EmoPrompt-ECPE: Emotion knowledge-aware Prompt-tuning for Emotion-Cause Pair Extraction

Xue Gu[1,2], **Zhihan Zhou**[2], **Ziyao Meng**[1,2], **Jian Li**[2], **Tiago Gomes**[1], **Adriano Tavares**[1], **Hao Xu**[2,*]

[1]School of Engineering, Department of Industrial Electronics, University of Minho, Portugal
id9267, id9272@alunos.uminho.pt, mr.gomes, atavares@dei.uminho.pt

[2]College of Computer Science and Technology, Jilin University, Changchun, China
zhzhou22, lijian20@mails.jlu.edu.cn, xuhao@jlu.edu.cn

Abstract

Emotion-cause pair extraction (ECPE) main focus is on extracting all potential emotion clauses and corresponding cause clauses from unannotated documents. Existing methods achieve promising results with the help of fine-tuning and prompt paradigms, but they present three downsides. First, most approaches cannot distinguish between the emotion-cause pairs that belong to different types of emotions, limiting the existing approaches' applicability. Second, existing prompt methods utilize a one-to-one mapping relation to achieve label words to category mapping, which brings considerable bias to the results. Third, existing methods achieve the cause extraction task supported by explicit semantic understanding or basic prompt templates, ignoring the implicit information contained in the cause clauses themselves. To solve these issues, we propose an Emotion knowledge-aware Prompt-tuning for Emotion-Cause Pair Extraction (EmoPrompt-ECPE) method, which integrate the knowledge of emotion categories in the ECPE task and mine the implicit knowledge of cause clauses. Specifically, we inject the latent knowledge of the cause clauses and the emotion types into the prompt template. Besides, we extend the emotion labels for many-to-one mapping of label words to categories with an external emotion word base. Furthermore, we utilize the cosine similarity filtering of the label word base to reduce the noise caused by knowledge introduction. Experiments on both Chinese and English benchmark datasets show that our approach can achieve state-of-the-art results. Our code and data can be found at: https://github.com/xy-xiaotudou/EmoPrompt-ECPE.

Keywords: Emotion-cause pair extraction, Prompt-tuning, Emotion knowledge

1. Introduction

The purpose of sentiment analysis is to classify texts into positive, neutral, negative, or more fine-grained emotional categories by analyzing the content of given text, which plays a vital role in decision-making and behavior analysis. However, recent research is focused on identifying the reasons behind the emotions instead on discovering emotional polarity (Gui et al., 2016a,b; Xu et al., 2017; Cheng et al., 2017). Emotion cause analysis (ECA) helps to provide insight into the emotions expressed in a text and the reasons behind their expression, and it has a wide range of applications including but not limited to opinion monitoring and customer service.

One of the most well-known tasks for ECA is Emotion Cause Extraction (ECE), which main objective is to determine the cause of a given emotion expressed in text. However, it requires pre-annotation of the emotion before performing cause extraction, which limits its application in real-world scenarios. Moreover, it disregards the association between emotion and cause clauses. In order to tackle these problems, (Xia and Ding, 2019) described a new task, Emotion-Cause Pair Extraction (ECPE), which can directly extract all potential emotion-cause pairs (ECPs) from unannotated text, as illustrated in Figure 1. The document has six clauses, and C4 shows the emotion "happy" caused by C2 and C3. Clause C5 expresses a feeling of "worried," which relates to

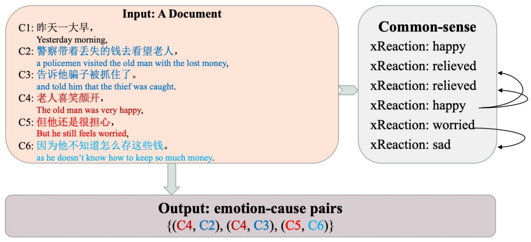

Figure 1: Here is a sample document that incorporates common-sense knowledge gathered from COMET's xReaction relationship. The xReaction relationship refers to the emotions experienced by the primary entity in the occurrence. For example, emotion clause C4's xReaction is "happy", the xReaction for the cause clauses C3 and C2 of C4 are both "relieved", and it is clear that the xReaction's corresponding to cause and emotion clauses are different.

*Corresponding author

LREC-COLING 2024, pages 5678–5688
20-25 May, 2024.

C6. Therefore, the output includes a set of ECPs: $\{(C4, C2), (C4, C3), (C5, C6)\}$.

Intuitively, ECPE is achieved by solving three individual tasks: Emotion clause Extraction (EE), Cause clause Extraction (CE), and Pair extraction (ECPE). Based on this intuition, (Xia and Ding, 2019) first presented a 2-step pipeline framework where the emotion and cause sets are first extracted independently, and then their elements are paired and filtered one by one. However, this 2-step method is prone to cascading errors, allowing errors occurrence in the first step to be passed on to the second step. Researchers attempted to construct a one-step framework for the ECPE task and proposed several compelling end-to-end works to solve this issue. Current literature approaches ECPE task (1) as a multi-task joint learning of three tasks (Huang et al., 2023; Li et al., 2023), (2) using graph constructions to represent the relations between emotions and causes (Chen et al., 2020b; Liu et al., 2022; Chen and Mao, 2023), (3) as a sequence tagging problem (Fan et al., 2020; Yuan et al., 2020; Cheng et al., 2021) and (4) as a machine reading comprehension problem (Zhou et al., 2022; Nguyen and Nguyen, 2023).

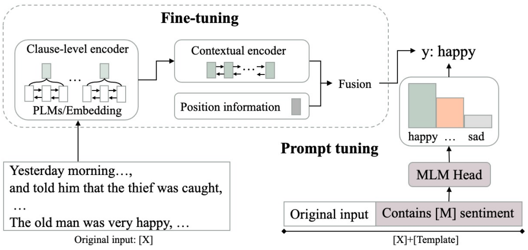

Figure 2: The generic forms of processing for emotion classification tasks are fine-tuning and prompt-tuning, respectively.

