Transition-based Opinion Generation for Aspect-based Sentiment Analysis

Anonymous ACL submission

Abstract

Recently, the use of pre-trained generation models for extracting sentiment elements has resulted in significant advancements in aspect-based sentiment analysis benchmarks. However, these approaches often overlook the importance of explicitly modeling structure among sentiment elements. To address this limitation, we present a study that aims to integrate general pre-trained sequence-to-sequence language models with a structure-aware transition-based approach. Therefore, we depart from a transition system for opinion tree generation, designed to better exploit pre-trained language models for structured fine-tuning. Extensive experiments show that our model significantly advances the state-of-the-art performance on several benchmark datasets. In addition, the empirical studies also indicate that the proposed opinion tree generation with transition system is more effective in capturing the sentiment structure than other generation models.

1 Introduction

Aspect-based sentiment analysis (ABSA) has been garnering increasing interest within the community. This area encompasses four key subtasks: aspect term extraction, opinion term extraction, aspect term category classification, and aspect-level sentiment classification. The initial two subtasks focus on extracting aspect terms and opinion terms from within a given sentence. The subsequent two subtasks aim to identify the category of the extracted aspect term and determine its sentiment polarity. Through these subtasks, ABSA offers a comprehensive approach to analyzing sentiment at a more granular, aspect-based level.

Previously, most ABSA tasks were formulated as either sequence-level (Qiu et al., 2011; Peng et al., 2020; Cai et al., 2021) or token-level classification problems (Tang et al., 2016b). However, these approaches often encountered significant challenges due to error propagation, as the overall prediction performance was heavily dependent on the accuracy of each individual step (Peng et al., 2020). As a result, recent studies have shifted towards tackling the ABSA problem with a unified generative approach (Yan et al., 2021; Mao et al., 2022; Hu et al., 2022b; Bao et al., 2023; Zhou et al., 2023). This new approach offers a promising direction for ABSA research, as it aims to mitigate the issues caused by error propagation in traditional methods.

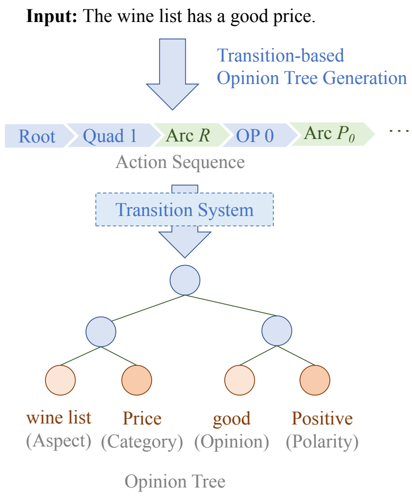

Figure 1: An example of transition-based opinion tree generation.

Despite their promise, these unified generative approaches for ABSA suffer from certain limitations. A notable limitation is lack of a structural guarantee for sentiment elements. This absence means that the model may generate string outputs that do not conform to a valid opinion tree structure, necessitating additional post-processing steps to establish the necessary relationships between sentiment elements.

In this study, we introduce a novel approach named the *transition-based opinion generation model* to address above challenges. This model is designed to fully harness the generative power of pre-trained language models while simultaneously capturing the explicit structure of sentiment elements. As shown in Figure 1, we design a transition system composed of a concise set of fundamental actions. This transition system serves as the backbone of our model, enabling it to generate structurally sound outputs. Utilizing this transition system, we develop a neural transition-based opinion tree generation model. This model takes a review sentence as input and is tasked with generating an action sequence that adheres to the predefined transition system. Once this action sequence is generated, it can be seamlessly utilized to reconstruct the opinion tree and sentiment elements, ensuring structural integrity and alignment with the original sentiment structure.

The detailed evaluation shows that our model significantly advances the state-of-the-art performance on several benchmark datasets. In addition, the empirical studies also indicate that the proposed transition-based opinion tree generation is more effective in capturing the sentiment structure than generative models.

2 Related Work

Aspect-based sentiment analysis (ABSA) has drawn wide attention during the last decade. Early studies focus on the prediction of a single element, such as extracting the aspect term (Qiu et al., 2011), detecting the mentioned aspect category (Bu et al., 2021), and predicting the sentiment polarity for a given aspect (Tang et al., 2016a; Chen et al., 2022; Cao et al., 2022).

