OEE-CFC: A Dataset for Open Event Extraction from Chinese Financial Commentary

Qizhi Wan, Changxuan Wan[*]**, Rong Hu, Dexi Liu**[*]**, Wenwu Xu**
Kang Xu, Meihua Zou, Tao Liu, Jie Yang, Zhenwei Xiong
School of Computer and Artificial Intelligence
Jiangxi University of Finance and Economics
Jiangxi Key Laboratory of Data and Knowledge Engineering
`wanqizhi@jxufe.edu.cn, wanchangxuan@263.net`
`hurong2014@126.com, dexi.liu@163.com`

Abstract

To meet application needs, event extraction has shifted from simple entities to unconventional entities serving as event arguments. However, current corpora with unconventional entities as event arguments are limited in event types and lack rich multi-events and shared arguments. Financial commentary not only describes the basic elements of an event but also states the background, scope, manner, condition, result, and tool used for the event, as well as the tense, intensity, and emotions of actions or state changes. Therefore, it is not suitable to develop event types that include only a few specific roles, as these cannot comprehensively capture the event's semantics. Also, there are affluent complex entities serving as event arguments, multiple events, and shared event arguments. To advance the practicality of event extraction technology, this paper first develops a general open event template from the perspective of understanding the meaning of events, aiming to comprehensively reveal useful information about events. This template includes 21 event argument roles, divided into three categories: core event roles, situational event roles, and adverbial roles. Then, based on the constructed event template, Chinese financial commentaries are collected and manually annotated to create a corpus OEE-CFC supporting open event extraction. This corpus includes 17,469 events, 44,221 arguments, 3,644 complex arguments, and 5,898 shared arguments. Finally, based on the characteristics of OEE-CFC, we design four types of prompts, and two models for event argument extraction are developed, with experiments conducted on the prompts.

1 Introduction

Recently, with the impressive advancement of event extraction task (Wan et al., 2023a, 2024a), event extraction corpora have attracted considerable attention (Wan et al., 2022, 2023b, 2024b).

Some scholars have annotated and released several corpora, including Title2Event (Deng et al., 2022) for financial news headlines, document-level corpora (e.g., ChFinAnn (Zheng et al., 2019), RAMS (Ebner et al., 2020), WikiEvent (Li et al., 2021), DocEE (Tong et al., 2022), etc.), supporting multiple language MEE (Veyseh et al., 2022), CFinEE (Wu et al., 2023) based on stock data, etc. Also, there are event detection and argument extraction corpora (e.g., MAVEN (Wang et al., 2020), GENEVA (Parekh et al., 2023), etc.).

Compared to the common event extraction corpus ACE2005, the aforementioned corpora not only expand the scope but also feature specific task characteristics or domain-specific traits. Title2Event focuses on open events in the form of triplet (subject, predicate, object). ChFinAnn is used for document-level multi-event extraction from Chinese financial announcements, while RAMS and WikiEvent aim to support event extraction where unconventional entities serve as arguments. CFinEE focuses on Chinese stock headline data, adding content such as event tense, subjects, and negative phrase expressions. Nevertheless, there are still some shortcomings in the existing event extraction corpora.

Firstly, the event ontology is not sufficiently rich. In addition to the roles of common triplet (subject, predicate, object), time, place, and tools, the background, scope, manner, conditions, and results that trigger actions or state changes, as well as the tense, intensity, and emotions of the actions or state changes, all reveal event-related information that people are interested in when events occur. This information plays an important role in understanding the event's semantics. Figure 1 shows part of a financial commentary article. From Figure 1, it is clear that for sentence S_6, the occurrence of the "低于" event is established under a certain background and measurement standard; lacking this information will directly affect the original semantic meaning of the sentence. Similarly, in

---

[*] Corresponding Author.

...

S_1: 上世纪80年代末和90年代初，工业消费品价格在居民消费价格指数中的权重为40%，目前这一比重已经降到了30%左右；四是构成居民消费价格指数的服务价格多数仍由政府管制，受生产资料价格上涨的影响相对较小。
In the late 1980s and early 1990s, the weight of industrial consumer goods prices in the Consumer Price Index (CPI) was 40%, but this proportion has now dropped to around 30%. Additionally, most of the service prices that make up the CPI are still regulated by the government, making them relatively less affected by the rise in the prices of production materials.

S_2: 当前居民消费价格指数上升主要是国内市场粮食价格大幅度上扬引起的。
The current rise in the CPI is mainly due to a significant increase in domestic grain prices.

S_3: 今年前三个月的粮食价格分别比去年同期上涨了14.5%、17.5%、30%，其中3月份是1995年7月以来的最高升幅。
In the first three months of this year, grain prices rose by 14.5%, 17.5%, and 30% compared to the same period last year, with March seeing the highest increase since July 1995.

S_4: 粮食价格上涨又带动了肉、禽、蛋、水产品等其他食品价格的上涨。
The increase in grain prices has also driven up the prices of other foods such as meat, poultry, eggs, and aquatic products.

S_5: 从CPI构成要素上看，2004年第一季度食品价格上升7.1%，拉动CPI上升2.4%，对居民消费价格上升的贡献率达到了85.7%。
Looking at the components of the CPI, food prices in the first quarter of 2004 rose by 7.1%, pushing the CPI up by 2.4%, with a contribution rate of 85.7% to the increase in the CPI.

S_6: 若扣除能源和食品的价格，以西方发达国家通常采用的核心通货膨胀率衡量，今年第一季度我国的核心通货膨胀率低于1%，表明通货紧缩的威胁在一定程度上依然存在，需要继续大力刺激国内消费需求的增长。
Excluding the prices of energy and food, and measuring by the core inflation rate commonly used in developed Western countries, China's core inflation rate in the first quarter of this year was below 1%. This indicates that the threat of deflation still exists to some extent, necessitating continued strong stimulation of domestic consumption demand.

