Unified Representation and Interaction for Joint Relational Triple Extraction

Wei Tang [1,2], **Benfeng Xu** [1,2], **Yuyue Zhao** [1,2], **Zhendong Mao** [1,2], **Yifeng Liu** [3]
Yong Liao [1,2] *, **Haiyong Xie** [1,2]
[1] University of Science and Technology of China, Anhui, China
[2] CCCD Key Lab of Ministry of Culture and Tourism, Anhui, China
[3] National Engineering Laboratory for Risk Perception and Prevention (RPP), Beijing, China
{weitang, benfeng, yyzha0}@mail.ustc.edu.cn, zdmao@ustc.edu.cn, liuyifeng3@cetc.com.cn
{yliao, hxie}@ustc.edu.cn

2022. 12. 03 • ChongQing

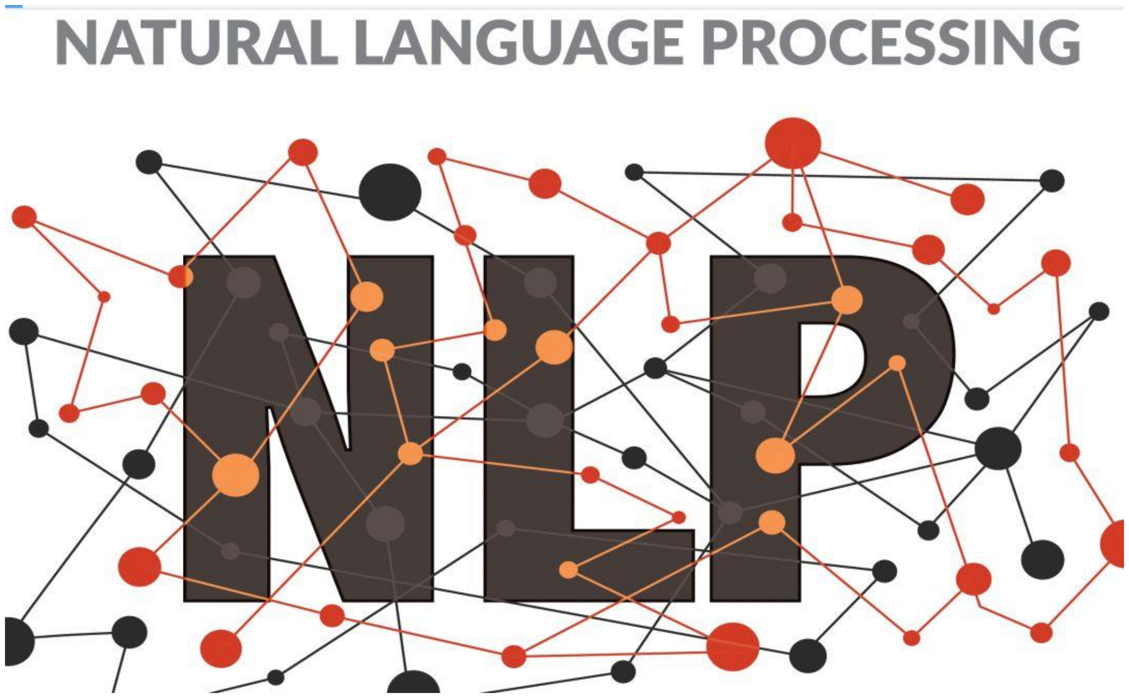
2022_EMNLP

Leibniz-Institut für Sozialwissenschaften

Forschungszentrum · Research Center

Reported by Yidan Liu

https://github.com/wtangdev/UniRel



gesis
Leibniz-Institut für Sozialwissenschaften

Introduction

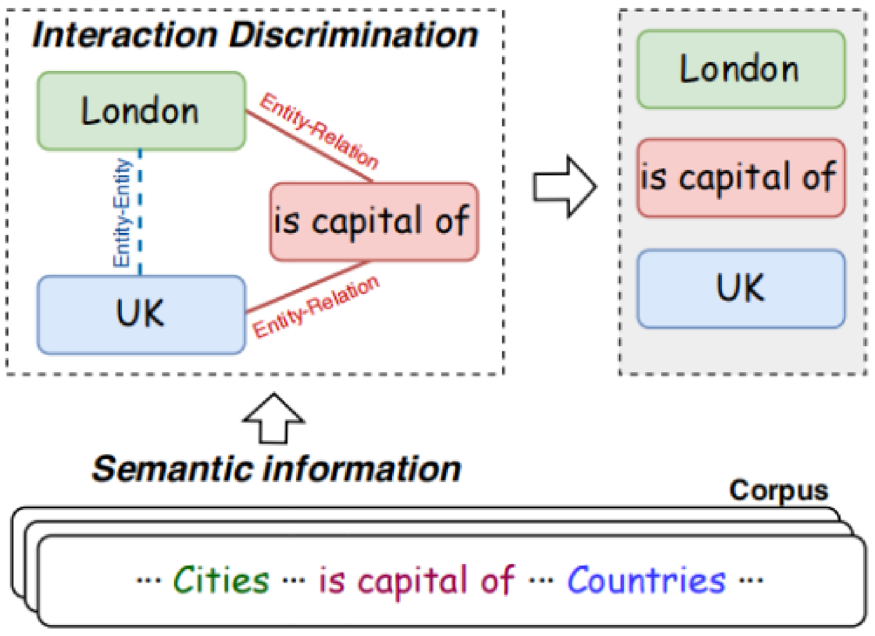


Figure 1: We leverage semantic information to unify the representation of entities and relations. Relational triples are extracted by modeling the entity-entity interaction (blue dashed line) and entity-relation interaction (red solid line) in a unified way.

- heterogeneous representations of entities and relations.
- heterogeneous modeling of entity-entity interactions and entity-relation interactions.

