Synergetic Interaction Network with Cross-task Attention for Joint Relational Triple Extraction

Da Luo ⋆, Run Lin⋆, Qiao Liu*, Yuxiang Cai, Xueyi Liu, Yanglei Gan, Rui Hou

University of Electronic Science and Technology of China

(⋆ Co-First Author, * Corresponding author)

May, 2024

Outline

1. Introduction
2. Proposed Framework
3. Experiments
4. Future Work

Outline

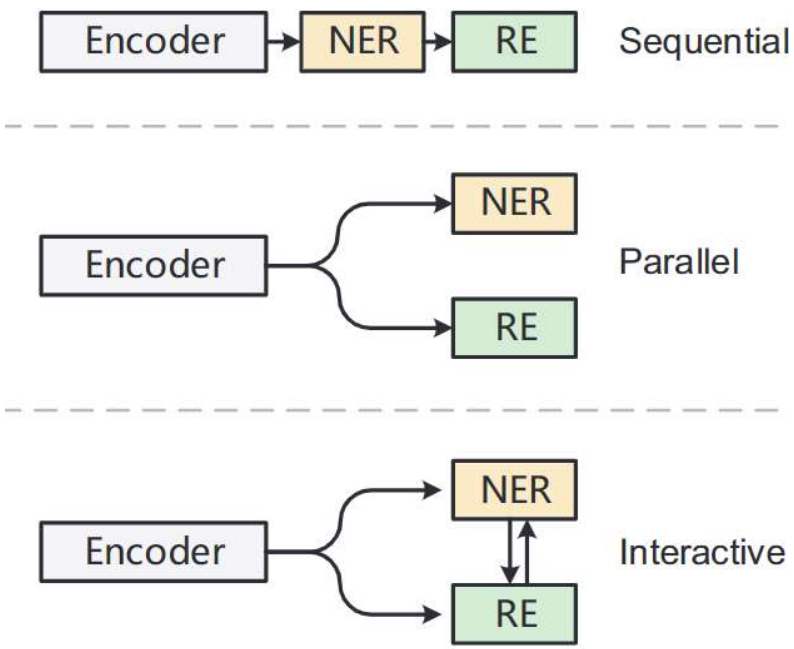
Task Definition

Calvin had a delicious **lasagna** near the University of Turin, Italy		
Subject	Relation	Object
Calvin	Eat	lasagna

Joint relational triple extraction simultaneously identifies entities and their relations in text to create relational triples **<Subject, Relation, Object>**.

Related Works

Recent methods have explored three main categories:



- Sequence encoding method
  - MHSM (Bekoulis et al., 2018b), CASREL (Wei et al.,2020), AT (Bekoulis et al., 2018a)
- Parallel encoding method
  - Graphrel (Fu et al., 2019), DYGIE++ (Wadden et al., 2019)
- Interactive encoding method
  - RIN (Sun et al., 2020), MD-RNN (Wang and Lu,2020),
  - PFN (Yan et al., 2021), MGE (Xiong et al., 2022)

Contributions

- We present a novel synergetic interaction network that enables effective bi-directional interaction between NER and RE sub-tasks by leveraging contextual association.
- Our proposed method leverages a cross-task attention mechanism to boost interaction between NER and RE sub-tasks, enhancing contextual comprehension and inference capabilities while ensuring efficient computational and memory use.
- The experimental results on three standard benchmarks indicate that our method performs better than state-of-the-art baselines.

