

# Learning to Automatically Solve Algebra Word Problems

**Nate Kushman**

Yoav Artzi, Luke Zettlemoyer, Regina Barzilay

# Task

## Automatically Solve Algebra Word Problems

*An amusement park sells 2 kinds of tickets. Tickets for children cost \$1.50. Adult tickets cost \$4. On a certain day, 278 people entered the park. On that same day the admission fees collected totaled \$792. How many children were admitted on that day? How many adults were admitted?*

Goal:

**Generate Numerical Answers**

128

150

Two Training Scenarios:

**Full Equations**

$$X + Y = 278$$

$$1.5 * X + 4 * Y = 729$$

**Numerical Answers**

128

150

# Wide Variety of Problems

## Interest

An investor will invest a total of 15000 dollars in 2 accounts , one paying 4 % annual simple interest and the other 3 % . If he wants to earn 550 dollars annual interest , how much should he invest at 4 %? How much at 3 %?

$$3.0*0.01*X+4.0*0.01*Y=550.0$$
$$X+Y=15000$$

## Traveling Apart

Two airplanes left the same airport traveling in opposite directions. If one airplane averages 400 miles per hour and the other 250 miles per hour , how many hours will it take for the distance between them be 1625 miles?

$$(250.0*X)+(400.0*X)=1625.0$$

## Math Problems

A math test is worth 100 points and has 30 problems. Each problem is worth either 3 points or 4 points. How many 4 point problems are there?

$$X + Y = 30$$
$$3*X + 4*Y = 100$$

## Ticket Purchase

An amusement park sells 2 kinds of tickets. Tickets for children cost \$1.50. Adult tickets cost \$4. On a certain day, 278 people entered the park. On that same day the admission fees collected totaled \$792. How many children were admitted..

$$X + Y = 278$$
$$1.5*X + 4*Y = 729$$

## Ratio

A writing workshop enrolls novelists and poets in a ratio of 5 to 3. There are 24 people at the workshop. How many novelists are there? How many poets are there?

$$24 = X+Y$$
$$3.0*X=5.0*Y$$

## Fixed+Variable

Sunshine Car Rentals rents a basic car at a daily rate of 17.99 dollars plus 0.18 per mile. City Rentals rents a basic car at 18.95 dollars plus 0.16 per mile. For what mileage is the cost the same?

$$17.99 + 0.18*X = 18.95 + 0.16*X$$

## Coffee Beans

Colombian coffee beans cost 5.50 dollars per pound, while Peruvian coffee beans cost 4.25 dollars per pound. We want to mix the beans together so as to produce a 40-pound bag , costing 4.60 dollars per pound. How many pounds of Colombian...

$$(5.5*X)+(4.25*Y)=40.0*4.6$$
$$X+Y=40.0$$

## Height Compare

A physician 's assistant measures a child and finds that his height is 41.5 inches. At his last visit to the doctor's office , the child was 38.5 inches tall. How much did the child grow , in inches?

$$X=41.5-38.5$$

## Value of Coins

Jill has 3.50 dollars in nickels and dimes. If she has 50 coins, how many nickels does she have? How many dimes?

$$X+Y=50.0$$
$$0.05*X+0.1*Y=3.5$$

## Mixture

Arianne is mixing a solution for Chemistry class. She has a 25 % copper solution and a 50 % copper solution. How many milliliters of the 25 % solution and 50 % solution should she mix to make 10 milliliters of a 45 % solution?

$$10 = X + Y$$
$$25.0*.01*X + 50.0*.01*Y = 45.0*.01*10$$

## Row Upstream

It takes a boat 4 hours to travel 24 miles down a river and 6 hours to return upstream to its starting point. What is the rate of the current in the river?

$$(X+Y)*4.0=24.0$$
$$(X-Y)*6.0=24.0$$

## Animals

There are 11 animals in a barnyard. Some are chickens and some are cows. There are 38 legs in all. How many chickens and cows are in the barnyard?

$$(2.0*X)+(4.0*Y)=38$$
$$X+Y=11.0$$

# Eventually Solve More Difficult Problems

## Finance Problems

You decide that you want to save 1,528,717 dollars for retirement. Assuming that you are 25 years old today, will retire at the age of 65, and can earn a 6 percent annual interest rate on your deposits, how much must you deposit each year to meet your retirement goal?

$$X = \frac{1528717}{Y}$$
$$Z = 65 - 25$$
$$Y = \frac{(1 + 0.01 * 6)^Z - 1}{0.01 * 6}$$

## Physics Problems

A block of mass  $m$  is pushed across a rough surface by an applied force,  $F$ , directed at an angle  $\theta$  relative to the horizontal. The block experiences a friction force,  $f$ , in the opposite direction. What is the coefficient of friction between the block and the surface?

$$X = \frac{f}{Y}$$
$$-Z - mg + Y = 0$$
$$Z = F * \sin \theta$$

# Challenge 1:

## Complexity of Semantic Inference

An amusement park sells 2 kinds of tickets. Tickets for children cost \$1.50. Adult tickets cost \$4. On a certain day, 278 people entered the park. On that same day the admission fees collected totaled \$792. How many children were admitted on that day? How many adults were admitted?