Method

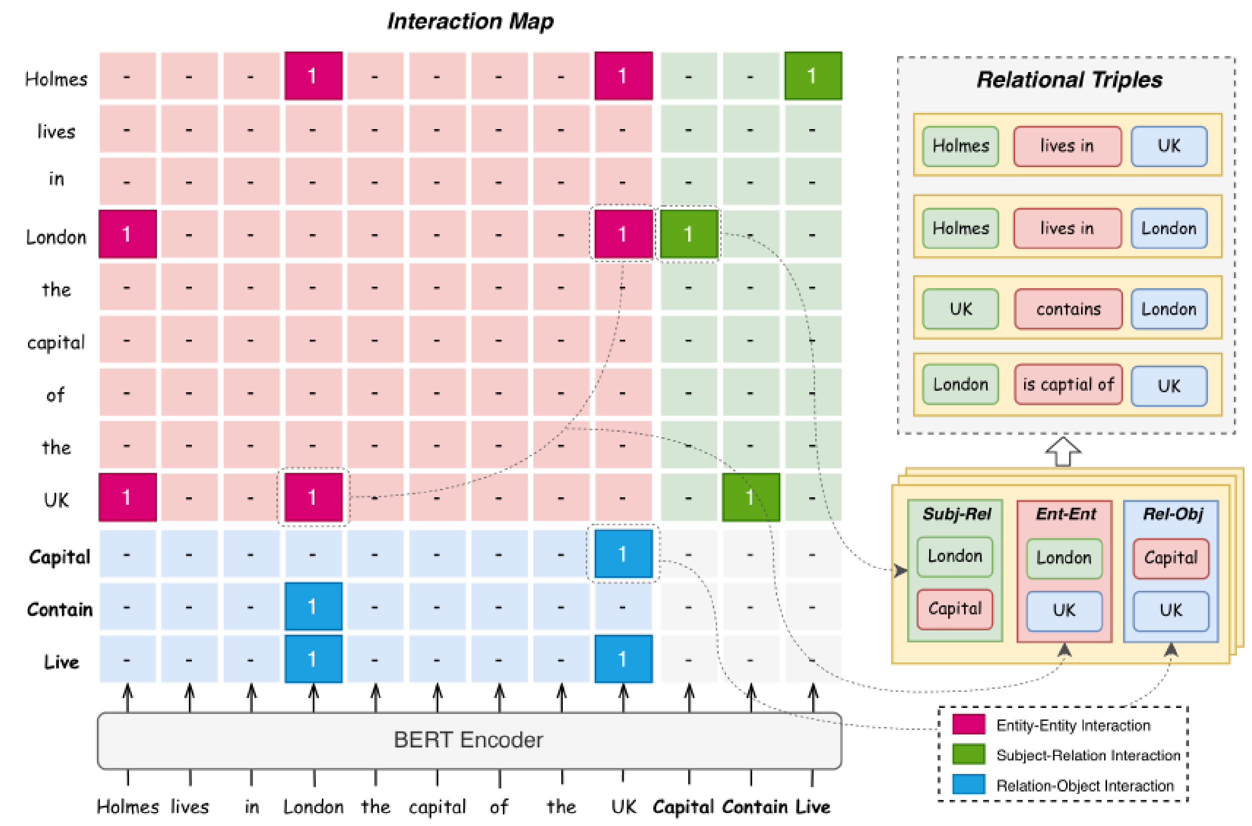
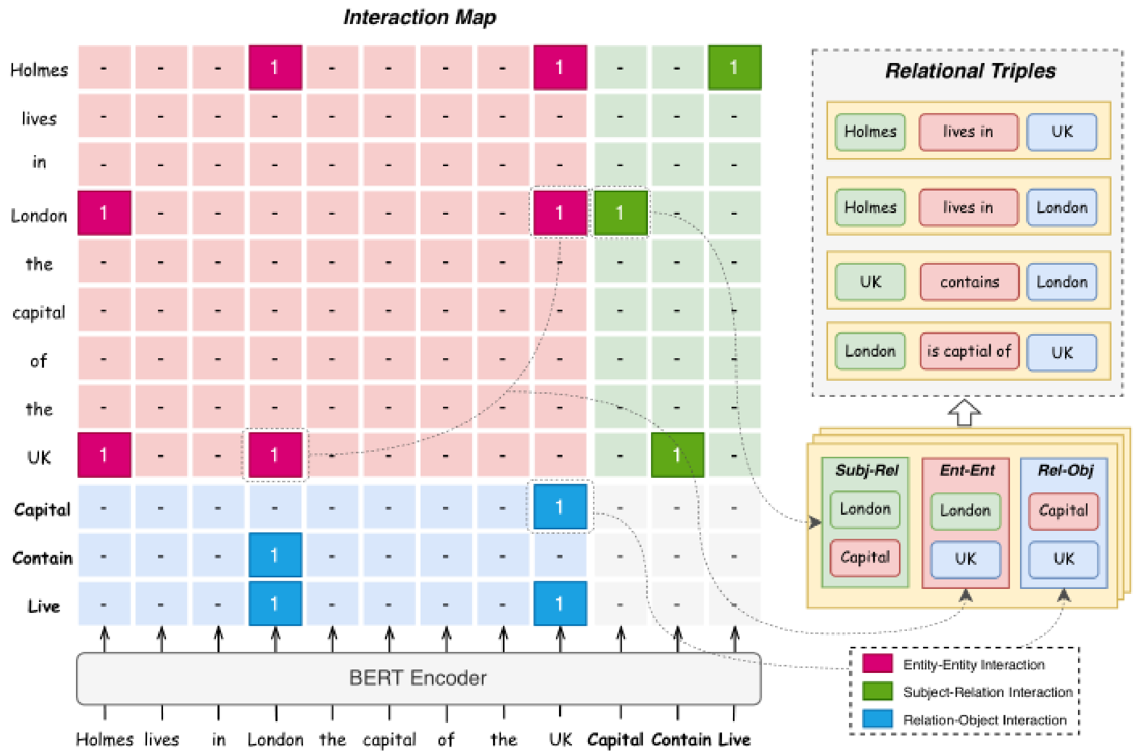


Figure 2: We input the concatenation of the input sentence and the natural language texts of relations (in bold). The Interaction Map is learned from the attention map inside the 12th layer of BERT Encoder, which consists of Entity-Entity Interaction (red rectangle) and Entity-Relation Interaction (green rectangle for subject and blue rectangle for object). Relational triples are extracted intuitively from the map.

