Haili LI, Zhiliang TIAN, Xiaodong WANG, Yunyan ZHOU, Shilong PAN, Jie ZHOU, Qiubo XU, Dongsheng LI, 2025. Handling polysemous triggers and arguments in event extraction: an adaptive semantics learning strategy with reward–penalty mechanism. *Frontiers of Information Technology & Electronic Engineering*, 26(4):534-555. https://doi.org/10.1631/FITEE.2400220

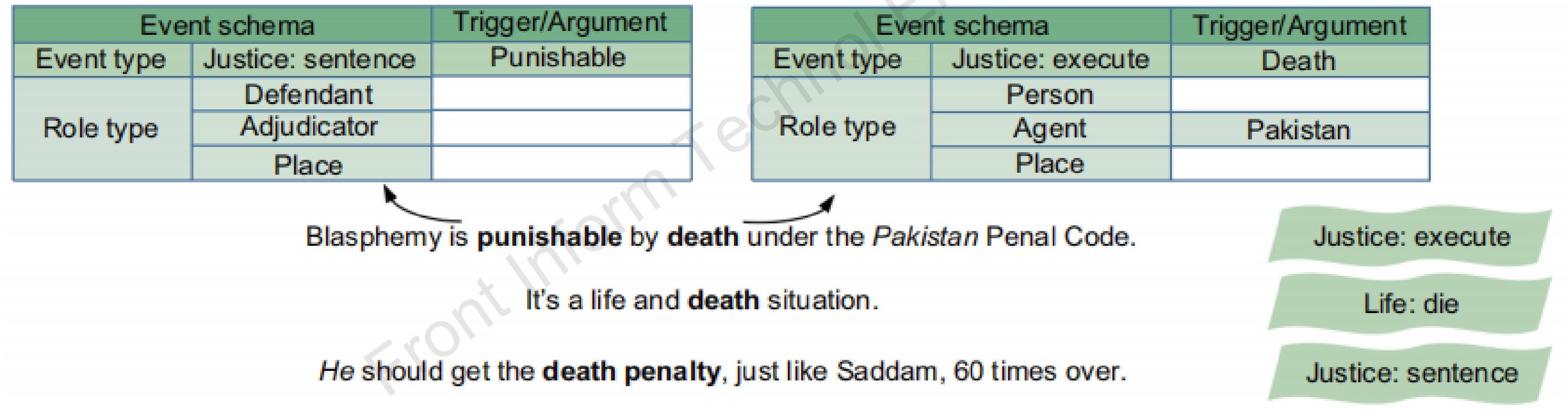
Handling polysemous triggers and arguments in event extraction: an adaptive semantics learning strategy with reward–penalty mechanism

Key words: Event extraction; Polysemous triggers; Polysemous arguments; Semantic imbalance; Reward–penalty mechanism

Corresponding authors: Zhiliang TIAN, Dongsheng LI
E-mail: tianzhiliang@nudt.edu.cn; lidongsheng@nudt.edu.cn
ORCID: https://orcid.org/0000-0002-8906-5198
https://orcid.org/0000-0001-9743-2034

Motivation

- The polysemy of triggers and arguments poses significant challenges for event extraction. Various methods have been proposed to mitigate ambiguity by enhancing the semantics.



Event schema		Trigger/Argument
Event type	Justice: sentence	Punishable
Role type	Defendant	
	Adjudicator	
	Place	

Event schema		Trigger/Argument
Event type	Justice: execute	Death
Role type	Person	
	Agent	Pakistan
	Place	

Blasphemy is **punishable** by **death** under the *Pakistan* Penal Code.

It's a life and **death** situation.

He should get the **death penalty**, just like Saddam, 60 times over.