Infer:

part\_of(people, children)

part\_of(people, adults)

size(y)=sum parts(y)

size(people) = size(children)+size(adults)

B/g:

1 ticket per person

size(s:chld-tk) = size(children)

cost(s:x)=size(s:x)\*cost(x)

cost(s:chld-tk)=size(s:chld-tk)\*cost(chld-tk)

Infer:

part\_of(\$792, cost(s:chld-tk))

part\_of(\$792, cost(s:adult-tk))

B/g:

size(y)=sum parts(y)

\$792 = cost(s:child-tk) + cost(s:adult-tk)

# Challenge 1:

## Complexity of Semantic Inference

*An amusement park sells 2 kinds of tickets. Tickets for children cost \$1.50. Adult tickets cost \$4. On a certain day, 278 people entered the park. On that same day the admission fees collected totaled \$792. How many children were admitted on that day? How many adults were admitted?*

*Infer:*

part\_of(people, children)

part\_of(people, adults)

**Solution:**

Abstract to a restricted semantic representation – equations:

$$\begin{aligned} X + Y &= 278 \\ 1.5 * X + 4 * Y &= 792 \end{aligned}$$

Space of relations defined by equations seen in training data

# Challenge 2:

## Complex Cross Sentence Relationships

An amusement park sells 2 kinds of tickets. Tickets for children cost \$1.50. Adult tickets cost \$4. On a certain day, 278 people entered the park. On that same day the admission fees collected totaled \$792. How many children were admitted on that day? How many adults were admitted?

$$\text{Tickets for children} * \$1.50 + \text{Adult tickets} * \$4 = \$792$$

### Solution:

Explore a very general space of alignments between the variables in an equation and the natural language

# Challenge 3:

## Significant Domain Variation

### Ticket Sales

*An amusement park sells 2 kinds of tickets. Tickets for children cost \$1.50. Adult tickets cost \$4. On a certain day, 278 people entered the park. On that same day the admission fees collected totaled \$792. How many children were admitted on that day? How many adults were admitted?*

$$\begin{aligned}X + Y &= 278 \\1.5 * X + 4 * Y &= 792\end{aligned}$$

### Math Problems

*A math test is worth 100 points and has 30 problems. Each problem is worth either 3 points or 4 points. How many 4 point problems are there?*

$$\begin{aligned}X + Y &= 30 \\3 * X + 4 * Y &= 100\end{aligned}$$



# Challenge 3:

## Significant Domain Variation

### Ticket Sales

*An amusement park sells 2 kinds of tickets. Tickets for children cost \$1.50. Adult tickets cost \$4. On a certain day, 278 people entered the park. On that same day the admission fees collected totaled \$792. How many children were admitted on that day? How many adults were admitted?*

### Math Problems

*A math test is worth 100 points and has 30 problems. Each problem is worth either 3 points or 4 points. How many 4 point problems are there?*

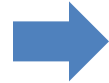
### Solution:

Move beyond lexicalized properties, e.g. syntax, discourse

# Overview: Representation

Space of possible **Equation Types** defined by generalizing labeled equations

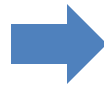
$$\begin{aligned} X + Y &= 278 \\ 1.5 * X + 4 * Y &= 792 \end{aligned}$$



$$\begin{aligned} u_1 + u_2 &= n_1 \\ n_3 * u_1 + n_4 * u_2 &= n_5 \end{aligned}$$

For each word problem choose:

System of equation types



Alignment of equation variables to text



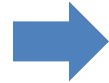
$$u_1 + u_2 = n_1 \qquad n_3 * u_1 + n_4 * u_2 = n_5$$

An amusement park sells 2 kinds of tickets. Tickets for children cost \$1.50. Adult tickets cost \$4. On a certain day, 278 people entered the park. On that same day the admission fees collected totaled \$792. How many children were admitted on that day? How many adults were admitted?

# Overview: Representation

Space of possible **Equation Types** defined by generalizing labeled equations

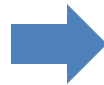
$$\begin{aligned} X + Y &= 278 \\ 1.5 * X + 4 * Y &= 792 \end{aligned}$$



$$\begin{aligned} u_1 + u_2 &= n_1 \\ n_3 * u_1 + n_4 * u_2 &= n_5 \end{aligned}$$

For each word problem choose:

System of equation types



Alignment of equation variables to text

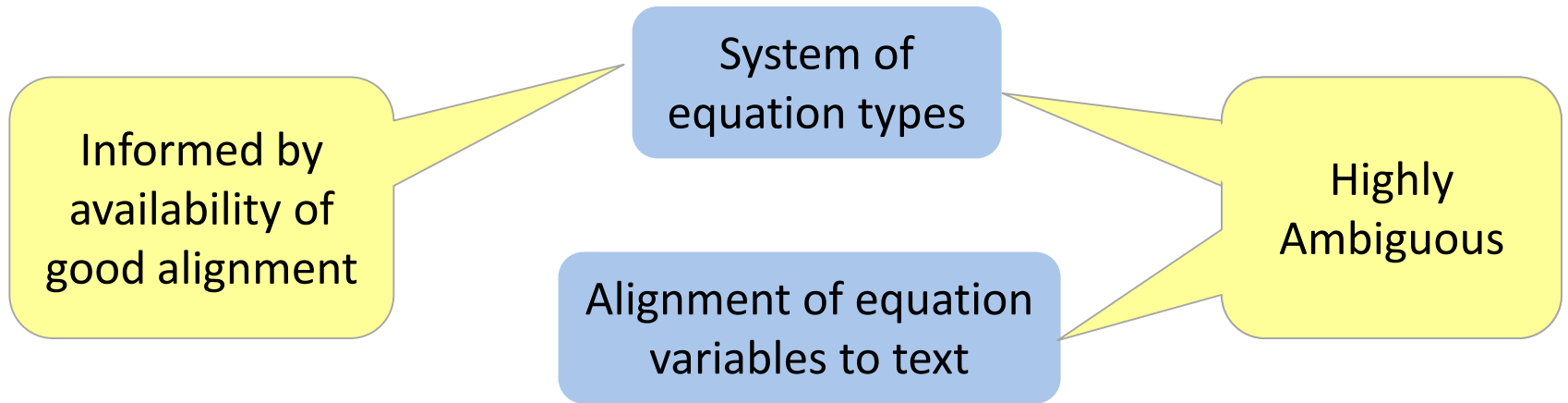


$$u_1 + u_2 = n_1 \quad n_3 * u_1 + n_4 * u_2 = n_5$$

An amusement park sells 2 kinds of tickets. Tickets for children cost \$1.50. Adult tickets cost \$4. On a certain day, 278 people entered the park. On that same day the admission fees collected totaled \$792. How many children were admitted on that day? How many adults were admitted?

