

# Appendix

## Anonymous ACL submission

### 1 Implementation Details

For conducting our experiments we have used two Boston SYS-7048GR-TR nodes equipped with NVIDIA GeForce GTX 1080 Ti computational GPU's having 11GB of GDDR5X RAM. All implementations of training and testing is coded in Python with Pytorch framework. The number of parameters range from 20M to 130M for different models. We use negative log likelihood as the loss criterion. Hyper-parameter values were not modified, and we follow the recommendations of the respective models. To reduce carbon footprint from our experiments, we run the models only on a single fold for searching hyperparameter values. We chose the number of base candidates after primary stage  $n$  as 7. Generating augmentation examples using Paraphrasing Methods took around 12 minutes on average for MaWPS and 8 minutes for ASDiv-A datasets. Substitution methods took around 5 minutes on average for both MaWPS and ASDiv-A dataset. The experiments conducted by us are not computation heavy. Each of the state-of-the-art models get trained within 5 hrs of time, with Graph2Tree taking the maximum time.

### 2 Additional Augmented Examples

In this section, we present some additional valid as well as invalid augmented examples. Additionally, we also show some more examples with their attention weights. Table 1 shows some additional examples with their attention weight distribution. These weights have been shown for the base model trained before augmentation and after augmentation on MaWPS dataset. Table 2 illustrates some additional problem statements for all the techniques in paraphrasing methods and substitution methods. In Table 3, we present some invalid augmented examples which do not satisfy our human evaluation criteria. These examples are such that they alter the semantics of the original problem statement.

---

**Problem:** A magician **was** selling magic card **decks** for 2 dollars **each**. If he started with 25 decks and by the end of the day he had 4 left, how much money did he earn ?

**Mean attention values:** 0.34 0.11 0.09

**Problem:** A magician was selling magic **card** decks for 2 **dollars** each. If he started with 25 decks and by the end of the day he had 4 left, how **much** money did he earn ?

**Augmented mean attention values :** 0.19 0.18 0.15

---

**Problem:** There **are** 18 pencils in the drawer and 6 pencils **on** the desk. Dan placed 4 **pencils** on the desk. How many pencils are now there in total ?

**Mean attention values:** 0.21 0.16 0.06

**Problem:** There are 18 pencils **in** the drawer and 6 pencils on the desk. **Dan** placed 4 pencils on the desk. How many pencils are now there in **total** ?

**Augmented mean attention values :** 0.29 0.19 0.12

---

**Problem:** Dan has 12 violet marbles, he **gave** Mary 4 of the **marbles**. How many violet marbles does **he** now have ?

**Mean attention values:** 0.23 0.21 0.17

**Problem:** Dan has 12 **violet** marbles , he **gave** Mary 4 of the marbles. How **many** violet marbles does he now have ?

**Augmented mean attention values :** 0.23 0.18 0.11

---

**Problem:** Angela has 7 tickets . **Annie** gives Angela 5 **more** . How many **tickets** does Angela have in all ?

**Mean attention values:** 0.30 0.19 0.15

**Problem:** Angela has 7 tickets . Annie **gives** Angela 5 **more** . How many tickets does Angela have in **all** ?

**Augmented mean attention values :** 0.29 0.21 0.14

---

**Problem:** Maria **had** 5 **bottles** of water **in** her fridge . If she drank 1 of them and then bought 2 more , how many bottles would she have ?

**Mean attention values:** 0.48 0.14 0.04

**Problem:** Maria had 5 bottles of water in her fridge . If she drank 1 of them and **then** bought 2 **more** , how **many** bottles would she have ?

**Augmented mean attention values :** 0.23 0.17 0.11

---

Table 1: Examples illustrating distribution of top three attention weights before and after training on the full augmented dataset.