Some works further consider the joint detection of two sentiment elements, including the pairwise extraction of aspect and opinion term (Xu et al., 2020; Li et al., 2022); the prediction of aspect term and its corresponding sentiment polarity (Zhang and Qian, 2020); and the co-extraction of aspect category and sentiment polarity (Cai et al., 2020). Recently, aspect sentiment triplet and quadruple prediction tasks are proposed in ABSA, they employ end-to-end models to predict the sentiment elements in triplet or quadruple format (Peng et al., 2020; Wan et al., 2020; Cai et al., 2021; Zhang et al., 2021a; Bao et al., 2022; Zhou et al., 2023; Bao et al., 2023).

More recently, there are some attempts on tackling ABSA problem in a sequence-to-sequence manner (Zhang et al., 2021a), either treating the class index (Yan et al., 2021) or the desired sentiment element sequence (Zhang et al., 2021b) as the target of the generation model. For example, Yan et al. (2021) treated the ABSA as a text generation problem, and employ a sequence-to-sequence pre-trained model to generate the sequence of aspect terms and opinion words directly. Zhang et al. (2021a) proposed a paraphrase model that utilized the knowledge of the pre-trained model via casting the original task to a paraphrase generation process. They employed the paraphrase to represent aspect-based quads. Bao et al. (2022) employed a generation model to generate all the sentiment elements as a tree structure. Zhou et al. (2023) simultaneously detected aspect categories and co-extract aspect-opinion-sentiment triplets, can absorb deeper interactions between sentiment elements without error propagation.

Our study differs from previous research in that we integrate pre-trained sequence-to-sequence language models with a transition-based approach for opinion tree parsing. This integration allows us to explore the complementarity between these two powerful techniques and assess their combined potential for enhancing sentiment analysis.

3 Preliminaries

As shown in Figure 2, our proposed approach involves several key steps. Firstly, we introduce a transition system that serves to normalize sentiment elements into an opinion tree structure. Next, we employ a neural transition-based opinion tree generation model to generate an action sequence from a given review text. Following the generation of the action sequence, we construct the opinion tree based on this sequence and the transition system. Finally, since all sentiment elements are normalized into the opinion tree, it becomes straightforward to recover them from the tree.

In this section, we give the definition of the aspect-based sentiment analysis task and the construction process of opinion tree. The transition system and the transition-based generation model will be discussed in the next two sections.

3.1 Task Definition

Given a review sentence $x = \{x_1, x_2, ..., x_n\}$, the ABSA task aims to predict all aspect-level sentiment quadruples (a, c, o, s), which corresponds to the aspect term, aspect category, opinion term, and

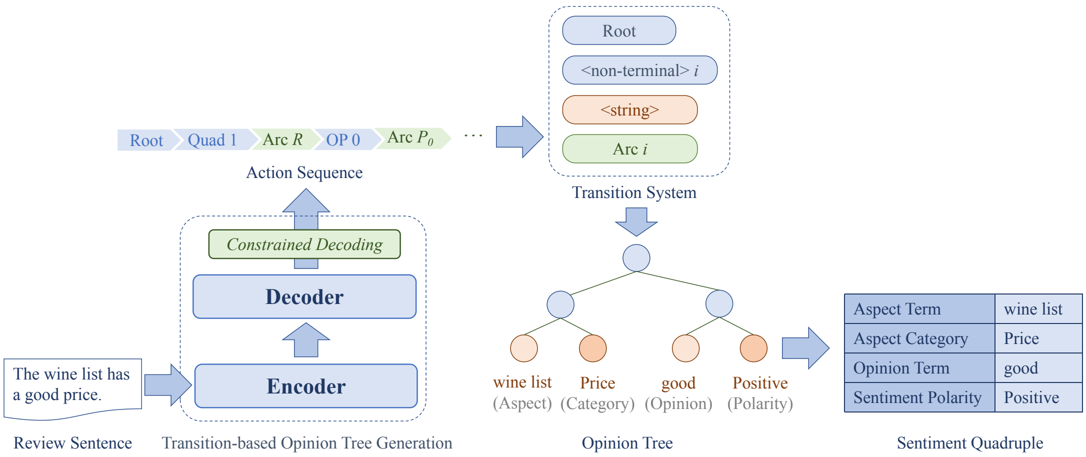

Figure 2: An overview of the proposed model.

Type	Name	Description
Non -terminal	ROOT	root of opinion tree.
	QUAD	sentiment quadruple.
	AP	virtual node of aspect.
	OP	virtual node of opinion.
	AT	aspect term.
	CA	category of aspect.
	OT	opinion term.
	SP	sentiment polarity.
Terminal	Aspect	e.g., wine list, snacks
	Category	e.g., food quality
	Opinion	e.g., good, delicious
	Polarity	e.g., positive, negative

Table 1: The notation of all symbols.

sentiment polarity, respectively. The *aspect category* c belongs to a category set $\mathcal{C}$; the *aspect term* a and the *opinion term* o are typically text spans in x while they can be null if the target is not explicitly mentioned. The *sentiment polarity* s is one of the sentiment classes $\mathcal{S}$, which corresponds to the positive, neutral, and negative sentiment, respectively.