Although existing approaches differ in model structures, they all use a fine-tuning paradigm. As illustrated in Figure 2, the fine-tuning approach first obtains the feature representation of each clause from a pre-trained language model (PLM) or word embedding. After that, the contextual representation is obtained from the context encoder and combined with the clause-level position information or dataset bias (i.e., nearly 87% of cause clauses are located near emotion clause in the Chinese dataset). Finally, the merged representation is fed into the fully connected layer for task prediction. However, the context encoder in fine-tuning paradigm is often designed for a specific task goal, requiring redesign and training for different tasks. The position information makes the model sensitive to data distribution, which may cause performance degradation when biases exist in the data distribution.

Unlike the fine-tuning paradigm, the prompt paradigm is naturally applicable to multi-task learning, which enables explicit modeling of multiple-task objectives by converting a specific fine-tuning task into a pre-trained task. Based on this, (Zheng et al., 2022) proposed UECA-Prompt, which provides an effective solution for ECPE tasks by constructing prompts to guide the modeling of the three sub-tasks. However, existing prompt-tuning approaches suffer from the following issues: (1) they do not know which emotion is contained in the extracted ECPs, (2) the one-to-one verbalizer limits the coverage of labeled words, affecting the model's generalization ability, and (3) the "is/isn't a cause clause" template does not learn the latent semantic information of the cause clause well.

In general, the emotion clauses usually contain an emotion keyword, such as "happy" and "worried" in this case, and cause clauses generally represent an event or an action. This feature alone is not good enough to effectively identify the cause clause. For this reason, we extract and use the implicit knowledge in the text, such as common-sense, to assist Cause Clause Extraction task. With the help of Atomic[1], we obtained the xReaction knowledge of each clause in the document, as shown in Figure 1. xReaction of the emotion clause C4 is "happy", and the xReaction of the corresponding cause clauses C2 and C3 are both "relieved". C5 has an xReaction of "worried". The xReaction of its cause sentence C6 is "sad". We noticed that: the cause clause always appears in the text where the xReaction is different from the emotion sentence, and this phenomenon accounts for 90% of the overall benchmark data.

Motivated by the above problem, we propose an Emotion Knowledge-aware Prompt-Tuning for Emotion-Cause Pair Extraction method (EmoPrompt-ECPE). Specifically, in EmoPrompt-ECPE the three sub-tasks of ECPE are designed as three sub-prompts and explicitly models the relationship among them by combining the sub-prompts. We convert the CE task into sub-prompts: "*i feel things changed/unchanged*" to exploit the latent knowledge of the text. For the EE task, we integrate the verbalizer to expand label words with the help of an external knowledge base to achieve many-to-one category mapping. To cope with the noise caused by label word expansion, we design a relevance-based selection method to improve the accuracy and stability of the model.

The main contributions of this paper are as follows:

- We propose an EmoPrompt-ECPE task by incorporating an external knowledge base to reduce the bias due to single emotion label words, while easily distinguishing ECPs of dif-

[1] https://mosaickg.apps.allenai.org/kg_atomic2020

ferent emotion types, and effectively exploiting textual implicit knowledge to improve model performance.

- We construct a complete knowledge base of sentiment labels containing six categories: Happiness, Sadness, Anger, Disgust, Surprise, and Fear.
- EmoPrompt-ECPE improves state-of-the-art results on two typical datasets.

2. Related Work

(Xia and Ding, 2019) first presented the ECPE task and designed a pipeline model consisting of two classifiers. The first classifier is trained to extract emotion clauses and cause clauses separately, and then the second one is trained to eliminate negative samples. Although it yielded promising performance, it still suffers from cascading errors. In addition, the pair extraction model and the clause extraction model are independent of each other, leading to the inability of the method to fully exploit the correlation between tasks.

To avoid these problems, numerous end-to-end methods have emerged. (Wei et al., 2020) propose RankCP to handle the ECPE task from a ranking perspective, modeling the relationships between clauses using graph attention networks. (Ding et al., 2020a) propose ECPE-2D to denote the ECPs in a 2D representation scheme and uses a 2D transformer to model the interactions between different ECPs. (Song et al., 2020) propose E2EECPE, which treats the ECPE task as predicting directional links among emotions and causes through biaffine attention. (Wu et al., 2020) present MTNECP that combines EE and CE to learn emotion-cause relationship classification and share useful features across tasks. (Cheng et al., 2020) propose SLSN consisting of the emotion and cause sub-networks to enable simultaneous detection and matching. (Chen et al., 2022a) present a recurrent synchronization network to model the interaction across different tasks. (Hu et al., 2022) transforms ECPE into a probabilistic problem to leverage joint learning of emotion and cause clauses with the help of mutual information. (Chen et al., 2022b) propose an A^2Net to model multi-tasks. (Huang et al., 2023) design an interactive attention module (IA-ECPE) for ECPE tasks. (Li et al., 2023) design an end-to-end multi-task model containing a shared and task-specific module. (Chen et al., 2020b) present a PairGCN to model dependency relations among candidate pairs, while (Liu et al., 2022) contrarily propose PBJE, a pair-based joint coding network. (Chen and Mao, 2023) present a multi-granular graph to solve the Emotion-Cause Pair Extraction task. (Ding et al., 2020b) propose a multi-label learning method that transforms the ECPE task into the problem of extracting the emotion/cause-pivot cause/emotion in the sliding window. (Chen et al., 2020a) and (Cheng et al., 2021) allocate the emotion category labels to emotion and cause clauses, and by capturing the emotion consistency, it easily distinguish the ECPs of different emotion types. (Yuan et al., 2020) and (Fan et al., 2021) carry out the emotion-cause pairs extraction by encoding the distance between the emotion and cause clauses into labels, capturing the positional bias to extract the emotion and the corresponding cause simultaneously. (Chang et al., 2022) and (Zhou et al., 2022) tackle the Emotion-Cause Pair Extraction task as machine reading comprehension problem. (Nguyen and Nguyen, 2023) cast the Emotion-Cause Pair Extraction task to the question answering problem.