S_7: 去年第四季度以来国内粮食价格上涨主要不是粮食供应不足引起的，而是我国粮食连续几年减产和国际市场粮食供应趋紧影响了市场预期。
Since the fourth quarter of last year, the rise in domestic grain prices has mainly been due to several years of reduced grain production in China and tighter grain supply in the international market, which have affected market expectations, rather than a shortage of grain supply.

S_8: 去年央行把准备金率提高了1个百分点。
Last year, the central bank raised the reserve requirement ratio by 1 percentage point.

...

Figure 1: Example of Chinese financial commentary. Red tokens refer to triggers, and the tokens of blue, green, purple, orange, pink, light blue, pale blue, and light purple denote subject, object, manner, scope, background, time, adverb, and result, respectively. The underscore represents referring terms (annotate the original content during annotation), and the italicized texts indicate prepositional objects. Detailed explanations of each role are provided in Section 2.

sentence S_3, the "上涨了" event loses its semantic clarity without the component "分别比去年同期" which serves as the manner role. Furthermore, disregarding the adverb "不是" modifying the event "引起" in sentence S_7 will misinterpret the event's meaning, and the absence of "准备金率" or "1个百分点" in sentence S_8 is also inappropriate.

Secondly, the argument components are relatively simple. Most corpora only consider entities as event arguments, but in practice, many event components in textual data (e.g., financial commentaries) are not simply tokens or entities. These components can be tokens (phrases) or entities with qualifying descriptions, or even a continuous sequence of characters. For example, a text segment containing event information can serve as an argument for another event (e.g., "肉、禽、蛋、水产品等其他食品价格" in S_4), or it can be a phrase or event containing an action, such as the subject of the "拉动" event in S_5. This paper collectively refers to these non-simple tokens (phrases) or entities as event arguments of **unconventional entities**, including an event (referred to as a complex entity in this paper) serving as the argument of another event. This type of corpus can effectively support the training of models for extracting unconventional entities, addressing real-world needs in practical applications.

Thirdly, the event types are fixed. The corpora with specified event types lack flexibility and cannot cover diverse event types. Existing open event corpora do not distinguish meaningful events based on needs, resulting in models being unable to precisely extract events of interest. As shown in Figure 1, the verbs "受" in S_1, "是" in S_2, and "需要" in S_6, although syntactically serving as core tokens, do not reveal meaningful information and are thus not considered as events.

Lastly, the format of text content is generally standardized and fixed. Existing Chinese event extraction corpora mostly use announcements as data sources, which have a standardized format that does not reflect the actual usage scenarios in practical applications. This limits the transferability of models. Financial commentary, due to the need for analysis, reasoning, and evaluation of facts, has strong logic, making the events strongly interconnected (e.g., the "引起" event in S_2 connects the "上涨" and "上扬" events, and the first "上升" event in S_5 leads to the "拉动" event). The most notable phenomenon is the nesting of event components, including complete nesting (shared components), partial nesting (an event component is part of another event's component, or parts of an event component are part of another event's components), and an event as a whole serving as a component of another event. Additionally, Chinese is a paratactic language with certain linguistic features, such as omissions, fronting of objects, and special sentence structures (e.g., sentences with "把", "被",

4447

and "将"), making the components complex. For instance, the "把" structure in S_8 of Figure 1 separates "准备金率" and "1个百分点", making the sentence components difficult to understand.

To better align with practical scenarios and effectively support research on event extraction methods, this paper aims to reveal the semantic meaning of event sentences. Inspired by the idea of semantic role labeling, we first designed a comprehensive event template (ontology). Then, based on this event template, we annotated an <u>O</u>pen <u>E</u>vent <u>E</u>xtraction corpus tailored for <u>C</u>hinese <u>F</u>inancial <u>C</u>ommentary, OEE-CFC. Finally, four prompts were designed and two models were constructed on OEE-CFC for experiments to validate the significance of the corpus.

The main contributions of this paper are as follows:

- Combining knowledge from the financial domain, basic linguistic theories, and the idea of semantic role labeling, we design a general-domain open event template. This template includes not only common roles such as subject, trigger, object, time, location, and tool but also roles like scope, background, manner, source, and result, totaling 21 argument roles.

- We annotate an open event extraction corpus OEE-CFC, which can effectively support various complex event extraction tasks. These tasks include extraction of arguments for unconventional entities, extraction of complex relationships (partial overlap, sharing), extraction of open meaningful events, identification of complex triggers (where a text segment serves as a trigger, or a trigger triggers multiple events), and construction of event graphs, etc. Due to the highly logical nature of financial commentaries, the annotated corpus can effectively reveal various relationships between events, making the event graphs constructed based on this corpus more practically meaningful.

- We design four types of prompts and conducted the corresponding experiments for event argument extraction based on OEE-CFC. The prompt focusing on explicit role order achieves the best performance, confirming the significance of OEE-CFC. The corpus and experimental code can be accessed at https://github.com/view98/OEE-CFC.

No.	Role	No.	Role
1	Subject	12	Reason
2	Object	13	Purpose
3	Time	14	State
4	Location	15	degreeAdverb
5	Background	16	scopeAdverb
6	Manner	17	timeAdverb
7	Scope	18	placeAdverb
8	Material	19	affirmativeAdverb
9	Tool	20	negativeAdverb
10	Source	21	modalAdverb
11	Result	22	–

Table 1: General open event template

2 General Open Event Template

The OEE-CFC corpus targets Chinese financial commentaries and aims to manually annotate events that people find interesting. Considering the characteristics of events in financial commentaries and the diversity of events interesting people, it is challenging to summarize templates that reflect different types of events. Therefore, from a linguistic perspective, this paper develops a general open event template based on the semantics and roles of token segments expressed in sentences. By consulting linguistics experts and team discussions, and considering the characteristics (e.g., including many background, manner, scope, and result components) of financial commentaries, a general open event template is formulated, as shown in Table 1.

According to the event roles defined in Table 1, this paper categorizes the roles into three types: **core event roles**, **situational event roles**, and **adverbial roles**. Below, we introduce the meaning of each role. The setting motivation of each role can be found in Appendix A.

