

Logical Form Generation via Multi-task Learning for Complex Question Answering over Knowledge Bases

Xixin Hu, Xuan Wu, Yiheng Shu, Yuzhong Qu

State Key Laboratory for Novel Software Technology,
Nanjing University, Nanjing, China

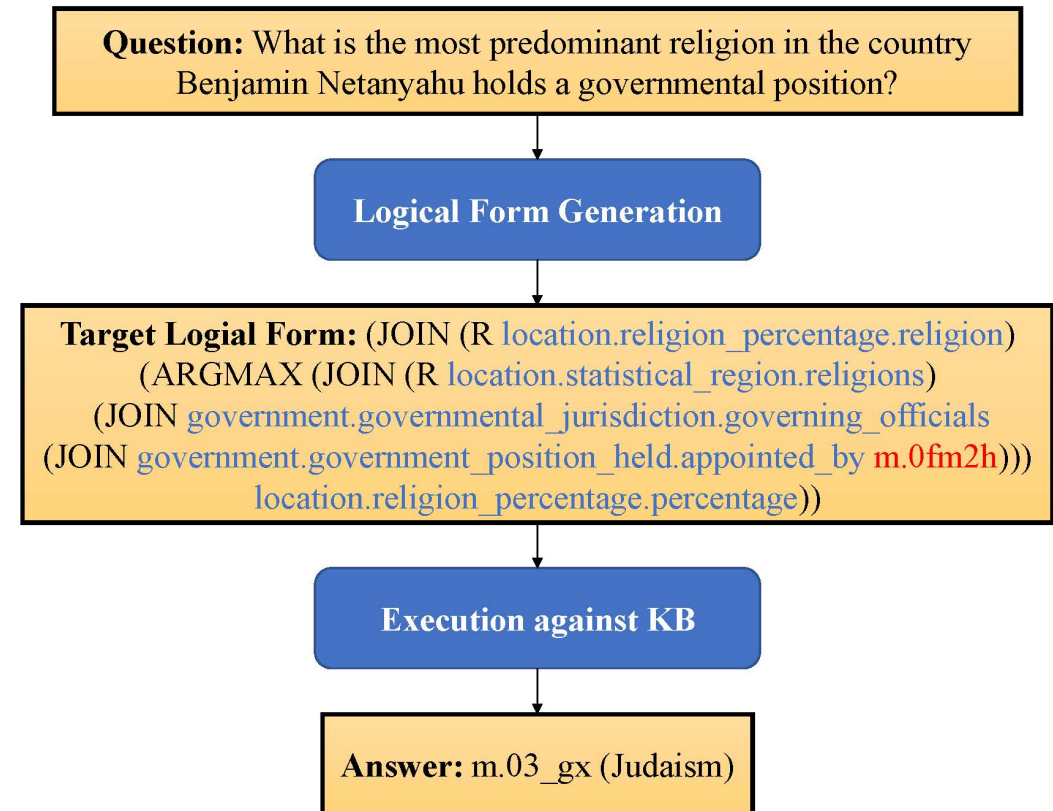
Email: xuanwu@smail.nju.edu.cn

/ CONTENT /

- Background
- Methods: the details of GMT-KBQA
- Experiments
- Summary

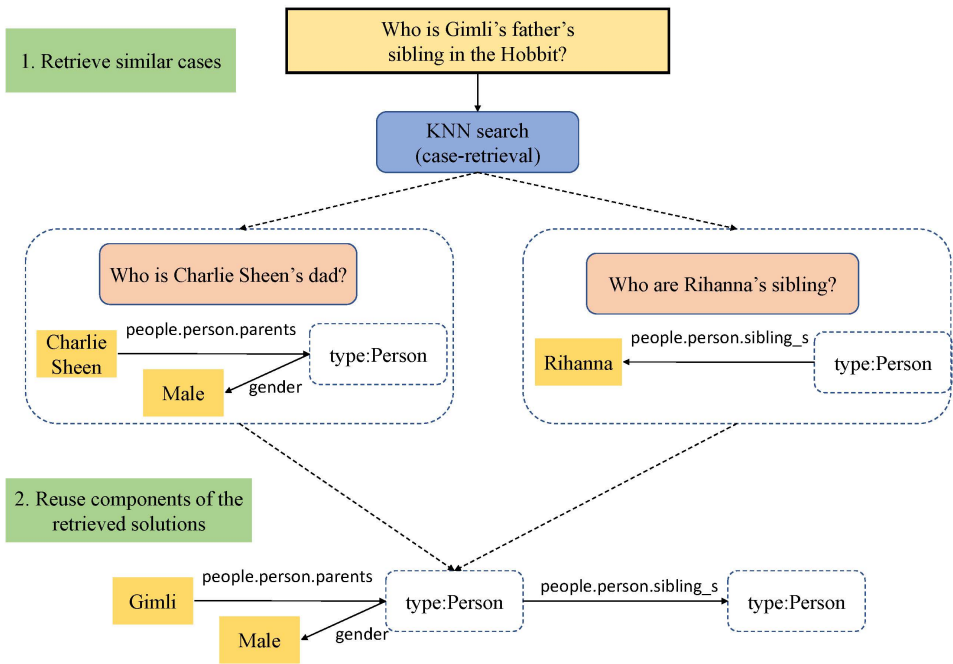
Complex Question Answering over KBs