Unlike the above fine-tuning approaches, (Zheng et al., 2022) propose a prompt learning based Emotion-Cause Pair Extraction approach. This method transforms the three sub-tasks of ECPE into three sub-objectives and realizes the extraction of multiple sub-tasks by constructing special prompts. However, existing methods cannot distinguish the emotion categories of the extracted pairs and cannot also utilize potential semantic information of the cause clauses, such as the emotion state of the primary entity in the clause. Contrarily, our approach leverages the potential semantics of the cause clauses and distinguish pairs of different emotion categories. In doing so, it uses specific prompt structure while extending the mapping between labels to categories supported by an additional emotion knowledge base to enhance the performance of the ECPE task.

3. Problem Formulation

ECPE aims to extract pairs of emotions and its corresponding causes from a given document (see Figure 1). The input is a document with multiple clauses, represented as $D = (c_1, c_2, \dots, c_n)$, where n is the number of clauses. Each clause represented as $c_i = (w_1^i, w_2^i, \dots, w_m^i)$, where m is the length of the clause. The outcome is an extracted set P containing all ECPs from D: $P = \{(c_{emo^1}, c_{cau^1}), \dots, (c_{emo^l}, c_{cau^l})\}$. Here, c_{emo^l} and c_{cau^l} are the emotion and cause clauses in the l-th pair, respectively.

4. Approach

The proposed EmoPrompt-ECPE mainly contains four parts, sub-prompts for task decomposition, emotion verbalizer construction, verbalizer adjustment, and task implementation, as shown in Figure 3. We transform the classification tasks into

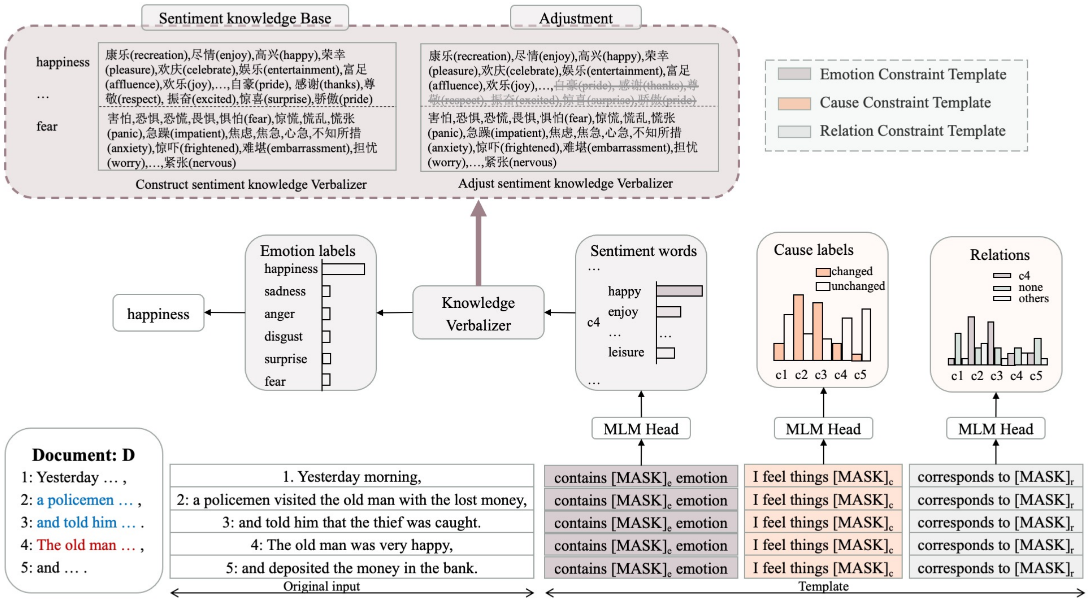

Figure 3: The overview of the EmoPrompt-ECPE. $[MASK]_e$, $[MASK]_c$, $[MASK]_r$ are the sub-prompts for EE, CE, and ECPE. The knowledge Verbalizer maps predictions over emotion words into emotion category labels.

masked language modeling (MLM) problems with the help of prompt-tuning blocks enhanced by emotional knowledge. EmoPrompt-ECPE first wraps the input documents according to the templates. Then, the PLM obtains the probability of each word in the vocabulary being filled into the [MASK]. After that, the label mapping of CE and ECPE tasks is obtained directly from the probabilities, while the EE task performs the mapping using the pre-constructed emotion label word set.

4.1. Sub-prompts for ECPE

For the ECPE task, we design three sub-prompts for three sub-tasks: emotion constraint template, cause constraint template, and relationship constraint template.

Emotion Constraint Template, shown as the purple rectangle in Figure 3, is used to determine the emotion category of the clauses. Specifically, each clause c_i in the document D is classified into an emotion category $Y_e = \{$Happiness, Sadness, Anger, Disgust, Surprise, Fear, None$\}$. The template "contains [MASK] emotion", and the candidate label words used to predict answer [MASK] is given by the label word set V_e consisting of additional sentiment words. For clause c_4 = "4:The old man was very happy.", we follow the template to wrap it as c_4^p = "[CLS] 4:The old man was very happy. contains [MASK] emotion." Then the PLM M gives the probability $P_M([MASK] = v_e|c_4^p)$ of the words that will be filled in the [MASK]. Then the probability of label $y_e \in Y_e$ given by:

$$P(y_e|c_4^p) = g(P_M([MASK] = v_e|c_4^p)|v_e \in V_e^y)$$

where V_e^y denote the subset of V_e that is mapped into a specific label y_e.

Cause Constraint Template, shown as the orange rectangle in Figure 3, determines whether a clause c_i in document D is a cause clause. The template is "I feel things [MASK]", and the candidate tag words for predicting answer [MASK] include "*changed*" and "*unchanged*". The "*changed*" indicates that the current clause is a cause clause, and the "*unchanged*" says it is not. Unlike the many-to-one label mapping in the emotion module, the cause module uses a one-to-one mapping relationship.

