SMART: A Situation Model for Algebra Story Problems via Attributed Grammar

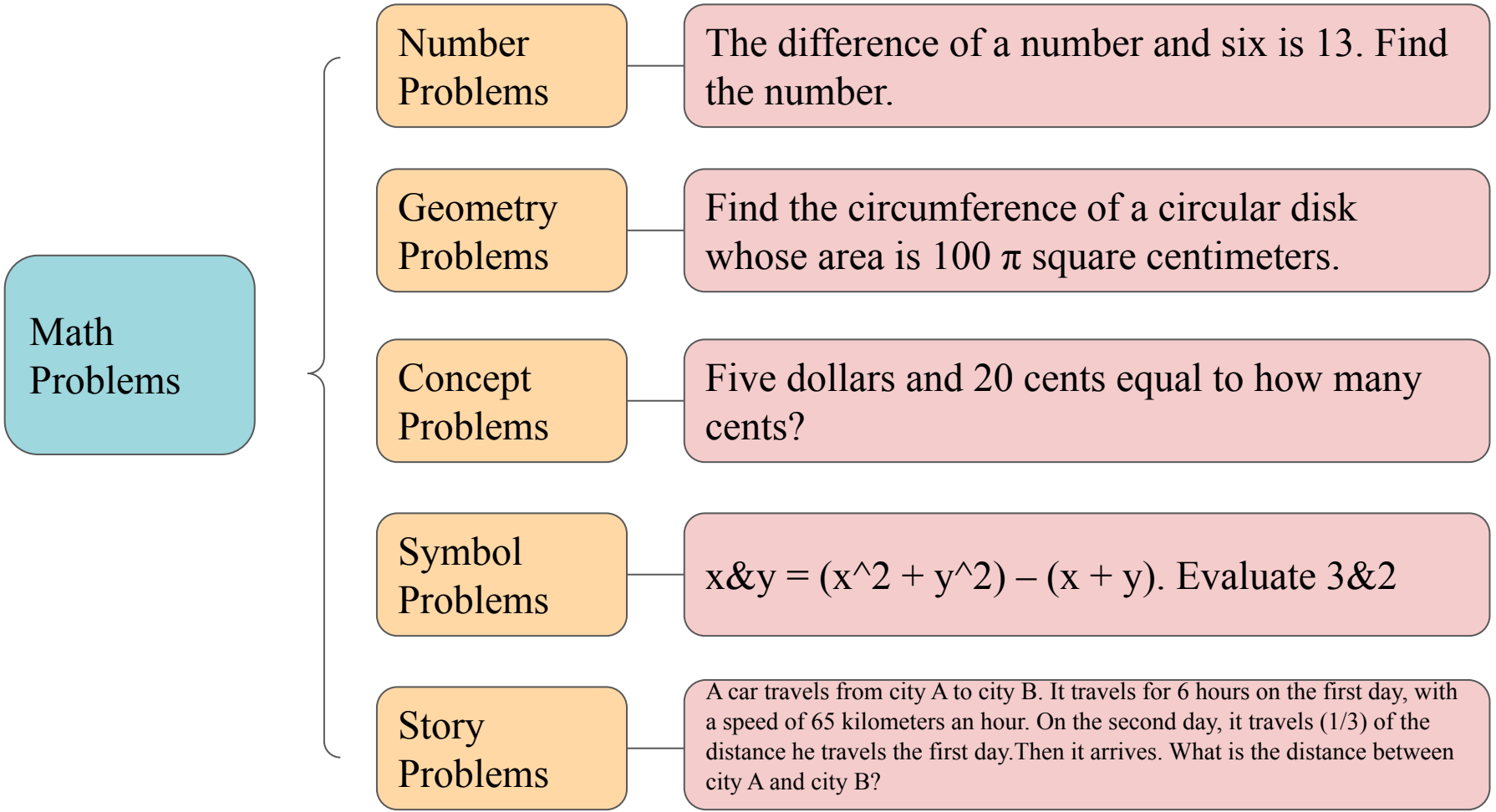





Yining Hong[1] Qing Li[2] Ran Gong[1] Daniel Ciao[1] Siyuan Huang[2] Song-Chun Zhu[1,2]

[1] Department of Computer Science, UCLA

[2] Department of Statistics, UCLA


Math Problems
Number Problems
The difference of a number and six is 13. Find the number.
Geometry Problems
Find the circumference of a circular disk whose area is 100 π square centimeters.
Concept Problems
Five dollars and 20 cents equal to how many cents?
Symbol Problems
x&y = (x^2 + y^2) – (x + y). Evaluate 3&2
Story Problems
A car travels from city A to city B. It travels for 6 hours on the first day, with a speed of 65 kilometers an hour. On the second day, it travels (1/3) of the distance he travels the first day.Then it arrives. What is the distance between city A and city B?