Outline

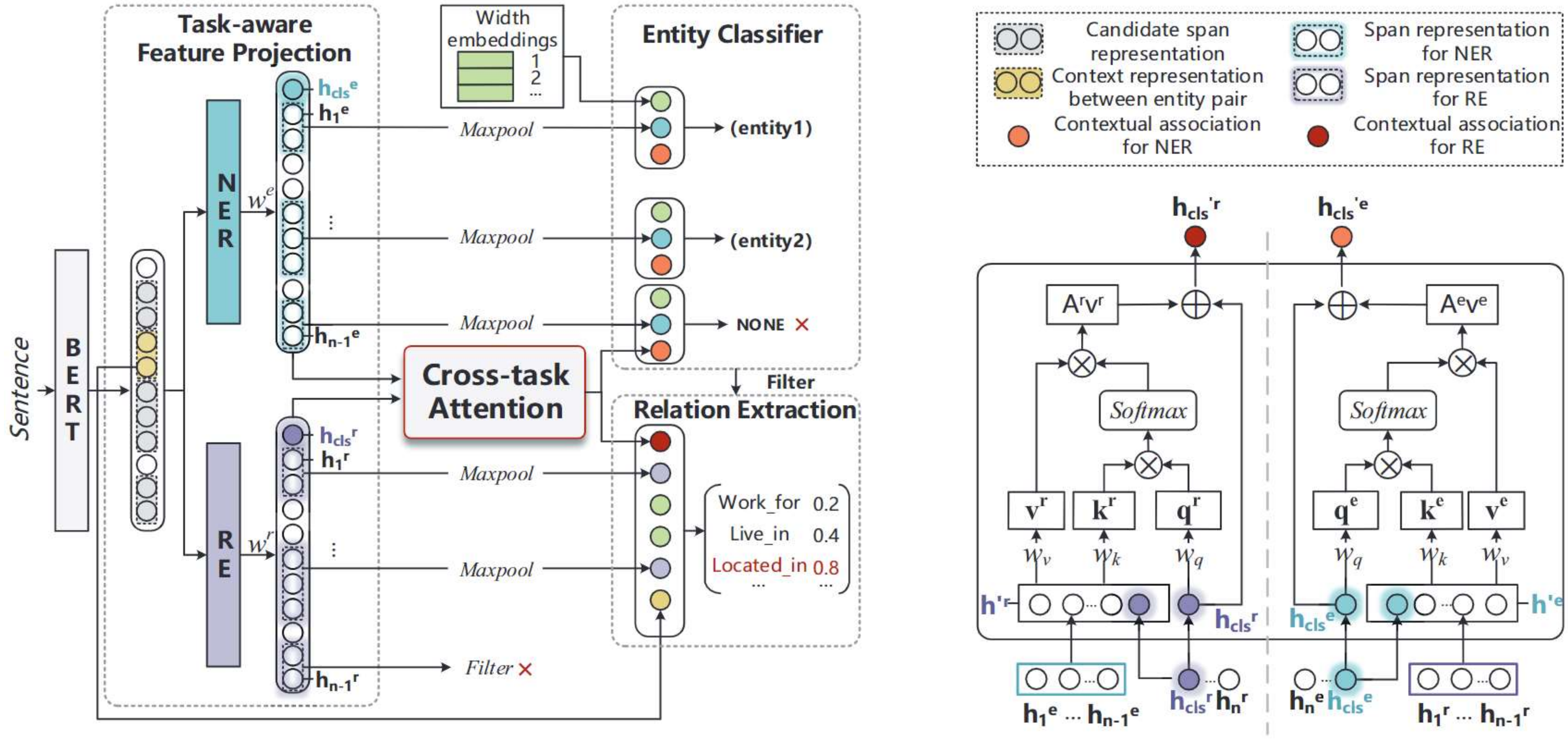
1 Introduction

2 Proposed Framework

3 Experiments

4 Future Work

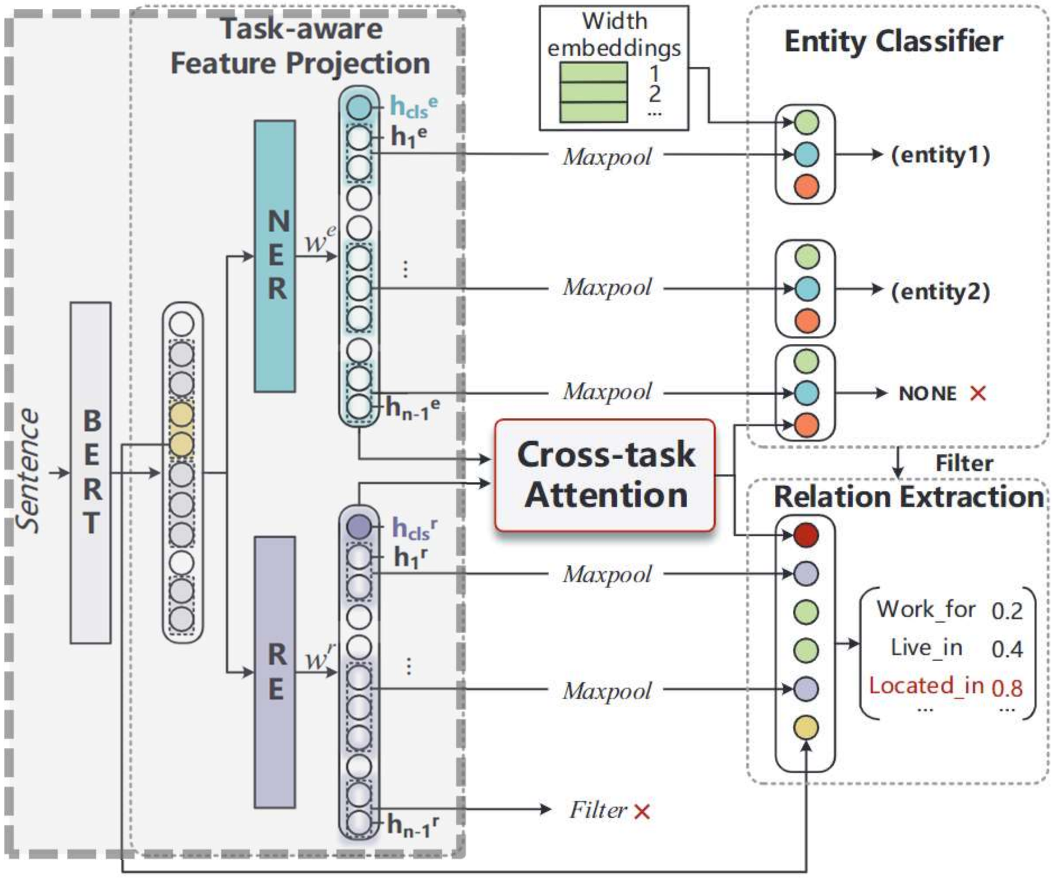
Model Overview