2.1 Core Event Roles

Subject role. The *Subject* role is used to describe the entity performing an action or possessing a state, acting as the subject or agent of the action in an event. In Figure 1, blue token segments are the subject of the events. For instance, in the S_1 sentence, "工业消费品价格在居民消费价格指数中的权重" is the subject of the event "为". Additionally, the text within the dotted boxes in Figure 1 (parts of an event) also serve as subjects, such as the first "上升" event in S_5, which serves as the subject of the "拉动" event.

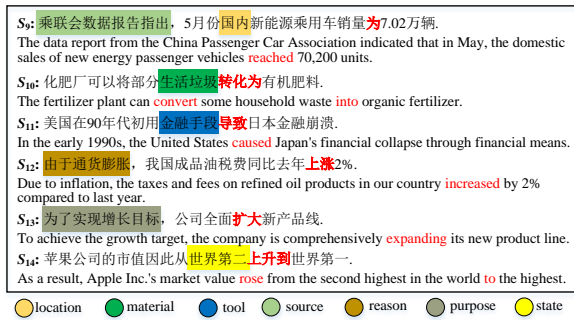Moreover, the annotated event subjects might

Figure 2: Example of situational event roles. Event triggers are highlighted in bold red.

lack some content. For example, the subject of the event "达到了" is marked, but it misses the entire content of the "拉动" event. The complete subject of the "达到了" event is the combination of the "拉动" event and the currently marked subject "对居民...". Since there is a comma between the "拉动" event and the currently marked subject, it separates the two parts, complicating the annotation. Eventually, event subjects can be token segments with modifiers, simple connected tokens, an entire event, or token segments across sentences.

Object role. The *Object* role describes entities in events that are affected by actions or states, serving as the recipient of the action or the bearer of the state. Similar to the *Subject*, arguments assuming the *Object* role are complex, involving various lexical fragments. For instance, the "上涨" event in S_7 serves as the object of the "引起" event. Also, objects can appear in a preposed position, as seen in the green italicized tokens of S_1 in Figure 1.

2.2 Situational Event Roles

This paper collectively refers to roles 3∼14 in Table 1 as situational event roles. Next, we explain the details of each role. Since the examples in Figure 1 do not include all situational event roles, we have added Figure 2, which includes examples from OEE-CFC.

Time role. The *Time* is used to describe the specific time or duration involved in an event. Note that the modifying tokens indicating time or location range in the *Subject/Object* are marked both as part of the complete *Subject/Object* content and separately as the arguments of *Time* or *Location* roles. In Figure 1, the light blue content represents the arguments of the *Time* role, such as "今年前三个月" in sentence S_3. This phrase functions both as an argument of the *Time* role, revealing the time span of the "上涨了" event, and as part of the sub-

ject of the "上涨了" event, indicating overlapping of event arguments. Additionally, the *Time* role often results in complete overlapping of event arguments, where the same phrase indicating the time can simultaneously serve as arguments for multiple *Time* roles in different or the same events.

Location role. The *Location* is used to describe the specific location or area where an event occurs. It can involve cities, countries, regions, or even specific coordinates, depending on the spatial extent of the event. In Figure 2, the token "国内" in S_9 is an argument for the *Location* role in the "为" event.

Background role. The *Background* describes relevant conditions, environments, and prerequisites existing when the event occurs, including the premises leading to the event, background information, and related historical context. In Figure 1, the pink content in sentence S_6, "若扣除能源和食品的价格", serves as the background for the "低于" event, indicating that the "低于" event occurs under the assumed conditions in the pink portion.

Manner role. The *Manner* describes the specific method or way in which an action or behavior in the event is carried out, such as the method, speed, or scale of the action, conveying the specific characteristics of the action in the event. In sentence S_3 of Figure 1, the "上涨了" event occurs in the manner of "分别比去年同期" (see purple content), indicating that the increases of 14.5%, 17.5%, and 30% are conclusions reached in this manner.

Scope role. The *Scope* describes the domain, range, aspect, or angle covered by the event. In sentence S_5 of Figure 1, "从CPI构成要素上看" (see orange content) expresses the scope of the "上升" event, and in sentence S_6, "在一定程度上" expresses the scope of the "存在" event; also, "以西方···通货膨胀率衡量" in sentence S_6 reveals the scope of the "低于" event. Note that "以西方···通货膨胀率衡量" is not modifying the trigger of the "低于" event, so it is considered the scope of the event rather than its manner.

Material role. The *Material* describes the substances used in the event, typically various specific materials or equipment. These materials are usually used to achieve specific actions, processes, or goals, playing a significant role in the development and outcome of the event. For example, "生活垃圾" in sentence S_{10} of Figure 2 is the raw material for producing organic fertilizer at a fertilizer plant.

Tool role. The *Tool* refers to the tools or abstract objects used, involved, or playing a key role in the event to achieve specific goals, complete tasks, or

influence the event's development. In sentence S_{11} of Figure 2, "金融手段" is an abstract tool in the "导致日本金融崩溃" event.

Source role. The *Source* determines the origin of the event information, supporting the authenticity and credibility of the event description. In sentence S_9, "乘联会数据报告指出" describes the source of the "新能源乘用车销量为7.02万辆" event, improving its credibility.

Result role. The *Result* describes supplementary explanations of the outcomes or effects related to the event's actions or behaviors. In sentence S_8 of Figure 1, "1个百分点" is the result of the "提高了" event, where the object is the reserve requirement ratio, indicating the specific result of the increase in the reserve requirement ratio.

Reason role. The *Reason* describes the reasons for the occurrence or development of the event. For example, "由于通货膨胀" in sentence S_{12} of Figure 2 is the cause of the "上涨" event.

Purpose role. The *Purpose* reveals the goal the event aims to achieve, explicitly expressing the event's intent or objective, often directly linked to the main action of the event, forming a causal relationship. In sentence S_{13} of Figure 2, the purpose of the "扩大" event is "为了实现增长目标", and without this role, the meaning of the "扩大" event would be unclear.

State role. The *State* describes changes, evolutions, or transitions of entities or attributes in the event. In sentence S_{14} of Figure 2, the market value of Apple Inc. being the "世界第二" is the state before the "上升" event.