Algebra Story Problems

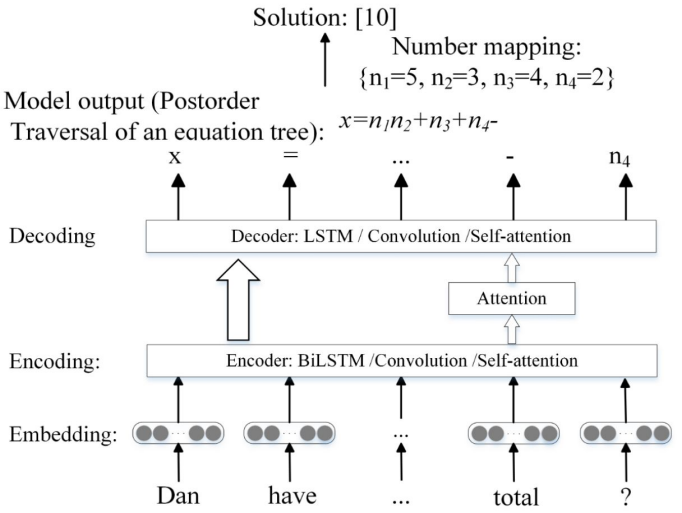
An Exemplar Problem: A car travels from city A to city B. It travels for 6 hours on the first day, with a speed of 65 kilometers an hour. On the second day, it travels (1/3) of the distance he travels the first day.Then it arrives. What is the distance between city A and city B?

- Describe a real-world situation
- Inquire about an unknown attribute in the situation
- The problem has a story-line consisting of characters, objects, and actions. (Definitions by [1])

Although algebra story problems are distinguished, related works often mix algebra story problems with other types of problems into one whole task.

[1] Mitchell J. Nathan , Walter Kintsch & Emilie Young. A Theory of Algebra-Word-Problem Comprehension and Its Implications for the Design of Learning Environments.

Neural Networks for Math Word Problems

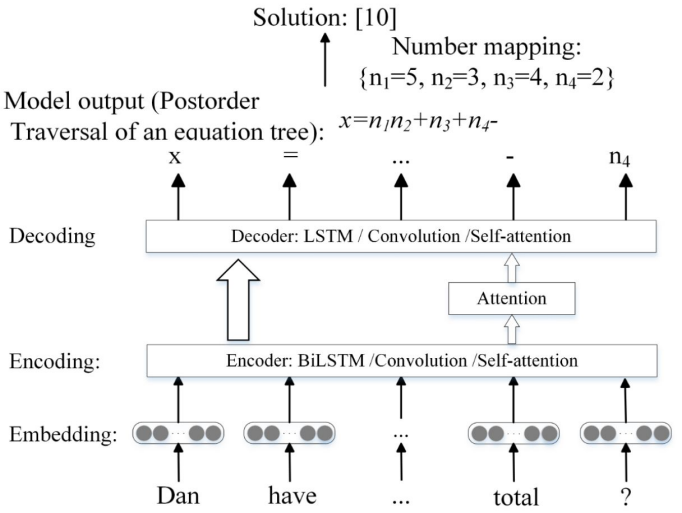


Model input (Problem text after number mapping): Dan has n_1 pens and n_2 pencils, Jessica has n_3 more pens and n_4 less pencils than him. How many pens and pencils does Jessica have in total?

Lack of generalization ability. These neural solvers usually fail in scenarios that are more sophisticated than those in training.

Problem	GTS
A road repair team needs to repair a section of road. The first day it repairs (1/5) kilometers, the second day they repairs (3/10) kilometers, and the road they already repairs is (1/7) of the total length. How many kilometers is the total length of the road?	((1/5)+(3/10))/(1/7) ✓
An engineering team is building a the road. On the first day, it builds (2/5) of the total length. On the next day, it builds 24 meters more than (3/10) of the remaining length. On the third day, it builds 60 meters (3/ 4) the length of the first day, and then they finish. How many meters is the length of the road?	(60-24)/ ((1-(2/5)) -(3/4))-24 ✗

Neural Networks for Math Word Problems



Model input (Problem text after number mapping): Dan has n_1 pens and n_2 pencils, Jessica has n_3 more pens and n_4 less pencils than him. How many pens and pencils does Jessica have in total?

