Google

TableFormer: Robust Transformer Modeling for Table-Text Encoding

Jingfeng Yang, Aditya Gupta, Shyam Upadhyay, Luheng He, Rahul Goel, Shachi Paul

Georgia Tech College of Computing
Computational Science and Engineering

Table-Text Understanding

Original intent:
What super hero from Earth appeared most recently?

1. Who are all of the super heroes?

2. Which of them come from Earth?

3. Of those, who appeared most recently?

Legion of Super Heroes Post-*Infinite Crisis*

Character	*First Appeared*	*Home World*	*Powers*
Night Girl	2007	Kathoon	Super strength
Dragonwing	2010	Earth	Fire breath
Gates	2009	Vyrga	Teleporting
XS	2009	Aarok	Super speed
Harmonia	2011	Earth	Elemental

Sequential QA dataset (SQA) (Iyyer et al., 2017)

Recent Approaches

- General Recipe
  - Step 1: Pretraining on text-table pairs
    - Pretraining on existing table-text corpus (Wikipedia, ToTTo etc.):
      - TaBERT (Yin et al., 2020)
      - TAPAS (Herzig et al., 2020)
      - StruG (Deng et al., 2021)
    - Data augmentation for pretraining
      - Intermediate pretraining (Eisenschlos et al., 2020)
      - GRAPPA (Yu et al., 2021)
      - TaPEx (Liu et al. 2022)
  - Step 2: Fine-tuning on specific dataset (e.g. SQA)

Problem 1: Non-Robust Modeling

Question: Of all song lengths, which one is the longest?
Gold Answer: 5:02

Title	Producers	Length
Screwed Up	Mr. Lee	5:02
Smile	Sean T	4:32
Ghetto Queen	I.N.F.O. & NOVA	5:00

Problem 1: Non-Robust Modeling

Question: Of all song lengths, which one is the longest?
Gold Answer: 5:02
TAPAS Predicted Answer: 5:00

Title	Producers	Length
Screwed Up	**Mr. Lee**	**5:02**
Smile	Sean T	4:32
Ghetto Queen	**I.N.F.O. & NOVA**	**5:00**

Problem 1: Non-Robust Modeling

Question: Of all song lengths, which one is the longest?
Gold Answer: 5:02
TAPAS Predicted Answer: 5:00

Title	Producers	Length
Screwed Up	**Mr. Lee**	**5:02**
Smile	Sean T	4:32
Ghetto Queen	**I.N.F.O. & NOVA**	**5:00**

TAPAS Predicted Answer After Perturbation: 5:02

Title	Producers	Length
Smile	Sean T	4:32
Ghetto Queen	I.N.F.O. & NOVA	5:00
Screwed Up	**Mr. Lee**	**5:02**

Model is not robust to row/column order changes!

Accuracy drops from 66.8 to 60.5 on SQA dataset after perturbation.

Problem 2: Lack of Structural Biases

Question: Which nation received 2 silver medals?
Gold Answer: Spain, Ukraine
TAPAS Predicted Answer: Spain

Nation	Gold	Silver	Bronze
Great Britain	2	1	2
Spain	1	2	0
Norway	1	0	0
Ukraine	0	2	0

Problem 2: Lack of Structural Biases

Question: Which nation received 2 silver medals?
Gold Answer: Spain, Ukraine
TAPAS Predicted Answer: Spain

Nation	Gold	**Silver**	Bronze
Great Britain	2	1	2
Spain	1	**2**	0
Norway	1	0	0
Ukraine	0	**2**	0

Identify “Silver” column and “2” cells in this column

Problem 2: Lack of Structural Biases

Question: Which nation received 2 silver medals?
Gold Answer: Spain, Ukraine
TAPAS Predicted Answer: Spain

Nation	Gold	**Silver**	Bronze
Great Britain	2	1	2
Spain	1	**2**	0
Norway	1	0	0
Ukraine	0	**2**	0

Output contents of the same rows in “Nation” column

TableFormer

Robust Table+Text Modeling

Table-Text (Relative) Attention Bias Types

Question: Which nation received 2 silver medals?

Nation	Silver
Spain	2
Norway	0
Ukraine	2

Relative Attention:

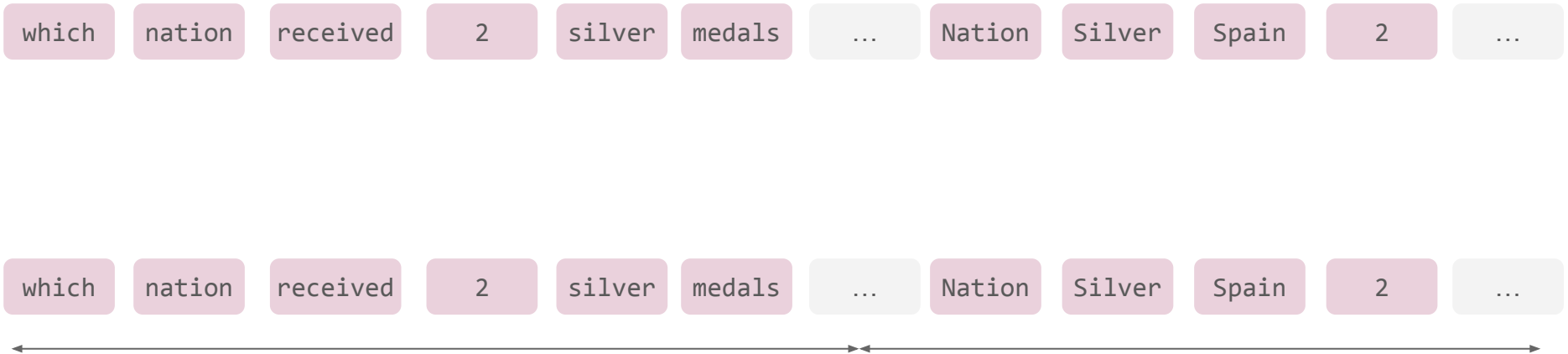


Table-Text (Relative) Attention Bias Types

Question: Which nation received 2 silver medals?