Fig. 1 Event extraction (EE) results of the sentence "Blasphemy is punishable by death under the Pakistan Penal Code." and examples of the semantics for the triggers "death" and "punishable." Words in bold are triggers, while those italicized are arguments. Event "justice: sentence" is triggered by the trigger "punishable" without arguments playing any roles. Trigger "death" triggers the "justice: execute" event, where "Pakistan" plays the "agent" role in the event schema of "justice: execute"

Motivation

- However, the semantic distribution of polysemous triggers and arguments is imbalanced or biased. The imbalanced semantic distribution results in the misidentification of relevant and irrelevant semantics and makes it difficult to identify the boundary of triggers.

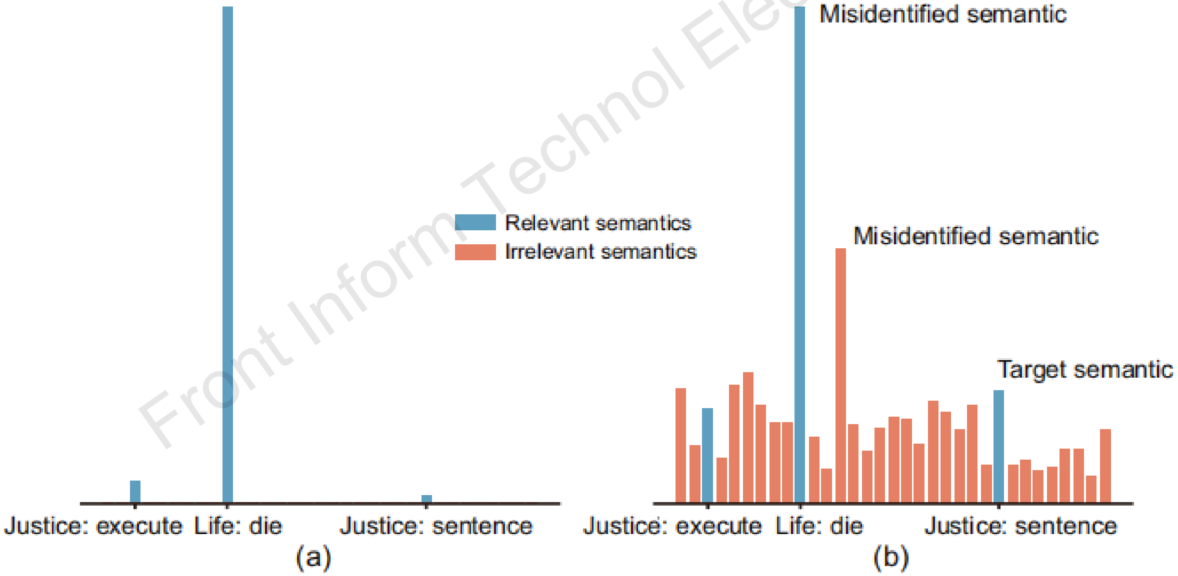


Fig. 2 Original semantic distribution of “death” (a) vs. semantic distribution of “death” after balancing (b). (a) shows the original imbalanced semantic distribution in which the number of samples with the semantics “life: die” is significantly higher than the numbers of samples with the semantics “justice: execute” and “justice: sentence.” (b) shows the semantic distribution after balancing, in which the probabilities of both relevant and irrelevant semantics being detected are higher than that of the target semantic, leading to two false positives (FPs) and one false negative (FN)

Main idea

- To balance the biased distribution of semantics, we leverage the reward–penalty mechanism to adapt the different semantics of polysemous triggers and arguments by weakening the high-frequency semantics and amplifying the low-frequency semantics.
- To reduce false positives for irrelevant semantics and nontarget relevant semantics, we propose to use the sentence's event to enhance the semantics of triggers and arguments. To ensure the accuracy of the sentence's event semantics for avoiding error propagation, the proposed sentence-level event situation awareness mechanism uses a sentence event classification task for precisely modeling the sentence's event semantics.