Lack of interpretability. The annotated expressions (as well as the expressions generated by neural networks) are hard to interpret without the intermediate problem-solving process.

An engineering team is building a the road. On the first day, it builds (2/5) of the total length. On the next day, it builds 24 meters more than (3/10) of the remaining length. On the third day, it builds 60 meters (3/ 4) the length of the first day, and then they finish. How many meters is the length of the road?	(24+60)/[1-(1 -(2/5))*(3/10) -(2/5)*(3/4)-(2/5)]

Dataset: *ASP6.6k*

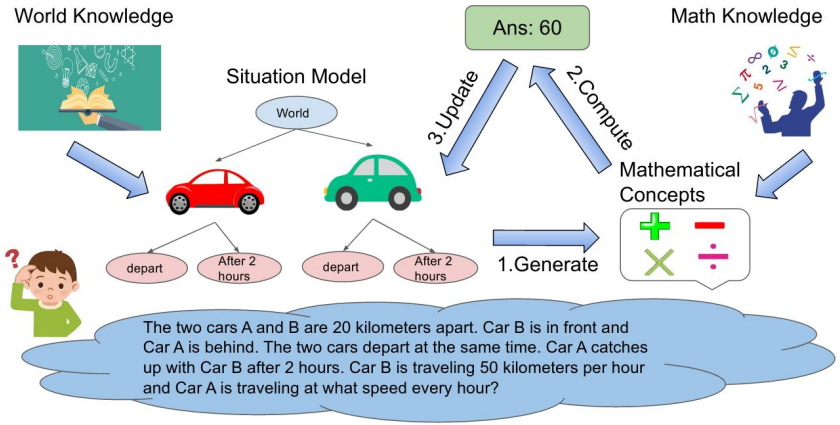
We curate a dataset from Math23K[2] following the criterion of [1,3]. We obtain a dataset of ASP6.6k 6666 problems spanning from 4 typical types of algebra story problems: motion, price, relation and task completion.

Problem Type	Sample Problem
Task Completion	The engineering team built a viewing trail and completed 30% of the full length in the first week and 45% of the full length in the second week. 150 meters in two weeks, how long is the length of this trail?
Motion	Mingming's family went to travel, they took a 14-hour train ride, and then a 5-hour car ride before reaching their destination. It is known that the speed of the train is 120 kilometers/hour and the speed of the car is 60 kilometers/hour. How long is this journey?
Relation	Xiaogang's weight is 28.4 kg, Xiaoqiang's weight is 1.4 times that of Xiaogang, Xiaoqiang's weight = how many kilograms?
Price	The school bought 45 sets of desks and chairs at 128 yuan per desk and 52 yuan per chair. How much did it spend?

[2]Yan Wang and Xiaojiang Liu and Shuming Shi, Deep Neural Solver for Math Word Problems. 2017
[3]Mayer, Richard E, Frequency norms and structural analysis of algebra story problems into families, categories, and templates. 1981