Nation	Silver
Spain	2
Norway	0
Ukraine	2

Relative Attention:

- **Header to Sentence**

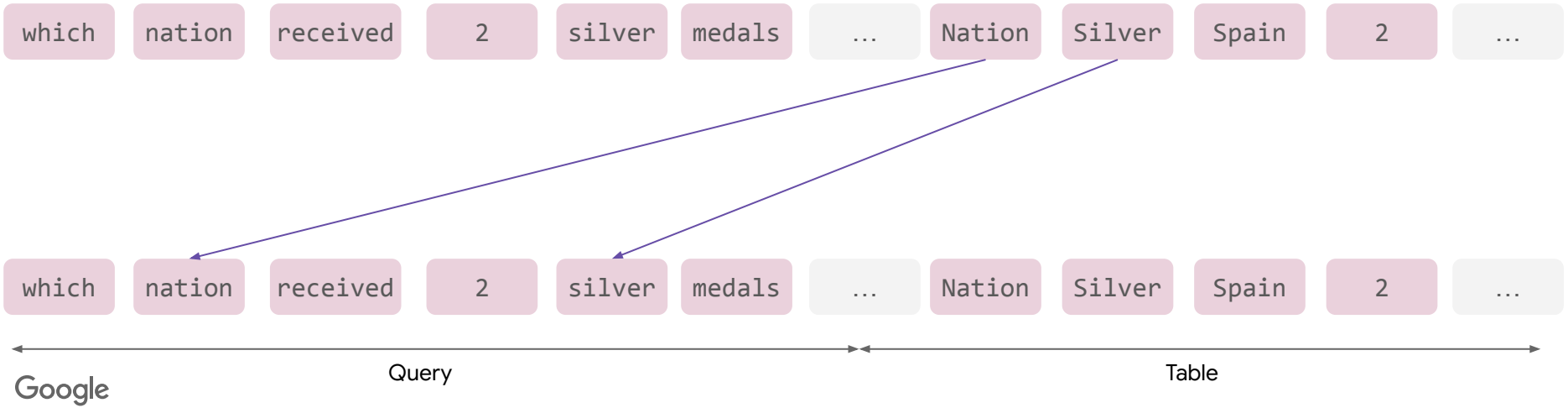


Table-Text (Relative) Attention Bias Types

Question: Which nation received 2 silver medals?

Nation	Silver
Spain	2
Norway	0
Ukraine	2

Relative Attention:

- **Header to Sentence**
- **Cell to Sentence**

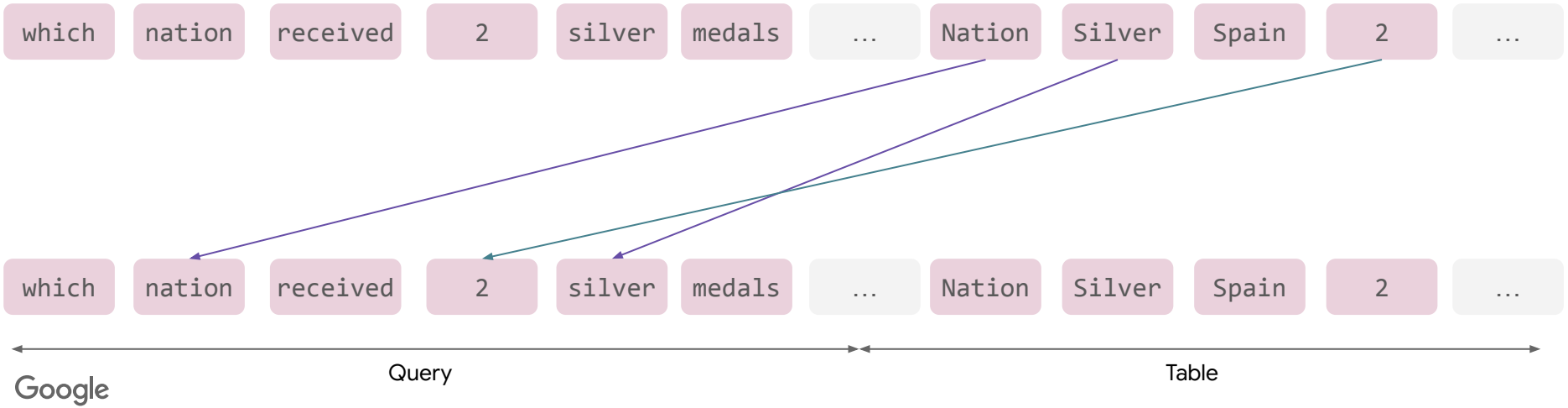


Table-Text (Relative) Attention Bias Types

Question: Which nation received 2 silver medals?

Nation	Silver
Spain	2
Norway	0
Ukraine	2

Relative Attention:

- Header to Sentence
- Cell to Sentence
- Cell to Column Header

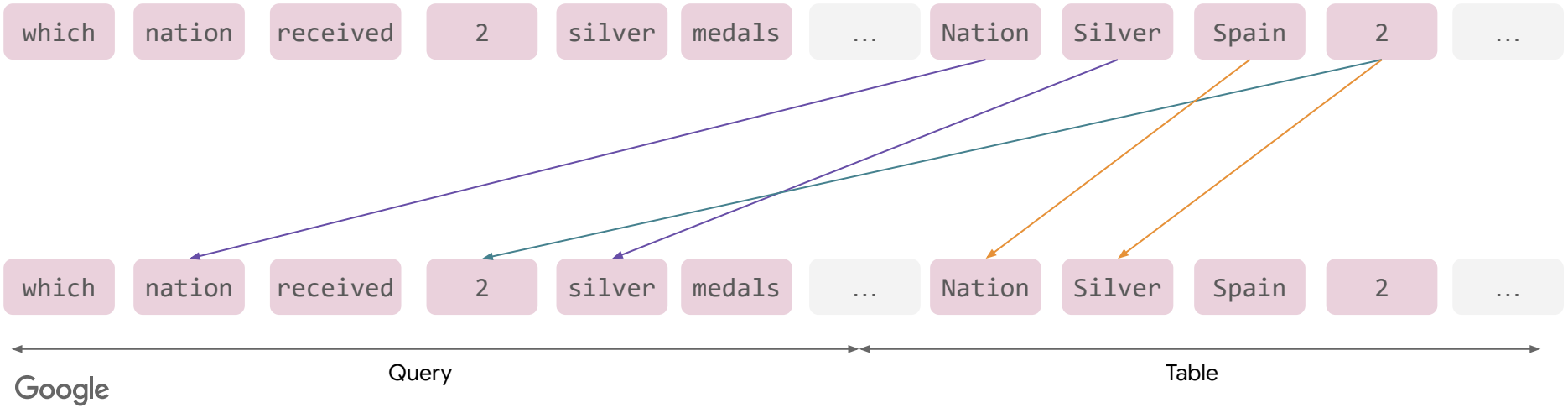


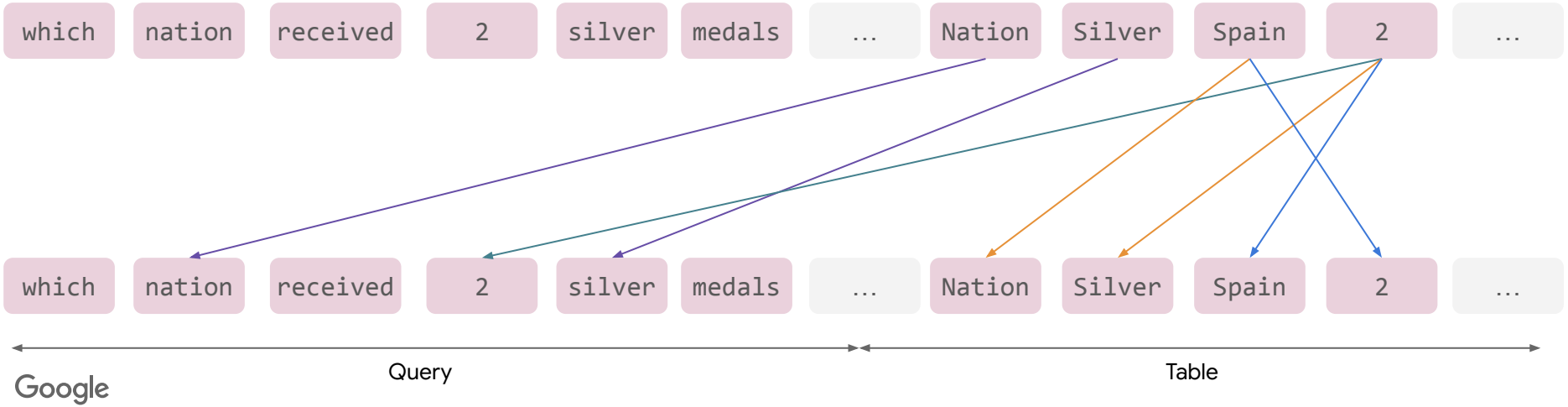
Table-Text (Relative) Attention Bias Types

Question: Which nation received 2 silver medals?

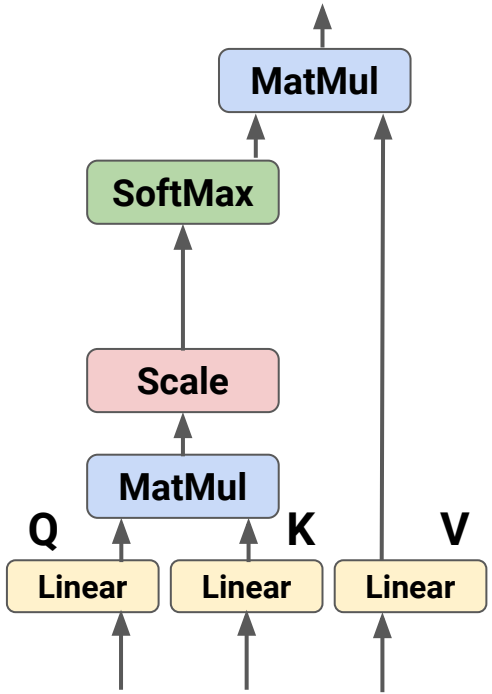
Nation	Silver
Spain	2
Norway	0
Ukraine	2

Relative Attention:

- Header to Sentence
- Cell to Sentence
- Cell to Column Header
- Same Row
- ...



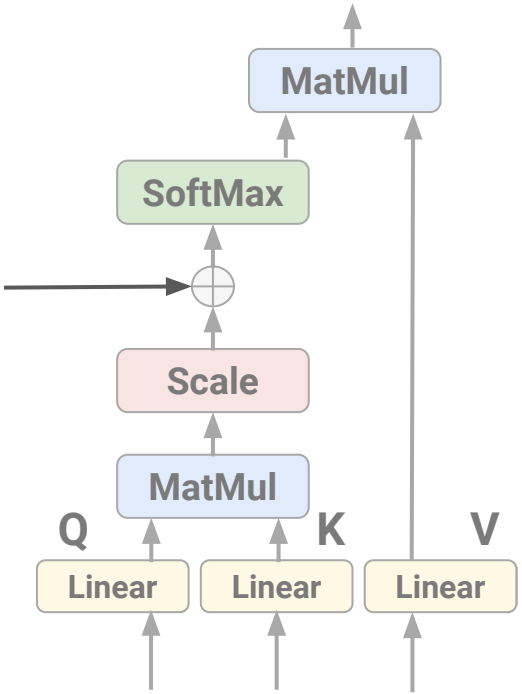
Transformer (Vaswani et al. 2017)



$$\text{Attn}(H) = \text{softmax}(\frac{QK^\top}{\sqrt{d_K}})V$$

TableFormer (our work)