Relationship Constraint Template, shown as the gray rectangle in Figure 3, is used to determine the clause associated with the current clause. The template is "corresponds to [MASK]", while the answer slots [MASK] are "{1, 2, i, . . . , n, none}", where n is the number of clauses in the document and "none" means there is no clause related to the current clause. The i indicates that the i-th clause is related to the current clause. To better learn the order of clauses in the document, we use the unique identifier value representing each clause as input to the model with the textual information. Put it simply, the serial number of the clause is learned directly as part of the input.

4.2. Emotion Verbalizer Construction

The verbalizer is a particularly critical component in the process of prompt-tuning mask prediction, and its main role is mapping label words to categories to promote better emotion classification. This process is not a single-choice process as multiple candidates can be generated based on the context. It also dictates the construction of a label word base with broad coverage and slight subjective bias for this many-to-one mapping. Therefore, we construct a reasonable and reliable emotion label word base for the EE sub-task by introducing an external emotion knowledge base.

The key role of EE is to mine the cue words associated with the emotion category from all subclauses, and these cue words are considered as projections of label words. In other words, we can identify various emotions quickly if we have a large and complete label word base of emotions containing all emotion categories.

Therefore, we introduce the Sentiment Ontology Library[2] as an additional Chinese emotion word base for Chinese emotion categories. Similarly, we use NRC Emotion Lex[3] as an external English emotion word base. The collected emotion word base consists of the additional emotion word base and the cue words from the ECPE benchmark dataset. Table 1 shows some examples of the emotion word base.

4.3. Emotion Verbalizer Adjustment

The Chinese and English sentiment words we collected cover most of the label words involved in the six emotion classification systems proposed by Ekman. Although this is a sentiment word base containing a comprehensive vocabulary, the external word base is not tailored for the ECPE task, which introduces some noise (such as ambiguity, as shown in Figure 3, 'SURPRISE' is not in the 'HAPPINESS' emotion). For this reason, it is necessary to adapt the collected word base to retain large-scale, high-quality label words.

We mainly filter the label word database based on the correlation between label words and emotion categories to obtain the relevance of label words to each emotion category. We use the predicted probability of label words on the support set $\tilde{S}$ as a vector representation q^v of label words with q^v's i-th element given by:

$$q_i^v = P_M([MASK] = v|c_i^p), c_i \in \tilde{S}$$

where c_i^p denotes the clause x_i combined with the template p.

For the vector representation of emotion categories, we assume that the name of each category is v_{e0}, e.g., *happiness* for the *Happiness* category. Then we use the vector representation of these names $q^{v_{e0}}$ as the vector q_e^y of these emotion categories. In doing so, the relevance of the label word and the emotion category can be expressed as the cosine similarity between the two vectors:

$$r(v_e, y_e) = \cos(q^{v_e}, q^{y_e}) = \cos(q^{v_e}, q^{v_{e0}})$$

Additionally, some labeled words have high correlations with multiple categories, which may easily lead to class confusion. For example, the label "tearful" of the category "Happiness" may also be assigned to the category "Sadness" with high probability. In order to address this issue, we designed a metric that prefers label words with high relevance to the category to which it belongs and low relevance to other categories:

$$R(v_e) = r(v_e, f(v_e)) \frac{|Y_e| - 1}{\sum_{y_e \in Y_e, y \neq f(v_e)} (r(v_e, y_e))}$$

where $f(v_e)$ is the corresponding category of v_e.

Ideally, the relevance score of the category to which a good label word belongs should be at least higher than the average relevance score of other categories. Therefore, we removed tag words with $R(v_e) < 1$.

We perform the above operations separately on the introduced additional emotion word base and the task-related database to obtain a two word base with relatively high intra-class relevance, denoted as the "BASIC" and the "INTRA" word base, respectively. The primary statistical information of the emotion word base is shown in Table 2.

It can be seen clearly from Table 2 that the "BASIC" word base is imbalanced, e.g., there are 25 label words in the "Happiness", while the "Surprise" category has only 3. This can lead to bias in the semantic understanding of the model and affect the accuracy of the model. In order to narrow the gap, we complement the "BASIC" class with the "INTRA" one to obtain a "HYBRID" emotion word base with relatively balanced categories. The "HYBRID" can ensure a certain degree of generalization with internal knowledge.

4.4. ECPE Implementation

The composite prompt function for the ECPE task is:

$$f_{ECPE}(c_i) = c_i[M]_e[M]_c[M]_r$$

where $[M]_e$, $[M]_c$, and $[M]_r$ are the sub-prompts for EE, CE, and ECPE, respectively.

For the mapping of EE, we consider that the corresponding label words of an emotion category contribute equally to the predicted label. So, we utilize the average of the predicted scores on V_e^y as the predicted score for label y_e, $y_e \in Y_e$ and the predicted label $\hat{y_e}$ given by:

$$\hat{y_e} = argmax \frac{\sum_{v_e \in V_y^e} \tilde{P}_M([MASK] = v_e|x_i^p)}{|V_y^e|}$$

[2]https://github.com/ZaneMuir/DLUT-Emotionontology

[3]http://saifmohammad.com/WebPages/lexicons.html

Dataset	Label	Label Words
Sentiment Ontology Library	happiness	康乐(recreation), 荣幸(pleasure), 欢庆(celebrate), 娱乐(entertainment), 富足(affluence), 欢乐(joy), ……
Chinese benchmark dataset	fear	慌张(panic), 急躁(impatient), 惊吓(frightened), 难堪(embarrassment), 不知所措(anxiety), ……
NRC Emotion Lex	happiness	admiration, adorable, ecstasy, memorable, perfect, remarkable, sympathetic, thankful, ……
NTCIR-13	disgust	awful, rebellion, backwards, deprivation, perfect, torture, toxic, frowning, fugitive, ……

Table 1: Examples of the expanded label words. Here, we only show the label words for happiness and fear sentiment. The Chinese sentiment label words comprise the external Sentiment Ontology and Chinese benchmark corpus. The English sentiment label words include the NRC Emotion Lex and NTCIR-13 corpus.