2.3 Adverbial Roles

Adverbs in Chinese texts are often used to modify actions or behaviors in events, providing more information about the actions or behaviors, such as degree, frequency, and time. Previous event templates have overlooked adverbial roles, while adverbial arguments play important roles in events due to their close relationship with the event's triggers.

Adverbial arguments typically precede the event's triggers, modifying and limiting different aspects of the event, enriching the details and context of the event description, making it more specific, accurate, and vivid. The meanings of the adverbial event roles in Table 1 align with the conventional meanings of adverbs, which are not elaborated here. The following examples illustrate the significance of annotating adverbial event roles.

Example 1: "短期供给侧的超调在随后几个月可能会缓和。"

Short-term supply-side over-adjustment may ease in the following months.

Example 2: "随着我国人口结构的变化，新增劳动力总量压力已经明显地减弱。"

With the change in China's population structure, the pressure on the total amount of new labor force has significantly decreased.

In Example 1, the adverb "可能" modifies the trigger "缓和", demonstrating the writer's speculation about the likelihood of the event. Omitting this adverb would imply that the "缓和" event has already occurred, distorting the actual meaning of the text. In Example 2, two adverbs appear: "已经" and "明显地". "已经" indicates that the "减弱" event has occurred, while "明显地" reflects the degree of the "减弱", indicating that this reduction is neither vague nor slight but is very noticeable and clear.

In summary, setting various adverbial roles for events in a general event template can further enrich event descriptions, providing more detailed and accurate information, and helping readers better understand the specific contexts, characteristics, details, and impacts when events occur.

3 Dataset OEE-CFC

3.1 Collection and Pre-processing

This paper selects financial commentary articles from the Forum of 50 Economists[1] as the source of annotated data. During the selection process, we focused on including diverse data. The corpus selection covers various aspects within the financial domain, automotive production and sales, industrial production, and macroeconomics, etc. Additionally, we filtered out low-quality, duplicate, or irrelevant texts to ensure the accuracy and reliability of the dataset. Detailed collection is reported in Table 2, and pre-processing is provided in Appendix B.

3.2 Annotation

Annotation Process. This paper selected BRAT (BRAT Rapid Annotation Tool) as our annotation tool (Stenetorp et al., 2012), since it is a powerful and user-friendly open-source annotation tool, particularly suitable for text annotation tasks. Six native Chinese graduate students are selected as annotators and provided with detailed training mate-

---

[1] http://www.50forum.org.cn/

Attribute	Count
Total Documents	4,253
Total Sentences	10,402
Total Tokens	311,466
Maximum Sentence per Document	12
Average Number of Sentence per Document	2.45
Average Number of Token per Document	73.23

Table 2: Statistics of original dataset

rials, including annotation guidelines and example annotation tasks.

During the annotation training process, detailed definitions of each role in the general open event template, annotation rules, and instructions on how to use the annotation tools were emphasized. Additionally, we conducted practical exercises and simulated annotation tasks to familiarize the annotators with the annotation process and tool, enhancing their annotation skills and accuracy. Finally, we reasonably allocated annotation tasks, managed the annotators' work progress, and implemented strict quality control measures for the annotation. By adopting methods such as mutual annotation among multiple annotators, cross-checking, and correcting annotation results, we aimed to improve the quality of the annotations.

Special Annotation Guidelines. Due to the flexibility of the Chinese language, annotators encountered some complex situations during the actual annotation process that were difficult to handle. After multiple rounds of discussion, we finally unified the annotation strategies. Some special annotation guidelines are reported in Appendix C.

3.3 Analysis

Characteristics and argument distribution. Due to the length of individual financial commentaries, we split the annotated articles into documents to alleviate the window limitations of pre-trained language models. The splitting process was manually performed to ensure the semantic structure of each document remained intact. Finally, the OEE-CFC comprises a total of 115 articles and 4,253 documents. Detailed statistics of the OEE-CFC corpus are shown in Table 3. The count of complex arguments refers to the number of event arguments that are served by complex entities. Other attributes have similar meanings.

As shown in Table 3, complex arguments account for 8.24% of the total event arguments, and most of the event arguments are unconventional

entities. Shared arguments account for 13.34% of the total event arguments, demonstrating that OEE-CFC also features a rich phenomenon of argument sharing. Additionally, the dataset contains some shared triggers (a token triggering multiple events).

In addition to the overall analysis of the characteristics of OEE-CFC, Figure 3 also shows the argument distribution of OEE-CFC under each role. As seen in Figure 3, since the subject and object are the core event roles, they constitute the majority of event arguments, accounting for 60.43%. Among the situational event roles, arguments serving as time, background, scope, manner, etc., are notably prominent, making up 17.57% of the total event arguments and 77.45% of the situational event arguments. Arguments in adverbial roles constitute 16.88% of the total event arguments, with a significant number serving as timeAdverb, degreeAdverb, and scopeAdverb. This indicates that in financial commentaries, information that modifies the event action in terms of adverbs of time, degree, and scope is crucial and common.

Overall, compared to corpora such as ACE2005 and RAMS (Ebner et al., 2020), the argument distribution across roles in OEE-CFC is relatively balanced, with only two roles (i.e., material and placeAdverb) having fewer arguments.

Comparison with RAMS. Each document in RAMS contains only one event, totaling 9,124 events, with a complexity of language description that is less than OEE-CFC, and without multi-valued arguments. Additionally, the annotation quality of RAMS is relatively poor. Many events are forcibly annotated according to event ontology, which fails to reveal the semantic expression of the text. Figure A1 in Appendix demonstrates partial annotated examples in RAMS. Therefore, there is a significant difference in the gold events between RAMS and OEE-CFC that aims to reveal events' semantics according the meanings of sentences.

For example 3 in Figure A1 of Appendix, the text content aims to discuss the information provided to Congress and involves some rules of firefight related to the troops. The core is not about the firefight event. Therefore, annotating a firefight event from this content cannot reflect the semantics of the text. That is, in fact, there are no firefight events in example 3.