Method

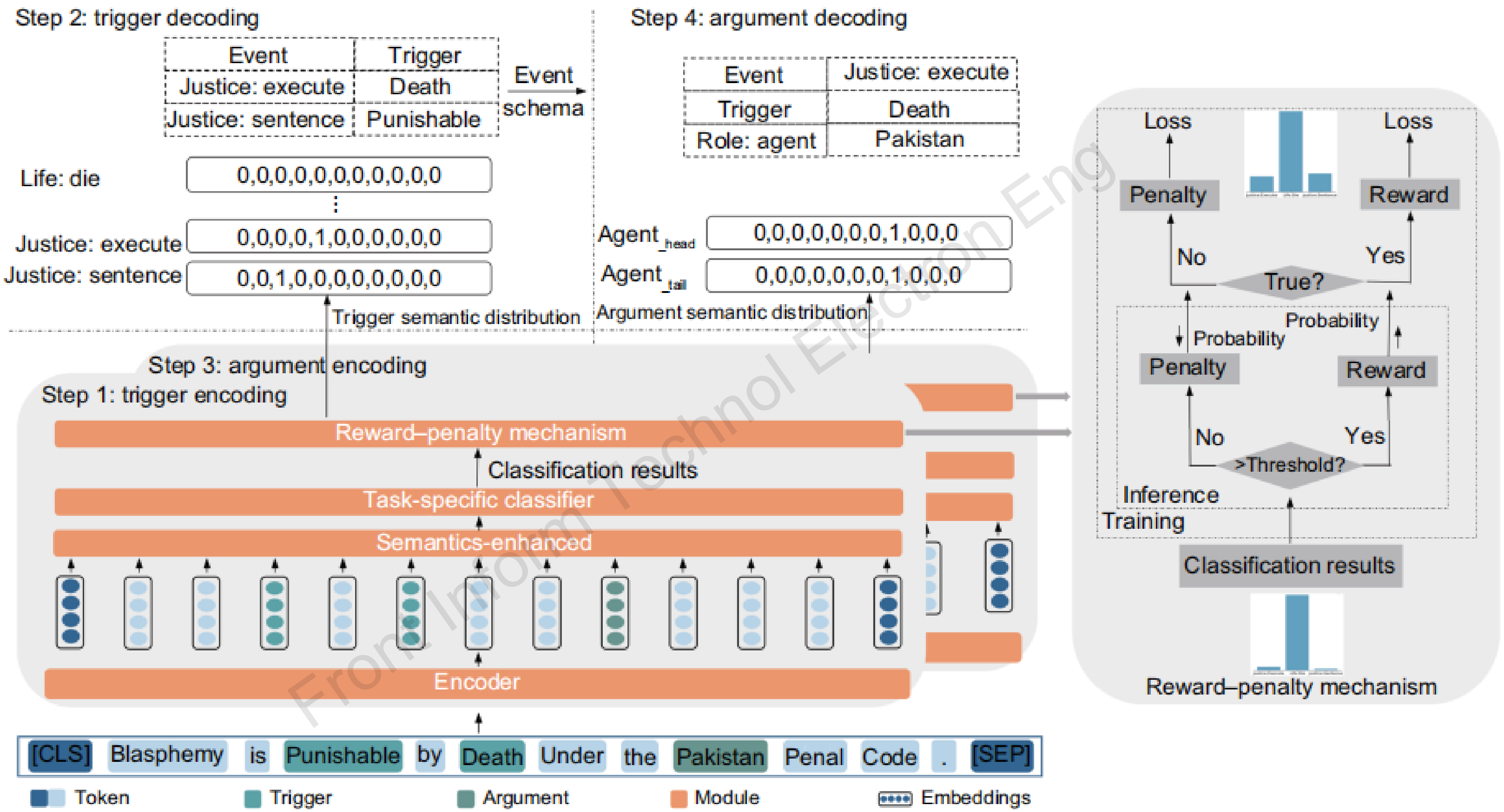
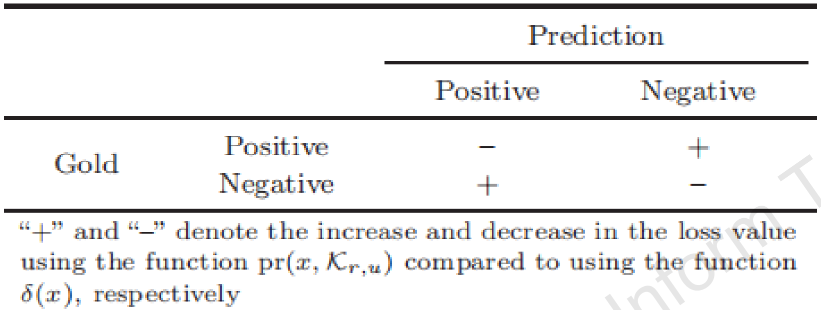