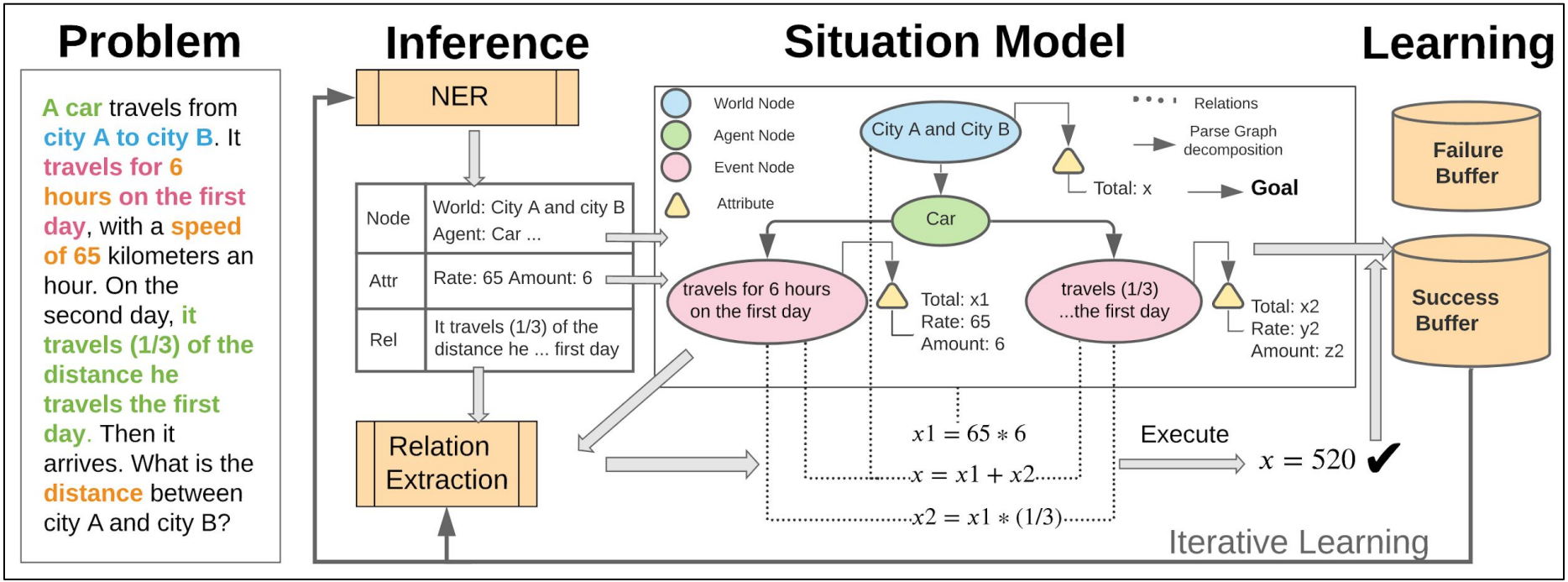
Situation Model

Definition: The situation model is frequently used in psychology and cognitive science for algebra story problem solving. The situation model[4] is used to represent agents, actions, and events when humans solve algebra story problems [5,6,7,8,9]. It is believed that problem-solving techniques, such as mathematics and logic, are applied to the situation model instead of the problem text. The situation model often interacts with a math conceptual model, which in turn embeds relations in the situation model.

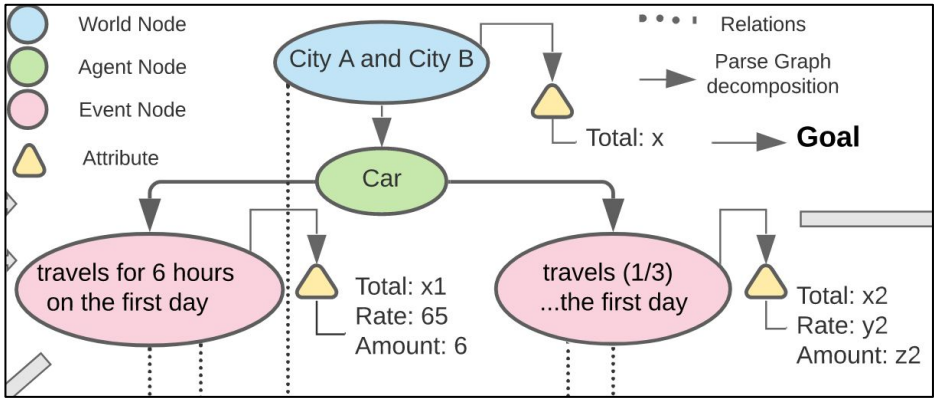


[4]van Dijk, T. A. and Kintsch, W, Strategies of discourse comprehension. 1983
[5]Reusser1990FromTT, From text to situation to equation: cognitive simulation of understanding and solving mathematical word problems. 1990
[6]Greeno, J. G., Situation models, mental models, and generative knowledge,.1989
[7]Nathan, M. J. and Young, E, Thinking situationally: Results with an unintelligent tutor for word algebra problems. 1991
[8]Daniele Coquin-Viennot and S. Moreau Arithmetic problems at school: when there is an apparent contradiction between the situation model and the problem model, 2007
[9]D. Leiss and S. Schukajlow and W. Blum and R. Messner and R. Pekrun, The Role of the Situation Model in Mathematical Modelling—Task Analyses, Student Competencies, and Teacher Interventions. 2010

SMART: Situation Model implemented by Attributed Grammar



Attributed Grammar



$G = (S, V, A, E, R)$

S is the start symbol.