3.2 Opinion Tree Structure

As shown in Figure 2, we convert all aspect-level sentiment quadruples into an opinion tree (Bao et al., 2023). The opinion tree explicitly delineates intricate connections among vital sentiment components (i.e., aspect term and aspect category). This deliberate structuring aims to unveil a more comprehensive and intricate aspect-level semantic framework, enhancing the efficiency of sentiment element extraction.

To standardize the structure of the opinion tree, we introduce a formal representation (N, Σ, P), comprising finite, disjoint sets of non-terminal symboles N, terminal symbols Σ, and a set of conditional rules denoted as P. The notation for all symbols is detailed in Table 1, where each recurrent non-terminal symbol is accompanied by a numerical label indicating its current occurrence count. Notably, we position the category and polarity elements to the right of the aspect and opinion terms, respectively. This structured organization facilitates the generation of category and polarity nodes. Additionally, we introduce a virtual node, labeled $NULL$ representing implicit aspect or opinion items. Unrelated nodes are systematically omitted, enhancing the conciseness and intuitiveness of the opinion tree. Each rule in P follows the form $A \rightarrow \alpha$, where $A \in N$, $\alpha \in N \cup \Sigma$. For example, $OP \rightarrow OT|SP$ signifies that arcs originating from virtual node OP exclusively connect to either opinion term OT or sentiment polarity SP in the opinion tree.

4 Transition System

In this section, we present a transition system specifically designed for opinion tree generation. Figure 3 shows a parsed example of a transition action sequence and the graph structure from opinion tree.

Diverging from previous transition-based approaches, our transition system operates without employing traditional data structures like stacks or buffers. In particular, receiving a source sentence $x = \{x_1, x_2, ..., x_n\}$, our transition system undertakes a left-to-right scan of the sentence using a cursor c_t, where $t \in \{1, 2, ..., n\}$. Transition actions involve either shifting the cursor by one token forward or generating multiple nodes and edges while the cursor points to the same token. The transition process concludes when the final word in the sentence is shifted, signifying the completion of parsing processing. Available actions for our transition system are as follows:

- **ROOT** creates the root node of opinion tree.
- **<non-terminal>-<i>** creates a non-terminal node of name <non-terminal> labeled i. (i.e. QUAD-0). A non-terminal node is one of the high-level notation set.
- **<string>** creates a terminal or a non-terminal node of name <string>. A terminal node could be the word under current cursor c_t as a part of aspect or opinion term, an aspect category c or a sentiment polarity s.
- **ARC<i>-<j>** creates an arc from last generated node to the corresponding non-terminal node in layer i labeled j. Note that we can only point to past node generating actions in the action history.

Within a transition system, nodes are exclusively generated through <non-terminal> and <string> actions. This offers an opportunity to leverage the pre-trained vocabulary on the target side of the generation model, thereby maximizing the utilization of linguistic knowledge acquired during pre-training. Furthermore, the use of a cursor variable in the transition system disentangles node referencing from source tokens, enabling the generation of multiple nodes and edges under the same token, even constructing the entire opinion tree structure if necessary. This imparts more expressiveness and flexibility to opinion tree generation, particularly when aspect term or opinion term is implicit.

In summary, our proposed transition system ensures the structural integrity of the generated opinion tree. By leveraging pre-trained generation models and simplifying the transition set, we are able to maximize the efficiency and accuracy of opinion tree generation. This innovative approach paves the way for more effective sentiment analysis.

5 Transition-based Opinion Tree Generation

In this section, we employ a sequence-to-sequence model to generate the *action sequence* of the transition system via a transformer-based encoder-decoder architecture. As discussed in the above

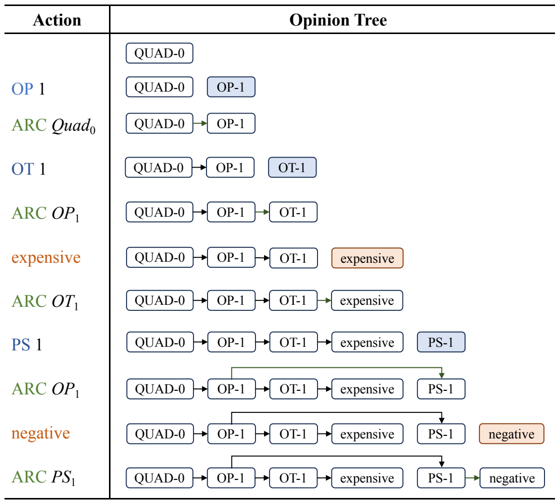

Figure 3: An example of generating the *opinion pair* (OP) sub-tree with proposed transition system.

sections, the opinion tree can be generated based on the action sequence and the transition system.

