Dolphin18K for Math Word Problem Solving

1. Introduction

Dolphin18K is a dataset created for math word problem solving, containing 18,460 problems posted by users on the community question answering site, Yahoo! Answers.

2. Data Description

(1) File Format

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>id</td>
<td>Unique id of a problem.</td>
<td>yahoo.answers. 20061213041448AAxoy5z</td>
</tr>
</tbody>
</table>
| original  | Original problem text from yahoo answer.        | what is 30 divided by half plus 10? 1st correct answer gets 10 points!!!?
| text      | Cleaned problem text by removing problem unrelated text, used in our paper. | what is 30 divided by half plus 10? 1st correct answer gets 10 points!!!?
| equations | The equations used to solve the problem.         | unkn: x equ: x=30/(1/2)+10                                              |
| ans       | Problem answer.                                  | 70                                                                     |
| flag      | Annotation status of the problem                | 0 – annotation done                                                     |

(2) Equation Annotation

- “unkn”: the variables whose values will be problem answers
- “equ”: <the first equation>
- “equ”: <the second equation>
...
- “equ”: <the last equation>

(3) Answer Annotation
- “,”: separate the values of different variables.
- “{}`: any order of variables is allowed.
- “|”: different answer formats.
- “or”: different solutions to the problem.

See the examples below:

<table>
<thead>
<tr>
<th>Problem</th>
<th>Equations</th>
<th>Answer</th>
</tr>
</thead>
</table>
| 3 times the sum of twice a number and 5 is 4 times 2 less than the number. Find the number. | unkn: x 
equ: 3*(2x+6)=4(x-2) | -13 |
| Find two consecutive integers whose sum is 27. | para: x 
unkn: x,x+1 
equ: x+(x+1)=27 | 13; 14 |
| Find 2 numbers whose sum AND product are 11. | unkn: x,y 
equ: x + y = 11 
equ: x*y = 11 | {1.113; 9.887} |
| 60 is 140% of what number? | unkn: x 
equ: 60 = 140/100 * x | 300/7 | 42.857 |

3. Dataset Statistic
(1) Manual/Auto

Manual: equations and answers are annotated by human
Auto: equations and answers are annotated automatically
(method of automatic labeling is introduced in [1])

(2) Template Size

Template: Unique form of equation system.
For example, the following equation x=3+5 corresponds to
template x=n_1+n_2.
Template size = number of problems corresponding to the
template.

<table>
<thead>
<tr>
<th></th>
<th>Equations + Answer</th>
<th>Answer only</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dev</td>
<td>Manual</td>
<td>909</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>Auto</td>
<td>2,245</td>
<td>507</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>3,154</td>
<td>574</td>
</tr>
<tr>
<td>Eval</td>
<td>Manual</td>
<td>3,605</td>
<td>321</td>
</tr>
<tr>
<td></td>
<td>Auto</td>
<td>8,754</td>
<td>2,052</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>12,359</td>
<td>2,373</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Template size</th>
<th>Manual</th>
<th>Auto</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;= 1</td>
<td>2,675</td>
<td>7,969</td>
<td>10,644</td>
</tr>
<tr>
<td>&gt;= 2</td>
<td>2,036</td>
<td>5,956</td>
<td>8,229</td>
</tr>
<tr>
<td>&gt;= 6</td>
<td>1,578</td>
<td>4,826</td>
<td>6,827</td>
</tr>
</tbody>
</table>

4. Get the dataset

(1) Download original text from Yahoo! Answers

Run OriginalDataExtractor.py
• Arguments:
  -i input_url_file –o output_file –t number_of_threads (default 10)
  • Example command line:
  python OriginalDataExtractor.py –i dev_urls.json –o
dev_original.json –t 10

(2) Get the cleaned-text version
  Run CleanVersionExtractor.py
  • Arguments:
    -i input_orginial_file –d input_diff_file –o output_cleaned_file
  • Example command line:
    python CleanVersionExtractor.py –i dev_original.json –d
dev_diff.pkl –o dev_cleaned.json

(3) Get different subsets
  Run SubsetExtractor.py
  • Arguments:
    -i input_cleaned_file –s subset_id_file –o output_subset_file
  • Example command line:
    python SubsetExtractor.py –i dev_cleaned.json –s
dev_ids\dev_manual.txt –o dev_ids\dev_manual.json

Reference:
[1] Danqing Huang, Shuming Shi, Chin-Yew Lin, Jian Yin and Wei-Ying
Ma. How Well Do Computers Solve Math Word Problems? Large-Scale
Dataset Construction and Evaluation. ACL 2016.