- Question answering over knowledge bases (KBQA) aims at answering natural language questions based on facts stored in KBs
- Complex questions: requiring multiple KB facts to answer
- Recent KBQA methods cast question answering to Seq2Seq logical form generation task.

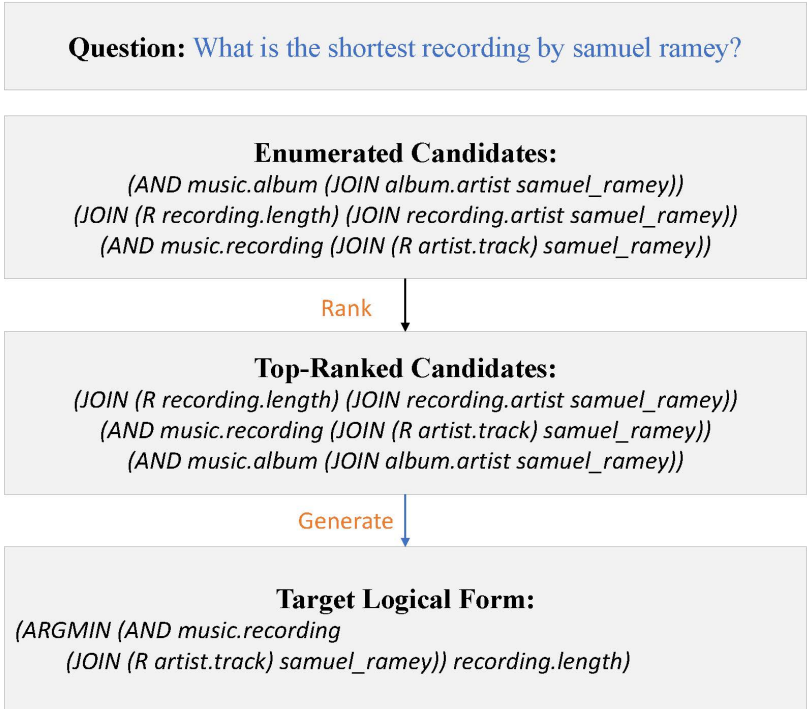


Auxiliary information augmented generation-based KBQA methods

- Auxiliary information is widely used in generation-based methods.
- Generate unseen KB items or novel combinations.



(a) CBR-KBQA (Das et al., 2021)



(b) RnG-KBQA (Ye et al., 2021)