Given the example 4 in Figure A1 of Appendix, it can be seen that the text describes two women documented their lives on the al-Raqqah for several weeks, making it difficult to reveal the anno-

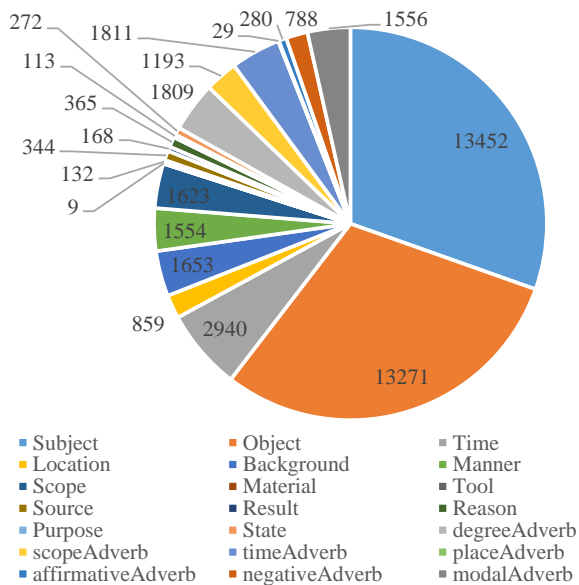| Attribute	Count	Attribute	Count
Total Documents	4,253	Total Shared Complex Entities	220
Total Events	17,469	Total Shared Complex Arguments	495
Average Number of Event per Document	4.11	Total Shared Entities	2,522
Maximum Event Number in Document	19	Total Shared Arguments	5,898
Total Arguments	44,221	Shared Trigger	56
Average Number of Argument per Event	2.53	Total Shared Triggering Events	124
Maximum Number of Argument in Event	8	Number of Multi-valued Arguments	58
Total Complex Entities	3,435	Total Multi-valued Arguments	125
Total Complex Arguments	3,644	–	–

Table 3: Statistics of OEE-CFC

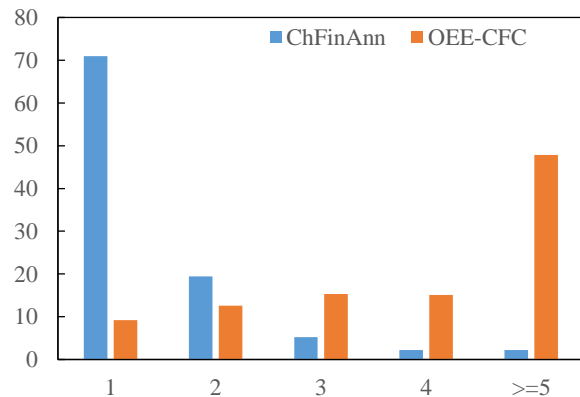Figure 3: The argument distribution under each role.

Figure 4: The event distribution in documents. The horizontal axis represents the number of events per document, while the vertical axis indicates the proportion of documents containing the specified number of events.

tated smuggling incident. There are many more examples like this. Thus, RAMS cannot effectively reveal event information and event-related tasks.

Event distribution. To better reflect the event distribution within documents in the OEE-CFC, we analyzed the number of events contained in documents in both the OEE-CFC and ChFinAnn (Zheng et al., 2019) corpora, as shown in Figure 4. According to Zheng et al. (2019), ChFinAnn is primarily used for multi-event recognition and cross-sentence argument extraction tasks. Despite this, most documents in the ChFinAnn contain only one event, accounting for 70.96%, with documents containing four or more events making up only 4.44%. The situation is the opposite for the OEE-CFC, where docuemnts containing one event are the least common, at just 9.18%, and documents containing four or more events make up 62.92%.

Notably, documents with five or more events account for 47.86%. Therefore, compared to ChFinAnn, OEE-CFC is better suited for training and evaluating tasks related to multi-event recognition and shared argument extraction.

Due to the subjectivity inherent in the manual annotation process, we use the widely adopted kappa score (Sidney, 1957) to measure the agreement among our annotators on the dataset annotations. Six annotators independently annotated 100 common documents in this paper. The average agreement across all roles is 84.88% (0.81 to 0.99 representing the highest level of agreement), and the annotation agreement for triggers is 91.2%.

4 Models

In order to verify the significance of OEE-CFC, this paper constructs two event argument extraction frameworks based on the span selection strategy, including PAIE (Ma et al., 2022) and TabEAE (He et al., 2023). PAIE is a prompt-based event

Non-continuous prompt: background source scope time location subject manner material tool reason state degreeAdverb scopeAdverb timeAdverb placeAdverb affirmativeAdverb negativeAdverbs modalAdverb object purpose result

Explicit continuous prompt: event background description event source description in the scope involved in the event in a certain time range a certain location range the subject of an event that uses some manner material or tool for some reason or change of state in degreeAdverb in different scopeAdverb at different timeAdverb in different placeAdverb affirmativeAdverb negativeAdverbs modalAdverb and the object of the event achieves some purpose and produces some result

General continuous prompt: SEP background SEP source SEP scope SEP time SEP location SEP subject SEP manner SEP material SEP tool SEP reason SEP state SEP degreeAdverb SEP scopeAdverb SEP timeAdverb SEP placeAdverb SEP affirmativeAdverb SEP negativeAdverbs SEP modalAdverb SEP object SEP purpose SEP result

Argument quantity-based prompt: SEP subject SEP object SEP time SEP timeAdverb SEP degreeAdverb SEP background SEP scope SEP modalAdverb SEP manner SEP scopeAdverb SEP location SEP negativeAdverbs SEP reason SEP source SEP affirmativeAdverb SEP state SEP result SEP tool SEP purpose SEP placeAdverb SEP material

Figure 5: Prompts based on role order and quantity. Blue tokens refer to roles, SEP indicates the general symbols.