V = {S, World, Agents, Agent, Events, Event}

A = {rate, amount, total}

E = {e: e is a valid equation on attributes.}

R = {$S \rightarrow$ World

World $\rightarrow$ Agents

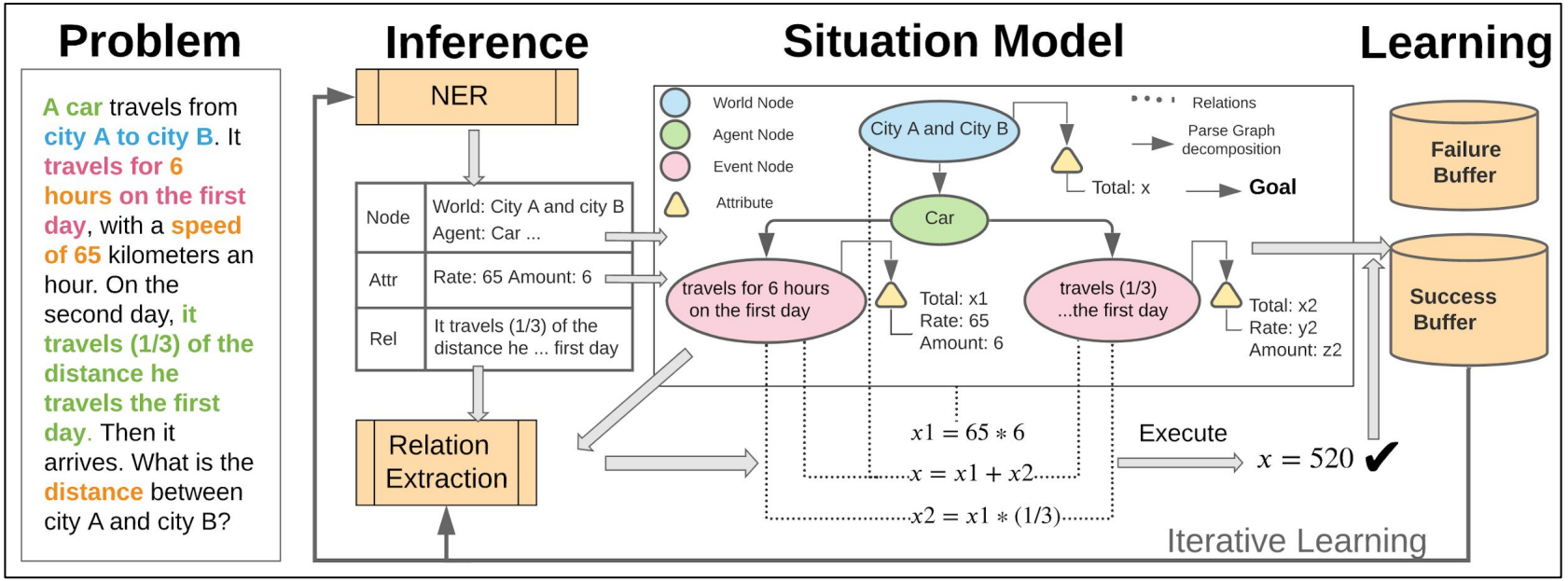
Agents $\rightarrow$ Agents Agent | Agent

Agent $\rightarrow$ Events

Events $\rightarrow$ Events Event | Event}

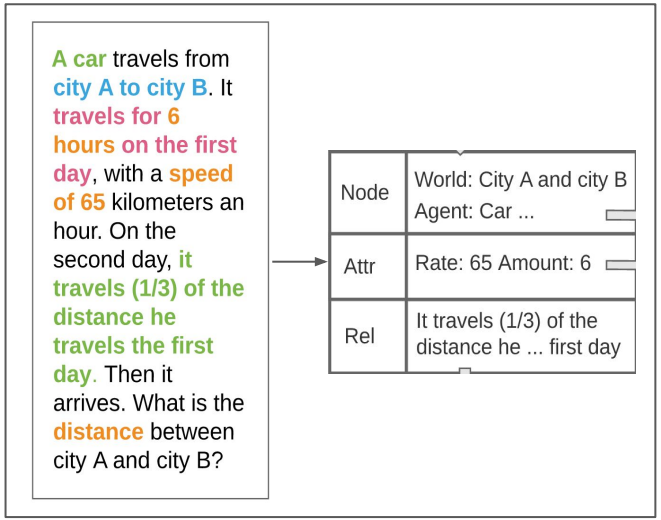
$$pg^* = \arg \max_{pg \in \mathcal{L}(G)} p(pg \mid x)$$

Grammar Parsing



$$p(pg \mid x) = p(V_{pg}, A_{pg}, E_{pg}|x)$$
$$= p(V_{pg}|x) \cdot p(A_{pg}|x) \cdot p(E_{pg}|x)$$