(a) Framework of SINET

(b) Cross-task Attention Mechanism

Input Layer



(a) Framework of SINET

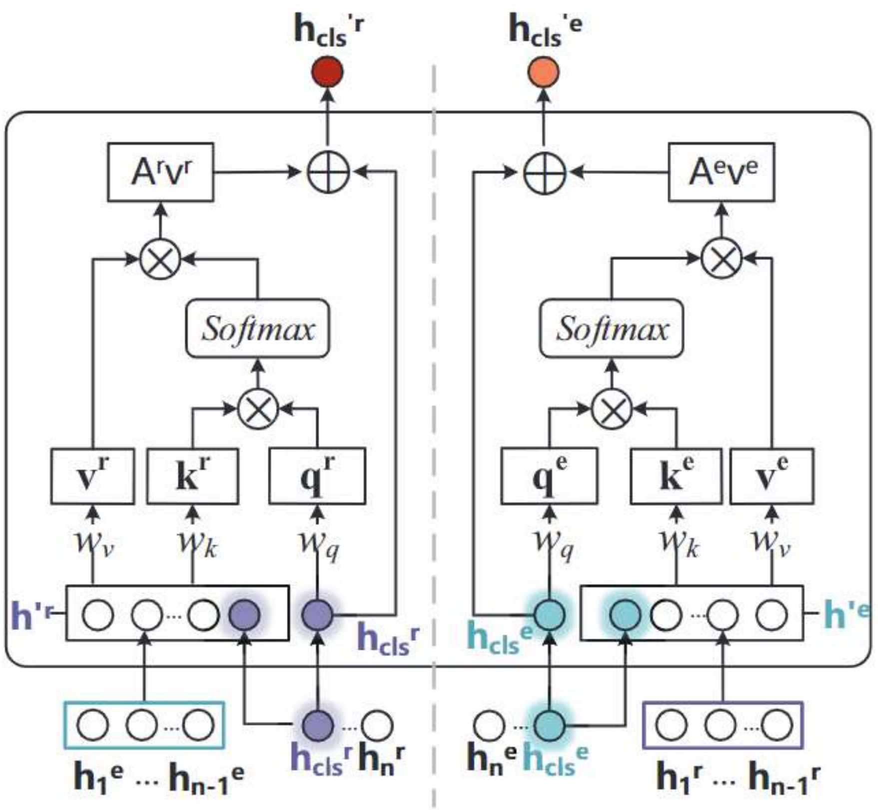
- the input sentence X of BERT encoder :