Fig. 3 The overview of our joint event extraction (EE) model. The encoder converts tokens into high-dimensional vectors. Before this, the sentence-level event situation awareness (SA) mechanism fine-tunes the encoder to guarantee the precise representation of the sentence events. The encoder then uses this representation $\boldsymbol{S}$ to enhance token semantics during encoding. Subsequently, our model calculates the probability distribution $P(x_i)$ and employs the reward–penalty mechanism to amplify the correct semantics and diminish the incorrect semantics, widening the gap between them. Finally, using $P(x_i)$, the trigger decoder and argument decoder use task-specific thresholds to identify and classify all candidates

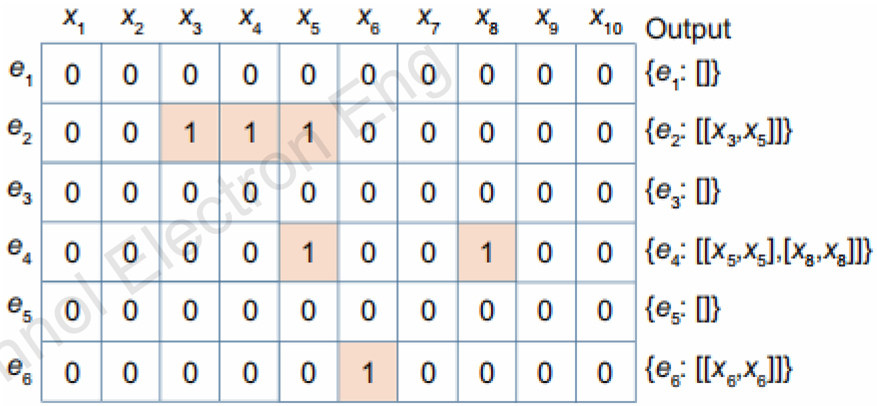
Method

Table 1 Changes in the loss value



		Prediction	
		Positive	Negative
Gold	Positive	−	+
	Negative	+	−

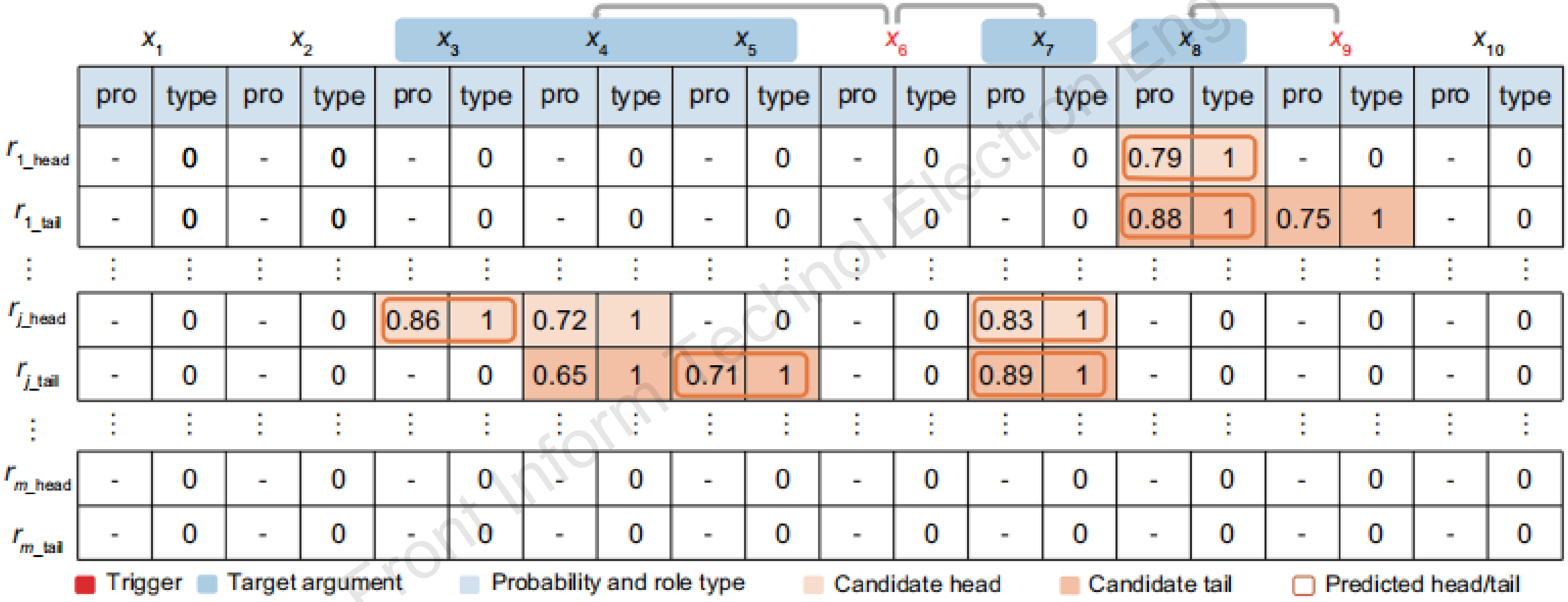
“+” and “−” denote the increase and decrease in the loss value using the function $\mathrm{pr}(x, \mathcal{K}_{r,u})$ compared to using the function $\delta(x)$, respectively



	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	Output
e_1	0	0	0	0	0	0	0	0	0	0	$\{e_1: []\}$
e_2	0	0	1	1	1	0	0	0	0	0	$\{e_2: [[x_3,x_5]]\}$