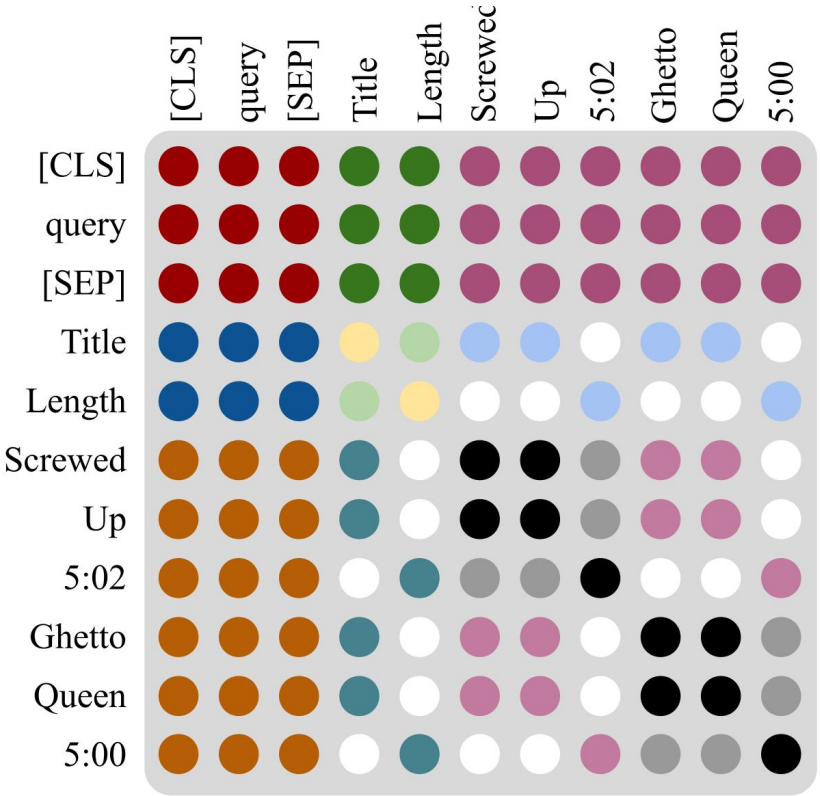
$\hat{A}$ **Learnable Table-Text Attention Bias Matrix (13 types of attention biases)**



$$\text{Attn}(H) = \text{softmax}(\frac{QK^\top}{\sqrt{d_K}})V$$

$$\bar{A} = \frac{QK^\top}{\sqrt{d_K}}, \quad A = \bar{A} + \hat{A}$$

Table-Text (Relative) Attention Bias Types

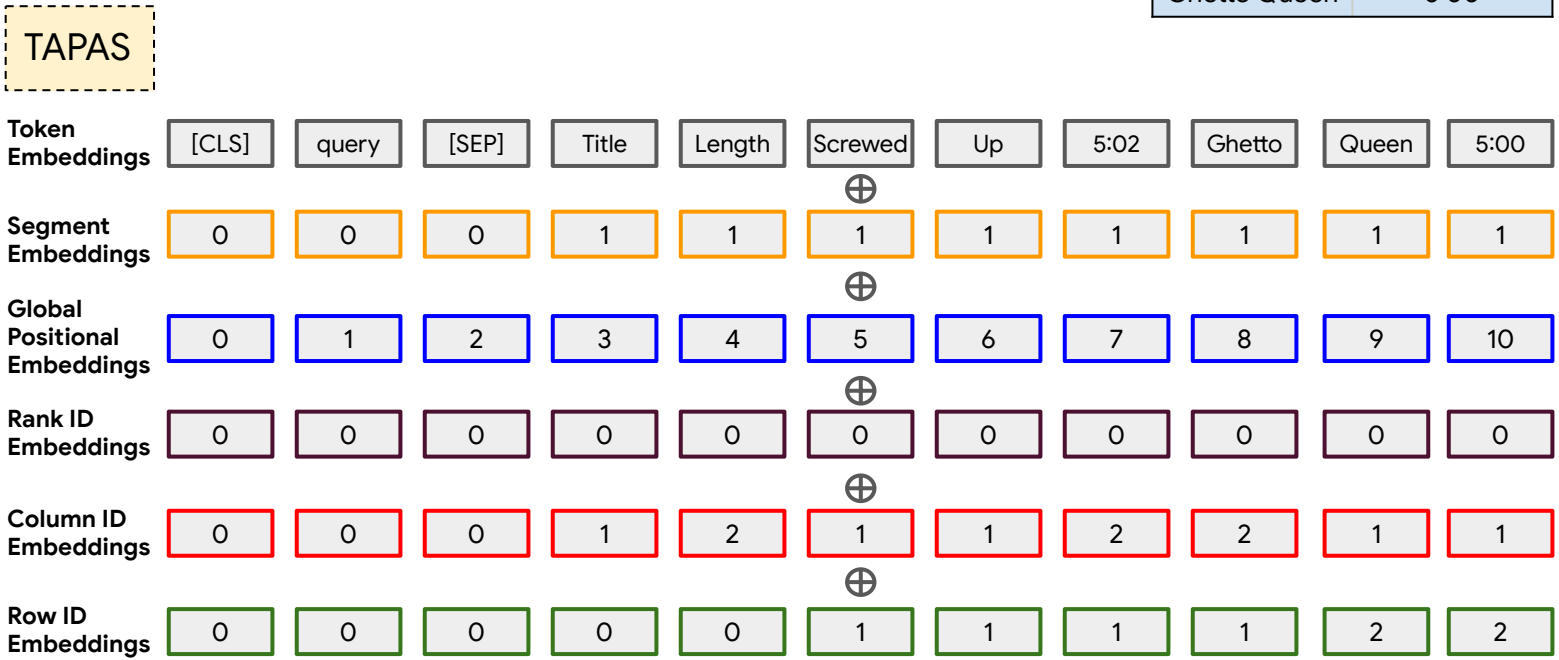


	Attention Bias Type
	header to sentence
	cell to sentence
	cell to its column header
	same row bias
	same column bias
...	...