Refine auxiliary information

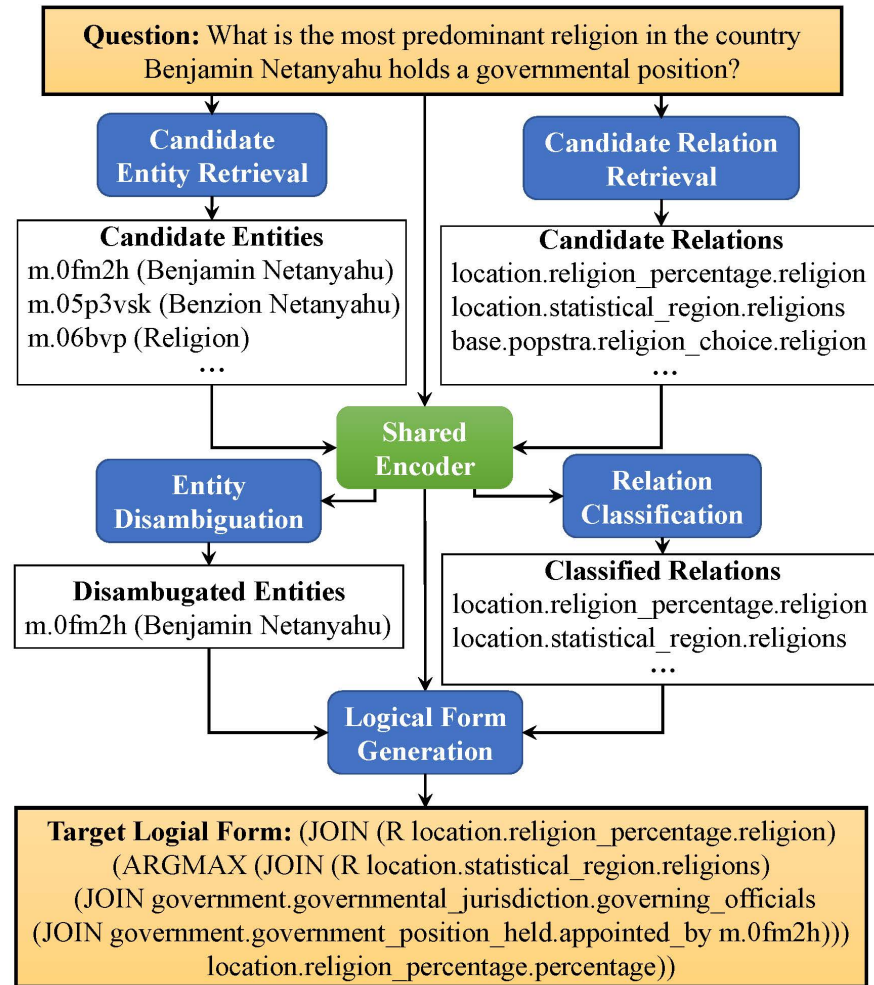
- Noisy information introduced can also lead to incorrect results.
- To refine auxiliary information:
We propose **Generation** via **Multi-Task** learning (GMT-KBQA).

Methods	CWQ		WebQSP	
	F1	Δ	F1	Δ
T5-base	74.0	-	74.7	-
w/ Retrieval	71.4	- 2.6	72.9	- 1.8
w/ Oracle	94.1	+20.1	94.8	+20.1

