

SISER: Semantic-Infused Selective Graph Reasoning for Fact Verification

Eunhwan Park* ♠, Jong-Hyeon Lee♠, Donghyeon Jeon, Seonhoon Kim,
Inho Kang, Seung-Hoon Na

judepark@jbnu.ac.kr*



NAVER



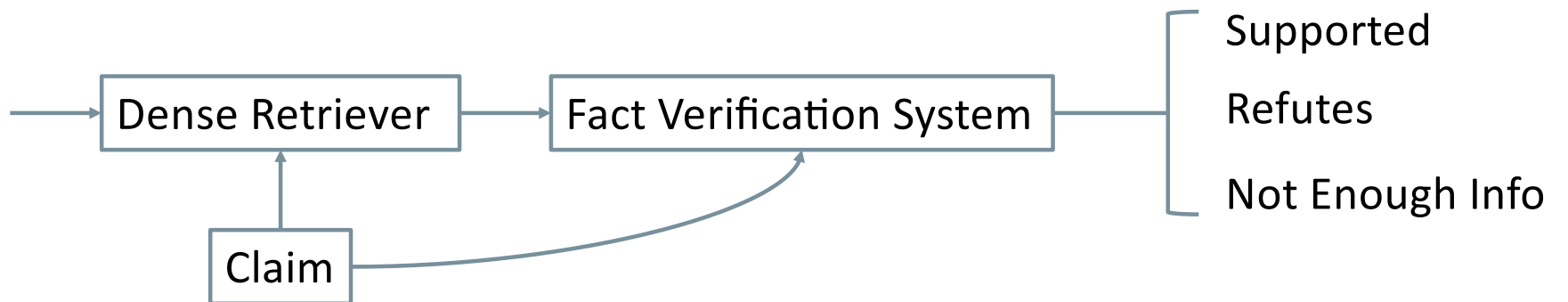
Jeonbuk National University*

NAVER Corporation,

NCSoft

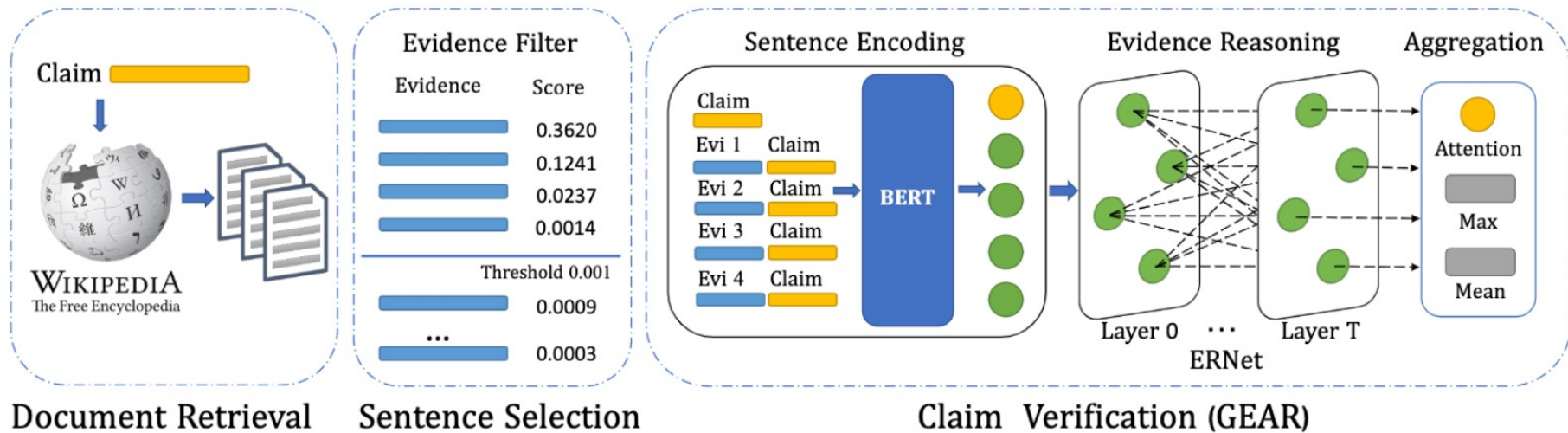
Fact Verification

Fact Verification

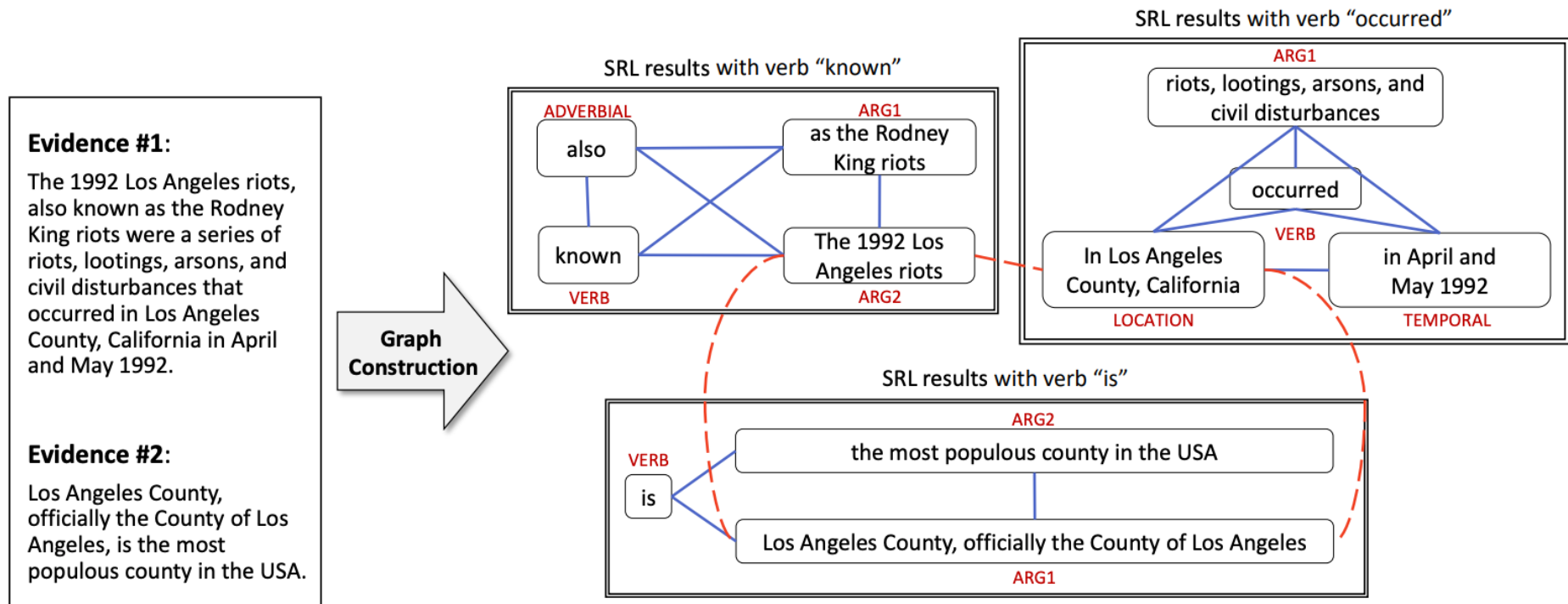


Prior work

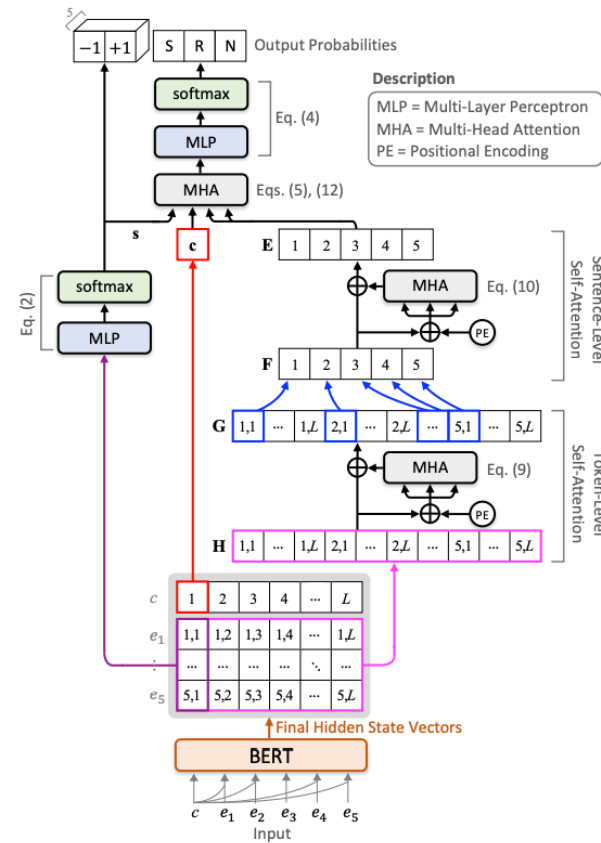
Prior work



Prior work



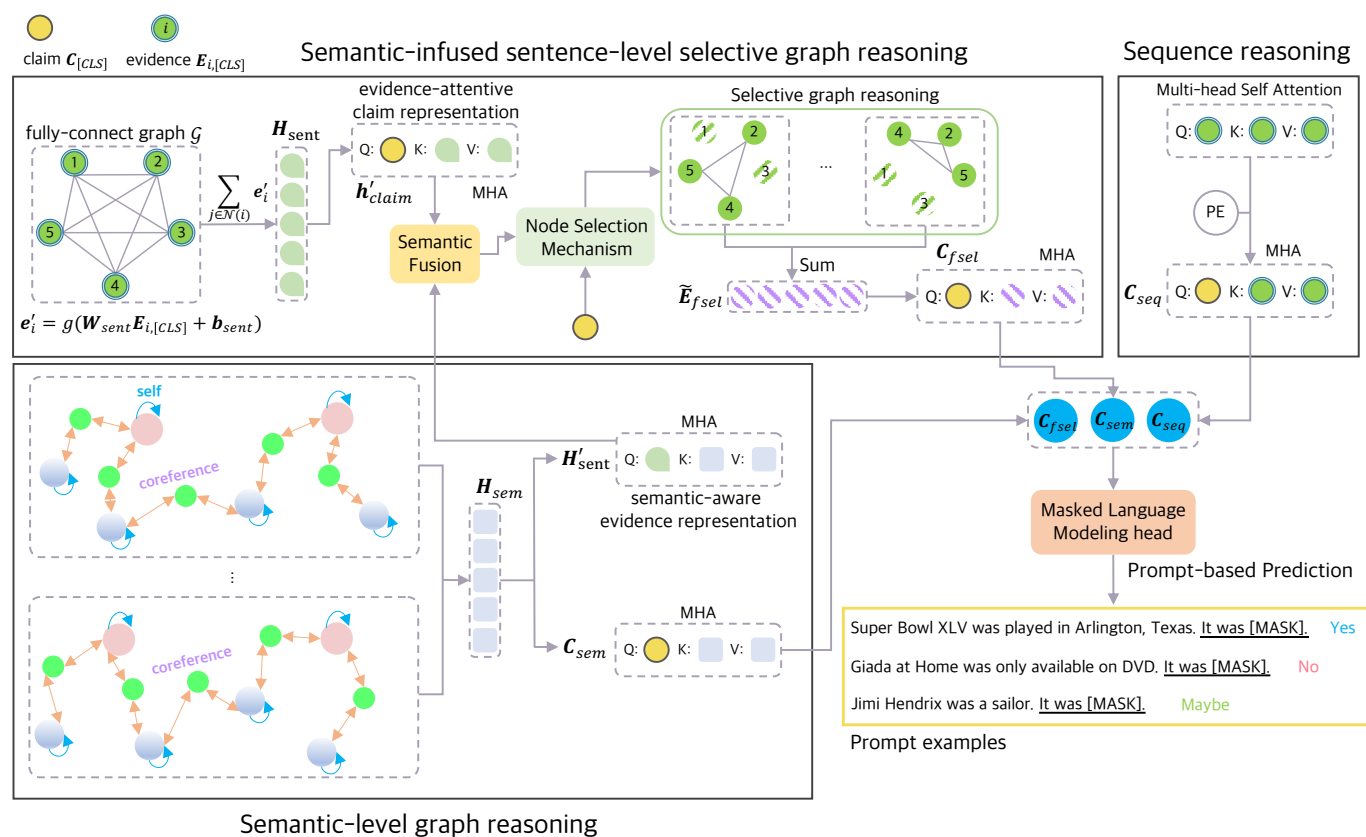
Prior work



A Multi-Level Attention Model for Evidence-Based Fact Checking, Kruengkrai et al, ACL 2021 Findings.

Our approach: SISER

