

We trained ARIS on problems from Indian books and then chose problems from the scheme in Table 2 for which animations could be generated perfectly and the solving accuracy by ARIS was 100%. We used the verbclass that ARIS generated to cluster semantically close problems. These clusters would be used to analyze the similar questions, which would then be used to design the visualizations. We show an example result with the entities and the containers in Figure 2.

Designing visualizations

Currently we have designed visualizations for 4 types of questions, after analyzing the clusters generated. An example of each is given in Table 1 and the animation generating algorithm is shown in Figure 2.

Simple transfer *Type1*
"Ramesh has 5 apples. Rinky has 3 apples. Ramesh gave 3 apples to Rinky. How many apples does Ramesh have now?"

Two verb transfer *Type2a*
"Last year, 10 people were born in a country, and 4 people immigrated to it. How many people began living in the country last year?"
(Change within activity, here born and immigrate)

Two-time transfer *Type2b*
"Last year at Newberg 's airport, 12 airplanes landed on time. Unfortunately, 4 airplanes landed late. In all, how many airplanes landed in Newberg last year?"
(Change in time, here late and on time)

Total subtraction *Type3*
"A treasure hunter found a buried treasure chest filled with a total of 12 gems. Of the gems, there were 5 diamonds, and the rest were rubies. How many of the gems were rubies?"

Table 2: Questions types, with code and example

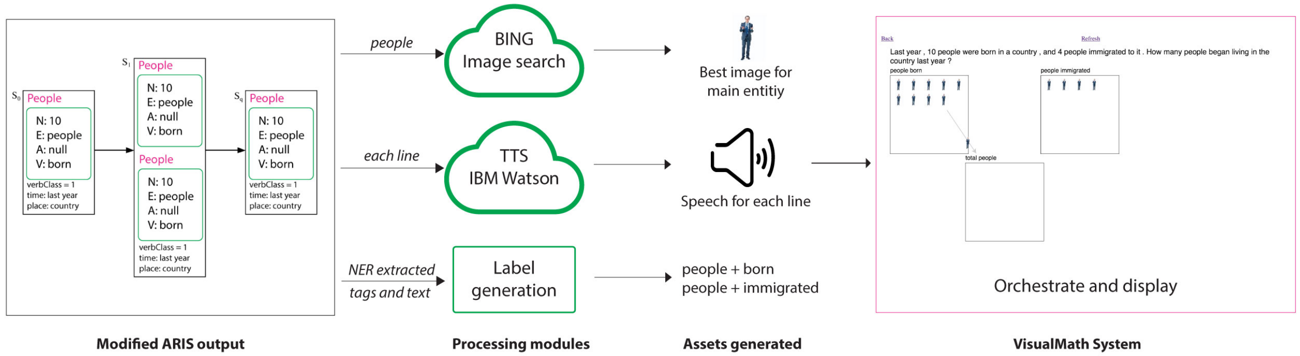

Figure 2: VisualMath uses ARIS to extract the main entity and the world states of the question, then get an image for the entity and audio for each line. It uses NER to generate label for each box. It reasons on the states, and orchestrates the audio and the animation of moving an entity, ie. all 'people' move from the 'born' and 'immigrate' boxes to 'total people' while each line been spoken aloud.

Experiment

To test the effectiveness of our system and the animation it generates automatically, we conducted an interview based study and assess student's understanding of certain types of word problems.

Users and Interface design

We run the experiments with 12 students in grades 2 and 3, with their ages ranging from 8 to 10 years. The average age was 9.5 years (5 males, 7 females). Informed consent for the study was received from the teachers. All the children spoke and read English and had faced math word problems recently. A coding scheme was developed to track the changes in the student's interpretation of the question and solution, described in Table 1.

The first screen is a question menu with a list of questions for the Visual Math system. Students can click on them to view the audio-visual rendering of the question. The system then synchronizes the animation and audio speak out with each line, as explained in Figure 2 and shown in Figure 1. Students can replay the visualization as many times as they want, go back to the menu or pause the animation with the spacebar.