Word Base	Happiness	Sadness	Anger	Disgust	Surprise	Fear
BASIC	25	19	3	22	3	5
INTRA	36	46	10	27	8	15
HYBRID	25	19	13	22	10	19

Table 2: The basic statistical information of the obtained emotion word base.

The prediction of CE can be mapped from the intermediate result of $[M]_c$ directly from the indicator function, formalized as:

$$P_c(c_i) = \begin{cases} 1, & \hat{m}_c =' changed' \\ 0, & \hat{m}_c =' unchanged' \end{cases}$$

where $\hat{m}_c$ is the highest scoring output of the search function:

$$\hat{m}_c = argmax\ p(f(x^{'}, m_c); \theta)$$

where $f(x^{'}, m_c)$ is employed to fill the answer $[M_c]$ in prompt $x^{'}$ with the potential answer m_c.

From the results for $[M]_c$ and $[M]_r$, the prediction of ECPs can be obtained as follows:

$$P_{pair}(c_i) = \begin{cases} (i,j), & \hat{m}_r =' j' \& \hat{m}_c =' changed' \\ null, & others \end{cases}$$

where (i, j) indicates that clause i and clause j are ECPs in this document, with i-th clause as the emotion clause and j-th as the cause clause. Moreover, "null" indicates no clause related to the current clause.

Summing up, the extraction of ECPs mainly involves emotion extraction, cause extraction, and relationship extraction. Emotion extraction will get the sentiment category of the clause. Cause extraction will determine whether it is a cause clause by judging "I feel things changed/unchanged". And relation extraction will find out the clauses related to the current one. If the current clause has $\hat{m}_c =' changed'$ and $\hat{m}_r = j$, $j! = null$, we associate the clause with j to form an ECP.

5. Experiments

5.1. Datasets and Metrics

We conducted experiments on two benchmark datasets: one is the Chinese benchmark dataset published by Rui Xia (2019), which consists of Sina City News and the other is the English dataset NTCIR-13 Workshop Qinghong Gao (2017), composed of English novels. For both datasets, we follow the setup of (Xia and Ding, 2019). For the data split strategy, we utilize the 10-fold cross-validation method. We then evaluate our results using precision (P), recall (R), and F1-score metrics on three tasks: ECPE, EE, and CE. To minimize the effect of randomization, we execute the task 10 times to obtain the average results.

5.2. Implementation Details

We implement EmoPrompt-ECPE based on Transformers[4]. Using (1) *WoBERT* as our encoding backbone for Chinese data, which is a Chinese PLM based on word granularity and (2) *BERT-base-uncased* as the encoding backbone for English data. As for the training, we apply the adaptive moment estimation (ADAM) optimizer to optimize the loss function with the batch size and the learning rate set to 2 and $1e-5$, respectively. For the loss function setting, we use the CrossEntropy loss function. The experiments are run on the Ubuntu Operating System using an NVIDIA GeForce RTX 3090 Ti 24G GPU.

5.3. Baseline Models

We compare our proposed method in this paper with the following methods for the Chinese dataset: a 2-step framework Indep, Indep-EC, Indep-CE (Xia and Ding, 2019); End-to-end methods, RANKCP (Wei et al., 2020), ECPE-2D (Ding et al., 2020a), PairGCN (Chen et al., 2020b), ECPE-MLL (Ding et al., 2020a), RSN (Chen et al., 2022a), A^{2}Net (Chen et al., 2022b), PBJE (Liu et al., 2022), ECPE-MTL (Li et al., 2023), IA-ECPE (Huang et al., 2023). The comparison was also extended to sequence labeling methods, such as TransECPE (Fan et al., 2020), Tagging (Yuan et al., 2020), UTOS (Cheng et al., 2021), machine reading comprehension methods, such as MM-R (Zhou et al., 2022), Guided-QA (Nguyen and

[4] https://github.com/huggingface/transformers

	Main Task			Auxiliary Task					
	Emotion-Cause Pair Extraction			Emotion Clause Extraction			Cause Clause Extraction		
Approach	F1 (%)	P (%)	R (%)	F1 (%)	P (%)	R (%)	F1 (%)	P (%)	R (%)
2-step framework									
Indep	58.18	68.32	50.82	82.10	83.75	80.71	62.05	69.02	56.73
Inter-CE	59.01	69.02	51.53	83.00	84.94	81.22	61.51	68.09	56.34
Inter-EC	61.28	67.21	57.05	82.30	83.64	81.07	65.07	70.41	60.83
End-to-end Methods									
RANKCP	73.60	71.19	76.30	90.57	91.23	89.99	76.15	74.61	77.88
ECPE-2D	68.89	72.92	65.44	89.10	86.27	92.21	71.23	73.36	69.34
PairGCN	72.02	76.92	67.91	83.75	88.57	79.58	73.75	79.07	69.28
ECPE-MLL	74.52	77.00	72.35	88.86	86.08	91.91	76.30	73.82	79.12
RSN	73.93	76.01	72.19	87.55	86.14	89.22	75.45	77.27	73.98
MGSAG	75.21	77.43	73.21	87.17	92.08	82.11	77.12	79.79	74.68
A^2Net	76.34	75.03	77.80	90.80	90.67	90.98	78.35	77.62	79.20
PBJE	76.37	79.22	73.84	88.76	90.77	86.91	78.78	81.79	76.09
ECPE-MTL	75.03	75.48	75.57	90.04	90.93	89.22	77.49	77.69	77.39
IA-ECPE	64.78	69.80	60.56	83.23	85.24	81.38	67.43	72.53	63.22
Sequence labeling Methods									
TransECPE	70.72	77.08	65.32	85.88	88.79	83.15	72.33	78.74	66.89
Tagging	67.76	72.43	63.66	77.39	81.96	73.29	70.18	74.90	66.02
UTOS	72.03	73.89	70.62	85.56	88.15	83.21	74.71	76.71	73.20
Machine Reading Comprehension Methods									
Guided-QA	75.40	71.90	79.20	72.90	77.10	69.20	87.60	84.70	90.80
MM-R	80.62	82.18	79.27	93.70	97.38	90.38	81.35	83.28	79.64
Prompt-tuning based Methods									
UECA-Prompt	74.70	71.82	77.99	88.16	84.75	91.95	77.55	76.24	79.16
EmoPrompt-ECPE(BERT)	92.39	93.15	92.19	97.43	97.85	97.25	94.25	94.70	94.34
EmoPrompt-ECPE(WoBERT)	**93.71**	**94.19**	**93.65**	**97.58**	**98.01**	**97.42**	**95.13**	**95.40**	**95.37**

Table 3: The main results compare our EmoCPrompt-ECPE model with the existing Chinese benchmark dataset benchmark methods.