Solve resulting equations to get final answer

# Overview: Model



# Overview: Model

System of  
equation types

Alignment of equation  
variables to text

Joint Log-Linear Model

# Key Departures

## Simultaneously interpret multiple sentences

*Branavan et al. 2009; Artzi & Zettlemoyer, 2011, 2013; Zettlemoyer & Collins, 2009; Kwiatkowski et al. 2010; Lei et. al., 2013; Kushman & Barzilay, 2013;*

*Semantic Parsing: Process one sentence at a time*

## Semantics grounded in math; Domain specific meanings not predefined

*Grishman et al., 2005; Maslennikov and Chua, 2007; Ji & Grishman, 2008; Reichart & Barzilay, 2012*

*Information Extraction: Meanings are well defined*

## Learn entirely from data

*Mukherjee & Garain, 2008; Lev et al., 2004*

*Word Problems: Largely hand coded for specific domains*

# Representation

## System of Equation Types

*An amusement park sells 2 kinds of tickets. Tickets for children cost \$1.50. Adult tickets cost \$4. On a certain day, 278 people entered the park. On that same day the admission fees collected totaled \$792. How many children were admitted on that day? How many adults were admitted?*

$$u_1 + u_2 = n_1$$

$$n_3 * u_1 + n_4 * u_2 = n_5$$

n = number variable

u = unknown variable

# Representation

## Aligning Equation Variables

An amusement park sells 2 kinds of tickets. Tickets for children cost \$1.50. Adult tickets cost \$4. On a certain day, 278 people entered the park. On that same day the admission fees collected totaled \$792. How many children were admitted on that day? How many adults were admitted?

$$u_1 + u_2 = n_1$$

$$n_3 * u_1 + n_4 * u_2 = n_5$$

**n = number variable**

**u = unknown variable**



# Representation

## Aligning Equation Variables

An amusement park sells 2 kinds of tickets. Tickets for children cost \$1.50. Adult tickets cost \$4. On a certain day, 278 people entered the park. On that same day the admission fees collected totaled \$792. How many children were admitted on that day? How many adults were admitted?

$$u_1 + u_2 = n_1$$

$$n_3 * u_1 + n_4 * u_2 = n_5$$

**n = number variable**

**u = unknown variable**

# Representation

## Aligning Equation Variables

An amusement park sells 2 kinds of tickets. Tickets for children cost \$1.50. Adult tickets cost \$4. On a certain day, 278 people entered the park. On that same day the admission fees collected totaled \$792. How many children were admitted on that day? How many adults were admitted?

$$u_1 + u_2 = n_1$$

$$n_3 * u_1 + n_4 * u_2 = n_5$$

n = number variable

u = unknown variable

# Representation

## Aligning Equation Variables

An amusement park sells 2 kinds of tickets. **Tickets** for children cost \$1.50. Adult tickets cost \$4. On a certain day, 278 people entered the park. On that same day the admission fees collected totaled \$792. How many **children** were admitted on that day? How many adults were admitted?

$$u_1 + u_2 = n_1$$

$$n_3 * u_1 + n_4 * u_2 = n_5$$

n = number variable

**u = unknown variable**

# Representation

## Aligning Equation Variables

An amusement park sells 2 kinds of tickets. **Tickets** for children cost \$1.50. Adult **tickets** cost \$4. On a certain day, 278 people entered the park. On that same day the admission fees collected totaled \$792. How many **children** were admitted on that day? How many **adults** were admitted?

$$u_1 + u_2 = n_1$$

$$n_3 * u_1 + n_4 * u_2 = n_5$$

n = number variable

u = unknown variable

# Probabilistic Model

$T$  = equation types

$v$  = alignment

$y$  = solution derivation =  $\langle T, v \rangle$

Probability of derivation  $y$  given problem text  $x$

$$p(y|x; \Theta) = \frac{e^{\Theta \cdot \phi(x,y)}}{\sum_{y' \in Y} e^{\Theta \cdot \phi(x,y')}}$$

Probability of numerical answer  $a$  given problem text  $x$

$$p(a|x; \Theta) = \sum_{\substack{y \in Y \\ \text{s.t. ANS}(y)=a}} p(y|x; \Theta)$$

# Example Features

## Domain Independent Alignment Cues

### Shared Nouns:

A discount store sold plastic cups for \$3.25 each and ceramic cups for \$4.50 each. 500 cups were sold.

$$X + Y = 500$$

### Dependency Path:

The lab has 16 workstations. Some are set up for 2 students and the others are set up for 3 students.

$$X + Y = 16$$

Some are set up for 2 students and the others are set up for 3 students.

conj

# Example Features

## Domain Independent Alignment Cues

### Shared Dependency Path Relationships

A grain warehouse has a total of 15 bins. Some hold 20 tons of grain. The rest hold 15 tons of grain. The capacity of the warehouse is 510 tons.

$$X * 15 + Y * 20 = 510$$

Tickets for children cost \$1.50. Adult tickets cost \$4. ...  
On that same day the admission fees collected totaled \$792.

$$X * 1.5 + Y * 4 = 792$$

# Example Features

## Taking Advantage of Grounding to Math

### Compare Numbers

$$n_1 + u = n_2 \longrightarrow n_1 < n_2$$

### Compute Answers

- Positive
- Integer



# Feature Set

## Alignment Pairs/Quadruples

- Dep path contains: Word
- Dep path contains: Dep. Type
- Dep path contains: Word X Dep
- Same word instance
- Same lemma
- Same sentence
- Same phrase
- Connected by a preposition
- Numbers are equal
- Numerical comparison
- Equivalent verb relationship
- Equivalent preposition relationship

## Single Alignment

- Same lemma as question object
- Is in a question sentence
- Is equal to one or two
- Word lemma X nearby constant