Category	Augmentation Method	Example
Paraphrasing Methods	Round trip Translation	<b>Original:</b> Alyssa’s dog had puppies . She gave 2 to her friends . She now has 3 puppies . How many puppies did she have to start with ? <b>Augmented:</b> Alyssa’s dog had puppies . <b>She gave her friends 2</b> . She now has 3 puppies . <b>How many puppies did she start ?</b>
	Problem Reordering	<b>Original:</b> Rachel was organizing her book case making sure each of the shelves had exactly 3 books on it. If she had 4 shelves of mystery books and 2 shelves of picture books , how many books did she have total ? <b>Augmented:</b> <b>How many books did she have given that</b> rachel was organizing her book case making sure each of the shelves had exactly 3 books on it and she had 4 shelves of mystery books and 2 shelves of picture books .
Substitution Methods	Fill Masking	<b>Original:</b> A cell phone company has a total of 1000 customers across the world . If 740 of its customers live in the United States , how many of its customers live in other countries ? <b>Augmented:</b> A <b>mobile</b> phone <b>firm</b> has a network of 1000 customers across the world . If 740 of its customers live in the <b>violetUS</b> , <b>How many customers live in other locations?</b>
	Named-Entity Replacement	<b>Original:</b> Daniel had some noodles. He gave 20 noodles to William. Now Daniel only has 11 noodles. How many noodles did Daniel have to begin with ? <b>Augmented:</b> Matt had some noodles. He gave 20 noodles to Zeal. Now Matt only has 11 noodles. How many noodles did Matt have initially ? <b>Edd</b> found 7 seashells , <b>Alan</b> found 12 seashells , and <b>Royal</b> found 5 seashells on the beach . <b>How many seashells were found together ?</b>
	Synonym Replacement	<b>Original:</b> There are 5 rulers in the drawer. Tim took 3 rulers from the drawer. How many rulers are now in the drawer ? <b>Augmented:</b> There are 5 <b>consonants</b> in the drawer. Tim went 3 <b>consonants</b> from the drawer. <b>How many other consonants are in the drawer now ?</b>

Table 2: Valid Augmentation examples from all proposed methods. Coloured text represents the changes in problem statement.

Category	Augmentation Method	Example
Paraphrasing Methods	Round trip Translation	<b>Original:</b> Kimberly went to the store 6 times last month . She buys 9 peanuts each time she goes to the store . How many peanuts did Kimberly buy last month ? <b>Augmented:</b> Kimberly <b>travelled to club six times last month</b> . She buys 9 peanuts every time she goes to the club . <b>How many peanuts did Kimberly buy last year ?</b>
	Problem Reordering	<b>Original:</b> Fred has 10 blue marbles . Fred has number1 times more blue marbles than Tim . How many blue marbles does Tim have ? <b>Augmented:</b> <b>If</b> fred has 10 blue marbles and fred has number1 more blue marbles <b>than Tim then how many blue marbles does tim have ?</b>
Substitution Methods	Fill Masking	<b>Original:</b> Sarah had 7 homework problems . She finished 2 of them but still had 3 pages of problems to do . If each page has the same number of problems on it , how many problems are on each page ? <b>Augmented:</b> Sarah had 7 of <b>them</b> . She had 2 of <b>them</b> but still had 3 more of <b>them</b> to do . If each more has the same number of <b>them</b> on it, <b>How many them are on each more ?</b>
	Named-Entity Replacement	<b>Original:</b> Beverly had 10 dimes in his bank . His sister Maria borrowed 2 of his dimes . How many dimes does Beverly have now ? <b>Augmented:</b> <b>Silva</b> had 10 dimes in his bank . His sister <b>Jeanie</b> borrowed a <b>pair</b> of his dimes . <b>How many dimes does Jeanie have now ?</b>
	Synonym Replacement	<b>Original:</b> Shawn’s team won their dodgeball game and scored 25 points total . If Shawn scored 13 of the points and everyone else scored 4 points each , how many players were on his team ? <b>Augmented:</b> Shawn’s <b>group</b> won their <b>rumble</b> game and scored 25 points total . If Shawn scored 13 of the points and everyone else scored <b>quarter</b> points each, <b>how many people were there ?</b>

Table 3: Invalid Augmentation examples from all proposed methods. Coloured text represents the changes in problem statement.