Method



$$X = \{x_1, x_2, \cdots, x_N\} \quad T = [(s_l, r_l, o_l)]_{l=1}^{L}$$

$$R = \{R_1, R_2, \cdots, R_M\}$$

$$T = \text{Concat}(T_s, T_p) \qquad (1)$$

$$H = E[T] \qquad (2) \qquad H \in \mathbb{R}^{(N+M)\times d_h}$$

$$\text{Attention}(Q, K, V) = \text{softmax}(\frac{QK^T}{\sqrt{d_h}})V. \qquad (3)$$

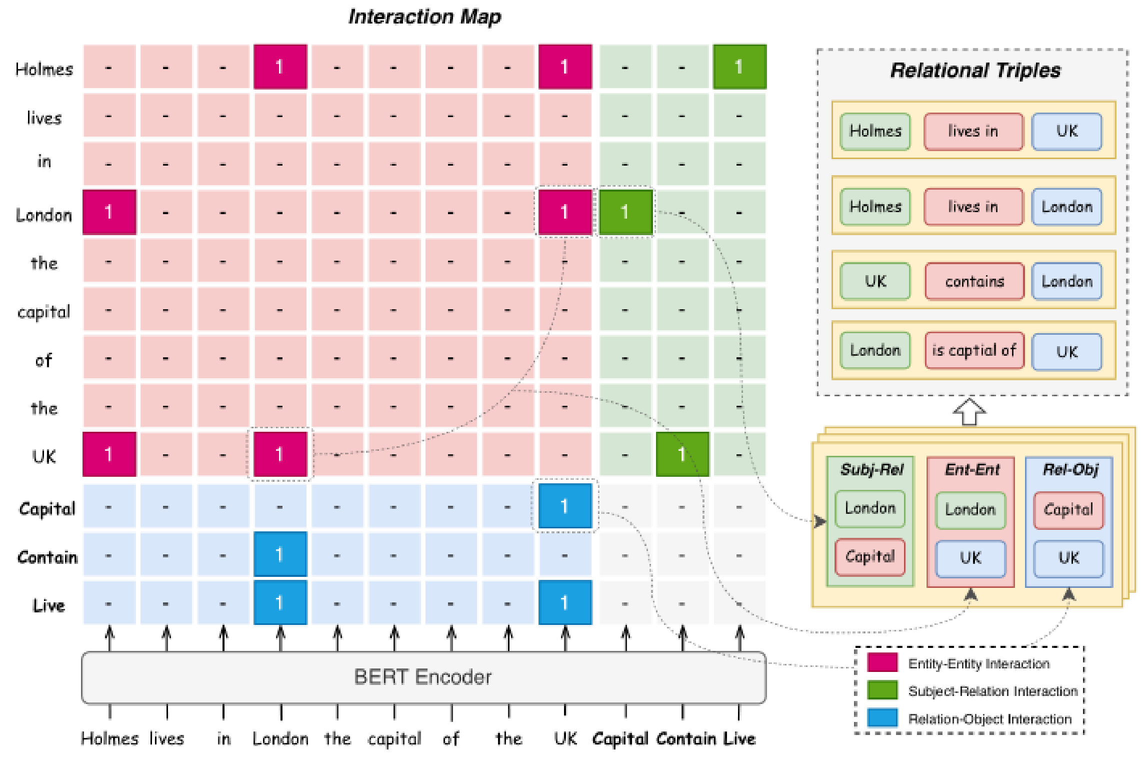
Entity-Entity Interaction

$$I_e(e_a, e_b) = \begin{cases} \mathit{True} & (e_a, r, e_b) \in T \text{ or} \\ & (e_b, r, e_a) \in T, \exists r \in R\,, \\ \mathit{False} & \text{otherwise} \end{cases} \qquad (4)$$

$I_e(e_b, e_a) = I_e(e_a, e_b)$, as entity-entity interaction is symmetrical.