## Answers

- Positive Number
- Integer Number

## Document Level

- Unigrams
- Bigrams
- Bias features

# Parameter Estimation

## Learn from either

Full Equations:

$$V(y) = \begin{cases} 1 & \text{if EQ}(y) = \text{correct system of equations} \\ 0 & \text{otherwise} \end{cases}$$

Numerical Answers:

$$V(y) = \begin{cases} 1 & \text{if AN}(y) = \text{correct numerical answer} \\ 0 & \text{otherwise} \end{cases}$$

## Objective

$$O = \sum_i \log \sum_{y \in Y} p(y|x_i; \theta)$$

*s.t.*  $V_i(y) = 1$

# Inference

Exact Inference is NP-hard

Long problems: >100B derivations

Exact Inference is computationally intractable

## Joint Beam Search

Initialize with unaligned equation types

Align one variable at a time

Prune beam after each single variable alignment:

- Limit total beam size
- Limit beam entries per equation type

# Inference

Exact Inference is NP-hard

Long problems: >100B derivations

Exact Inference is computationally intractable

## Joint Beam Search

Initialize with unaligned equation types

Align one variable at a time

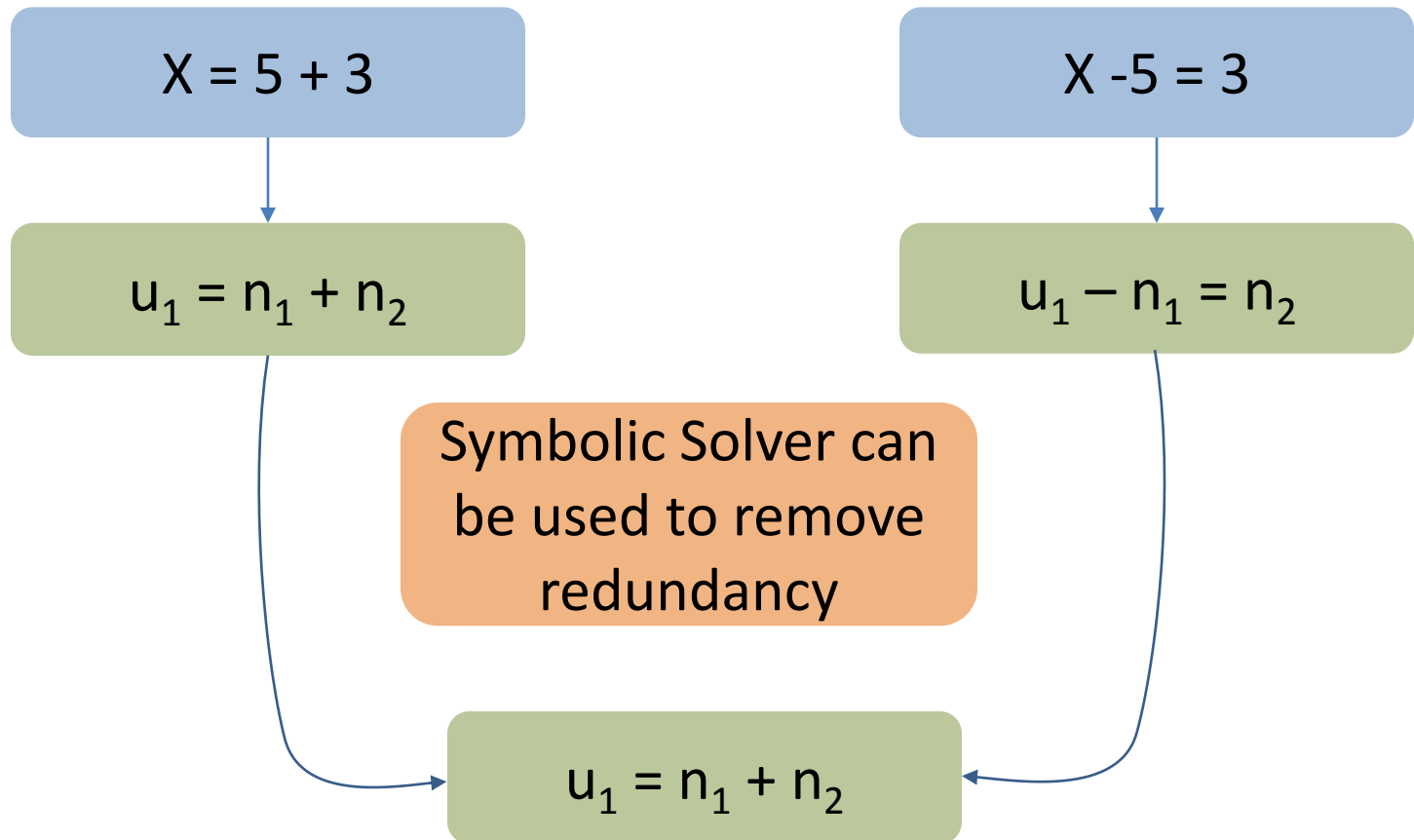
Prune beam after each single variable alignment:

- Limit total beam size

Joint search improves accuracy by 15%

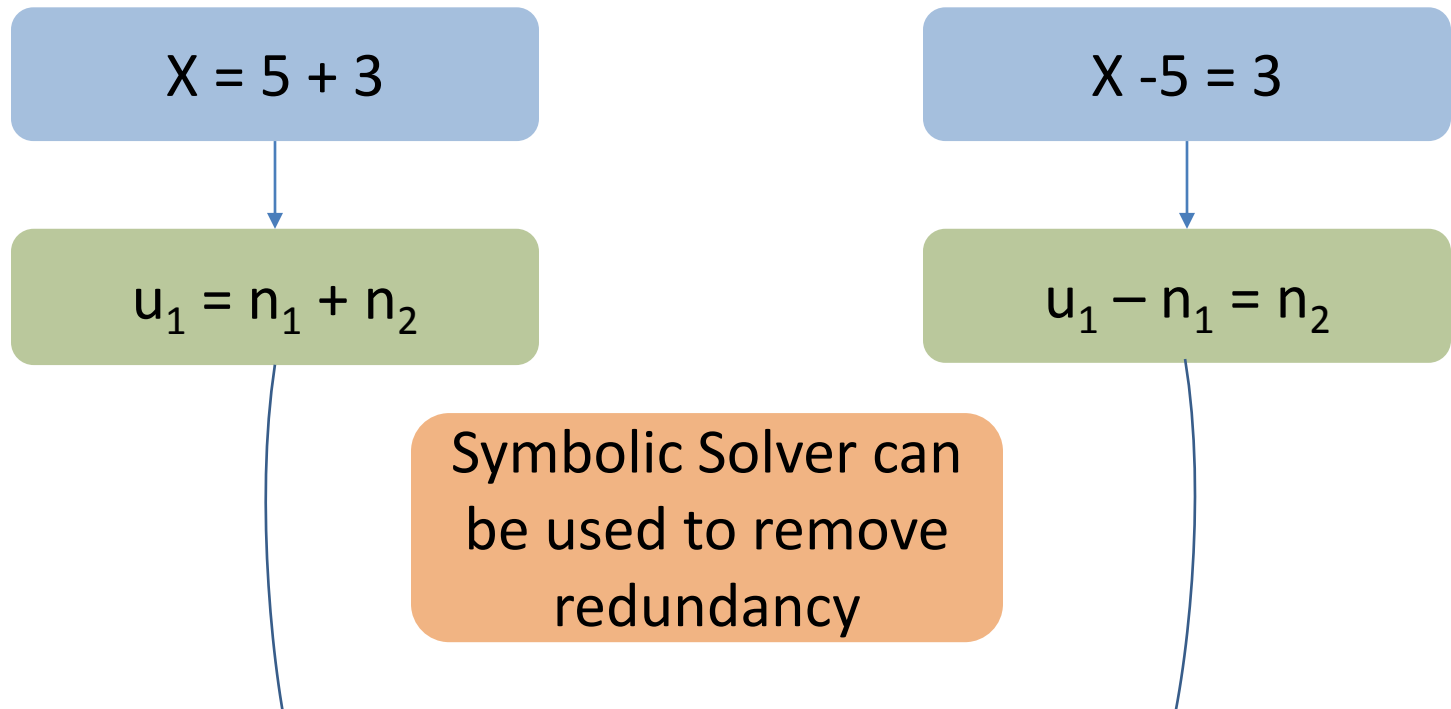
# Utilizing Equational Inference

Naïve equation type generation inefficient



# Utilizing Equational Inference

Naïve equation type generation inefficient



Reduces space of equation types by a factor of 3  
Improves overall accuracy by 7%

# Experiments

# Dataset

Collected from *algebra.com*

<i>Total # of problems</i>	<i>512</i>
<i>Vocabulary size</i>	<i>2352</i>
<i>Avg. words per problem</i>	<i>37</i>
<i>Avg. sentences per problem</i>	<i>3.1</i>