Study design

We present each student with a pre-test of 8 problems 2 for each question type from NCERT math books, all students could solve atleast 60% of problems. Students who could solve the questions via keyword spotting or guesswork but had difficulty in articulating their understanding or reading were asked questions on quantities do they hold and how are those quantities changing. This is the interview strategy followed throughout all the experiments to get an understanding of the student's reasoning. Students solved though 8 questions with equitable distributions of each type through the VisualMath system. An extra Type1 question was used to demonstrate the interface. Before

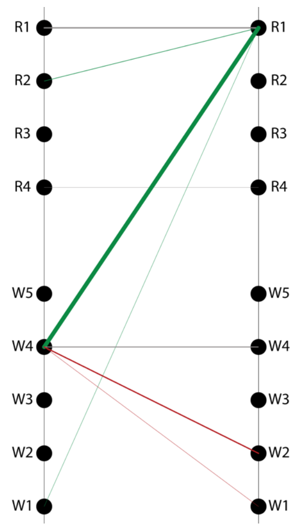

Figure 3: Change in understanding for Type 3 questions.

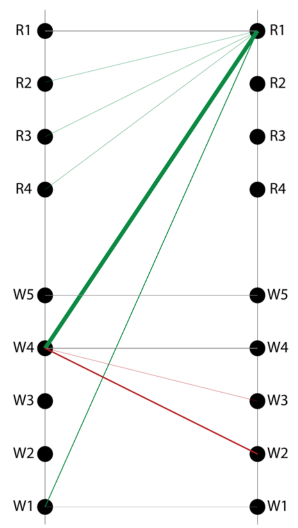

Figure 4: Change in understanding for Type 2 questions.

starting the system, students solved the question and their interpretation was coded as their initial understanding. Maximum of 5 trials for each question were performed or until the student solved. They gave a post-test of 8 questions with the same distribution, keeping the same interview strategy and were asked questions on their experience.

Results

We show an aggregate transition through each line in understanding of all questions of Type3 in Figure 3, Type2a and Type2b in Figure 4. Green lines denote *improved understanding*, red for *decrease* and grey for *no change*. The thickness of the line is the *percentage of students who made the transition.* Thus the total of green lines show a net increase in understanding.

We observed that most students could correct their interpretation of the Type3 questions after a few runs of the visual math system without any help (47% students for W4 to R1). Total students for Type2a, Type2b and Type3 that improved their interpretation and correctly answered, i.e the average of all green lines in both graphs was 55%.

Type 1 visualizations were solved easily (95%). For Type 2 questions 25% students could already guess the type of operation to be addition or guessed (R2, R3, R4 to R1 initially) but corrected their interpretation. More importantly 42% corrected themselves completely (W4 to R1). In the post-test they showed significant improvements ($p<0.02$) for Type2a questions, transitioning from any W code to R1.

Students reported an understanding of terms like ‘immigrated’. Type 3 visualizations were beneficial in making student understand what entities were asked for in the question, by identifying the possession within boxes for texts like ‘rest were oranges’. These are interesting observations for further exploration.

Conclusion and Future Work

These qualitative results encourage more investigation through this system that students can be benefited by such an automatic visualization approach. We will do a comparative study with plain text book visuals, add novel input modalities to the system and improve the aptness of the labels in the future.

Acknowledgements

We thank the staff and teachers at Tamana School in helping us conduct the user study.

References

1. Hosseini, M. J., Hajishirzi, H., Etzioni, O., & Kushman, N. (2014). Learning to Solve Arithmetic Word Problems with Verb Categorization. In *EMNLP* (pp. 523-533).
2. Koedinger, K. R., and Nathan, M. J. The real story behind story problems: Effects of representations on quantitative reasoning. *Journal of the Learning Sciences 13*, 2 (2004), 129–164.
3. Yung, H. I., and Paas, F. Effects of computer-based visual representation on mathematics learning and cognitive load. *Journal of Educational Technology & Society 18*, 4 (2015), 70.
4. Gooding, S. (2009). Children's Difficulties with Mathematical Word Problems. *Proceedings of the British Society for Research into Learning Mathematics*, *29*(3), 31-36.
5. Mayer, R. E., & Moreno, R. (2002). Animation as an aid to multimedia learning. *Educational psychology review*, *14*(1), 87-99.