/ CONTENT /

- Background
- **Methods: the details of GMT-KBQA**
- Experiments
- Summary

Overview of GMT-KBQA

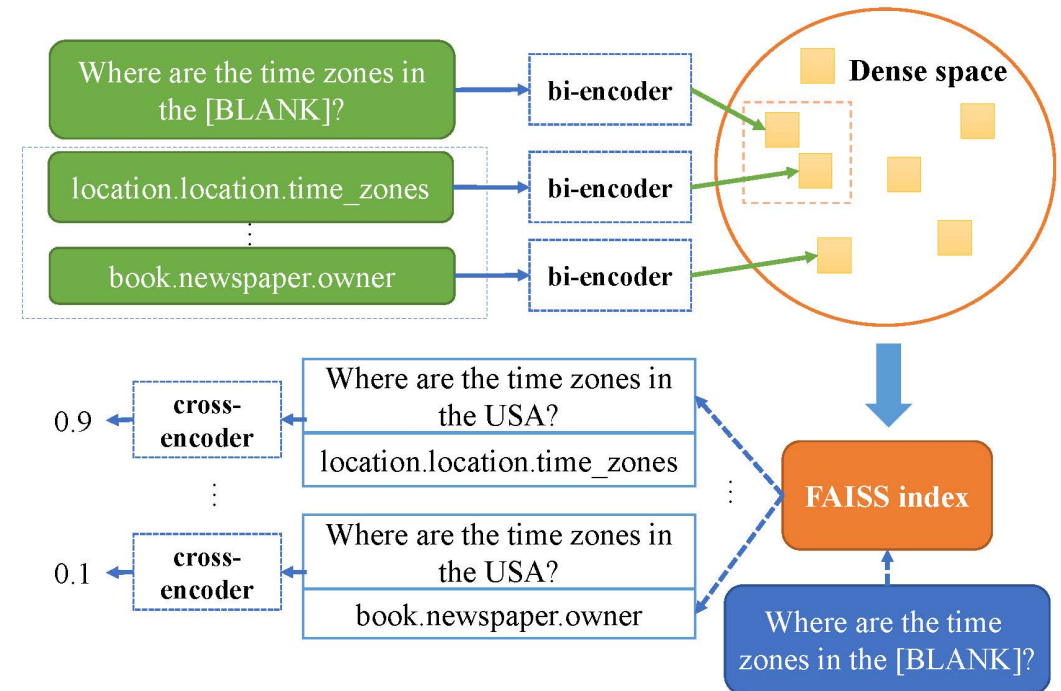


- Retrieve candidate entities and relations in dense space
- Multi-task: **Logical Form Generation**, Entity Disambiguation, Relation Classification
 - Shared parameters
 - Jointly learning

Retrieval of auxiliary information

- Candidate entity retrieval:
 - merge linking results from ELQ (Li et al., 2020) and FACC1 (Gabrilovich et al., 2013).
- Candidate relation retrieval:
 - Bi-encoder:
 - Embeds questions and relations into the same dense space.
 - efficient relations retrieval over KB.

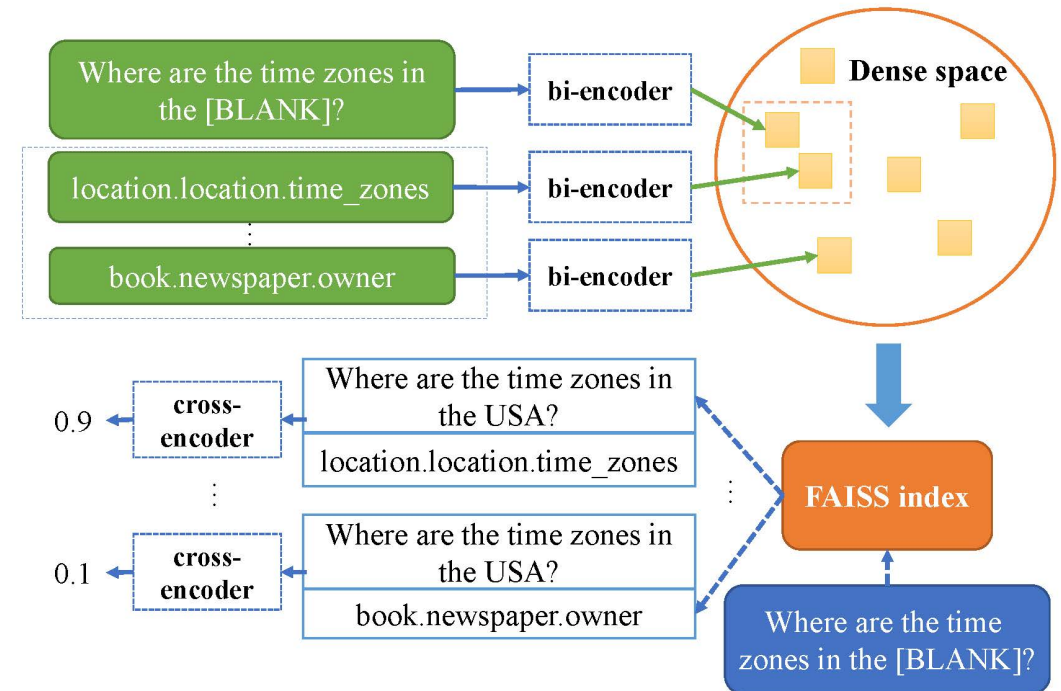
$$\mathcal{L}(q, r_i) = -s_b(q, r_i) + \log \sum_{j=1}^B \exp(s_b(q, r_j))$$



Retrieval of auxiliary information

- Candidate relation retrieval:
 - Cross-encoder: Re-ranking with cross-attention

$$s_c(q, r) = \text{LINEAR}(\text{BERTCLS}([q; r]))$$



Logical Form Generation via Multi-task Learning

Q: Where are the time zones in the USA?