**For each problem collected**

Problem text

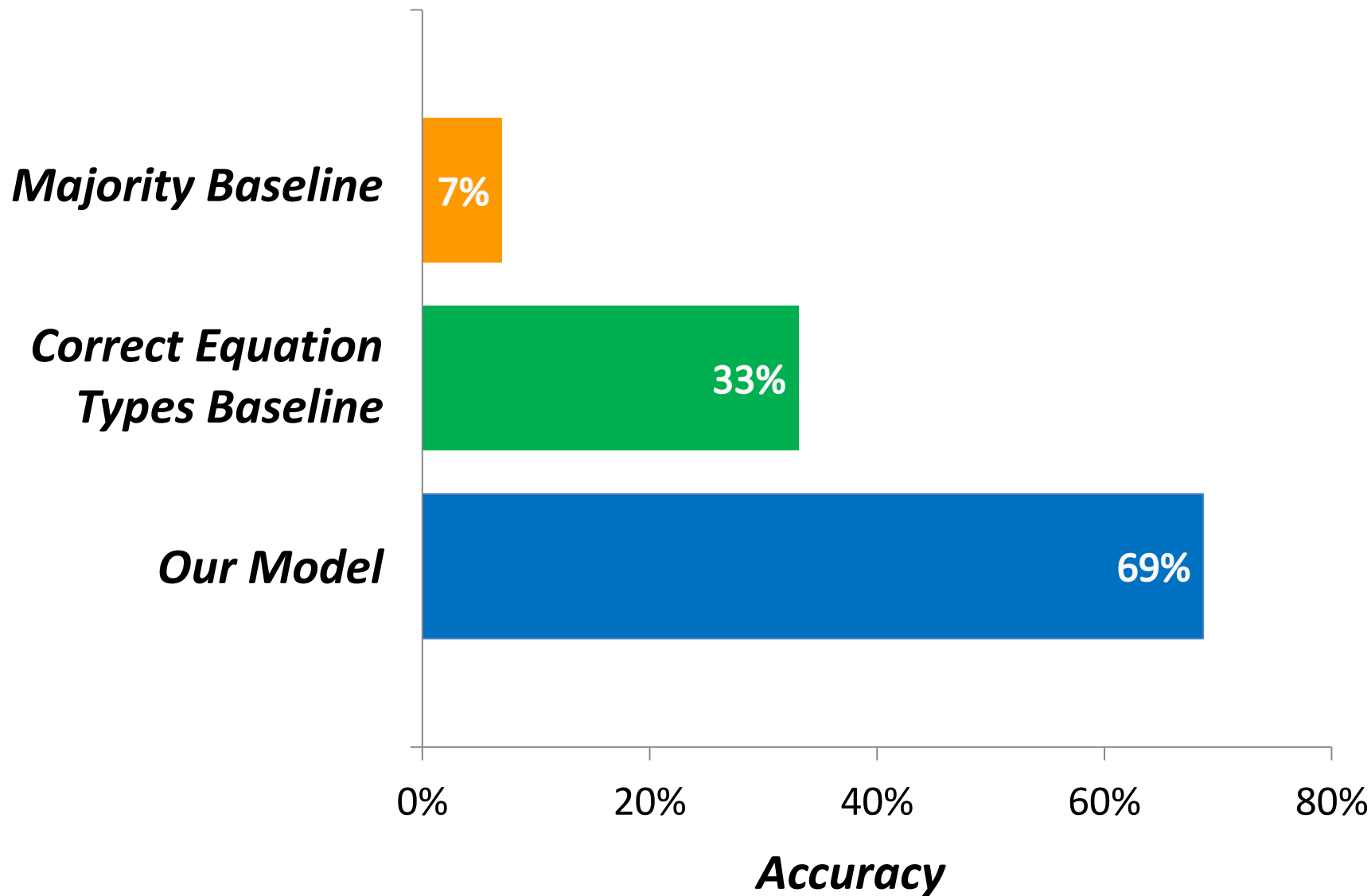
Correct System of Equations



# Fully Supervised Baselines

	Majority Baseline	Correct Equation Types Baseline
Equation Types	<b>Most Common</b>	<b>Correct</b>
Alignment	<b>Most Common Ordering in Text</b>	