TAPAS Input

Table:

Title	Length
Screwed Up	5:02
Ghetto Queen	5:00

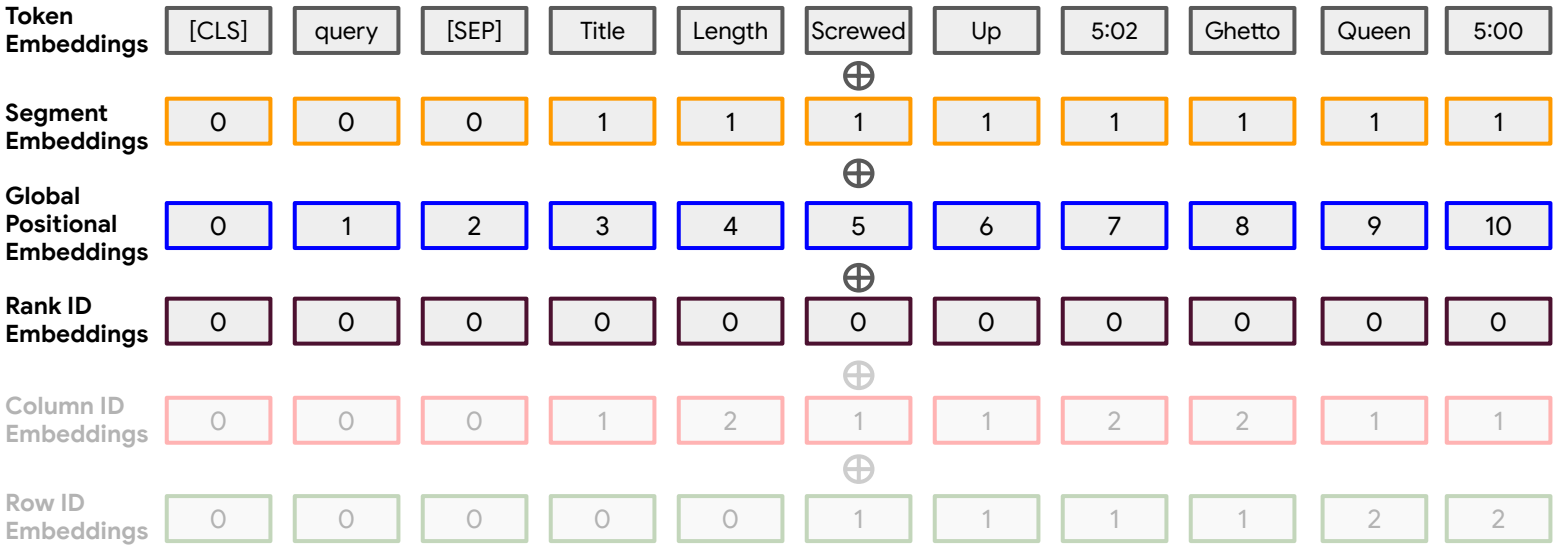


TableFormer Input

Table:

Title	Length
Screwed Up	5:02
Ghetto Queen	5:00

TAPAS - Row/Column IDs

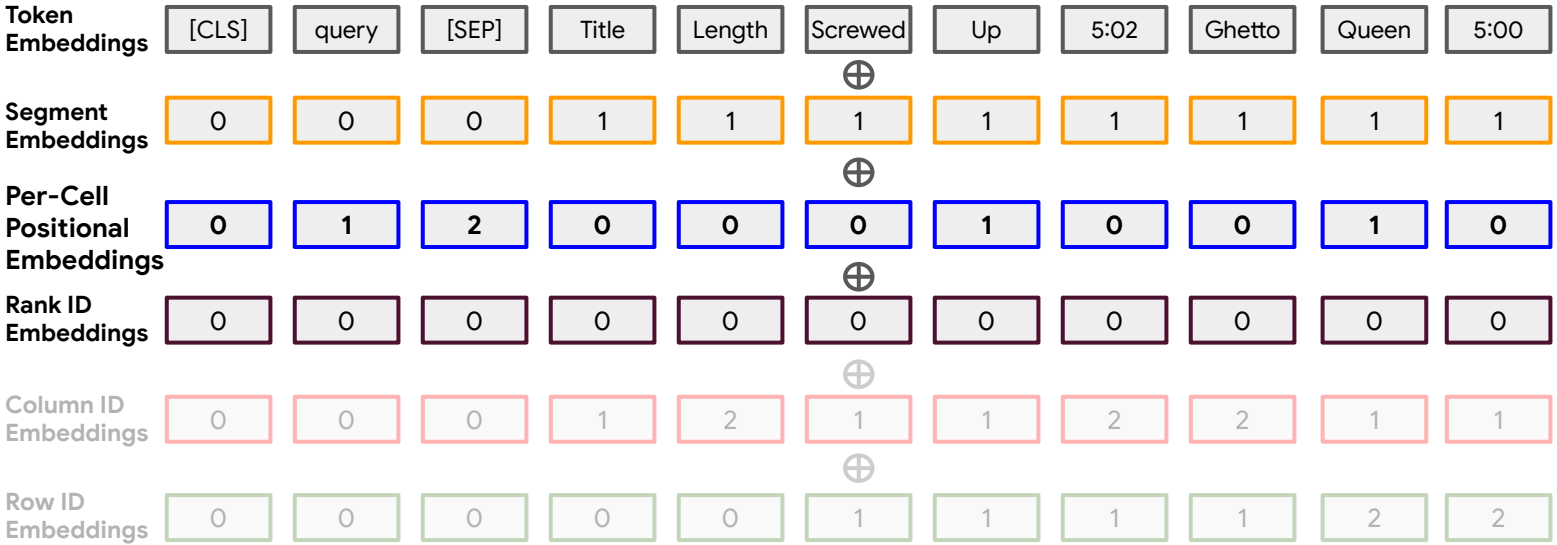


TableFormer Input

Table:

Title	Length
Screwed Up	5:02
Ghetto Queen	5:00

TAPAS - Row/Column IDs + Per Cell Positional IDs



Results

Experimental Setup

1. Reasoning Tasks
    a. Wikipedia Table based QA
    b. Table and Text Entailment

2. Evaluation Settings and Metrics
    a. Accuracy in Standard Evaluation
    b. Accuracy in Perturbation Evaluation: Randomly shuffle rows and columns of tables on test set without changing table contents
    c. Variation Percentage (VP) after Perturbation:

$$\text{VP} = \frac{\text{\# incorrect predictions that were corrected + \# correct predictions that became incorrect}}{\text{\# total}}$$

Table-based Sequential QA: SQA (Iyyer et al., 2017)

Original intent:
What super hero from Earth appeared most recently?