5.1 Encoder

Given the token sequence $x = \{x_1, x_2.., x_n\}$ as input, the sequence-to-sequence model outputs the target action sequence $y = \{y_1, y_2.., y_m\}$. To this end, the sequence-to-sequence model first computes the hidden vector representation $H = \{h_1, h_2.., h_n\}$ of the input via a multi-layer transformer encoder:

$$H = \text{Encoder}(\{x_1, ..., x_n\}) \quad (1)$$

where each layer of Encoder is a transformer block with the multi-head attention mechanism.

5.2 Decoder

After the input token sequence is encoded, the decoder predicts the output action sequence token-by-token with the sequential input tokens' hidden vectors. At the i-th step of generation, the self-attention decoder predicts the i-th token y_i in the linearized form, and the decoder state h_i^d as:

$$t_i, \hat{z}_i^l = \text{Decoder}([Z; \hat{z}_1^l, ..., \hat{z}_{i-1}^l], t_{i-1}) \quad (2)$$

where each layer of Decoder is a transformer block that contains self-attention with decoder state h_i^d and cross-attention with encoder state H. The generated output structured sequence starts from the start token "⟨bos⟩" and ends with the end token "⟨eos⟩". The conditional probability of the whole output sequence $p(T|X)$ is progressively combined by the probability of each step

$p(t_i|t_{<i}, X)$:

$$p(T|X) = \prod_{i}^{m} p(t_i|t_{<i}, X) \tag{3}$$

where $t_{<i} = \{t_1...t_{i-1}\}$, and $p(t_i|t_{<i}, X)$ is the probability over target vocabulary V normalized by $\mathrm{softmax}$.

5.3 Constrained Decoding

In this study, we leverage a constrained decoding method (Chen et al., 2020; Cao et al., 2021) to guide the generation of action sequences.

Specifically, instead of exhaustively searching the entire vocabulary, our constrained decoding method dynamically selects and prunes a candidate vocabulary $V_t \in V$ based on the current generated state, where V represents the set of all possible actions in our transition system. The valid actions at each generated step are defined by the following rules: 1) Generating a token aligned with the cursor's word to serve as the node in the opinion tree, and 2) Generating a valid arc immediately after generating a node.

As a result, the constrained rules of the transition system are injected as prompts into the decoder, ensuring the generation of a valid action sequence during decoding.

5.4 Objective Functions and Training

In this subsection, we show the objective function and training process of the proposed model.

The goal is to maximize the target action sequence T probability given the review text X. Therefore, we optimize the negative log-likelihood loss function:

$$\mathcal{L} = -\frac{1}{|\tau|} \sum_{(X,T)\in\tau} \log p(T|X;\theta) \tag{4}$$

where θ is the model parameters, and (X, T) is a (*input,output*) pair in training set τ, then

$$\begin{aligned} &\log p(T|X;\theta) = \\ &= \sum_{i=1}^{m} \log p(t_i|t_1, t_2, ...t_{i-1}, X; \theta) \end{aligned} \tag{5}$$

where $p(t_i|t_1, t_2, ..., t_{i-1}, X; \theta)$ is calculated by the decoder.

Domain	Train	Dev.	Test
Restaurant	1,529	171	582
Laptop	2,929	326	816
Phone	4,986	1,068	1,061

Table 2: Distribution of three domains.

6 Experiments

In this section, we introduce the datasets used for evaluation and the baseline methods employed for comparison. We then report the experimental results conducted from different perspectives.

6.1 Setting

In this study, we use restaurant and laptop domains in ACOS dataset (Cai et al., 2021) and phone domain in Zhou et al. (2023)'s dataset for our experiments. There are 2,286 sentences in the restaurant domain, 4,076 sentences in the laptop domain and 7,115 sentences in the phone domain. The distribution of these three domains can be found in Table 2.

We tune the parameters of our models by grid searching on the validation dataset. We employ T5-large[1] (Raffel et al., 2020) and fine-tune its parameters for our proposed model, and the parameters are optimized by AdamW with a learning rate of 5e-5. The batch size is 16 with a maximum 256 token length. Our experiments are carried out with a Nvidia RTX 3090 GPU. The experimental results are obtained by averaging five runs with different random seeds.