Prompt	PAIE				TabEAE			
	Head-I	Head-C	Span-I	Span-C	Head-I	Head-C	Span-I	Span-C
Non-continuous	75.01	68.22	65.05	60.93	71.64	64.00	61.58	55.56
Explicit Continuous	75.92	69.51	66.35	63.47	70.38	62.21	60.48	56.24
General Continuous	75.24	69.23	65.85	62.22	67.83	61.24	57.02	53.99
Argument Quantity	75.48	69.21	66.01	62.29	69.21	61.35	58.89	54.03

Table 4: $F1$ score of models using different prompts on OEE-CFC

argument extraction model. It designed a prompt for each event type and selected the start and end positions from an input document for each role to identify unconventional arguments playing the role. TabEAE transformed the event argument extraction task into a problem of table generation and encoded event co-occurrence information based on the PAIE framework.

For event argument extraction, guiding pre-trained language models with prompts can effectively integrate semantic information between event roles, helping the model to understand and extract event argument information for specific roles. Therefore, based on the PAIE and TabEAE frameworks, this paper designs four prompts according to the proposed open event templates.

Due to the high generality of the open event template, where role definitions are broad and role names are not sufficiently clear in terms of semantics, this paper mainly constructs prompts from

two aspects: the order of roles and the number of arguments contained by the roles.

Role order prompt. Considering that people's sentence expressions have certain logic and style habits, this paper designs the role order prompt to capture these potential semantic information. The role order is divided into **non-continuous role order prompt** (direct combination of roles) and **explicit continuous role order prompt** (adding specific descriptive tokens between roles), based on whether there are descriptive connecting tokens between roles. Additionally, to make the continuous role order prompt more general, we also design a **general continuous role order prompt** that uses fixed symbols to represent specific tokens. According to the common expression habits of most people, the three designed prompts are shown in Figure 5.

Role quantity prompt. This strategy assumes that the more frequently a role appears in docu-

ments, the more significant its influence factor is, thus it should be placed earlier in the prompt. To this end, we constructed the last prompt in Figure 5 according to the argument distribution in Figure 3, arranging them in descending order of quantity.

5 Experiments

5.1 Dataset and Evaluation Metrics

The experiments were constructed based on the OEE-CFC corpus using the PAIE and TabEAE. We randomly divided the OEE-CFC into training, dev, and test sets in a 7:2:1 ratio, resulting in 2,977/851/425 samples, respectively. Following PAIE, experimental results are evaluated in terms of head-token and span with the $F1$ score. Most model hyper-parameters are consistent with PAIE, while due to the longer sentences in OEE-CFC, we modified the values of hyper-parameters such as the maximum input window and maximum span length for the encoder and decoder. The pre-trained model used is BART-Base-Chinese.

5.2 Main Results

Table 4 reports the argument extraction performance of PAIE and TabEAE on OEE-CFC using different prompts. Head-I/Head-C and Span-I/Span-C represent the identification/classification performance based on head-tokens and spans, respectively.

As shown in Table 4, using the explicit continuous prompt achieved the best results in PAIE, with the Span-C surpassing other prompts by 1.18 to 2.54 percentage points. As expected, explicit continuous prompt provides richer role semantics due to its specific context, whereas the noncontinuous prompt can only capture the basic semantic meaning of roles, resulting in the poorest performance. Compared to the explicit continuous prompt, the general prompt requires the model to learn the underlying general rules automatically, which is more demanding. Therefore, when the dataset size is not particularly large, the performance is slightly lower than that of the explicit continuous prompt. The argument quantity prompt also achieved good results, indicating that the role order in the prompt can influence the argument extraction performance. Given the TabEAE model, Non-continuous prompt achieves the second performance by employing the RoBERTa language model, suggesting that the role of prompts is related to the language model selected.

6 Conclusions

To advance the extraction of unconventional arguments in complex scenarios and to overcome the limitation of extracting only specified event types, this paper first develops a general open event template from a linguistic perspective. This template includes three categories of roles, encompassing a total of 21 event roles. Then, we annotated the event extraction corpus OEE-CFC based on Chinese financial commentaries. This corpus includes 4,253 documents, 17,469 events, and 44,221 arguments, with shared arguments accounting for 13.34% of the total event arguments, and documents containing five or more events making up 47.86%. Finally, given the PAIE and TabEAE, four types of prompts were designed, and the corresponding experiments were conducted. Among them, the explicit continuous prompt achieved the best performance, verifying the significance of OEE-CFC.

Limitations

The OEE-CFC corpus annotated in this paper is based on Chinese financial commentaries and can only support Chinese scenarios. The scale of events and arguments is moderate and could be further expanded. Additionally, the relationships between events are not annotated. For evaluation, only event argument extraction model has been constructed, and the benchmark for extracting events involving unconventional entities is not provided. All these are the limitations of this paper.

Acknowledgements

References

Haolin Deng, Yanan Zhang, Yangfan Zhang, Wangyang Ying, Changlong Yu, Jun Gao, Wei Wang, Xiaoling Bai, Nan Yang, Jin Ma, et al. 2022. Title2Event: Benchmarking open event extraction with a large-scale chinese title dataset. In *Proceedings of the*

2022 Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 6511–6524.

Seth Ebner, Patrick Xia, Ryan Culkin, Kyle Rawlins, and Benjamin Van Durme. 2020. Multi-sentence argument linking. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 8057–8077.

Yuxin He, Jingyue Hu, and Buzhou Tang. 2023. Revisiting event argument extraction: Can eae models learn better when being aware of event co-occurrences? In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 12542–12556.

Sha Li, Heng Ji, and Jiawei Han. 2021. Document-level event argument extraction by conditional generation. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, pages 894–908.

Yubo Ma, Zehao Wang, Yixin Cao, Mukai Li, Meiqi Chen, Kun Wang, and Jing Shao. 2022. Prompt for extraction? paie: Prompting argument interaction for event argument extraction. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 6759–6774.

Tanmay Parekh, I Hsu, Kuan-Hao Huang, Kai-Wei Chang, Nanyun Peng, et al. 2023. GENEVA: Benchmarking generalizability for event argument extraction with hundreds of event types and argument roles. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*, pages 3664–3686.