	Emotion-Cause Pair Extraction			Emotion Clause Extraction			Cause Clause Extraction		
Method	F1 (%)	P (%)	R (%)	F1 (%)	P (%)	R (%)	F1 (%)	P (%)	R (%)
Indep	43.67	46.94	41.02	69.40	67.41	71.60	53.01	60.39	47.34
ECPE-2D	50.73	60.49	43.84	71.89	74.35	69.68	58.55	64.91	53.53
ECPE-MLL	51.21	59.26	45.30	72.55	75.46	69.96	61.10	63.50	59.19
E2E-PExtE	50.17	51.34	49.29	69.43	71.63	67.49	52.26	66.36	43.75
IA-ECPE	50.05	60.14	43.03	71.80	73.98	69.85	58.80	63.87	54.55
EmoPrompt-ECPE	**79.20**	**79.71**	**80.02**	**85.44**	**85.74**	**85.72**	**86.09**	**86.37**	**87.03**

Table 4: The main results compare our EmoCPrompt-ECPE model with the existing English benchmark dataset benchmark methods.

Nguyen, 2023), and prompt-tuning based method, such as UECA-Prompt (Zheng et al., 2022).

For the English dataset, we compare the proposed method with Indep, E2E-PExtE (Singh et al., 2021), ECPE-2D (Ding et al., 2020a), ECPE-MLL (Ding et al., 2020a), and IA-ECPE (Huang et al., 2023).

5.4. Main Results

Results on Chinese data Table 3 shows the experimental results on the Chinese dataset. Our EmoPrompt-ECPE approach shows an obvious advantage over previous works for three tasks. The sub-prompt templates with emotional knowledge perception and the reflecting implicit knowledge of the clauses play an essential role, while the former fully explore the global information of the whole document instead of the local information of the individual clauses. Significantly, EmoPrompt-ECPE improves the F1 value on the main task of ECPE by 32.43% compared to Indep-EC. This is because the 2-step framework for EE and CE alone ignores the correlation relationship among emotion and cause, while our method achieves multi-task learning for three sub-tasks by prompt learning, which considers the relationship between tasks. Compared with end-to-end joint multitask learning methods, such as RANKCP, ECPE-MLL, and PBJE, our method improves F1 values on the ECPE main task by 20.11%, 19.19%, and 17.34%, respectively. Unlike the above-mentioned explicit modeling of the relationship between the three sub-tasks, our method models the relationship by constructing sub-prompts. Our approach is still competitive compared with TransECPE, Tagging, and UTOS, which address the ECPE task as a sequence tagging issue, or Guided-QA and MM-R methods, which

Method	Emotion-Cause Pair Extraction F1 (%)	P (%)	R (%)	Emotion Clause Extraction F1 (%)	P (%)	R (%)	Cause Clause Extraction F1 (%)	P (%)	R (%)
EmoPrompt-ECPE	93.71	94.19	93.65	97.58	98.01	97.42	95.13	95.40	95.37
w/o emotion-prompt	93.95	94.53	93.57	-	-	-	95.72	96.04	95.83
w/o cause-prompt	93.55	94.26	93.31	97.51	98.05	97.40	-	-	-
w/o relationship-prompt	-	-	-	97.77	98.29	97.49	95.23	95.61	95.38

Table 5: The experimental results of different sub-prompts for Emotion-Cause Pair Extraction task and two sub-tasks on the Chinese benchmark dataset.

Method	Emption-Cause Pair Extraction F1 (%)	P (%)	R (%)	Emotion Clause Extraction F1 (%)	P (%)	R (%)	Cause Clause Extraction F1 (%)	P (%)	R (%)
Chinese benchmark dataset									
BASIC	93.73	94.38	93.56	97.48	97.97	97.25	95.30	95.69	95.46
INTRA	93.71	94.19	93.65	97.59	98.01	97.42	95.13	95.40	95.37
HYBRID	93.72	94.25	93.63	97.58	98.08	97.33	95.18	95.33	95.50
English benchmark dataset									
BASIC	79.11	79.69	79.81	85.26	85.44	85.62	86.38	87.19	86.77
INTRA	78.93	79.51	79.70	85.43	85.69	85.76	86.01	86.41	86.87
HYBRID	79.20	79.71	80.02	85.44	85.74	85.72	86.09	86.37	87.03

Table 6: The results of different emotion label word bases for Emotion-Cause Pair Extraction and two sub-tasks on two benchmark datasets.

convert the ECPE task to a machine reading comprehension problem. While these methods are essentially classification problems, our EmoPrompt-ECPE is a mask prediction problem that can better exploit the contextual information in the text. Moreover, compared with UECA-Prompt, which is also based on prompt-tuning, our method improves F1 on the ECPE main task by 19.01%, mainly because it constructs sub-prompt templates that incorporate the emotion knowledge base and the implicit knowledge of the cause clause. Similar results were obtained on the EE and CE sub-tasks.

Results on English data We also evaluate our model on the English dataset, and the results are summarized in Table 4. Our method generally outperformed the fine-tuning-based works such as ECPE-MLL, and IA-ECPE on the ECPE task, improving the F1 values by 27.99%, and 29.15%, respectively. It further demonstrates that (1) the mask prediction model based on prompt learning can better access the contextual information of the text, (2) the fused additional emotion knowledge base can extract the label word information which is ignored by the fine-tuning paradigm, and (3), the utilization of latent semantic knowledge can help the model to better understand the task.