In evaluation, a quadruple is viewed as correct if and only if the four elements, as well as their combination, are exactly the same as those in the gold quadruple. On this basis, we calculate the Precision and Recall, and use F1 score as the final evaluation metric for aspect sentiment quadruple extraction (Cai et al., 2021; Zhang et al., 2021a).

6.2 Main Results

As shown in Table 3, We compare the proposed model with various strong baselines, where,

- **JET** (Xu et al., 2020) is an end-to-end framework which combines the identification of aspects, their corresponding opinions, and their sentiment polarities with a position-aware tagging scheme.

[1]`https://huggingface.co/t5-large`

Method	Restaurant			Laptop			Phone		
	P.	R.	F1.	P.	R.	F1.	P.	R.	F1.
JET	0.5981	0.2894	0.3901	0.4452	0.1625	0.2381	0.3845	0.2213	0.2809
TasBERT	0.2629	0.4629	0.3353	0.4715	0.1922	0.2731	0.3453	0.2207	0.2693
EClassify	0.3854	0.5296	0.4461	0.4556	0.2948	0.3580	0.3128	0.3323	0.3223
GAS	0.6127	0.5860	0.5959	0.4089	0.4219	0.4153	0.5072	0.4815	0.4940
DLO	0.5904	0.6029	0.5966	0.4359	0.4367	0.4363	0.5451	0.5173	0.5308
ILO	0.6071	0.6128	0.6099	0.4359	0.4297	0.4319	0.5307	0.5185	0.5245
Seq2Path	0.6029	0.5961	0.5995	0.4251	0.4317	0.4284	0.5263	0.4994	0.5125
OneASQP	**0.6591**	0.5624	0.6069	0.4380	0.3954	0.4156	0.5742	0.5096	0.5400
Ours	0.6432	**0.6248**	**0.6338**	**0.4532**	**0.4457**	**0.4494**	0.5441	**0.5607**	**0.5523**

Table 3: Comparison with baselines.

- **TasBERT** (Wan et al., 2020) integrates aspect category-based sentiment classification and aspect extraction in a unified framework by attaching the aspect category and the sentiment polarity to the review sentence and using it as the input of BERT.

- **EClassify** (Cai et al., 2021) firstly performs aspect-opinion co-extraction, and then predicts category-sentiment given the extracted aspect-opinion pairs.

- **GAS** (Zhang et al., 2021b) tackles all ABSA tasks in a unified generative framework, and formulates ABSA task as a sentiment element sequence generation problem.

- **DLO** and **ILO** (Hu et al., 2022b) first uses the pre-trained language model to select the template orders with minimal entropy, then fine-tunes generation model with these template orders to generate aspect-level sentiment quadruples.

- **Seq2Path** (Mao et al., 2022) generates sentiment tuples as paths of a tree, and calculate the average loss of over paths for training and inference.

- **OneASQP** (Zhou et al., 2023) simultaneously detect aspect categories and co-extract aspect-opinion-sentiment triplets, can absorb deeper interactions between sentiment elements without error propagation.

The results clearly demonstrate that pre-trained generation models, such as GAS, DLO, and OneASQP, consistently outperform pipeline-based methods used in previous research. This disparity highlights the inherent issues with pipeline-based approaches, which are prone to error propagation. Conversely, it underscores the effectiveness of a unified generation architecture in capturing the rich semantics of natural language labels by directly encoding them into the target output.

Method	Restaurant	Laptop	Phone
Sequence	0.5991	0.4153	0.4940
Paraphrase	0.6042	0.4197	0.4823
Tree	0.6122	0.4288	0.5307
Actions (Ours)	0.6338	0.4494	0.5523

Table 4: Results of different sentiment elements generation paradigms.

In comparison to earlier research, our proposed model demonstrates significant enhancement ($p < 0.05$) in all experimental settings, surpassing all preceding studies. This notable superiority underscores the preeminence of the opinion tree structure in comparison to other generation-based techniques. Additionally, the findings indicate that our transition system adeptly parses the opinion tree from the input sentence, while maintaining the integrity of sentiment structural information. These compelling results establish that our proposed model can ensure the structural well-formedness of the opinion tree.

6.3 Impact of Generation Paradigms

We analyze the impact of different sentiment element generation paradigms in Table 4, where *Sequence* (Zhang et al., 2021b) directly treat the quadruple sequence as the target for learning the generation model; *Paraphrase* (Zhang et al.,

Method	Restaurant	Laptop	Phone
GAS	0.5959	0.4217	0.4940
Chart	0.6271	0.4120	0.5347
Stack	0.5841	0.3745	0.4856
Seq2Seq	0.6014	0.3981	0.4931
Ours	0.6338	0.4532	0.5523

Table 5: Results of different opinion tree parsers.