Siegel Sidney. 1957. Nonparametric statistics for the behavioral sciences. *The Journal of Nervous and Mental Disease*, 125(3):497.

Pontus Stenetorp, Sampo Pyysalo, Goran Topić, Tomoko Ohta, Sophia Ananiadou, and Jun'ichi Tsujii. 2012. BRAT: A web-based tool for nlp-assisted text annotation. In *Proceedings of the Demonstrations at the 13th Conference of the European Chapter of the Association for Computational Linguistics*, pages 102–107.

MeiHan Tong, Bin Xu, Shuai Wang, Meihuan Han, Yixin Cao, Jiangqi Zhu, Siyu Chen, Lei Hou, and Juanzi Li. 2022. DocEE: A large-scale and fine-grained benchmark for document-level event extraction. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, pages 3970–3982.

Amir Pouran Ben Veyseh, Javid Ebrahimi, Franck Dernoncourt, and Thien Huu Nguyen. 2022. MEE: A novel multilingual event extraction dataset. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 9603–9613.

Qizhi Wan, Changxuan Wan, Keli Xiao, Rong Hu, and Dexi Liu. 2023a. A multi-channel hierarchical graph attention network for open event extraction. *ACM Transactions on Information Systems (TOIS)*, 41(1):1–27.

Qizhi Wan, Changxuan Wan, Keli Xiao, Dexi Liu, Chenliang Li, Bolong Zheng, Xiping Liu, and Rong Hu. 2023b. Joint document-level event extraction via token-token bidirectional event completed graph. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 10481–10492.

Qizhi Wan, Changxuan Wan, Keli Xiao, Dexi Liu, Qing Liu, Jiangling Deng, Wenkang Luo, and Rong Hu. 2022. Construction of a Chinese corpus for multi-type economic event relation. *Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 21(6):1–20.

Qizhi Wan, Changxuan Wan, Keli Xiao, Kun Lu, Chenliang Li, Xiping Liu, and Dexi Liu. 2024a. Dependency structure-enhanced graph attention networks for event detection. In *Proceedings of the Thirty-Eighth AAAI Conference on Artificial Intelligence (AAAI)*, pages 19098–19106.

Qizhi Wan, Changxuan Wan, Keli Xiao, Hui Xiong, Dexi Liu, Xiping Liu, and Rong Hu. 2024b. Token-event-role structure-based multi-channel document-level event extraction. *ACM Transactions on Information Systems (TOIS)*, 42(4):1–27.

Xiaozhi Wang, Ziqi Wang, Xu Han, Wangyi Jiang, Rong Han, Zhiyuan Liu, Juanzi Li, Peng Li, Yankai Lin, and Jie Zhou. 2020. MAVEN: A massive general domain event detection dataset. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1652–1671.

Mengjie Wu, Maofu Liu, Luyao Wang, and Huijun Hu. 2023. A Chinese fine-grained financial event extraction dataset. In *Companion Proceedings of the ACM Web Conference 2023 (WWW)*, pages 1229–1235.

Shun Zheng, Wei Cao, Wei Xu, and Jiang Bian. 2019. Doc2EDAG: An end-to-end document-level framework for chinese financial event extraction. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 337–346.

Appendix

A   Description about Event Roles

In the following, we introduce the event roles proposed in the paper.

(1) Subject role. The *subject* role can be subdivided into agent and experiencer[2]. The agent refers to the conscious subject that performs specific actions, continues actions, or exhibits psychological activities, while the experiencer refers to the non-acting initiator without subjective initiative.

(2) Object role. The *object* role describes entities in events that are affected by actions or states, serving as the recipient of the action or the bearer of the state. The *object* role can be further divided into four categories: patient, theme, dative, and associated objects. Patient refers to the direct object changed by the subject's action. Theme encompasses objects involved in the event but not altered by it, as well as new entities or outcomes produced by the action. Involved (also known as participant) denotes a non-active participant in the action, essentially a third-party participant in the semantic event. Associated objects are entities related to the participants of the event.

(3) Time role. The *Time* role is used to describe the specific time or duration involved in an event. This role helps to better understand the temporal background and sequence of the event, thereby constructing a timeline for the event and providing a deeper understanding of its process and context. Note that the modifying tokens indicating time or location range in the subject/object are marked both as part of the complete subject/object content and separately as the arguments of the event's *Time* or *Location* roles.

(4) Location. The *Location* role is used to describe the specific location or area where an event occurs. This role helps determine the geographic position of the event and provides spatial background information. It can involve cities, countries, regions, or even specific coordinates, depending on the spatial extent of the event.

(5) Background role. The *Background* role describes relevant conditions, environments, and prerequisites existing when the event occurs, including the premises leading to the event, background information, and related historical context. This role facilitates readers understand the background information of the event, offering crucial details for a comprehensive interpretation.

(6) Manner role. The *Manner* role describes the specific method or way in which an action or behavior in the event is carried out, such as the method, speed, or scale of the action, conveying the specific characteristics of the action in the event. Setting the *Manner* role helps to highlight the action process in the event, describe how the event is carried out, emphasize the specific execution method, and enrich the event description for greater clarity.

(7) Scope role. The *Scope* role describes the domain, range, aspect, or angle covered by the event. Setting this role provides a clearer definition of the event's scope and ensures a complete event description by connecting with other event components.

(8) Material role. The *Material* role describes the substances used in the event, typically various specific materials or equipment. These materials are usually used to achieve specific actions, processes, or goals, playing a significant role in the development and outcome of the event. Setting the *Material* role emphasize the core elements of the event, making the event description more concrete and practical.

(9) Tool role. The *Tool* role refers to the tools or abstract objects used, involved, or playing a key role in the event to achieve specific goals, complete tasks, or influence the event's development. Tools can be tangible items or abstract concepts, methods, or strategies. Setting the *Tool* role highlights the specific means or methods used in the event, making the event description more concrete and practical, providing richer details for a deeper understanding of the event's mechanism and implementation.

(10) Source role. The *Source* role determines the origin of the event information, supporting the authenticity and credibility of the event description. This helps readers better understand the basis and source of the information, enhancing its reliability and transparency.

