

# Trans-KBLSTM

Yerram Varun<sup>1\*</sup>, Aayush Sharma<sup>1\*</sup>, Vivek Gupta<sup>2\*</sup>

<sup>1</sup>Indian Institute of Technology, Guwahati; <sup>2</sup>University of Utah



## 1. Tabular Inference Problem

- Inference task where premises are tabular.
- Given a premise table determine hypothesis is true (entailment), false (contradiction), or undetermined (neutral), i.e. tabular natural language inference.

James Hetfield	
Birth Name	James Alan Hetfield
Born	Aug. 3, 1963(age 58), California, U.S.
Genres	Heavy metal, thrash metal, hard rock
Occupation(s)	Musician, Singer
Instruments	Vocals, Guitar
Years active	1978-present
Labels	Warner Bros, Elektra, MegaForce
Hypothesis	James Hetfield was born on the west coast of the USA.

- Example InfoTabS dataset (Gupta et al., 2020), here the Hypothesis is **Entailed**.

## 2. Knowledge Addition

- External knowledge is essential for model reasoning.

Hypothesis	James Hetfield was born on the west coast of the USA.
Focused Relation	coast $\xrightarrow{AtLocation}$ california
Human	Entailment
RoBERTa	Neutral
Trans-KBLSTM	Entailment

Through a *novel architecture*, *Trans-KBLSTM*, this work tackle challenges inherent in **prior methods** of **Knowledge Extraction, Addition, and Integration**.

We did a case study on **INFOTABS**, a Tabular NLI Dataset.

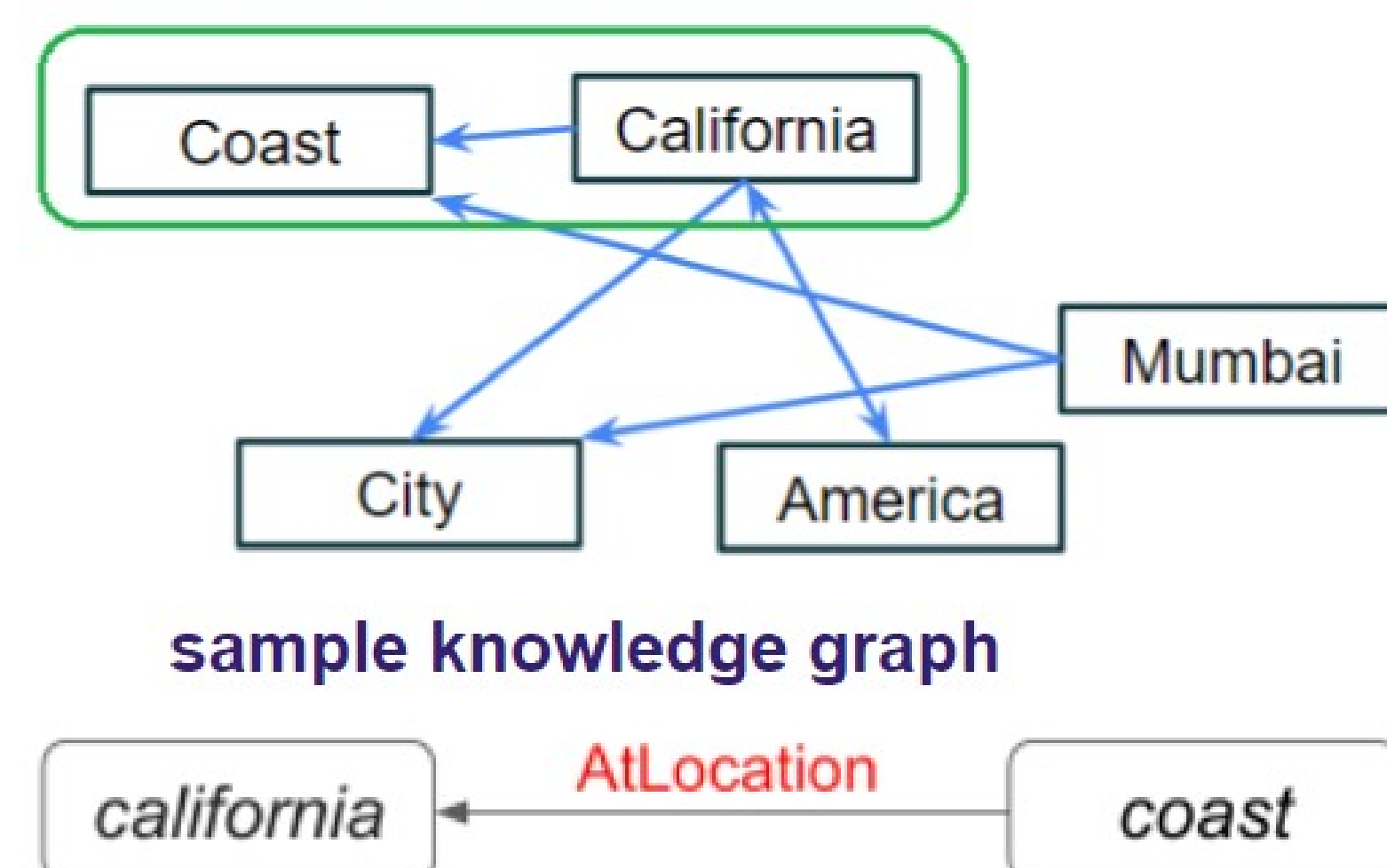
- External knowledge graphs supplement model.