$$Y_{enc}(X) = \{\boldsymbol{h_{cls}}, \boldsymbol{h_1}, \boldsymbol{h_2}, \ldots, \boldsymbol{h_n} \,|\boldsymbol{h_i} \in \mathbb{R}^{d\times 1}\}$$

- Task-aware representation:

$$\boldsymbol{h_i^e} = f^e(\boldsymbol{h_i}) = \boldsymbol{w^e h_i}, \boldsymbol{h_i^r} = f^r(\boldsymbol{h_i}) = \boldsymbol{w^r h_i}$$

Cross-task Attention Mechanism (CTAM)



(b) Cross-task Attention Mechanism

- Representation combination between task-aware features:

$$\boldsymbol{h}_i'^{e} = [\boldsymbol{h}_i^r; \boldsymbol{h}_{cls}^e],\ \boldsymbol{h}_i'^{r} = [\boldsymbol{h}_i^e; \boldsymbol{h}_{cls}^r]$$

- Model the contextual associations between NER and RE:

$$\boldsymbol{q}^e = \boldsymbol{w}_q \boldsymbol{h}_{cls}^e, \boldsymbol{k}^e = \boldsymbol{w}_k \boldsymbol{h}'^e, \boldsymbol{v}^e = \boldsymbol{w}_v \boldsymbol{h}'^e$$

$$\boldsymbol{q}^r = \boldsymbol{w}_q \boldsymbol{h}_{cls}^r, \boldsymbol{k}^r = \boldsymbol{w}_k \boldsymbol{h}'^r, \boldsymbol{v}^r = \boldsymbol{w}_v \boldsymbol{h}'^r$$

$$\boldsymbol{A}^e = softmax\left(\boldsymbol{q}^e \boldsymbol{k}^{eT} / \sqrt{d/h}\right)$$