SISER: Semantic-Infused Selective Graph Reasoning for Fact Verification



Main results

Main results

Model	Dev		Test	
	LA	F.S	LA	F.S
UNC NLP	69.72	66.49	68.21	64.21
GEAR (BERT _{base})	74.84	70.69	71.60	67.10
DREAM (XLNet _{large})	79.16	-	76.85	70.60
KGAT (BERT _{large})	77.91	75.86	73.61	70.24
⊥ (RoBERTa _{large})	78.29	76.11	74.07	70.38
LOREN (BERT _{large})	78.44	76.21	74.43	70.71
⊥ (RoBERTa _{large})	<u>81.14</u>	<u>78.83</u>	76.42	72.93
MLA (RoBERTa _{large})	79.31	75.96	<u>77.05</u>	<u>73.72</u>
Ours (RoBERTa _{large})	83.13	79.87	77.50	73.90

Ablation studies

The Effect of Using Semantic-Infused Sentence-level Selective Graph Reasoning and Sequence Reasoning

Model	Dev		Test	
	LA	F.S	LA	F.S
MLA	79.31	75.96	77.05	73.72
SISER★	83.13	79.85	76.82	73.18
SISERO ($\tau = 0.49$)	82.62	79.40	77.18	73.48
SISER ($\tau = 0.49$)	83.13	79.87	77.50	73.90

★ and ○ denote the run without the **semantic-infused sentence-level selective graph reasoning** and the **sequence reasoning**.

The Effect of Choosing Evidence Retrieval

Model	Dev		Test	
	LA	F.S	LA	F.S
MLA	79.31	75.96	77.05	73.72
SISER★ ($\tau = 0.49$)	79.88	75.04	77.96	73.06
SISER ($\tau = 0.49$)	83.13	79.87	77.50	73.90

★ denotes the run based on the evidence retrieval of MLA.

Evaluation of Node Selection Mechanism

Model	Dev		Test	
	LA	F.S	LA	F.S
$\tau = 0.0^\bullet$	83.07	79.84	77.07	73.65
$\tau = 0.35$	83.00	79.74	77.11	73.70
$\tau = 0.40$	83.05	79.84	77.00	73.63
$\tau = 0.45$	82.98	79.69	76.86	73.66
$\tau = 0.49$	83.13	79.87	77.50	73.90
$\tau = 0.60$	83.04	79.80	77.30	73.68

Ablation study of the node selection mechanism for varying values of the node masking rate τ . \bullet denotes the fully-connected setting.

Prompt-based Learning v.s. Conventional Fine-tuning

Model	Dev		Test	
	LA	F.S	LA	F.S
SISER★	83.05	79.77	76.82	73.18
SISER	83.13	79.87	77.50	73.90

Ablation study for the **prompt-based learning vs. the conventional fine-tuning** on the FEVER development set. ★ denotes the conventional fine-tuning.

Thanks!

- To my advisor, Seung-Hoon Na!
- To NAVER corporation!
- To the anonymous reviewers for their valuable feedback!