5.5. Ablation Experiments

We conducted ablation experiments to analyze the impact of different components on performance.

Sub-prompt template To explore the extent to which different sub-prompt templates affect the performance, we set up ablation experiments with no emotion-prompt, cause-prompt, or relationship-prompt. The results are shown in Table 5. Obviously, the absence of any of the sub-templates makes the performance decrease. It indicates that each sub-template has a different role in the feature learning process, demonstrating the efficacy of the three sub-prompts. More specifically, the model's performance (F1) decreases more significantly without the cause-template, proving that the cause-prompt dominates in the ECPE task. The lack of the emotion-prompt causes a decrease in the R of the ECPE task. In addition, the emotion-prompt can obtain the emotion category.

Emotion Word Base To verify the impact of emotion word base on performance, we train models using three-word bases, BASIC, INTRA, and HYBRID with the results as shown in Table 6. We found that for EE task, the HYBRID word base offers the best results. Through the analysis of Table 2, we observe that in HYBRID, the number of documents for each emotion category is relatively balanced. Based on this finding, we believe that HYBRID achieves optimal performance because it has a relatively balanced sentiment lexicon, which contributes to a more comprehensive coverage of the various sentiment categories, thus improving the accuracy and effectiveness of the task.

WoBERT VS BERT To determine the impact of PLM on performance, we trained EmoPrompt-ECPE using WoBERT and BERT, respectively. As shown in Table 3, we find that the method using WoBERT performs much better in the Emotion-Cause Pair Extraction (ECPE) task, with a 1.32% improvement in F1, 1.04% improvement in P, and 1.46% improvement in R compared to the method using BERT. This significant performance improvement can be attributed to WoBERT's Chinese-specific design and word-based pre-training method. WoBERT is more adaptive to the Chinese context than BERT and can process Chinese data more accurately. The word-based pre-training model structure makes it more capable of capturing the characteristics of the Chinese language, achieving a higher performance.

6. Case Analysis

To further validate the proposed EmoPrompt-ECPE's effectiveness, we conducted several case analyses to monitor the performance of the EmoPrompt-ECPE in the ECPE task.

Case One: *[c1]She was in poor health, [c2]and the baby had to be taken care of. [c3]When Zhang learned that his wife was going to donate a kidney to him, [c4]he was touched, [c5]but was also very worried. [...]*. Ground truths:{(c4, c3), (c5, c3)}. Predictions: {(c4, c3), (c5, c3)}

In Case One, the stimulus of multiple emotion clauses is an identical cause clause, but these emotion clauses express different emotion polarities. Our approach handles this situation easily and very well.

Case Two: *[...], [c9]with the help of her adoptive father's family, [c10]Wang led a happy life.[...].[c13]After Wang was abducted, [c14]her mother regretted it so much. [...]. [c16]But there was no news, [c17]her mother fell into endless miss for her daughter.*. Ground truths: {(c10, c9), (c14, c13), (c17, c16)}. Predictions: {(c10, c9), (c14, c13), (c17, c16)}.

This situation contains intricate multiple ECPs, and our approach still handles them well.

7. Conclusion

In this paper, we propose Emotion knowledge-aware Prompt-tuning for Emotion-Cause Pair Extraction, which models the task by transforming the three sub-task goals into sub-prompts. Among them, the emotion-prompt extends the verbalizer in prompt-tuning using an external emotion knowledge base, which realizes a many-to-one mapping of label words to emotion categories. The cause-prompt improves the performance of CE and ECPE tasks by determining the latent semantics of the cause clause. In future work, we will focus on solving the problem of unbalanced emotion categories in the data as well as analyzing cases where the system fails by incorporation more task specific data and knowledge.

8. Acknowledgements

This work was supported by the National Natural Science Foundation of China (NSFC) 581 under Grant 62077027 and 72091315 of 72091310, the Department of Science and Technology of Jilin 582 Province, China under Grant 20230201086GX and Industry University Research Innovation Fund of 583 the Ministry of Education under Grant 2022XF017.

9. Bibliographical References

Ting-Wei Chang, Yao-Chung Fan, and Arbee L. P. Chen. 2022. Emotion-cause pair extraction based on machine reading comprehension model. *Multim. Tools Appl.*, 81(28):40653–40673.

Fang Chen, Ziwei Shi, Zhongliang Yang, and Yongfeng Huang. 2022a. Recurrent synchronization network for emotion-cause pair extraction. *Knowledge Based System*, 238:107965.

Shunjie Chen, Xiaochuan Shi, Jingye Li, Shengqiong Wu, Hao Fei, Fei Li, and Donghong Ji. 2022b. Joint alignment of multi-task feature and label spaces for emotion cause pair extraction. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 6955–6965.

Siyuan Chen and Kezhi Mao. 2023. A graph attention network utilizing multi-granular information for emotion-cause pair extraction. *Neurocomputing*, 543:126252.

Xinhong Chen, Qing Li, and Jianping Wang. 2020a. A unified sequence labeling model for emotion cause pair extraction. In *Proceedings of the 28th International Conference on Computational Linguistics*, pages 208–218.

Ying Chen, Wenjun Hou, Shoushan Li, Caicong Wu, and Xiaoqiang Zhang. 2020b. End-to-end emotion-cause pair extraction with graph convolutional network. In *Proceedings of the 28th International Conference on Computational Linguistics*, pages 198–207.

Xiyao Cheng, Ying Chen, Bixiao Cheng, Shoushan Li, and Guodong Zhou. 2017. An emotion cause corpus for chinese microblogs with multiple-user structures. *ACM Trans. Asian Low Resour. Lang. Inf. Process.*, 17(1):6:1–6:19.

Zifeng Cheng, Zhiwei Jiang, Yafeng Yin, Na Li, and Qing Gu. 2021. A unified target-oriented sequence-to-sequence model for emotion-cause pair extraction. *IEEE ACM Trans. Audio Speech Lang. Process.*, 29:2779–2791.