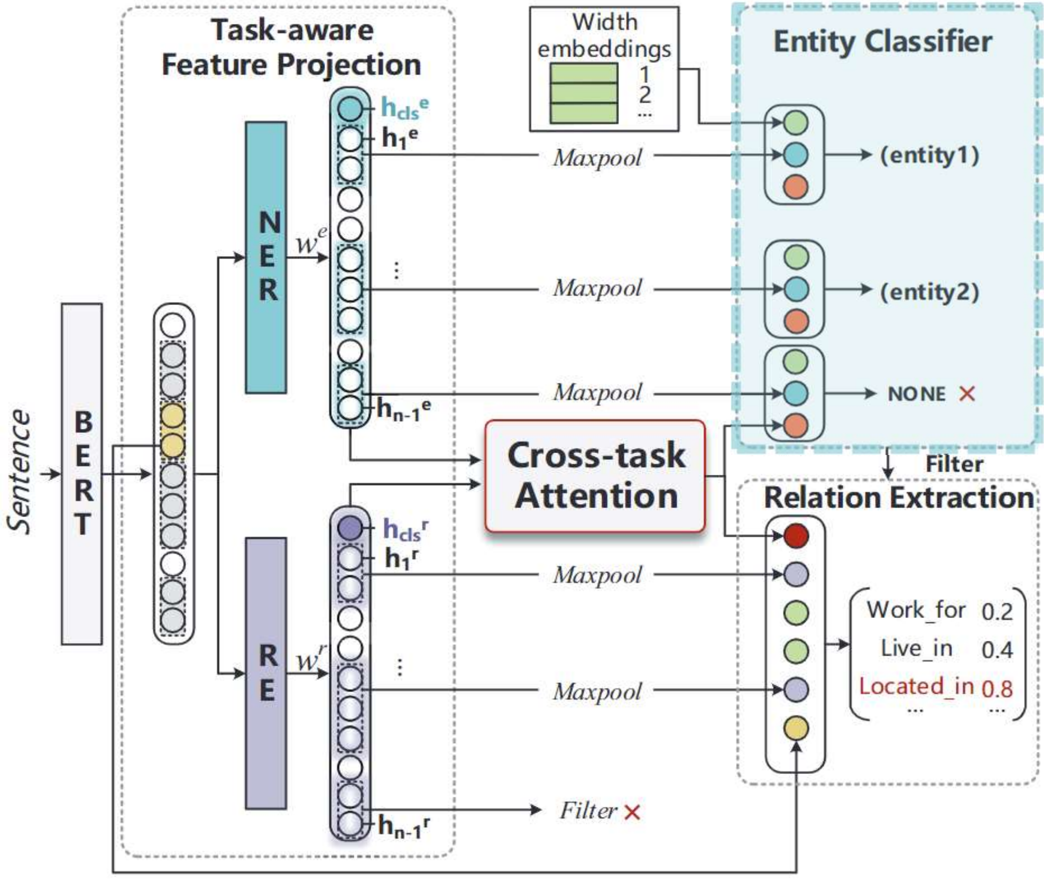
$$\boldsymbol{A}^r = softmax\left(\boldsymbol{q}^r \boldsymbol{k}^{rT} / \sqrt{d/h}\right)$$

$$\boldsymbol{h}'^e_{cls} = \boldsymbol{A}^e \boldsymbol{v}^e + \boldsymbol{h}_{cls}^e,\ \boldsymbol{h}'^r_{cls} = \boldsymbol{A}^r \boldsymbol{v}^r + \boldsymbol{h}_{cls}^r$$

- Output of CTAM with Layer normalization:

$$\boldsymbol{h}'^e_{cls} = LN\left(\boldsymbol{h}'^e_{cls}\right), \boldsymbol{h}'^r_{cls} = LN\left(\boldsymbol{h}'^r_{cls}\right)$$

Entity Classifier



(a) Framework of SINET

- The entity span representation:

$$s_i^e = Maxpool(h_i^e, h_{i+1}^e, \dots, h_j^e)$$

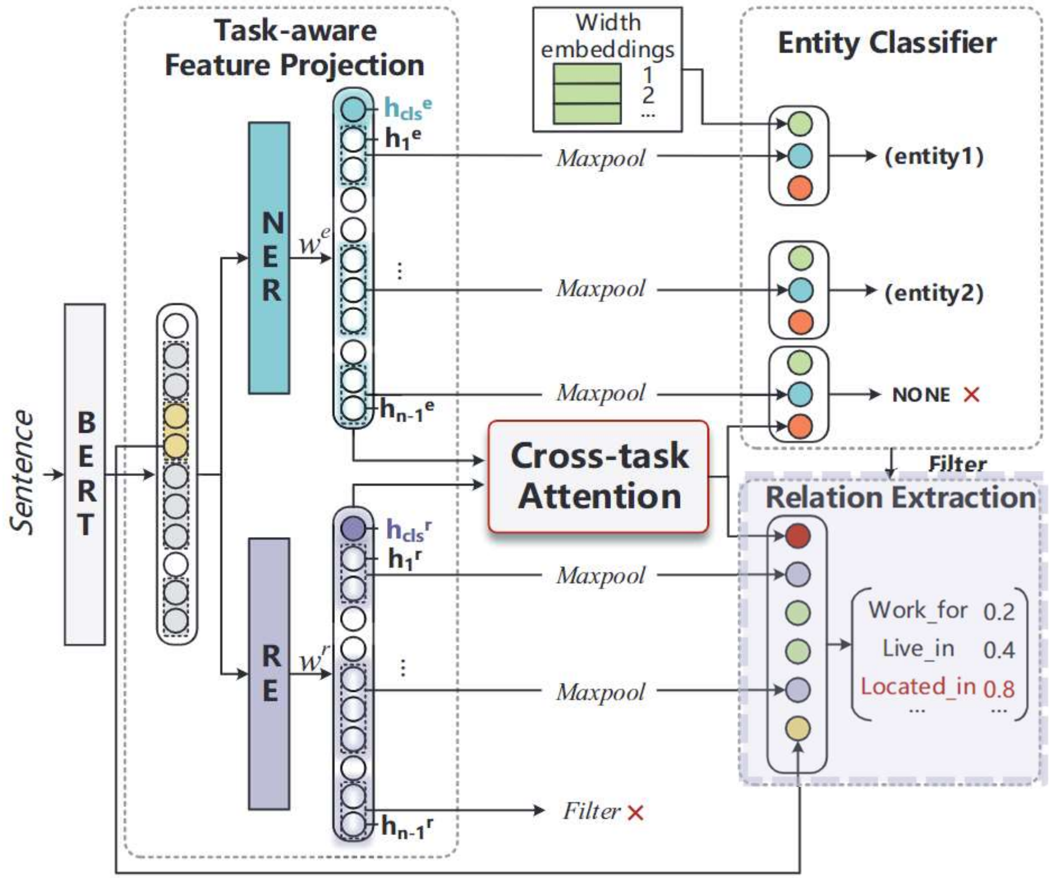
- The input to entity classifier:

$$e_i = [s_i^e;\ w_k;\ h'^e_{cls}]$$

- e_i is then fed to a Feed-forward Neural Network:

$$y^e = softmax(w_{ner} e_i + \ b^e)$$

Multi-label Relation Extraction



(a) Framework of SINET

- The relation span's embeddings:

$$\boldsymbol{s}_i^r = Maxpool(\boldsymbol{h}_i^r, \boldsymbol{h}_{i+1}^r, \ldots, \boldsymbol{h}_j^r)$$

$$\boldsymbol{e}(\boldsymbol{s_1}) = [\boldsymbol{s}_1^r;\ \boldsymbol{w_k}], \boldsymbol{e}(\boldsymbol{s_2}) = [\boldsymbol{s}_2^r;\ \boldsymbol{w_k}]$$

- The local context, $c(s_1, s_2)$, is derived from BERT embeddings of tokens between two entity spans using max pooling.

- The input to relation classifier:

$$\boldsymbol{r_1} = [\boldsymbol{e}(\boldsymbol{s_1}); c(\boldsymbol{s_1}, \boldsymbol{s_2}); \boldsymbol{e}(\boldsymbol{s_2}); \boldsymbol{h'}_{cls}^r]$$

$$\boldsymbol{r_2} = [\boldsymbol{e}(\boldsymbol{s_2}); c(\boldsymbol{s_2}, \boldsymbol{s_1}); \boldsymbol{e}(\boldsymbol{s_1}); \boldsymbol{h'}_{cls}^r]$$

- $\boldsymbol{r_1}$ and $\boldsymbol{r_2}$ are passed to the relation classifier:

$$\boldsymbol{y}^r = Sigmoid(\boldsymbol{w_{re}} \boldsymbol{r_{1/2}} + \boldsymbol{b}^r)$$

Outline

Datasets & Evaluation Metrics

- Datasets
  - ACE04, ACE05 (Walker et al., 2005)
  - SciERC (Luan et al., 2018)
- Evaluation Metrics
  - **micro F1-score**
  - One evaluation metric for NER
    - **Ent:** requires both correct type and boundary.
  - Two evaluation metrics for RE
    - **Boundaries evaluation (Rel):** considers a triple prediction correct only if the predicted relation and the boundaries of two spans are correct.
    - **Strict evaluation (Rel+):** requires predicted entity types to be correct in addition to satisfying the conditions of the boundaries evaluation.