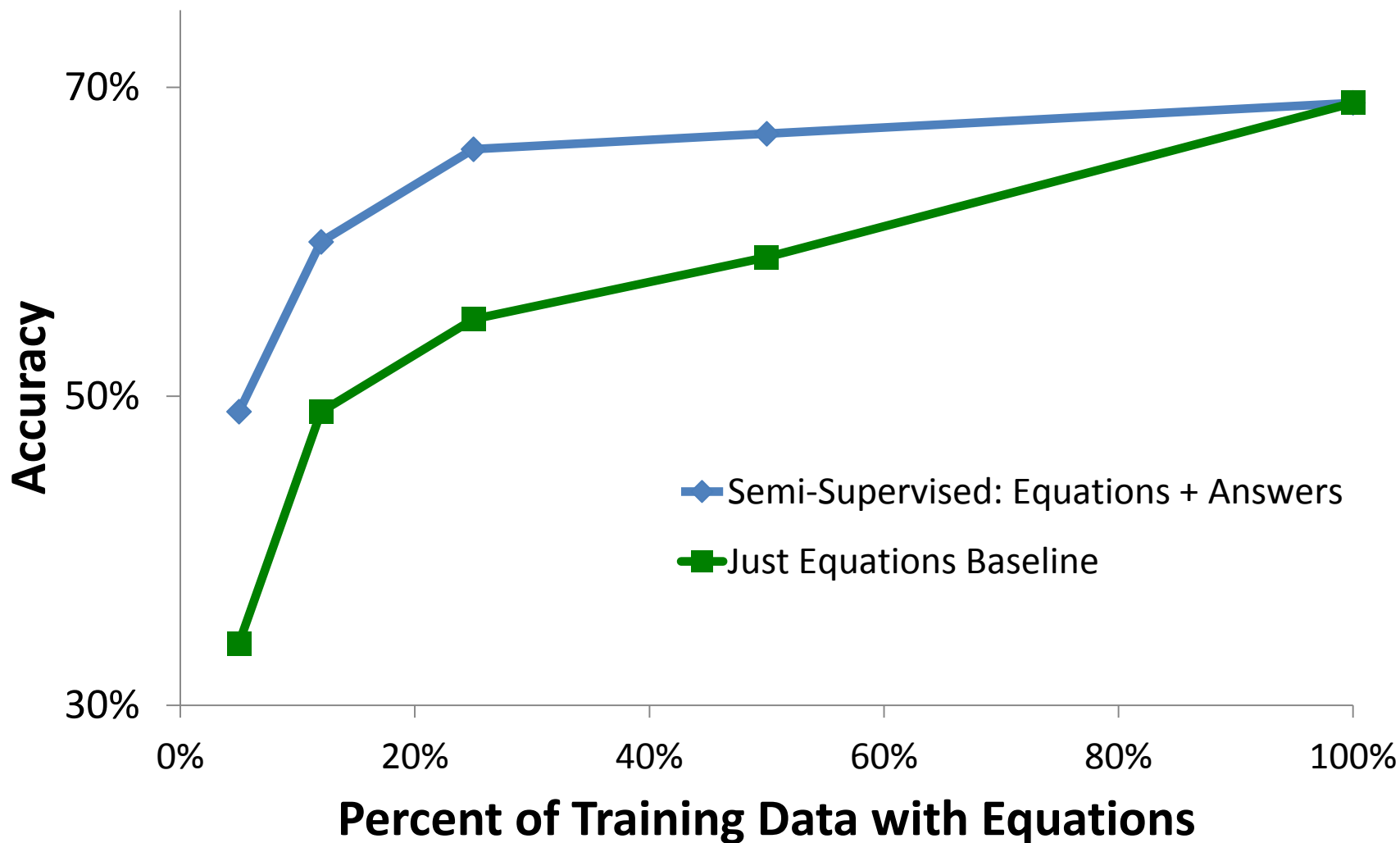
# Results: Fully Supervised Training



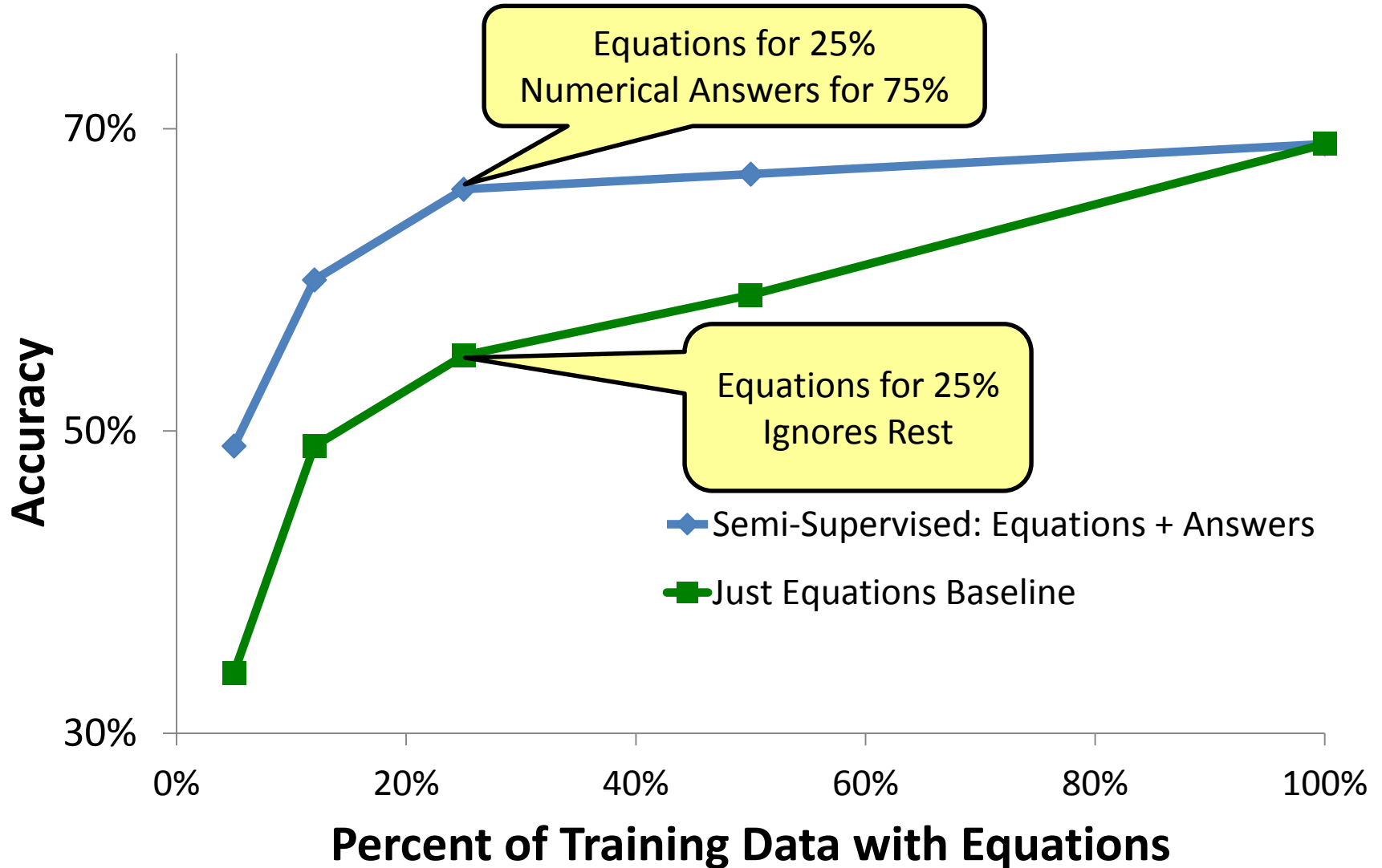
# Semi-Supervised Comparison

	Semi-Supervised Equations+Answers	Just Equations Baseline
Small Fraction of Data	<b>Full Equations</b>	
Rest of Data	<b>Numerical Answers</b>	<b>Ignored</b>

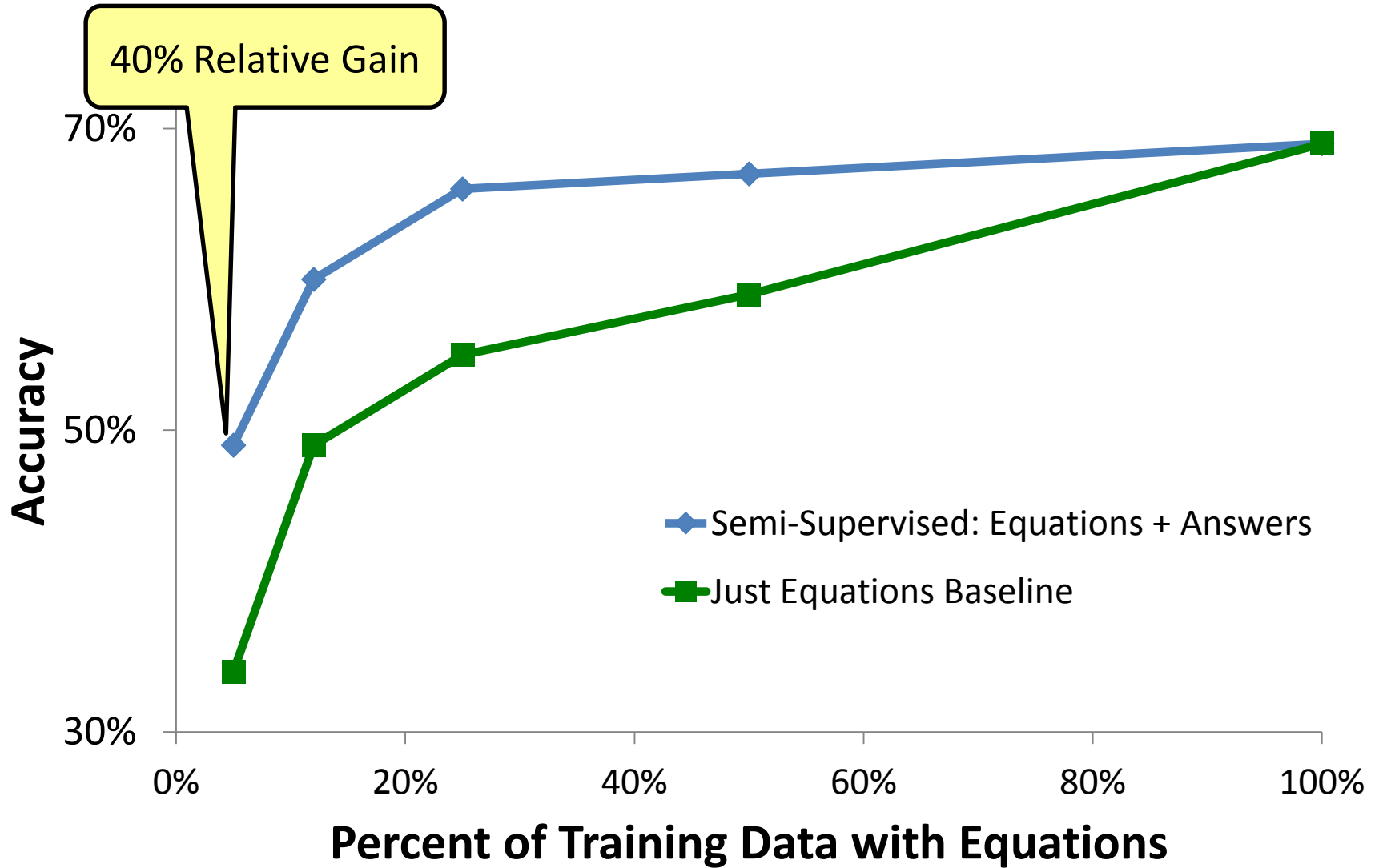
# Varying Percentage of Data with Equations