Zifeng Cheng, Zhiwei Jiang, Yafeng Yin, Hua Yu, and Qing Gu. 2020. A symmetric local search network for emotion-cause pair extraction. In *Proceedings of the 28th International Conference on Computational Linguistics*, pages 139–149.

Zixiang Ding, Rui Xia, and Jianfei Yu. 2020a. ECPE-2D: emotion-cause pair extraction based on joint two-dimensional representation, interaction and prediction. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 3161–3170.

Zixiang Ding, Rui Xia, and Jianfei Yu. 2020b. End-to-end emotion-cause pair extraction based on sliding window multi-label learning. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, pages 3574–3583.

Chuang Fan, Chaofa Yuan, Jiachen Du, Lin Gui, Min Yang, and Ruifeng Xu. 2020. Transition-based directed graph construction for emotion-cause pair extraction. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 3707–3717.

Chuang Fan, Chaofa Yuan, Lin Gui, Yue Zhang, and Ruifeng Xu. 2021. Multi-task sequence tagging for emotion-cause pair extraction via tag distribution refinement. *IEEE ACM Trans. Audio Speech Lang. Process.*, 29:2339–2350.

Qinghong Gao, Jiannan Hu, Ruifeng Xu, Lin Gui, Yulan He, Kam-Fai Wong, and Qin Lu. 2017. Overview of NTCIR-13 ECA task. In *The 13th NTCIR Conference*. National Institute of Informatics (NII).

Lin Gui, Dongyin Wu, Ruifeng Xu, Qin Lu, and Yu Zhou. 2016a. Event-driven emotion cause extraction with corpus construction. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, EMNLP 2016, Austin, Texas, USA, November 1-4, 2016*, pages 1639–1649. The Association for Computational Linguistics.

Lin Gui, Ruifeng Xu, Qin Lu, Dongyin Wu, and Yu Zhou. 2016b. Emotion cause extraction, A challenging task with corpus construction. In *Social Media Processing - 5th National Conference, SMP 2016, Nanchang, China, October 29-30, 2016, Proceedings*, volume 669 of *Communications in Computer and Information Science*, pages 98–109.

Guimin Hu, Yi Zhao, Guangming Lu, Fanghao Yin, and Jiashan Chen. 2022. An exploration of mutual information based on emotion-cause pair extraction. *Knowledge Based System*, 256:109822.

Weichun Huang, Yixue Yang, Xiaohui Huang, Zhiying Peng, and Liyan Xiong. 2023. Emotion-cause pair extraction based on interactive attention. *Appl. Intell.*, 53(9):10548–10558.

Chenbing Li, Jie Hu, Tianrui Li, Shengdong Du, and Fei Teng. 2023. An effective multi-task learning model for end-to-end emotion-cause pair extraction. *Appl. Intell.*, 53(3):3519–3529.

Junlong Liu, Xichen Shang, and Qianli Ma. 2022. Pair-based joint encoding with relational graph convolutional networks for emotion-cause pair extraction. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 5339–5351.

Huu-Hiep Nguyen and Minh-Tien Nguyen. 2023. Emotion-cause pair extraction as question answering. In *Proceedings of the 15th International Conference on Agents and Artificial Intelligence*, pages 988–995.

Aaditya Singh, Shreeshail Hingane, Saim Wani, and Ashutosh Modi. 2021. An end-to-end network for emotion-cause pair extraction. In *Proceedings of the Eleventh Workshop on Computational Approaches to Subjectivity*, pages 84–91.

Haolin Song, Chen Zhang, Qiuchi Li, and Dawei Song. 2020. End-to-end emotion-cause pair extraction via learning to link. *CoRR*, abs/2002.10710.

Qixuan Sun, Yaqi Yin, and Hong Yu. 2021. A dual-questioning attention network for emotion-cause pair extraction with context awareness. In *International Joint Conference on Neural Networks*, pages 1–8.

Penghui Wei, Jiahao Zhao, and Wenji Mao. 2020. Effective inter-clause modeling for end-to-end emotion-cause pair extraction. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 3171–3181.

Sixing Wu, Fang Chen, Fangzhao Wu, Yongfeng Huang, and Xing Li. 2020. A multi-task learning neural network for emotion-cause pair extraction. In *ECAI 2020 - 24th European Conference on Artificial Intelligence*, pages 2212–2219.

Rui Xia and Zixiang Ding. 2019. Emotion-cause pair extraction: A new task to emotion analysis in texts. In *Proceedings of the 57th Conference of the Association for Computational Linguistics*, pages 1003–1012.

Ruifeng Xu, Jiannan Hu, Qin Lu, Dongyin Wu, and Lin Gui. 2017. An ensemble approach for emotion cause detection with event extraction and multi-kernel svms. *Tsinghua Science and Technology*, 22(6):646–659.

Chaofa Yuan, Chuang Fan, Jianzhu Bao, and Ruifeng Xu. 2020. Emotion-cause pair extraction as sequence labeling based on A novel tagging scheme. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, pages 3568–3573.

Xiaopeng Zheng, Zhiyue Liu, Zizhen Zhang, Zhaoyang Wang, and Jiahai Wang. 2022. Uecaprompt: Universal prompt for emotion cause analysis. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 7031–7041. International Committee on Computational Linguistics.

Changzhi Zhou, Dandan Song, Jing Xu, and Zhijing Wu. 2022. A multi-turn machine reading comprehension framework with rethink mechanism for emotion-cause pair extraction. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 6726–6735.

10. Language Resource References

Qinghong Gao, Jiannan Hu, Ruifeng Xu, Yulan He, Kam-Fai Wong, Qin Lu. 2017. *Overview of NTCIR-13 ECA Task*. In Porceedings NTCIR-13. 2017, ISLRN I do not know. PID I do not know. Note: You don't really need both an ISLRN and another PID, but it can't hurt.

Rui Xia, Zixiang Ding. 2019. *Emotion-Cause Pair Extraction: A New Task to Emotion Analysis in Texts*. Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers, ISLRN I do not know. PID I do not know. Note: You don't really need both an ISLRN and another PID, but it can't hurt.