## 3 Challenges and Motivation

- Prior work on knowledge addition for tabular reasoning focus on explicit addition.
- Knowledge Extraction:** Extract contextually relevant knowledge from external source.
- Knowledge Representation:** Effectively represent external semantic knowledge relations.
- Knowledge Integration:** Schematically integrate external knowledge into model design?

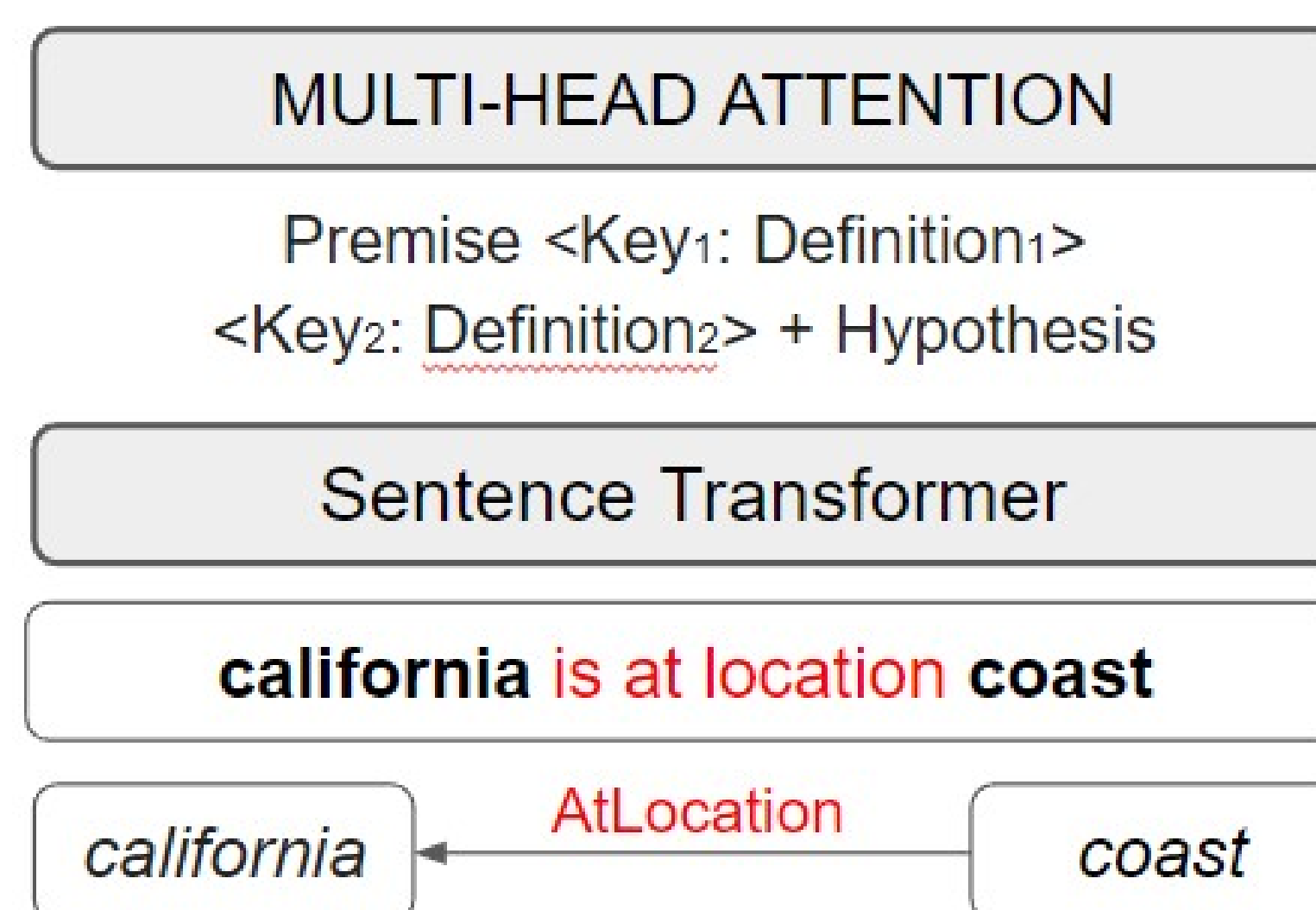
## 4. Knowledge Extraction

- Neeraja et al. 2021 augments the input with lengthy key definitions.
- Add noise and confusion caused by lengthy additions. At times definitions are incorrect.