Main Results

Model	Encoder	ACE04			ACE05			SciERC		
		Ent	Rel	Rel+	Ent	Rel	Rel+	Ent	Rel	Rel+
SPTree (Miwa and Bansal, 2016)	Bb	81.80	-	48.40	83.40	-	55.60	-	-	-
SciIE (Luan et al., 2018)	SciB	-	-	-	-	-	-	64.20	39.30	-
DYGIE (Luan et al., 2019)	Bb	87.40	59.70	-	-	-	-	-	-	-
DYGIE++ (Wadden et al., 2019)	Bb	-	-	-	88.60	63.40	-	-	-	-
DYGIE++ (Wadden et al., 2019)	SciB	-	-	-	-	-	-	67.50	48.40	-
Multi-turn QA (Li et al., 2019)	Bb	83.60	-	49.40	84.80	-	60.20	-	-	-
Two are Better than One (Wang and Lu, 2020)	ALB	88.60	-	59.60	89.50	-	64.30	-	-	-
SPE (Wang et al., 2020)	SciB	-	-	-	-	-	-	66.90	-	33.60
SpERT (Eberts and Ulges, 2020)	SciB	-	-	-	-	-	-	70.33	50.84	-
SPAN$_{Multi-Head}$ (Ji et al., 2020)	Bb	-	-	-	89.59	65.24	-	-	-	-
PURE (Zhong and Chen, 2021)	ALB	88.80	64.70	60.20	89.70	69.00	65.60	-	-	-
PURE (Zhong and Chen, 2021)	SciB	-	-	-	-	-	-	66.60	48.20	35.60
PFN (Yan et al., 2021)	ALB	89.30	-	62.50	89.00	-	66.80	-	-	-
PFN (Yan et al., 2021)	SciB	-	-	-	-	-	-	66.80	-	38.40
MGE (Xiong et al., 2022)	ALB	89.30	-	63.80	89.70	-	**68.20**	-	-	-
MGE (Xiong et al., 2022)	SciB	-	-	-	-	-	-	68.40	-	39.40
SINET (Ours)	Bb	88.27	60.86	57.34	88.58	66.18	62.83	-	-	-
	SciB	-	-	-	-	-	-	**72.59**	**51.01**	**40.13**
	ALB	**90.53**	**66.53**	**64.65**	**90.56**	**69.04**	65.71	-	-	-