Grammar Parsing



$$p(V_{pg}|x) = \frac{1}{|V_{pg}|} \sum_{w \in V_{pg}} p_{ner}(w)$$

$$p(A_{pg}|x) = \frac{1}{|A_{pg}|} \sum_{w \in A_{pg}} p_{ner}(w)$$

Attributes Extraction: BERT-base NER

Nodes Extraction & Relation Span Extraction: BERT-base Nested NER

Grammar Parsing

Relation Translation: Seq2seq

$$p(E_{pg}|x) = \sum_{e \in E_{pg}} p_{seq2seq}(e|\text{REL}, V_{pg}, A_{pg})$$

Example Inputs & Outputs for Seq2seq Model:

Problem	Seq2Seq Input	Seq2Seq Output
Ming reads a story book. The first day he reads (1/6) of the whole book. The second day he reads 24 pages. And the third day he reads 150% of the total number of page read in the previous two days. There's (1/4) left of the whole book, so how many pages does this book have?	The first day he reads $(1/6)$ of the whole book {{world: {a story book: {quantity: x}}, agent: {Ming}, event: {The first day...the whole book: {quantity: x_1}}}}	$x_1 = x * (1/6)$
	the third day he reads 150% of the total number of pages read in the previous two days.{{world: ..., agent: .., event: {The first day: {quantity: x_1}}, {The second day: {quantity: 24}}, {The third day: {quantity: x_3}}}}	$x_3 = (x_1 + 24) * 150\%$
	There's (1/4) left of the whole book {{world: ..., agent: .., event: {The first day: ..., {The second day: ..., {The third day: ...}}}	$x - (x_1 + 24 + x_3) = x * (1/4)$
A car drives from City A to City B, traveling 217 kilometers in 3.5 hours in the morning. At this speed, it will take another 4.5 hours to reach City B. How many kilometers is the distance between City A and City B?	At this speed {{world: City A and City B{quantity: x}}, agent: car, event: {traveling 217 kilometers in 3.5 hours in the morning: {quantity: 217, unit: 3.5, speed: y_1}}, {it will take another 4.5 hours to reach City B.: {quantity:x_2 unit: 4.5, speed:y_2}} }}	$y_1 = y_2$
	reach city B {{... event: {traveling ... morning: {quantity: 217, unit: 3.5, speed: y_1}}, {it will ... City B.: {quantity:x_2, unit: 4.5, speed:y_2}} }}	$x = 217 + x_2$

Initial Supervision

It’s time-consuming to annotate all the graphs. Therefore, we propose a rule-based initial parser to provide some pseudo annotations.

Initial Supervision

Attribute Extraction:

Problem: Each kilogram of pears cost 3.65 dollars. How many dollars does mom have to pay for 13 kilograms of pears?	***rate***	***amount***	***total***
	3.65	13	how many
	Num Unit	**Den Unit**	
	dollar	kilometer	

Node Extraction:

Based on Dependency Parsing, Pos Tagging and Regular Expressions.

Initial Supervision

Relation Extraction as First Order Logic:

- Variables: a node v.
- Functions: Rate(v), Amount(v), Total(v). Sum(Total(v_i), Total(v_j)); Left(Total(S), Total(v_i), Total(v_j)).
- Predicates: Equal(F(v_i),F(v_j)); More_than(F(v_i),F(v_j),n);Less_than(F(v_i),F(v_j),n); Times_of(F(v_i),F(v_j),n)

Predicate	Definition
Equal $(F(v_i), F(v_j))$	$F(v_i) = F(v_j)$
More_than$(F(v_i), F(v_j), n)$	$F(v_i) = F(v_j) + n$
Less_than$(F(v_i), F(v_j), n)$	$F(v_i) = F(v_j) - n$
Times_of$(F(v_i), F(v_j), n)$	$F(v_i) = F(v_j) * n$

Problem Solving

- The goal of a problem is usually an interrogative word extracted by NER (e.g., how many). It can be in an attribute or in REL. We transform the interrogative word into an unknown in the equation.

  e.g., What is the speed of car A -> Goal: y1

  How much greater is car A's speed than car B's speed? -> More_than(Rate(CarA), Rate(Car, B),n) -> Goal: n