Q </s> location ,
location , time
zones

Q </s> time , time
zone , dst offset
from utc

...

T5 Encoder

Relation
Classification

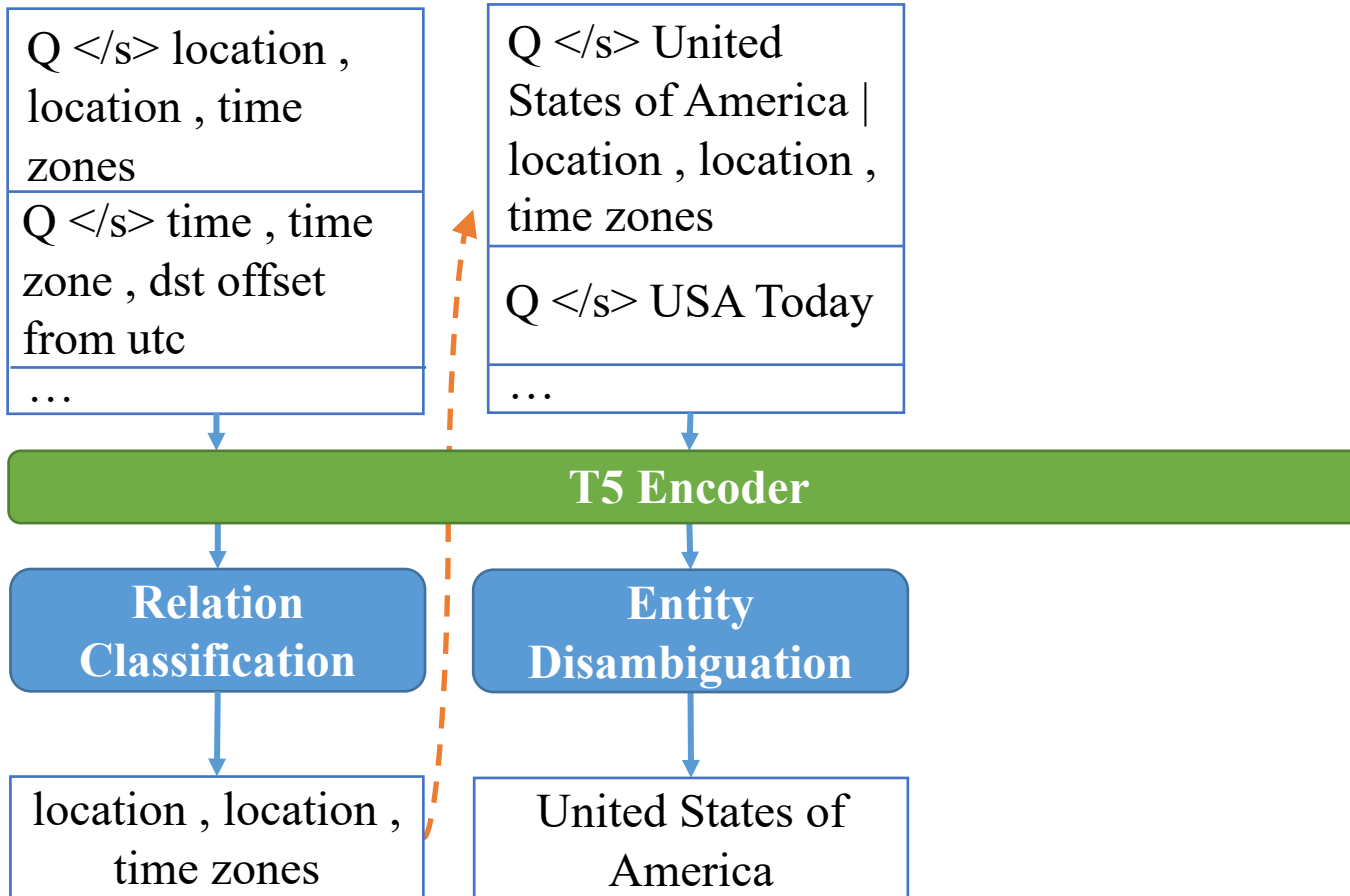
location , location ,
time zones

- Relation Classification:
cross-encoder architecture

$$\mathcal{L}^{\text{REL}} = -\frac{1}{k} \sum_{i=1}^k [u_i \cdot \log(s(q, r_i)) + (1 - u_i) \cdot \log(1 - s(q, r_i))]$$

Logical Form Generation via Multi-task Learning

Q: Where are the time zones in the USA?