1. Who are all of the super heroes?

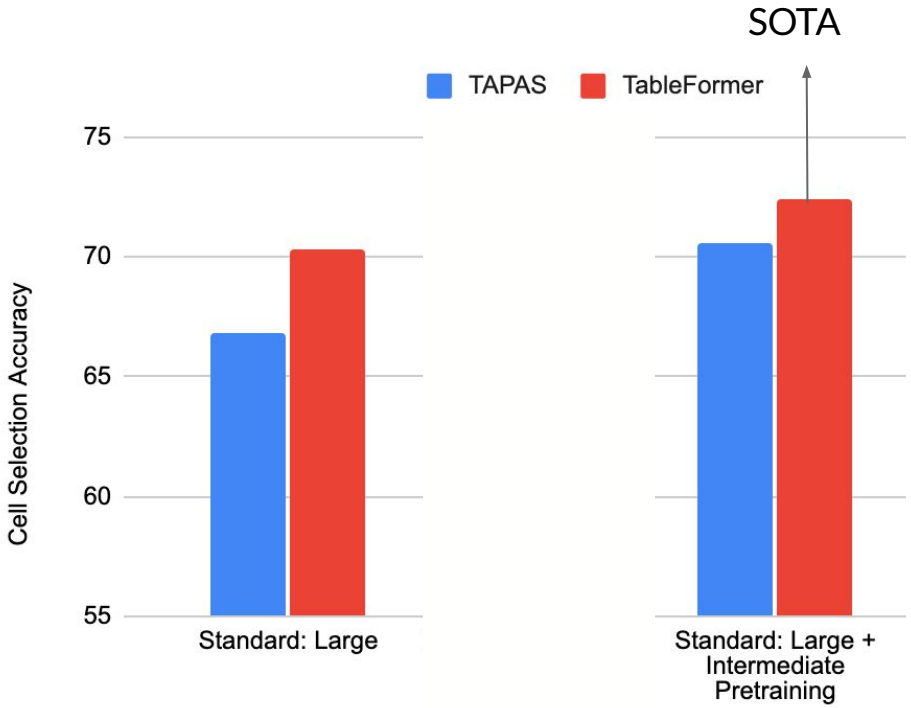
2. Which of them come from Earth?

3. Of those, who appeared most recently?

Legion of Super Heroes Post-*Infinite Crisis*

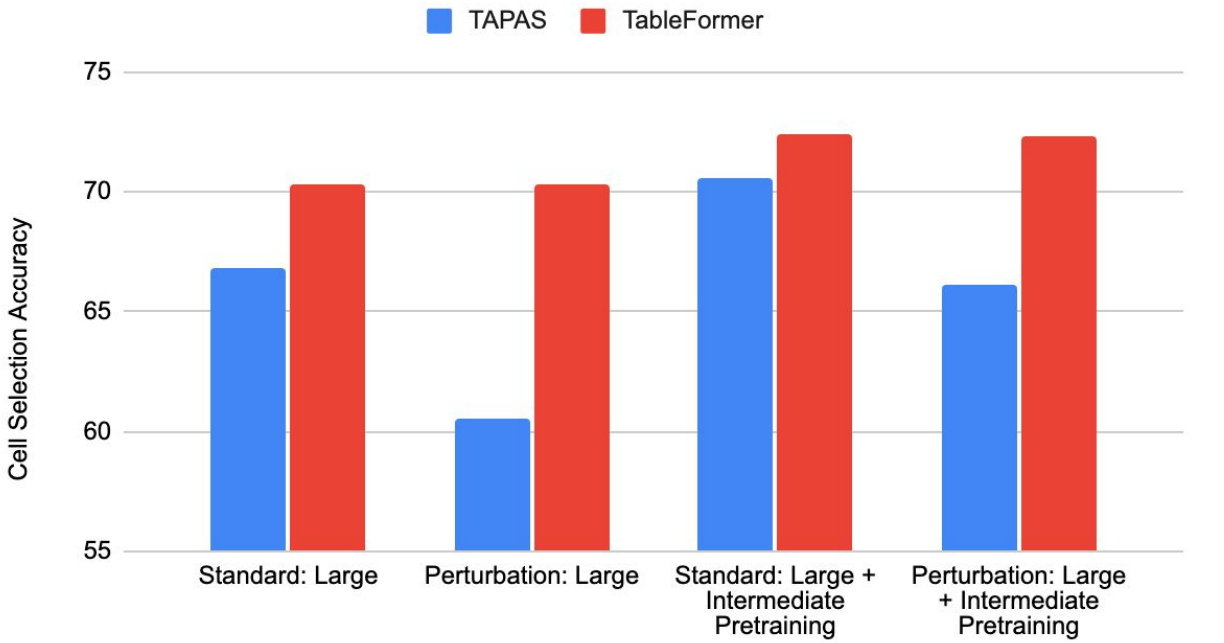
Character	*First Appeared*	*Home World*	*Powers*
Night Girl	2007	Kathoon	Super strength
Dragonwing	2010	Earth	Fire breath
Gates	2009	Vyrga	Teleporting
XS	2009	Aarok	Super speed
Harmonia	2011	Earth	Elemental

Results on SQA (Table-based Sequential QA)



Better overall performance with new SoTA!

Results on SQA (Table-based Sequential QA)



Invariant to perturbations which affect previous approaches!

Results on SQA (Instance-level Robustness)

Variation Percentage (VP) after Perturbation

$$VP = \frac{\#\text{ incorrect predictions that were corrected} + \#\text{ correct predictions that became incorrect}}{\#\text{ total}}$$

	TAPAS	TableFormer
Large	15.1%	0.0%
Large + Intermediate Pretraining	10.8%	0.0%

TableFormer prediction is strictly robust to perturbations in the instance level!