e_3	0	0	0	0	0	0	0	0	0	0	$\{e_3: []\}$
e_4	0	0	0	0	1	0	0	1	0	0	$\{e_4: [[x_5,x_5],[x_8,x_8]]\}$
e_5	0	0	0	0	0	0	0	0	0	0	$\{e_5: []\}$
e_6	0	0	0	0	0	1	0	0	0	0	$\{e_6: [[x_6,x_6]]\}$

Fig. 4 The process of trigger decoding with the semantic decoders. 1 in (i,j) indicates that the predicted type for token x_j is e_i, and 0 in (i,j) indicates that token x_j does not trigger type e_i. Each row in this figure is a semantic decoder that identifies the boundaries of all candidates for the semantic. Each column is the predicted type set of the token

Method



	x_1		x_2		x_3		x_4		x_5		x_6		x_7		x_8		x_9		x_{10}	
	pro	type	pro	type	pro	type	pro	type	pro	type	pro	type	pro	type	pro	type	pro	type	pro	type
r_{1_head}	-	0	-	0	-	0	-	0	-	0	-	0	-	0	0.79	1	-	0	-	0
r_{1_tail}	-	0	-	0	-	0	-	0	-	0	-	0	-	0	0.88	1	0.75	1	-	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
r_{j_head}	-	0	-	0	0.86	1	0.72	1	-	0	-	0	0.83	1	-	0	-	0	-	0
r_{j_tail}	-	0	-	0	-	0	0.65	1	0.71	1	-	0	0.89	1	-	0	-	0	-	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
r_{m_head}	-	0	-	0	-	0	-	0	-	0	-	0	-	0	-	0	-	0	-	0
r_{m_tail}	-	0	-	0	-	0	-	0	-	0	-	0	-	0	-	0	-	0	-	0

🟥 Trigger 🟦 Target argument ▢ Probability and role type ▨ Candidate head ▨ Candidate tail ☐ Predicted head/tail