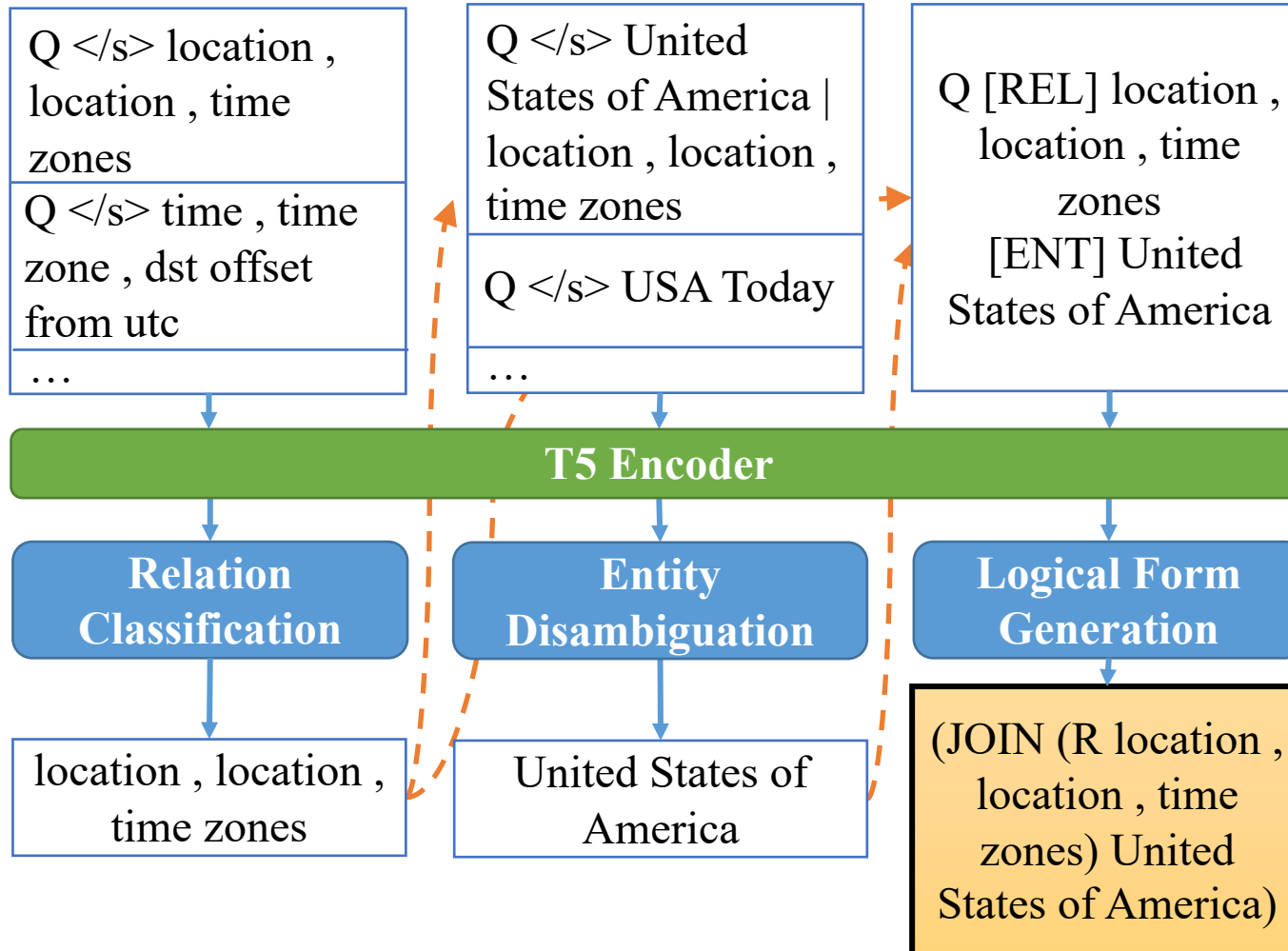


- Entity Disambiguation: cross-encoder, augmented with output of Relation Classification