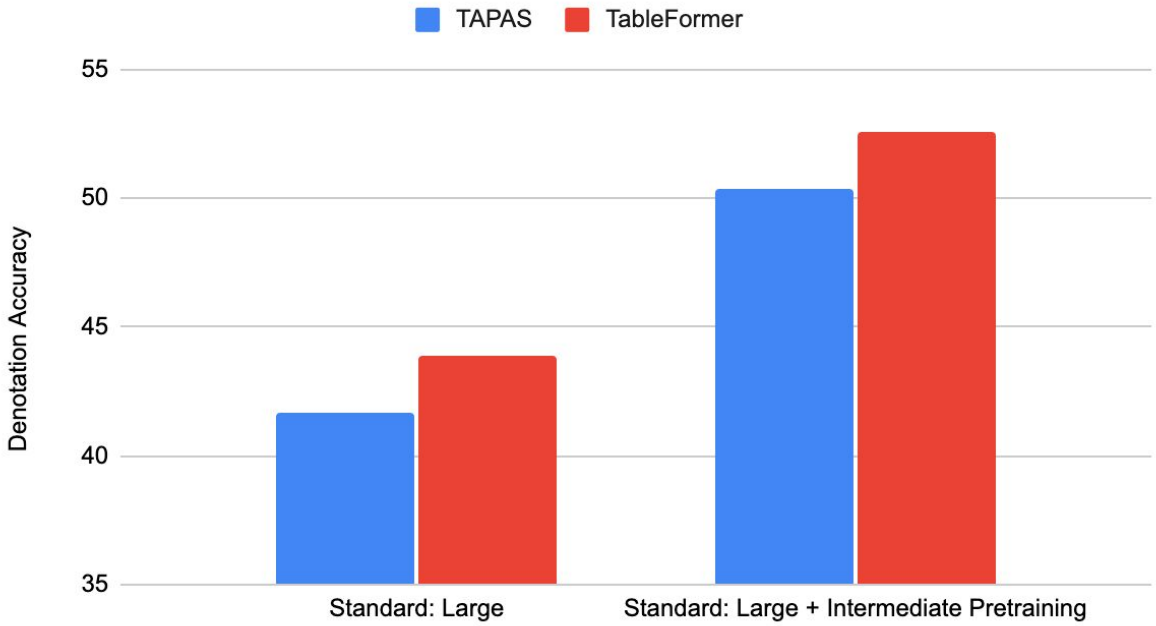
Table-based Complex QA: WikiTableQuestions (Pasupat et al., 2015)

Year	City	Country	Nations
1896	Athens	Greece	14
1900	Paris	France	24
1904	St. Louis	USA	12
...	...	...	...
2004	Athens	Greece	201
2008	Beijing	China	204
2012	London	UK	204

x_1: *"Greece held its last Summer Olympics in which year?"*
y_1: {2004}

x_2: *"In which city's the first time with at least 20 nations?"*
y_2: {Paris}

Results on WTQ (Table-based Complex QA)



Better overall performance

Table-Text Entailment: TabFact (Chen et al., 2020)

United States House of Representatives Elections, 1972

District	Incumbent	Party	Result	Candidates
California 3	John E. Moss	democratic	re-elected	John E. Moss (d) 69.9% John Rakus (r) 30.1%
California 5	Phillip Burton	democratic	re-elected	Phillip Burton (d) 81.8% Edlo E. Powell (r) 18.2%
California 8	George Paul Miller	democratic	lost renomination democratic hold	Pete Stark (d) 52.9% Lew M. Warden , Jr. (r) 47.1%
California 14	Jerome R. Waldie	republican	re-elected	Jerome R. Waldie (d) 77.6% Floyd E. Sims (r) 22.4%
California 15	John J. Mcfall	republican	re-elected	John J. Mcfall (d) unopposed

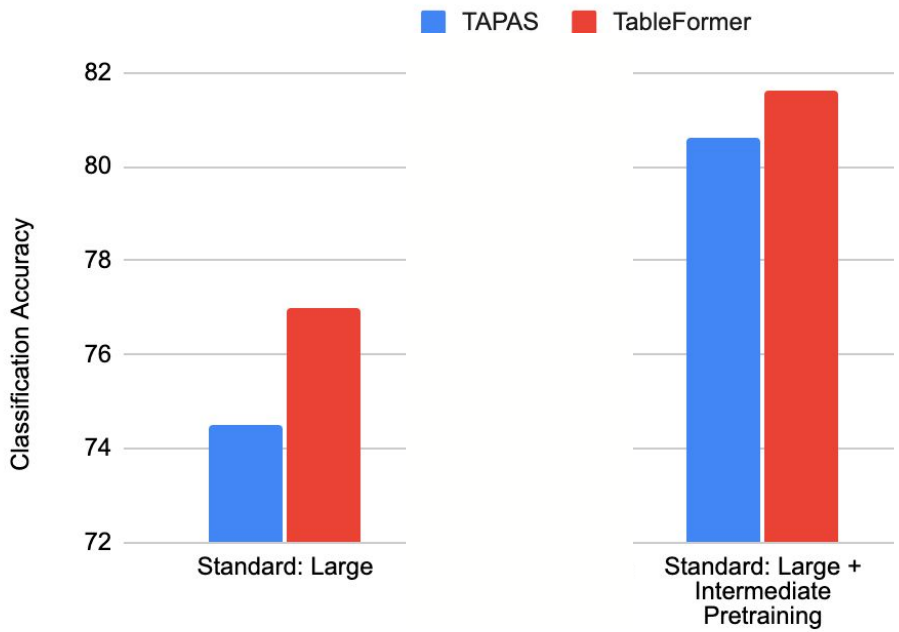
Entailed Statement

1. John E. Moss and Phillip Burton are both re-elected in the house of representative election.
2. John J. Mcfall is unopposed during the re-election.
3. There are three different incumbents from democratic.