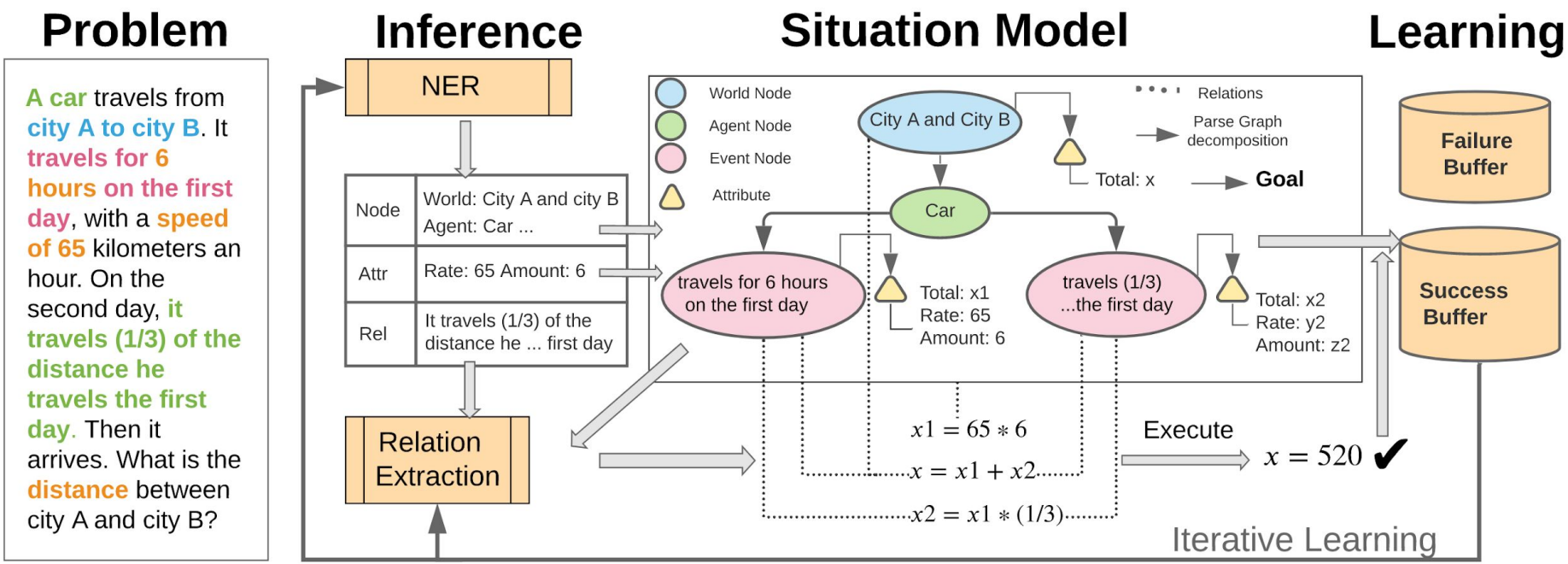
- Use SymPy to solve the equations.

Iterative Learning

Algorithm 1 Iterative Learning

Input: training set $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{N}$
Success buffer $\mathcal{B}$, Failure buffer $\mathcal{F}$, updated parser θ
▷*Parse Graph Proposal*
for $x_i, y_i \in \mathcal{D}$ **do**
    pg_i = initial_parser (x_i)
    if execute$(pg_i) = y_i$ **then**
        $\mathcal{B} \leftarrow \mathcal{B} \cup \{x_i, pg_i\}$
    else
        $\mathcal{F} \leftarrow \mathcal{F} \cup \{x_i, y_i\}$
▷*Iterative Learning*
while not converge **do**
    for $x_i, pg_i \in \mathcal{B}$ **do**
        $\theta = \theta - \nabla_\theta J(x_i, pg_i)$
    for $x_i, y_i \in \mathcal{F}$ **do**
        pg_i = updated_parser (x_i)
        if execute$(pg_i) = y_i$ **then**
            $\mathcal{B} \leftarrow \mathcal{B} \cup \{x_i, pg_i\}$
            remove $\{x_i, y_i\}$ from $\mathcal{F}$

SMART: Situation Model implemented by Attributed Grammar



Experiment

- Dataset: curated ASP 6.6k. 80/20 train/test split.
- Models:
  - MathEN[10]
  - Group-ATT[11]
  - GTS[12]
  - Graph2Tree[13]

[10] Translating a Math Word Problem to a Expression Tree. Lei Wang, Yan Wang, Deng Cai, Dongxiang Zhang and Xiaojiang Liu.
[11] Template-Based Math Word Problem Solvers with Recursive Neural Networks. Lei Wang, Dongxiang Zhang, Jipeng Zhang, Xing Xu, Lianli Gao, Bing Tian Dai and Heng Tao Shen.
[12] Goal-Driven Tree-Structured Neural Model for Math Word Problems. Zhipeng Xie and Shichao Sun.
[13] Graph-to-Tree Learning for Solving Math Word Problems. Jipeng Zhang, Lei Wang, Roy Ka-Wei Lee, Yi Bin, Yan Wang, Jie Shao and Ee-Peng Lim.