	OTG	OTP	Ours
AC	0.7680	0.7711	0.7763
AS	0.7677	0.7643	0.7874
OS	0.7663	0.7608	0.7721
ACS	0.6642	0.6752	0.6945
AOS	0.6605	0.6713	0.6836
ACOS	0.6164	0.6271	0.6338

Table 6: Results of sentiment elements combinations on *Restaurant* domain.

2021a) proposes a paraphrase modeling paradigm to cast the ABSA task to a paraphrase generation process, and joint extract all the sentiment elements; *Tree* (Bao et al., 2022) directly generates linearized opinion tree structures using generative models.

The results reveal that more complex structures yield better performance. For instance, *Paraphrase* outperforms *Sequence*, and *Tree* surpasses the other two baselines. Furthermore, our proposed action sequence generation paradigm (i.e., *Actions*) with transition system exhibits significant improvement ($p < 0.05$) compared to all baselines. This suggests that the action sequence with transition system is highly effective in capturing the explicit structure among sentiment elements. The results also indicate that the proposed transition-based opinion tree generation model provides a more nuanced understanding of sentiment elements and their relationship.

6.4 Impact of Opinion Tree Parsers

We then employ three representative mainstream parsers to evaluate the effective of them on opinion tree parsing and aspect-based sentiment analysis. Among them, the *Chart*-based parser (Bao et al., 2023) independently scores each span and conducts a global search across all possible trees to find the highest-scoring opinion tree; the *Stack*-based parser (Zhang et al., 2019) constructs a complex output structure holistically through a state-transition process with incremental output-building actions, relying on the implementation of a data structure stack; the *Seq2Seq* (Yang and Tu, 2022) parser employs a pointing mechanism for bottom-up parsing and use sequence-to-sequence backbone. The latter two parsers are designed for syntax parsing or information extraction tasks, we adopt them for aspect-based sentiment analysis.

As shown in Table 5, both the chat-based parser and the seq2seq parser surpass the basic GAS model in performance. This suggests that these opinion tree parsers are indeed effective in capturing the interdependencies among sentiment elements. However, the stack-based parser fails to yield satisfactory results. This could be attributed to the inherent complexity of traditional transition systems, which might hinder their ability to accurately model the opinion tree structure. Furthermore, our proposed model consistently outperforms all other baselines. This underscores the efficacy of both our proposed transition system and transition-based opinion tree generation model in capturing the intricate relationships among sentiment elements.

6.5 Influence of Sentiment Elements Combinations

We further investigate the capabilities of our proposed transition-based opinion tree generation model when dealing with different combinations of sentiment elements. In this context, *A* represents the aspect term, *C* denotes the category of the aspect term, *O* stands for the opinion term, and *S* signifies the sentiment polarity towards the aspect term. Each row corresponds to a specific combination. For instance, *ACS* indicates that the model should jointly generate the aspect term, aspect category, and sentiment polarity.

From the results on Table 6, we observe that as the complexity of the sentiment element combinations increases, the performance of the proposed model tends to decrease. However, it is noteworthy that our proposed transition-based opinion tree generation model consistently outperforms OTG (Bao et al., 2022) and OTP (Bao et al., 2023) in all combinations. This underscores the versatility and generality of our proposed model, indicating its applicability to various sentiment analysis tasks.

We then analyze the completeness of the tree structure generated by OTG, OTP and the pro-

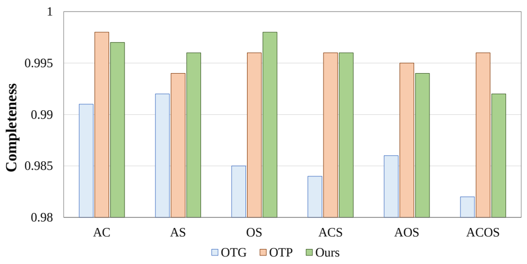

Figure 4: Tree structure completeness of different methods.

Models	Restaurant	Laptop	Phone
GAS	0.5959	0.4217	0.4940
BART-base	0.6097	0.4285	0.5310
BART-large	0.6241	0.4337	0.5455
T5-base	0.6217	0.4352	0.5439
T5-large	**0.6338**	**0.4532**	**0.5523**
LLaMA-7b	0.5827	0.3917	0.5214

Table 7: Results of different pre-trained language models.

posed model with different element combination settings. The completeness is calculated through the valid rate of a tree structure (Bao et al., 2023). As shown in Figure 4, the completeness of the proposed model is higher than OTG in all the schemas. This findings demonstrate that our model ensures structural well-formedness across all scenarios, regardless of the specific sentiment element combinations. Interestingly, our model's completeness is on par with that of OTP, which relies on an original chart-based parser. This comparison further emphasizes the robustness and adaptability of our approach, indicating its effectiveness in tackling a wide range of sentiment analysis challenges.