Refuted Statement

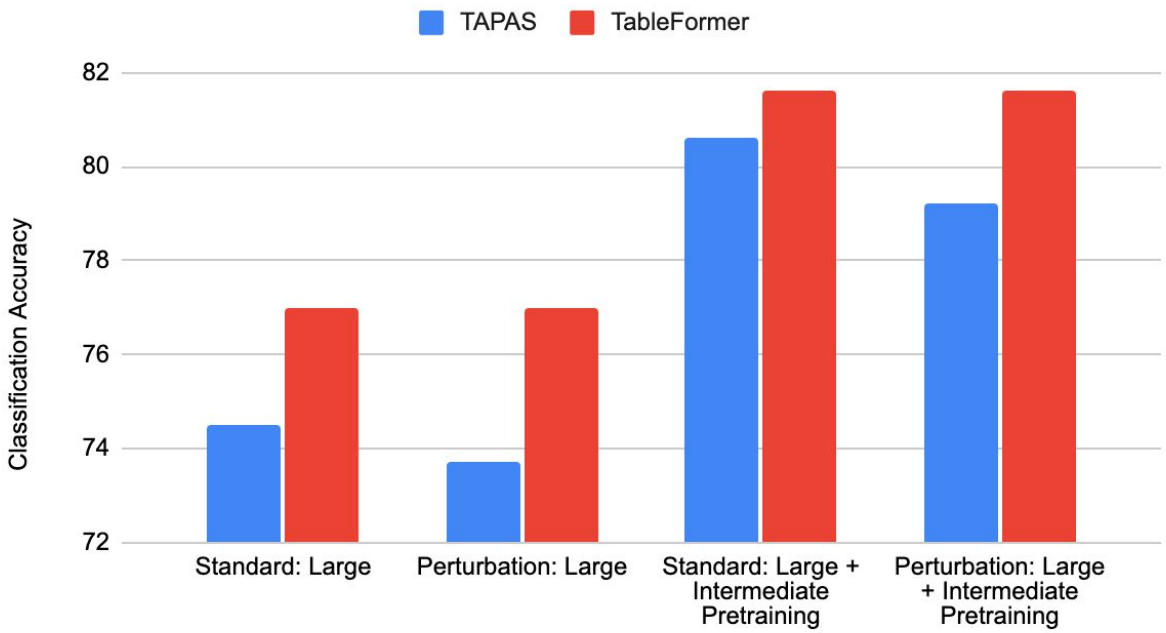
1. John E. Moss and George Paul Miller are both re-elected in the house of representative election.
2. John J. Mcfall failed to be re-elected though being unopposed.
3. There are five candidates in total, two of them are democrats and three of them are republicans.

Results on TabFact (Table-Text Entailment)



Better overall performance on wide range of tasks

Results on TabFact (Table-Text Entailment)



Invariant to perturbations which affect previous approaches!

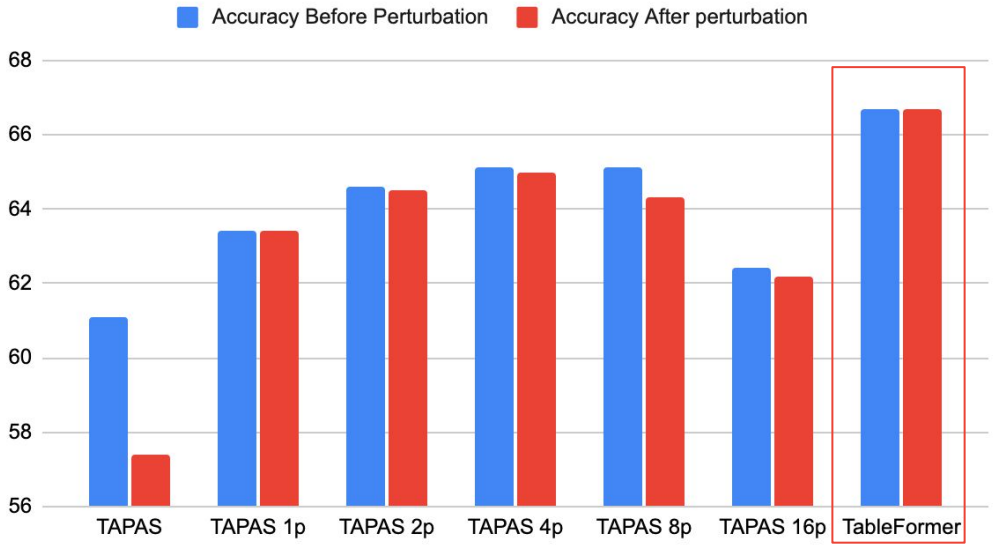
Model Size Comparison

	Number of parameters
TAPAS Base	110 M
TableFormer Base	110 M - 2*512*768 + 12*12*13 = 110 M **- 0.8 M + 0.002 M**
TAPAS Large	340 M
TableFormer Large	340 M - 2*512*1024 + 24*16*13 = 340 M **- 1.0 M + 0.005M**

Better Performance with even fewer parameters!

TableFormer v.s. Perturbed Data Augmentation

Experiment: Augment training data using {1, 2, 4, 8, 16} perturbations



Perturbed data augmentation can improve robustness to some extent, but the performance is still worse than TableFormer.

TableFormer v.s. Perturbed Data Augmentation

Experiment: Augment training data using {1, 2, 4, 8, 16} perturbations

Model	Variation Percentage
TAPAS	14.0%
TAPAS 1p	9.9%
TAPAS 2p	8.4%
TAPAS 4p	8.1%
TAPAS 8p	7.2%
TAPAS 16p	7.0%
TableFormer	**0.0%**

TableFormer has strict robustness in the instance level, while perturbed data augmentation do not have such a guarantee.

TableFormer Attention Bias Ablation Study

SQA dev result	ALL	SEQ
TableFormer base	**62.1**	**38.4**
- same row bias	**32.1**	**2.8**
- same column info	54.5	29.3
- cell to its column header	60.7	36.6
- cell to sentence	60.5	36.4
- header to sentence	61.1	36.3

Same row and column biases are the most important to encode table structures. Cell/header to sentence biases could help better table-text alignment.

Takeaways

- Structural attention biases in TableFormer help understand tables with relative attention and smaller model size.

- Current table encoding methods are not robust to table row and column order perturbation, while TableFormer is guaranteed to be robust to such perturbation.

- TableFormer has advantages over augmenting training data by row and column perturbation.

Thank You!