Experimental Result

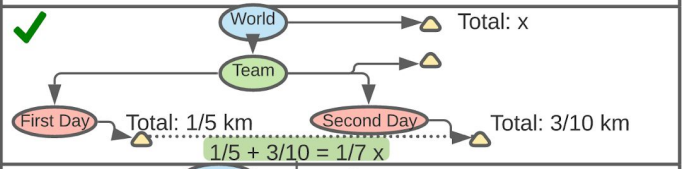
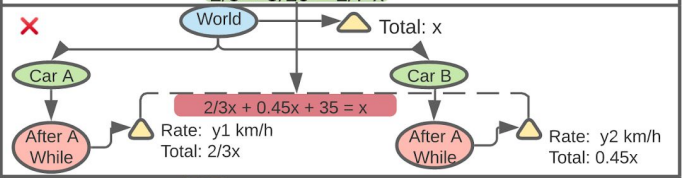
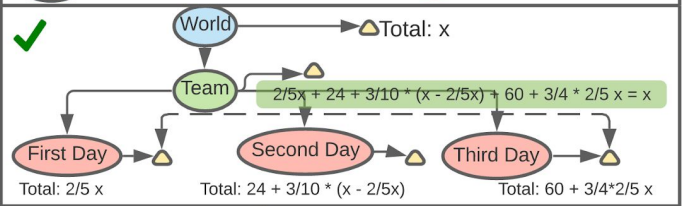
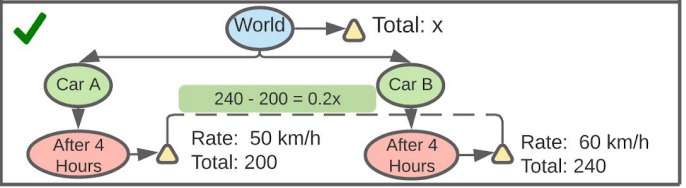
Model	Overall	Motion	Task	Relation	Price
MathEN	67.8	68.3	70.2	63.3	70.5
GroupATT	67.4	65.2	70.7	63.6	71.5
GTS	76.8	73.2	72.1	76.0	**83.6**
Graph2Tree	76.8	76.9	**79.0**	73.8	78.7
SMART	**79.5**	**79.8**	**79.0**	**77.9**	81.8

Table 5: The answer accuracy on the test set (%).

Model	Overall	Motion	Task	Relation	Price
MathEN	31.7	22.6	28.9	39.9	33.2
GroupATT	35.0	24.0	42.2	42.6	32.7
GTS	45.8	44.5	41.9	49.9	45.3
Graph2Tree	45.1	34.1	47.4	55.1	41.9
SMART	**63.2**	**65.0**	**64.8**	**62.9**	**60.8**

Table 6: The answer accuracy in the OOD evaluation (%). The test set is the 20% longest problems of each type.

Qualitative Study

Problem	Expression	GTS	SMART
A road repair team needs to repair a section of road. The first day it repairs (1/5) kilometers, the second day they repairs (3/10) kilometers, and the road they already repairs is (1/7) of the total length. How many kilometers is the total length of the road?	((1/5)+(3/10))/(1/7)	((1/5)+(3/10))/(1/7) ✓	
Car A and car B set off from city A and city B at the same time, and head toward each other. After a while Car A travels (2/3) of the entire journey, and Car B travels 45% of the entire journey. At this time, the two vehicles were 35 kilometers apart. How many kilometers are between two city A and city B?	35/((2/3)+45%-1)	35/(1-(2/3)) ✗	
An engineering team is building a the road. On the first day, it builds (2/5) of the total length. On the next day, it builds 24 meters more than (3/10) of the remaining length. On the third day, it builds 60 meters (3/ 4) the length of the first day, and then they finish. How many meters is the length of the road?	(24+60)/[1-(1 -(2/5))*(3/10) -(2/5)*(3/4)-(2/5)]	(60-24)/ ((1-(2/5)) -(3/4))-24 ✗	
Two cars drives out from city A and city B at the same time. Car A travels 50 kilometers per hour, and car B travels 60 kilometers per hour. After 4 hours, the distance between the two cars was 20% of the total distrance. How many kilometers are the total distance between city A and city B?	(50+60) *4/(1-20%)	50*4 ✗	

Conclusions

- We distinguish story problems from other types of math problems, and demonstrate why algebra story problems need to be studies alone.
- We introduce the cognitive concept of situation model, and propose a model based on attributed grammar to implement situation model.
- Our model achieves better interpretability and generalization ability than neural models.

You are welcomed to visit our website:

https://evelinehong.github.io/smart-site/