6.6 Influence of Pre-trained Language Models

We conducted an analysis to assess the impact of different pre-trained language models on the performance of our proposed transition-based opinion tree generation model. Specifically, we utilized the encoder-decoder style models BART (Lewis et al., 2020) and T5 (Raffel et al., 2020), as well as the decoder-only style large language model LLaMA-7b (Touvron et al., 2023), which was fine-tuned using Lora approach (Hu et al., 2022a). All these models were fine-tuned in the same GPU environment to ensure a consistent evaluation.

Upon analysis, we observed a general trend that models with more parameters tend to achieve better performance. In addition, most of them outperform the basic GAS model with T5-base pre-trained model, underscoring the robustness and effectiveness of our proposed transition-based architecture. These findings suggest that the integration of pre-trained language models with our transition-based approach not only enhances performance but also demonstrates adaptability and versatility, regardless of the specific pre-trained model used.

However, an interesting deviation was noted with LLaMA, which performed relatively poorer compared to other models. We attribute this disparity to the inherent challenges large language models face when generating action sequences that diverge from natural language expressions. This suggests that while larger models may offer improved language understanding and generation capabilities, they may not always be optimal for tasks requiring a high degree of precision and control over output sequences.

7 Conclusion

In this study, we introduce a novel approach named transition-based opinion tree generation, which seeks to bridge the gap between general pre-trained sequence-to-sequence language models and a structure-aware transition-based methodology. Our approach diverges from traditional methods by incorporating a transition system specifically tailored for opinion tree generation, designed to leverage the power of pre-trained language models through structured fine-tuning. Comprehensive experiments demonstrate that our model achieves substantial improvements over the state-of-the-art performance on multiple benchmark datasets. Furthermore, empirical studies reveal that our proposed transition-based opinion tree generation not only outperforms generative models but also excels in capturing the intricate sentiment structure within text.

Limitations

The limitations of our work can be stated from two perspectives. Firstly, it is necessary to evaluate our proposed transition-based opinion generation model to cross-domain settings and diverse languages. Furthermore, we also need to explore alternative parsing schemes for opinion generation, aiming to identify approaches that may further enhance the model's performance and versatility.

References

Xiaoyi Bao, Xiaotong Jiang, Zhongqing Wang, Yue Zhang, and Guodong Zhou. 2023. Opinion tree parsing for aspect-based sentiment analysis. In *Findings of the Association for Computational Linguistics: ACL 2023, Toronto, Canada, July 9-14, 2023*, pages 7971–7984. Association for Computational Linguistics.

Xiaoyi Bao, Zhongqing Wang, Xiaotong Jiang, Rong Xiao, and Shoushan Li. 2022. Aspect-based sentiment analysis with opinion tree generation. In *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI 2022, Vienna, Austria, 23-29 July 2022*, pages 4044–4050. ijcai.org.

Jiahao Bu, Lei Ren, Shuang Zheng, Yang Yang, Jingang Wang, Fuzheng Zhang, and Wei Wu. 2021. ASAP: A chinese review dataset towards aspect category sentiment analysis and rating prediction. In *NAACL-HLT 2021*, pages 2069–2079.

Hongjie Cai, Yaofeng Tu, Xiangsheng Zhou, Jianfei Yu, and Rui Xia. 2020. Aspect-category based sentiment analysis with hierarchical graph convolutional network. In *COLING 2020*, pages 833–843.

Hongjie Cai, Rui Xia, and Jianfei Yu. 2021. Aspect-category-opinion-sentiment quadruple extraction with implicit aspects and opinions. In *ACL/IJCNLP 2021*, pages 340–350.

Jiahao Cao, Rui Liu, Huailiang Peng, Lei Jiang, and Xu Bai. 2022. Aspect is not you need: No-aspect differential sentiment framework for aspect-based sentiment analysis. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1599–1609, Seattle, United States. Association for Computational Linguistics.

Nicola De Cao, Gautier Izacard, Sebastian Riedel, and Fabio Petroni. 2021. Autoregressive entity retrieval. In *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*. OpenReview.net.

Chenhua Chen, Zhiyang Teng, Zhongqing Wang, and Yue Zhang. 2022. Discrete opinion tree induction for aspect-based sentiment analysis. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2051–2064, Dublin, Ireland. Association for Computational Linguistics.