Fig. 5 The process of role decoding, where "pro" is the abbreviation of the probability indicating the probability of the token on the type, and the "type" value ($\in \{0,1\}$) means whether the type is the predicted one of the token. The results of role decoding are $\{e_{x_6}:\{r_j:[[x_3,x_5],[x_7,x_7]]\}, e_{x_9}:\{r_1:[x_8,x_8]\}\}$, where e_{x_i} is the event type for token x_i

Results

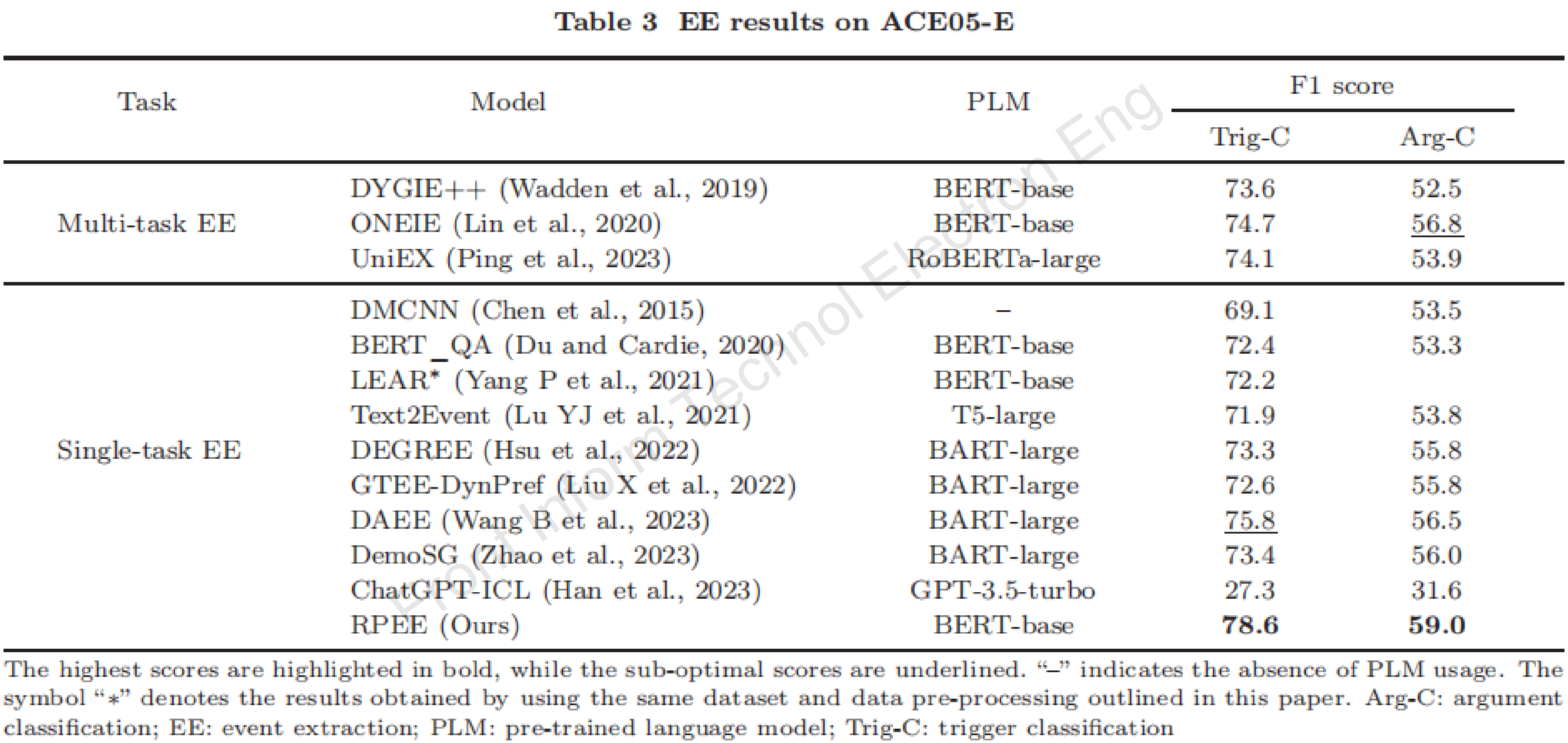


Table 3 EE results on ACE05-E

Task	Model	PLM	F1 score	
			Trig-C	Arg-C
Multi-task EE	DYGIE++ (Wadden et al., 2019)	BERT-base	73.6	52.5
	ONEIE (Lin et al., 2020)	BERT-base	74.7	56.8
	UniEX (Ping et al., 2023)	RoBERTa-large	74.1	53.9
Single-task EE	DMCNN (Chen et al., 2015)	–	69.1	53.5
	BERT_QA (Du and Cardie, 2020)	BERT-base	72.4	53.3
	LEAR* (Yang P et al., 2021)	BERT-base	72.2	
	Text2Event (Lu YJ et al., 2021)	T5-large	71.9	53.8
	DEGREE (Hsu et al., 2022)	BART-large	73.3	55.8
	GTEE-DynPref (Liu X et al., 2022)	BART-large	72.6	55.8
	DAEE (Wang B et al., 2023)	BART-large	75.8	56.5
	DemoSG (Zhao et al., 2023)	BART-large	73.4	56.0
	ChatGPT-ICL (Han et al., 2023)	GPT-3.5-turbo	27.3	31.6
	RPEE (Ours)	BERT-base	**78.6**	**59.0**

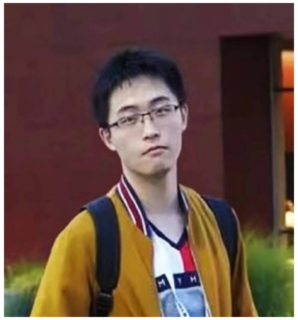
The highest scores are highlighted in bold, while the sub-optimal scores are underlined. "–" indicates the absence of PLM usage. The symbol "*" denotes the results obtained by using the same dataset and data pre-processing outlined in this paper. Arg-C: argument classification; EE: event extraction; PLM: pre-trained language model; Trig-C: trigger classification

Conclusions

In this paper, we propose an adaptive semantics learning strategy to mitigate the bias in the semantics distribution of polysemous triggers and arguments.

- We design a reward–penalty mechanism to enlarge the gap between the relevant and irrelevant semantics and diminish the gap between relevant semantics by rewarding the corrected classified semantics and punishing the misclassified semantics.