Method

Entity-Relation Interaction



$$I_r(e,r) = \begin{cases} True & (e,r,o) \in T, \exists o \in E \\ False & \text{otherwise} \end{cases}, \quad (5)$$

$$I_r(r,e) = \begin{cases} True & (s,r,e) \in T, \exists s \in E \\ False & \text{otherwise} \end{cases}, \quad (6)$$

$$\mathbf{I} = \text{sigmoid}\left(\frac{1}{T}\sum_{t}^{T} \frac{Q_t K_t^T}{\sqrt{d_h}}\right), \quad (7)$$

$$\mathcal{L} = -\frac{1}{(N+M)^2} \sum_{i}^{N+M} \sum_{j}^{N+M} (\mathbf{I}^*_{i,j} \log \mathbf{I}_{i,j} + (1-\mathbf{I}^*_{i,j}) \log(1-\mathbf{I}_{i,j})), \quad (8)$$

Method

Model	NYT			WebNLG		
	Prec.	Rec.	F1	Prec.	Rec.	F1
NovelTagging (Zheng et al., 2017)	62.4	31.7	42.0	52.5	19.3	28.3
CopyRE (Zeng et al., 2018)	61.0	56.6	58.7	37.7	36.4	37.1
GraphRel (Fu et al., 2019)	63.9	60.0	61.9	44.7	41.1	42.9
OrderCopyRE (Zeng et al., 2019)	77.9	67.2	72.1	63.3	59.9	61.6
CasRel_{BERT} (Wei et al., 2020)	89.7	89.5	89.6	93.4	90.1	91.8
TPlinker_{BERT} (Wang et al., 2020)	91.3	92.5	91.9	91.7	92.0	91.9
PRGC_{BERT} (Zheng et al., 2021a)	93.3	91.9	92.6	94.0	92.1	93.0
R-BPtrNet_{BERT} (Chen et al., 2021)	92.7	92.5	92.6	93.7	92.8	93.3
PFN (Yan et al., 2021)	-	-	92.4	-	-	93.6
TDEER_{BERT} (Li et al., 2021)	93.0	92.1	92.5	93.8	92.4	93.1
GRTE_{BERT} (Ren et al., 2021)	92.9	93.1	93.0	93.7	94.2	93.9
EmRel (Xu et al., 2022)	91.7	92.5	92.1	92.7	93.0	92.9
OneRel_{BERT} (Shang et al., 2022)	92.8	92.9	92.8	94.1	94.4	94.3
UniRel	**93.5**	**94.0**	**93.7**	**94.8**	**94.6**	**94.7**
UniRel_{unused}	93.1	93.2	93.1	75.1	68.6	71.7
$\text{UniRel}_{separate}$	92.6	93.7	93.1	93.5	94.4	93.9

Table 1: Main results. The highest scores are in bold.

Dataset	Train	Test	Overlapping Pattern			
			Normal	SEO	EPO	SOO
NYT	56195	5000	3266	1297	978	45
WebNLG	5019	703	245	457	26	84

Table 2: Statistics of evaluation datasets. Overlapping patterns are counted on test set.