Pinzhen Chen, Nikolay Bogoychev, Kenneth Heafield, and Faheem Kirefu. 2020. Parallel sentence mining by constrained decoding. In *ACL 2020*, pages 1672–1678.

Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. 2022a. Lora: Low-rank adaptation of large language models. In *The Tenth International Conference on Learning Representations, ICLR 2022, Virtual Event, April 25-29, 2022*. OpenReview.net.

Mengting Hu, Yike Wu, Hang Gao, Yinhao Bai, and Shiwan Zhao. 2022b. Improving aspect sentiment quad prediction via template-order data augmentation. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 7889–7900, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.

Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer. 2020. BART: denoising sequence-to-sequence pretraining for natural language generation, translation, and comprehension. In *ACL 2020*, pages 7871–7880.

Junjie Li, Jianfei Yu, and Rui Xia. 2022. Generative cross-domain data augmentation for aspect and opinion co-extraction. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 4219–4229, Seattle, United States. Association for Computational Linguistics.

Yue Mao, Yi Shen, Jingchao Yang, Xiaoying Zhu, and Longjun Cai. 2022. Seq2Path: Generating sentiment tuples as paths of a tree. In *Findings of the Association for Computational Linguistics: ACL 2022*, pages 2215–2225, Dublin, Ireland. Association for Computational Linguistics.

Haiyun Peng, Lu Xu, Lidong Bing, Fei Huang, Wei Lu, and Luo Si. 2020. Knowing what, how and why: A near complete solution for aspect-based sentiment analysis. In *AAAI 2020*, pages 8600–8607.

Guang Qiu, Bing Liu, Jiajun Bu, and Chun Chen. 2011. Opinion word expansion and target extraction through double propagation. *Comput. Linguistics*, 37(1):9–27.

Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. 2020. Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of Machine Learning Research*, 21(140):1–67.

Duyu Tang, Bing Qin, Xiaocheng Feng, and Ting Liu. 2016a. Effective lstms for target-dependent sentiment classification. In *COLING 2016*, pages 3298–3307.

Duyu Tang, Bing Qin, and Ting Liu. 2016b. Aspect level sentiment classification with deep memory network. In *EMNLP 2016*, pages 214–224.

Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, Aurélien Rodriguez, Armand Joulin, Edouard Grave, and Guillaume Lample. 2023. Llama: Open and efficient foundation language models. *CoRR*, abs/2302.13971.

Hai Wan, Yufei Yang, Jianfeng Du, Yanan Liu, Kunxun Qi, and Jeff Z. Pan. 2020. Target-aspect-sentiment joint detection for aspect-based sentiment analysis. In *AAAI 2020*, pages 9122–9129.

Lu Xu, Hao Li, Wei Lu, and Lidong Bing. 2020. Position-aware tagging for aspect sentiment triplet extraction. In *EMNLP 2020*, pages 2339–2349.

Hang Yan, Junqi Dai, Tuo Ji, Xipeng Qiu, and Zheng Zhang. 2021. A unified generative framework for aspect-based sentiment analysis. In *ACL/IJCNLP 2021*, pages 2416–2429.

Songlin Yang and Kewei Tu. 2022. Bottom-up constituency parsing and nested named entity recognition with pointer networks. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2022, Dublin, Ireland, May 22-27, 2022*, pages 2403–2416. Association for Computational Linguistics.

Junchi Zhang, Yanxia Qin, Yue Zhang, Mengchi Liu, and Donghong Ji. 2019. Extracting entities and events as a single task using a transition-based neural model. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI 2019, Macao, China, August 10-16, 2019*, pages 5422–5428. ijcai.org.

Mi Zhang and Tieyun Qian. 2020. Convolution over hierarchical syntactic and lexical graphs for aspect level sentiment analysis. In *EMNLP 2020*, pages 3540–3549.

Wenxuan Zhang, Yang Deng, Xin Li, Yifei Yuan, Lidong Bing, and Wai Lam. 2021a. Aspect sentiment quad prediction as paraphrase generation. In *EMNLP 2021*, pages 9209–9219.

Wenxuan Zhang, Xin Li, Yang Deng, Lidong Bing, and Wai Lam. 2021b. Towards generative aspect-based sentiment analysis. In *ACL/IJCNLP 2021*, pages 504–510.

Junxian Zhou, Haiqin Yang, Yuxuan He, Hao Mou, and Junbo Yang. 2023. A unified one-step solution for aspect sentiment quad prediction. In *Findings of the Association for Computational Linguistics: ACL 2023, Toronto, Canada, July 9-14, 2023*, pages 12249–12265. Association for Computational Linguistics.