Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

# Supplementary materials for

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. *Front Inform Technol Electron Eng*, 26(4):534-555.
https://doi.org/10.1631/FITEE.2400220

## 1 Case study

Table S1 presents the event detection (ED) results of our proposed method and other methods on a series of cases from multiple perspectives. The following case analysis further confirms that our work can effectively address the challenges introduced by the biased semantic distribution:

**Table S1 Case analysis of different types of semantic error classification for polysemous triggers**

| Types of case | Input sentence | Models | Results |
|---|---|---|---|
| Irrelevant semantics | Putin last spoke to Bush on April 5 at the US president's own initiative. | ONEIE | (Spoke, Contact:Meet) |
| | | DEGREE | (Spoke, Contact:Meet) |
| | | w/o Reward–Penalty | – |
| | | RPEE | – |
| Low-frequency semantics | This upcoming visit to Russia will be my first trip aboard since I became president of China. | ONEIE | (Visit,Movement:Transport) |
| | | DEGREE | – |
| | | w/o Reward–Penalty | (Visit,Movement:Transport)<br>(became,Personnel:Start-Position) |
| | | RPEE | (Visit,Movement:Transport)<br>(Became,Personnel:Start-Position) |
| High-frequency semantics | But few at the Kremlin forum suggested that Putin's own standing among voters will be hurt by Russia's apparent diplomacy failures. | ONEIE | (Forum,Contact:Meet) |
| | | DEGREE | (Forum,Contact:Meet) |
| | | w/o Reward-Penalty | (Forum,Contact:Meet) |
| | | RPEE | – |
| High-frequency semantics and low-frequency semantics | Faisalabad's Catholic Bishop John Joseph, who had been campaigning against the law, shot himself in the head outside a court in Sahiwal district when the judge convicted Christian Ayub Masih under the law in 1998. | ONEIE | (Shot,Conflict:Attack)<br>(Convicted,Justice:Convict) |
| | | DEGREE | (Shot,Conflict:Attack)<br>(Convicted,Justice:Convict) |
| | | w/o Reward-Penalty | (Shot,Conflict:Attack)<br>(Convicted,Justice:Convict) |
| | | RPEE | (Shot,Life:Die)<br>(Convicted,Justice:Convict) |

"–" indicates the absence of an event mention. Contents in the parentheses display the ED results, with the first item being the trigger and the second item being the event type. The yellow block indicates triggers of the sentence, while the green block indicates triggers or event types that are incorrectly recognized. ED: event detection.

1. FP of irrelevant semantics

Instance 1 does not state any information related to an event. In this instance, ONEIE and DEGREE erroneously identify "spoke" as a trigger and classify it as "Contact: Meet." However, after analyzing the samples of the training set and the development set, we confirm that "spoke" is not labeled as a sample of the event type "Contact: Meet," thus verifying that "Contact: Meet" is an irrelevant semantics of "spoke." In

contrast, RPEE and its variant w/o reward–penalty, which adopts the SESA mechanism, have successfully avoided the occurrence of irrelevant semantic misjudgments.

2. FN of low-frequency semantics

Table S1 indicates that ONEIE and DEGREE fail to recognize the trigger "became" in instance 2. After analyzing the training set data, we find that the number of samples of "become" as a trigger triggering event types "Personnel: Start-Position" and "Personnel: Elect" is relatively small, resulting in the model unable to adequately learn the low-frequency semantics in "became." The reward–penalty mechanism and sentence event situation awareness mechanism utilized by RPEE and its variant w/o reward–penalty facilitate the model to better learn low-frequency semantics.

3. FP of high-frequency semantics

From Table S1, it can be seen that ONEIE, DEGREE, and the w/o reward–penalty variant all identify "forum" as a trigger in instance 3, even though the instance does not mention any event-related information. Further analysis indicates that these models exhibit overfitting when learning the semantics of "forum." In contrast, our model and its variants adopt the SESA mechanism, effectively avoiding the issue of high-frequency semantic misclassification.

4. False positive (FP) of high-frequency semantics and false negative (FN) of low-frequency semantics

Although "shot" in instance 4 should be classified as "Life: Die," ONEIE, DEGREE, and the w/o reward–penalty variant misclassify it as "Conflict: Attack." Statistical analysis of training and validation samples reveals that "shot" triggers 23 instances of "Conflict: Attack" and only two instances of "Life: Die." This biased distribution causes the model to prioritize learning the high-frequency event type semantics of "Conflict: Attack" while neglecting the low-frequency event type semantics of "Life: Die." To address this issue, our RPEE model enhances the learning of the semantic "Life: Die" and suppresses the overlearning of the semantic "Conflict: Attack," achieving accurate identification and classification of the trigger "shot" while reducing model bias.

## 2 Discussions

1. Implications

First, this paper highlights the issue of biased semantic distributions of triggers and arguments due to polysemy, overlooked by most existing event extraction (EE) methods reliant on semantic enhancement. Existing approaches tend to assume balanced semantic distributions of triggers and arguments, resulting in an FN for low-frequency semantics and an FP for high-frequency semantics. The experiments in this paper, particularly on monosemous and polysemous triggers, demonstrate that polysemy hampers accurate extraction of events, and solving this problem can improve EE performance.

Second, we propose the reward–penalty mechanism to tackle the biased semantic distributions in polysemous triggers and arguments. This mechanism dynamically balances the semantic distributions of triggers and arguments based on their semantic distribution and the model's classification outcomes. Specifically, it guides the model to learn the semantic distribution of triggers and arguments in a balanced manner, considering the varying number of samples with different semantics, further constraining potential events for triggers with precise knowledge of the sentence's event. Extensive experiments illustrate that the reward–penalty mechanism and the knowledge of the sentence event mitigate the misclassification. RPEE surpasses all the baselines on ACE05-E and outperforms them in complex scenarios on ACE05-E+ and ERE-EN with multi-token triggers. It indicates that RPEE enhances EE performance and exhibits high scalability and strong robustness.

2. Limitations

To address the biased semantic distribution among polysemous triggers and arguments, we devise a semantic learning method dynamically balancing semantics utilizing the reward–penalty mechanism. Additionally, we identify and classify triggers and arguments within the semantic space of the pre-trained sentence's event. However, RPEE mitigates the problem of biased semantic distribution to some extent,

and further exploration is needed to solve this issue; e.g., large language models (LLMs) or other data-enhancement methods can generate low-frequency semantic samples without altering the types of triggers and arguments.

Moreover, RPEE utilizes task-specific thresholds to identify the boundaries of triggers and arguments during the decoding process. We determine the suitable threshold through a grid search on the validation set. However, the optimal threshold value varies with the task, and other boundary identification strategies, such as beam search and exhaustive search, can be explored. We will address these issues in our future work.