Ablation Study

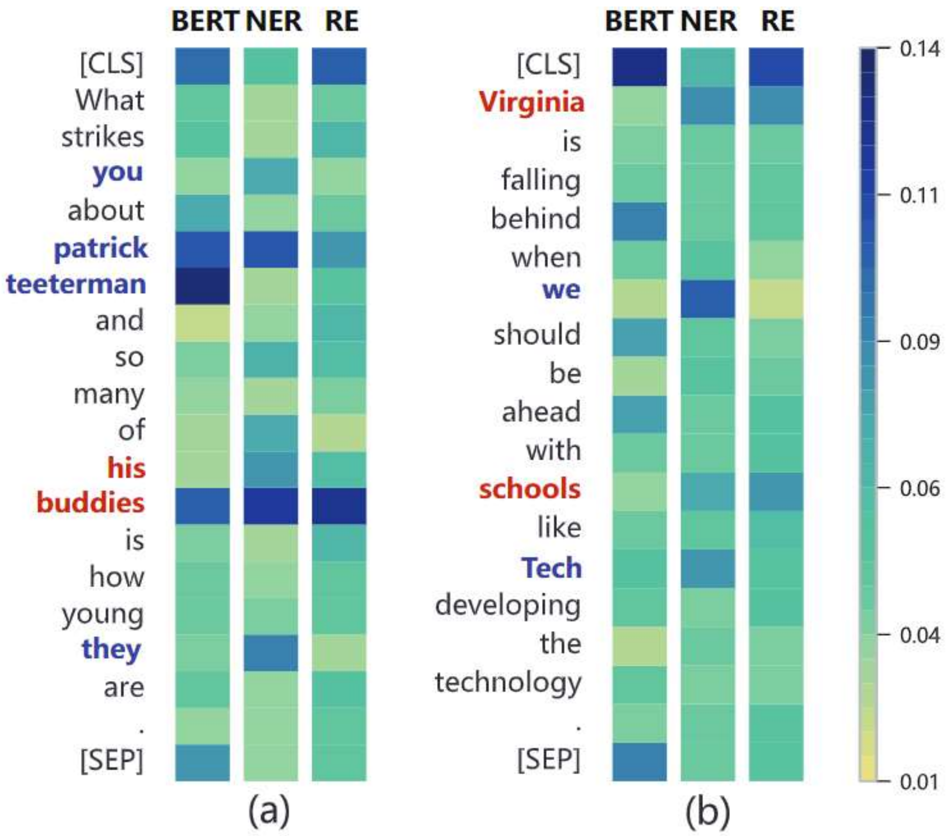
- Effects of synergetic interaction

Settings		Ent	Rel	Rel+
SciERC	Base	73.99	50.49	41.85
	Base+TAP	74.48	50.54	41.94
	Base+TAP+CTAM$_{NER}$	74.26	50.36	42.70
	Base+TAP+CTAM$_{RE}$	74.01	49.84	41.63
	Base+TAP+CTAM	**74.65**	**50.99**	**42.91**
ACE05	Base	86.57	63.70	60.42
	Base+TAP	86.75	63.83	60.56
	Base+TAP+CTAM$_{NER}$	86.76	63.80	60.35
	Base+TAP+CTAM$_{RE}$	86.83	64.09	60.98
	Base+TAP+CTAM	**86.85**	**64.67**	**61.42**

- **Base**: Directly feed the output from the BERT encoder into the entity and relation classifiers without additional encoding or interaction.
- **Base+TAP**: h^{e}_{cls} and h^{r}_{cls}, generated by the TAP, are utilized for NER and RE respectively.
- **Base+TAP+CTAM$_{\text{NER}}$**: h'^{e}_{cls} obtained from CTAM, and h^{r}_{cls} obtained from TAP are leveraged for NER and RE respectively.
- **Base+TAP+CTAM$_{\text{RE}}$**: h^{e}_{cls} obtained from TAP, and h'^{r}_{cls} obtained from CTAM are leveraged for NER and RE respectively.
- **Base+TAP+CTAM**: h'^{e}_{cls} and h'^{r}_{cls} obtained from CTAM are leveraged for NER and RE respectively.

Case Study

- Visualization



- **In-triple entities:** Entities and relations that form a valid relational triple, representing a meaningful relationship within the context.

  golden triple:

  <"his", "PER-SOC", "buddies">

  <"Virginia", "PART-WHOLE", "schools">

- **Out-of-triple entities**: Entities may be semantically linked in the text but do not form a valid relational triple.

Model Efficiency

- Efficiency

	Model	FLOPs (M)	Inference Time (s)
SciERC	SPAN_{Multi_Head}	3892.93	21
	PFN	1517.51	34
	SINET	**1279.25**	$\boxed{\textbf{11}}$
ACE05	Two are Better than One	**3867.13**	117
	PFN	26970.74	134
	SINET	26113.61	$\boxed{\textbf{65}}$

Comparison of model efficiency on Sci-ERC (SciBERT) and ACE05 (ALBERT-xxlarge-v1).

Outline

Future Work

- We will further optimize the effect of SINET in relational triple extraction and its computational efficiency, striving for the best balance between performance and efficiency.

- We attempt to apply and improve Cross-task Attention Mechanism to more NLP tasks, such as the Aspect Sentiment Triplet Extraction (ASTE) task and other interactive tasks.

THANKS FOR YOUR LISTENING

- Code available at https://github.com/AONE-NLP/RTE-SINET
- If you have problem, feel free to send email to
  runlin@uestc.edu.cn, qliu@uestc.edu.cn

Synergetic Interaction Network with Cross-task Attention
for Joint Relational Triple Extraction