- The sentence-level event semantics, pre-trained by using a sentence-level event SA mechanism to ensure accuracy, is integrated into token representations to narrow the target event scope of triggers.
- The model identifies the boundaries of triggers and arguments and classifies their types using task-specific semantic decoders.

Haili LI is currently pursuing her PhD degree at National University of Defense Technology (NUDT). She obtained her BS and MS degrees from Nanjing University of Science and Technology in 2011 and 2024, respectively. Her research interests include natural language processing and network security.



Zhiliang TIAN received his PhD degree in computer science and engineering from the Hong Kong University of Science and Technology in 2022. He is currently an associate professor in the College of Computer Science and Technology at NUDT. His research interests include natural language processing and machine learning, especially text generation, conversation systems, and privacy protection.

Xiaodong WANG is currently a professor with the National Key Laboratory of Parallel and Distributed Computing, NUDT, China. He received the BS, MS, and PhD degrees from NUDT in 1996, 1998, and 2002, respectively. His current research interests focus on deep learning, federal learning, natural language processing, knowledge graph, and large language models.

Dongsheng LI received his PhD degree in computer science and technology from NUDT in 2005. He is currently a professor and doctoral supervisor in the College of Computer Science and Technology at NUDT. He was a recipient of the Chinese National Excellent Doctoral Dissertation and was supported by the National Science Fund for Excellent Young Scholars. His research interests include distributed systems, cloud computing, and big data processing.