Experiment

Model	Normal	SEO	EPO	SOO	L= 1	L= 2	L= 3	L= 4	L≥ 5
CasRel	87.3	91.4	92.0	77.0	88.2	90.3	91.9	94.2	83.7
TPlinker	90.1	93.4	94.0	90.1	90.0	92.8	93.1	96.1	90.0
PRGC	91.0	94.0	94.5	81.8	91.1	93.0	93.5	95.5	93.0
R-BPtrNet	90.4	94.4	**95.2**	-	89.5	93.1	93.5	**96.7**	91.3
PFN	90.2	**95.3**	94.1	-	90.5	92.9	93.7	96.3	92.6
TDEER	90.8	94.1	94.5	-	90.8	92.8	94.1	95.9	92.8
GRTE	91.1	94.4	95.0	-	90.8	93.7	94.4	96.2	93.4
OneRel	90.6	95.1	94.8	**90.8**	90.5	93.4	93.9	96.5	**94.2**
UniRel	**91.6**$_{\pm 0.3}$	**95.3**$_{\pm 0.2}$	**95.2**$_{\pm 0.1}$	89.8$_{\pm 3.6}$	**91.5**$_{\pm 0.3}$	**94.3**$_{\pm 0.2}$	**94.5**$_{\pm 0.3}$	96.6$_{\pm 0.2}$	**94.2**$_{\pm 0.8}$

Table 3: F1-score on sentences with different overlapping patterns and different triple numbers. L is the number of triples in one sentence. All the compared models are implemented with BERT. We report the average results of UniRel of the five runs with different random seeds. The highest scores are in bold.

Experiment

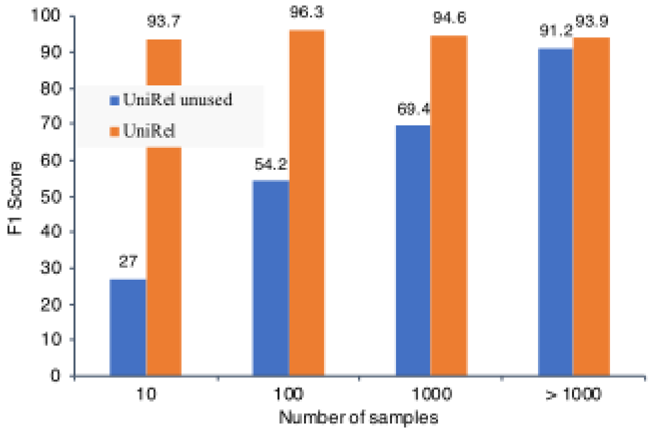


Figure 3: F1-score on relations with different orders of magnitude samples in training set for UniRel (in Orange) and UniRel$_{unused}$ (in Blue). 10/100/1000 means relations with less than or equal to 10/100/1000 samples.

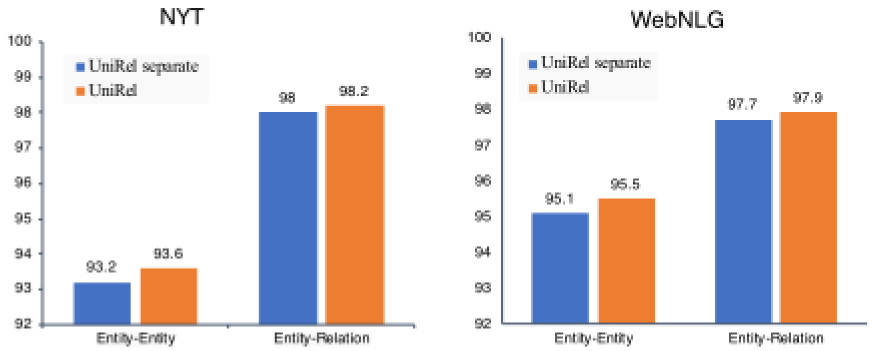


Figure 4: F1-score on Entity-Entity Interaction and Entity-Relation Interaction for UniRel (in Orange) and UniRel$_{separate}$ (in Blue).

Experiment

Model	Training Time		Inference Time	
	NYT	WebNLG	NYT	WebNLG
CasRel	1142	105	35	37
TPLinker	2951	810	48	57
PRGC	3632	498	13	13
OneRel	2998	186	20	21
UniRel	967	119	12	14

Table 4: Computational Efficiency. Training time represents the time (second) needed to train one epoch. Inference time represents the average time (millisecond) to predict one sample.

Figure 5: Visualization of Interaction Map with input sentence sampled from NYT. Relations are in bold.

Thank you!

Leibniz-Institut für Sozialwissenschaften