(11) Result role. The *Result* role describes supplementary explanations of the outcomes or effects related to the event's actions or behaviors. These results are directly related to the event's triggering action, explaining the specific changes or consequences resulting from the event. Setting the *Result* role helps provide detailed descriptions of the outcomes of the actions, enriching the event's details and comprehensiveness.

(12) Reason role. The *Reason* role describes the reasons for the occurrence or development of

the event. Clarifying the event's cause makes the description more complete, enabling readers to better understand the event's origin and the causal relationships between events.

(13) Purpose role. The *Purpose* role reveals the goal the event aims to achieve, explicitly expressing the event's intent or objective, often directly linked to the main action of the event, forming a causal relationship. The Purpose role can emphasizes the motivation behind the event, helping readers understand the event's driving factors.

(14) State role. The *State* role describes changes, evolutions, or transitions of entities or attributes in the event, usually directly related to the event's main action or triggering factor. The *State* role emphasizes the impact and significance of changes on entities or attributes before and after the event. By introducing the *State* role, the event description becomes more vivid and specific, clarifying the event's development process and its effects.

B Pre-Processing

We used two methods to obtain the raw texts: manual collection and web scraping. Then, we performed text cleaning, which included removing duplicate texts, HTML tags, special characters, and irrelevant formatting information to ensure the texts were clean and consistent. Finally, we segmented the documents into sentences. After the above processing, a total of 115 financial commentary articles were obtained for dataset annotation, which were divided into 4,253 documents. Detailed statistics are shown in Table 2, where the number of tokens is result after word Segmentation using LTP tool.

C Special Annotation Guidelines

Below are some important annotation guidelines.

(1) When both a light verb and a full verb appear in an event, annotate the full verb as the event trigger word.

Example A1: "电解铝行业125户生产企业中已经有39户停产，55户净亏损。"

Among the 125 electrolytic aluminum production enterprises, 39 have stopped production, and 55 are operating at a net loss.

In Example A1, there are two verbs: "有" and "停产". "有" is a light verb, while "停产" is a full verb. Light verbs usually do not directly express the executor or the recipient of the action but convey this information through subsequent components.

Therefore, in the actual annotation process, when both appear simultaneously, the full verb is always annotated as the event trigger word.

(2) When pronouns appear, identify their actual referents for accurate annotation.

Example A2: "中国的中等收入阶层人数占总人数的比重从1999年的15%上升到2003年的19%，到2020年，这个数字将有望达到40%左右。"

The proportion of the middle-income class in China increased from 15% in 1999 to 19% in 2003, and this figure is expected to reach about 40% by 2020.

In Example A2, "这个数字" actually refers to "中等收入阶层人数占总人数的比重". Thus, when annotating the arguments of the event "达到" in the latter part of the sentence, we stipulate that "中等收入阶层人数占总人数的比重" should be annotated as the actual subject of the event.

(3) When consecutive verbs or "verb-adverb-verb" combinations appear, they can be annotated as a whole.

Example A3: "苹果公司的市场份额在过去几个季度中保持增长。"

Apple's market share has continued to grow over the past few quarters.

Example A4: "投资者希望不断地提高他们的资产配置效率。"

Investors hope to continuously improve the efficiency of their asset allocation.

In Example A3, both "保持" and "增长" are verbs, and in this sentence, they jointly trigger the same event: Apple's market share continues to grow. Therefore, it is stipulated that these two verbs should be annotated as a whole. Similarly, in Example A4, "希望不断地提高" is a verb-adverb-verb combination, and it is also considered as a whole and annotated as the event trigger.

(4) Event trigger words may trigger multiple events and need to be annotated in detail.

Example A5: "XT4在5月和6月分别售出7,047辆和7,324辆。"

XT4 sold 7,047 units in May and 7,324 units in June.

In Example A5, the trigger "售出" actually triggers two different events: XT4 sold 7,047 units in May and XT4 sold 7,324 units in June. Therefore, during annotation, these two events should be annotated in detail, sharing the same trigger "售出" and the subject argument "XT4".

(5) Event arguments may overlap.

Example A6: "随着人均国内生产总值超过1,000美元，我国城乡居民收入水平大幅度提高。"

As the per capita GDP exceeds 1,000 USD, the income level of urban and rural residents in our country has increased significantly.

Example A6 is a complex case of overlapping event arguments. "随着人均国内生产总值超过1000美元" constitutes an event where the per capita GDP exceeds 1,000 USD, and it can also serve as the background argument for the "提高" event in the latter part of the sentence. During the annotation process, we detailed such situations.

Example 3:
[S_1] As members of Congress, it our constitutional duty to provide oversight on behalf of the American people.
[S_2] In order to provide accurate information for our constituents we request that a briefing be made available to Congress as soon as possible to answer the following questions:
[S_3] At any point during the firefight did the current rules of engagement restrict the immediate use of assets on hand?
[S_4] Has there been any consideration in altering the current rules of engagement to be less restrictive in the future?
[S_5] Did the U.S. or ANDSF have any air assets available in the region to support the forces that were engaged on the ground?
Event type: conflict.attack.firearmattack, **Trigger:** firefight, **Roles:** *target*: forces, *instrument*: firefight, place: ground

Example 4:
[S_1] 2014: ISIS marches into Raqqa and makes it the capital of the caliphate.
[S_2] Foto: Uncredited
[S_3] The two Syrian women who we shall call Om Omran and Om Mohammad, which are not their real names, were willing to wear Expressen's hidden cameras which have been smuggled in.
[S_4] We want the world to know, they say.
[S_5] Over the course of several weeks, they have documented life in the completely isolated city of al - Raqqah.
Event type: movement.transportartifact.hide, **Trigger:** smuggled, **Roles:** *transporter*: Om Omran and Om Mohammad, *artifact*: hidden cameras, *origin*: al - Raqqah

Figure A1: partial annotated examples in RAMS corpus. Red tokens refer to triggers, other color tokens are event arguments serving as different roles.