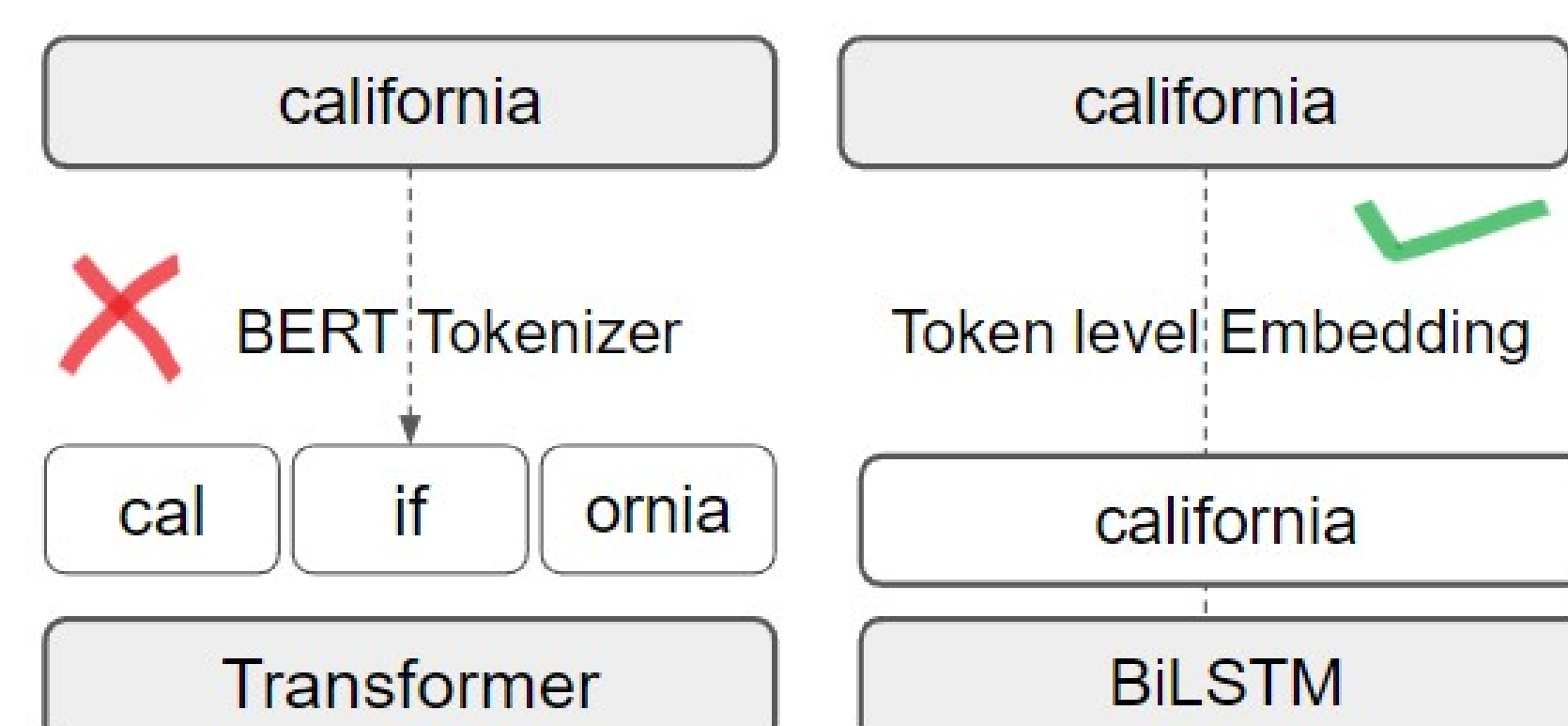
## 5. Knowledge Addition

- Knowledge Triplets are converted to sentences.
- Sentences are encoded using Sentence Transformers.