$$\mathcal{L}^{\text{ENT}} = -\frac{1}{k} \sum_{i=1}^k [v_i \cdot \log(s(q, e_i)) + (1 - v_i) \cdot \log(1 - s(q, e_i))]$$

Logical Form Generation via Multi-task Learning

Q: Where are the time zones in the USA?



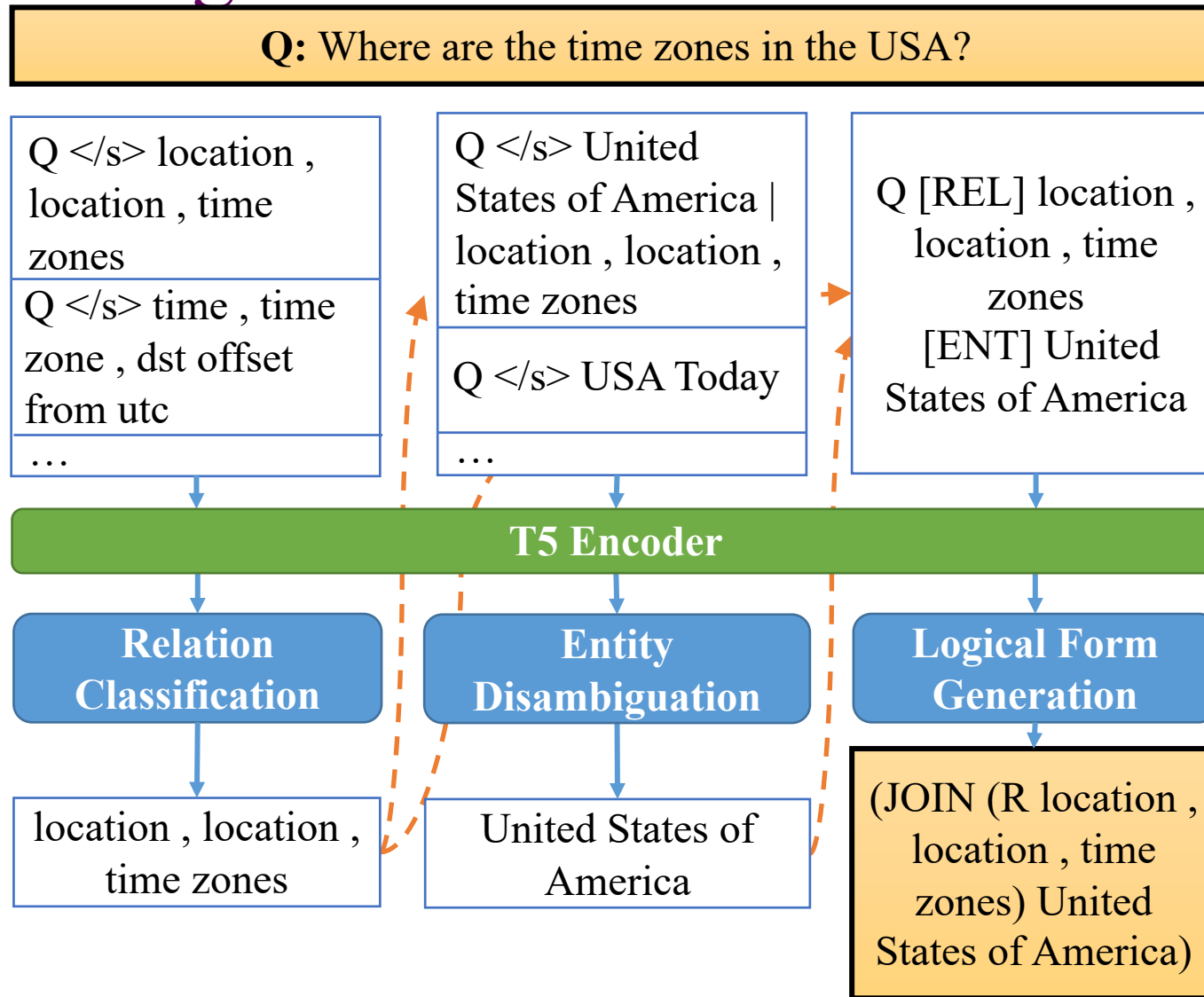
- Final Logical Form Generation
 - Augmented with refined entities and relations