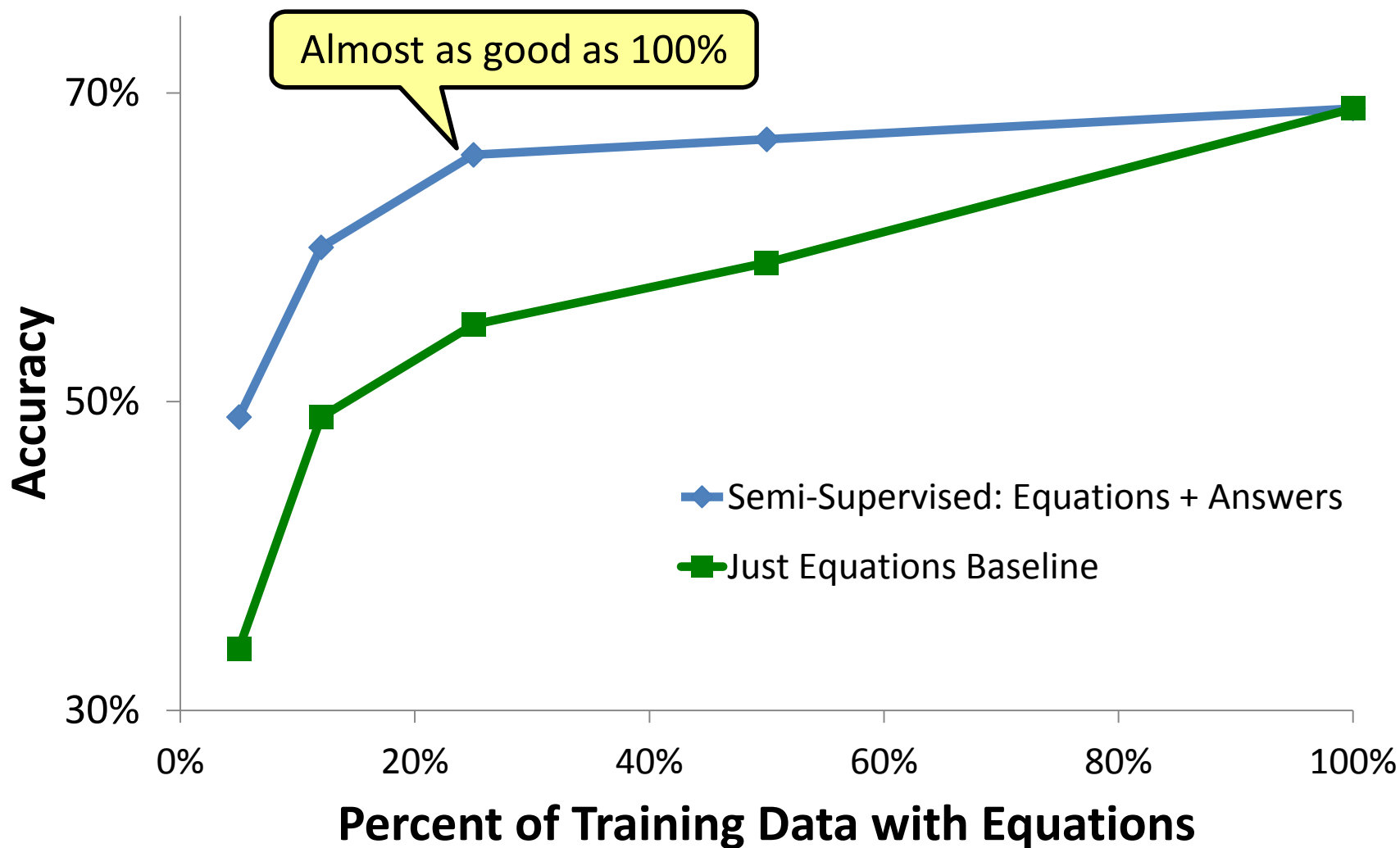
# Varying Percentage of Data with Equations



# Varying Percentage of Data with Equations



# Varying Percentage of Data with Equations



# Example Errors

A textbook costs a bookstore 44 dollars, and the store sells it for 55 dollars. Find the amount of profit based on the selling price.

Requires knowledge of profit and loss

A painting is 10 inches tall and 15 inches wide. A print of the painting is 25 inches tall, how wide is the print in inches?

Must know that print has same width to height ratio as original



# Conclusion

- We demonstrated the feasibility of learning to automatically solve algebra word problems
- Our method can learn effectively without alignments
  - Equations
  - Numeric Answers
- Utilizing the inference capabilities of the math domain improves performance of natural language interpretation

*Data and Code available at:*

***<http://groups.csail.mit.edu/rbg/code/wordprobs/>***