国际人工智能会议 AAAI 2021论文北京预讲会

A Bottom–Up DAG Structure Extraction Model for Math Word Problems

曹逸轩、洪峰、李宏伟、罗平中科院计算所中科院智能信息处理重点实验室

Math Word Problems

- Problem :
 - The ages of Tom and his father are in the ratio of 1: 5, half of their sum is 24. Find their ages.
- Equations :
 - $x \div y = 1 \div 5$, $\frac{1}{2} \times (x + y) = 24$
- Answer :

Is math word problem hard?





Solving Math Word Problems



Categories of Word Problem

• Arithmetic word problem (one variable, output expression)

```
Robin was making baggies of cookies with 6 cookies in each bag.
If she had 23 chocolate cookies and 25 oatmeal cookies, how
many baggies could she make?
```

(13 + 17) ÷ 5

• Algebra word problem (multiple variables, output equations)

The ages of Tom and his father are in the ratio of 1: 5, half of their sum is 24. Find their ages.

$$\begin{cases} x \div y = 1 \div 5 \\ \frac{1}{2} \times (x + y) = 24 \end{cases}$$

Related Work

Robin was making baggies of cookies with n_1 cookies in each bag. If she had n_2 chocolate cookies and n_3 oatmeal cookies, how many baggies could she make?





What is missed?

• 1. Algebra word problems (multiple variables)

The ages of Tom and his father are in the ratio of 1: 5, half of their sum is 24. Find their ages.

 $\begin{cases} x \div y = 1 \div 5 & \text{is a Direct Acyclic Graph (DAG)} \\ \frac{1}{2} \times (x + y) = 24 & \text{Tree structure is not enough} \end{cases}$



• 2. Mathematical Properties

commutative law: $a + b \Leftrightarrow b + a$

 $a \times b \Leftrightarrow b \times a$

Seq2Seq models does not satisfy commutative law

$$(n_2 + n_3)$$

 $(n_3 + n_2)$
different representations
with equivalent meaning
 $(n_2 + n_3) \div n_1$
 $(n_3 + n_2) \times n_1$
may result in different decisions

Bottom-Up DAG Structure Extraction



Bottom-Up DAG Structure Extraction

• In Bottom-Up fashion, layer by layer



Model



Experiments

DRAW1K

1,000 algebra word problems >200 templates

	Model	Acc.(%)
Similarity	SIM (Huang et al. 2016)	25.5*
Template	KAZB (Kushman et al. 2014)	43.2*
-	MixedSP (Upadhyay et al. 2016)	59.5
Deep	DNS (Wang, Liu, and Shi 2017)	31.0*
Learning	Seq2DAG	44.4
	w/o BERT	37.5

Table 2: Results on DRAW1K.

Math23K

23,161 arithmetic word problems >2,000 templates

Model	Acc.(%)
DNS (Wang, Liu, and Shi 2017)	64.7
T-RNN (Wang et al. 2019)	68.7
Ensemble (Wang et al. 2018)	68.4
Seq2Tree (Xie and Sun 2019)	74.3
Graph2Tree (Zhang et al. 2020)	75.5
Seq2DAG	77.1
-Attention	76.1
-BERT	72.5

Table 3: Results on Math23K

Commutative Law

	Acc. (%)
Consider CL	44.4
Not consider CL	41.9
Not consider CL & order disruption	41.0

Table 5: Results on whether considering the commutative law (CL).

Case Study

- Case 1
 - A park has 26 boats, and the rent revenue is \$910 per day. If we add 6 more boats, how much rent the park will get each day?
 - Ground truth: 910 \div 26 \times (26 + 6)
 - Prediction : $(910 \div 26) \times (26+6); 910+((910 \div 26) \times 6))$
- Case 2
 - Ben is 3 years younger than Dan. The sum of their ages is 53. How old is each?
 - Ground truth: y=x+3, x+y=53
 - Prediction : y=x-3, x+y=53
- Case 3
 - In Longgang orchard, the number of apple trees is (2/5) of pear trees, and (3/4) of peach trees. There are 480 pear trees in total. How many peach trees are there?
 - Ground truth: $480 \times (2/5) \div (3/4)$
 - Prediction : $480 \div (2/5) \div (3/4)$

A General Framework

- This bottom-up DAG extraction method is a **General** framework that has been used in other tasks:
 - Formula extraction

Towards Automatic Numerical Cross-Checking: Extracting Formulas from Text. In WWW 2018.

• Nested Causal Relation Extraction

Nested Relation Extraction with Iterative Neural Network, In CIKM 2019.

Future work

- Consider more mathematical properties
 - Associative law
 - Consider the order of variables

 $\begin{array}{c} \begin{array}{c} x = \text{Tom's age} \\ y = \text{Dad's age} \end{array} \begin{array}{c} \end{array} \begin{array}{c} \begin{array}{c} y = \text{Tom's age} \\ x = \text{Dad's age} \end{array} \end{array}$

- Incorporate with other encoders
 - GNN

国际人工智能会议 AAAI 2021论文北京预讲会

THANKS

2020.12.19

Case Study

- Case 1
 - •公园有26条船,每天的租金收入是910元,如果增加6条船,每天的租金是多少?
 - •标注结果: 910 ÷ 26 × (26 + 6)
 - 预测结果: (910 ÷ 26)×(26+ 6); 910+ ((910 ÷ 26)×6)
- Case 2
 - •小明比小红小3岁,他们年龄之和是53岁,他们的年龄分别是多少?
 - •标注结果:y=x+3,x+y=53
 - 预测结果: y=x-3, x+y=53
- Case 3
 - 在长龙果园里, 苹果树的数量是梨树的 2/5, 是桃树的3/4。果园共有梨树 480棵, 问有多少桃树?

 - •标注结果: $480 \times \frac{2}{5} \div \frac{3}{4}$ •预测结果: $480 \div \frac{2}{5} \div \frac{3}{4}$