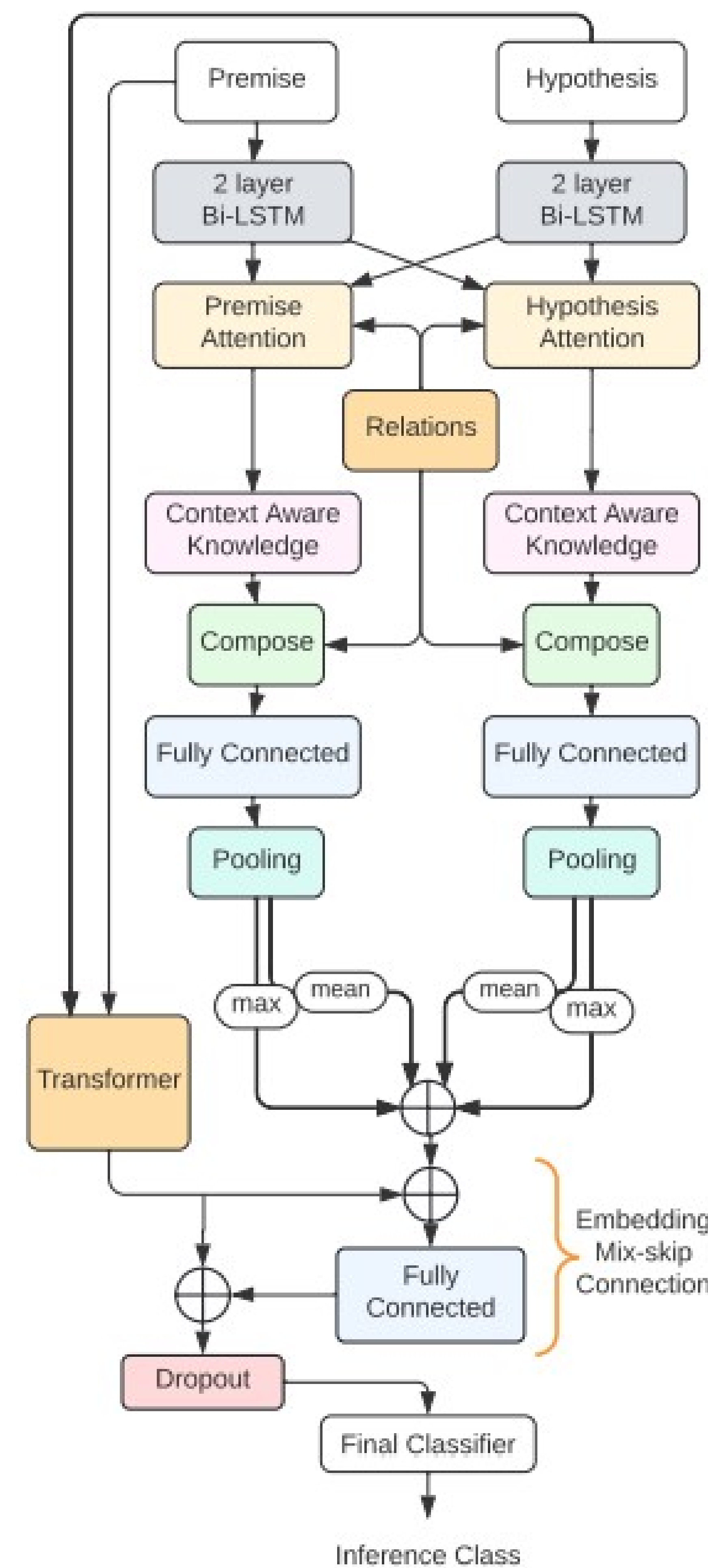
## 6. Knowledge Integration

- Word Pair External Knowledge Relations are not compatible with tokenized transformer inputs.



- BiLSTMs employ token level embeddings, thus complete word pair relations.

## 7. Trans-KBLSTM



## 8. Full Supervision

TRANS-KBLSTM outperform other baselines.

Model	Dev	$\alpha_1$	$\alpha_2$	$\alpha_3$
w/o Knowledge	77.30	76.44	70.49	69.05
Tok-KTrans	78.17	76.19	70.75	69.77
KG Explicit	78.97	77.84	71.13	69.58
Trans-KBLSTM	<b>79.92</b>	<b>79.62</b>	<b>72.10</b>	<b>70.21</b>