$$\mathbf{p}_j = \text{T5DECODER}(a_1, \dots, a_{j-1}, \mathbf{h}_1, \dots, \mathbf{h}_n)$$

$$\mathcal{L}^{\text{GEN}} = -\frac{1}{m} \sum_{j=1}^m \log \mathbf{p}_{j,a_j}$$

$$\mathcal{L} = \mathcal{L}^{\text{REL}} + \mathcal{L}^{\text{ENT}} + \mathcal{L}^{\text{GEN}}$$

Logical Form Generation via Multi-task Learning



- Refinement of auxiliary information
 - Relation Classification: cross-encoder architecture
 - Entity Disambiguation: cross-encoder, augmented with output of Relation Classification
- Final Logical Form Generation
 - Augmented with refined entities and relations

/ CONTENT /

- Background
- Methods: the details of GMT-KBQA
- Experiments
- Summary

KBQA Performance

■ Compared to other methods

Method	CWQ		WebQSP	
	Acc.	F1	Acc.	F1
QGG (Lan and Jiang, 2020)	-	40.4	-	74.0
BART-large (Huang et al., 2021)	-	68.2	-	74.6
CBR-KBQA (Das et al., 2021)	67.1	70.0	69.9	72.8
ReTraCk* (Chen et al., 2021)	-	-	-	74.7
RnG-KBQA (Ye et al., 2021)	-	-	71.1	75.6
GMT-KBQA (Ours)	72.2	77.0	73.1	76.6

Table 1: QA evaluation results (%) on CWQ test set and WebQSP test set. * denotes using oracle entity linking results. Acc.: Accuracy.

■ Compared to model variants

Methods	CWQ		WebQSP	
	F1	Δ	F1	Δ
GMT-KBQA	77.0	-	76.6	-
w/o Entity	72.7	-4.3	74.9	-1.7
w/o Relation	75.9	-1.1	75.1	-1.5
w/o Entity, Relation	74.0	-3.0	74.7	-1.9

Table 3: QA performance (%) of variants of our model. *Entity* and *Relation* are short for entity disambiguation task and relation classification task respectively.

Further Analysis

■ Refinement of auxiliary information

Methods	CWQ						WebQSP					
	Entity linking			Relation linking			Entity linking			Relation linking		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
Retriever	71.6	70.5	68.8	84.9	74.2	77.3	53.9	86.0	62.6	66.3	68.7	65.4
GMT-KBQA	78.8	69.4	72.3	85.2	81.0	81.4	79.9	75.3	76.1	74.1	73.9	72.1

Table 2: Entity and relation linking evaluation (%) of our retriever and multi-task model.

■ Performance on questions with unseen KB items

Methods	CWQ			WebQSP		
	P	R	F1	P	R	F1
T5-base	64.8	67.0	64.9	64.3	68.5	64.5
GMT-KBQA	68.3	70.9	68.5	66.7	69.8	66.6

Table 5: QA results (%) on test set questions with unseen KB items.

/ CONTENT /

- Background
- Methods: the details of GMT-KBQA
- Experiments
- **Summary**

Summary

- We propose a generation-based KBQA method via multi-task learning (GMT-KBQA).
 - GMT-KBQA retrieves candidate relations and entities through dense retrieval, which balances coverage and efficiency.
 - Refine noisy auxiliary information, enabling GMT-KBQA to generate unseen KB items.
- Experiment results on both CWQ and WebQSP demonstrates the effectiveness of our method. Further analysis shows that our method benefits from the multi-task learning framework.

Thanks for Listening! Q&A

Email: xuanwu@smail.nju.edu.cn

Repository: <https://github.com/HXX97/GMT-KBQA>