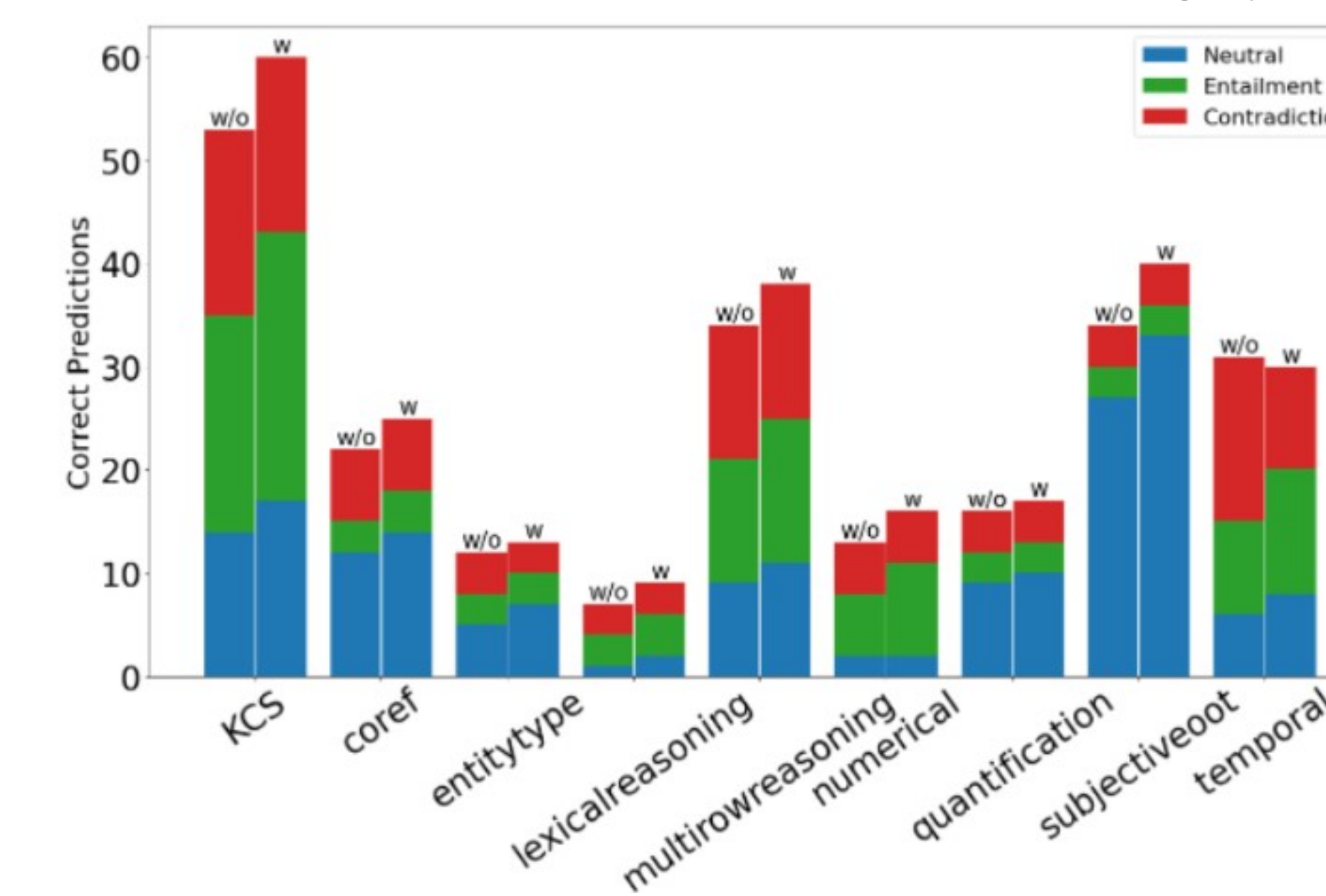
## 9. Limited Supervision

Improvement in this is more substantial.



## 10. Reasoning Types

TRANS-KBLSTM improve all reasoning types.



## 11. Ablation Analysis

Removing Skip connection and addition of random noise adversely affects model performance.

Ablations	Dev	$\alpha_1$	$\alpha_2$	$\alpha_3$
Trans-KBLSTM	<b>67.55</b>	<b>65.16</b>	<b>64.00</b>	<b>63.38</b>
- Skip Connect	65.72	62.83	60.00	61.55
- KB	60.44	61.88	56.94	55.55
-(KB + Skip Connect)	60.11	61.50	55.94	57.38

Joint training is better than Independent training

Ablations	Dev	$\alpha_1$	$\alpha_2$	$\alpha_3$
RoBERTa <sub>LARGE</sub>	77.30	76.44	70.49	69.05
+ KBLSTM (Independent)	79.22	78.38	71.00	69.22
+ KBLSTM (Joint Train)	<b>79.92</b>	<b>79.62</b>	<b>72.10</b>	<b>70.21</b>

Extensive analysis on InfoTabS dataset show TRANS-KBLSTM enhance performance.

Code: <https://trans-kblstm.github.io>