Perspective-driven Preference Optimization with Entropy Maximization for Diverse Argument Generation

Yilin Cao[1,2], **Ruike Zhang**[1], **Penghui Wei**[1*], **Qingchao Kong**[1,2], **Wenji Mao**[1,2*]
[1]State Key Laboratory of Multimodal Artificial Intelligence Systems,
Institute of Automation, Chinese Academy of Sciences
[2]School of Artificial Intelligence, University of Chinese Academy of Sciences
{caoyilin2022, zhangruike2020, qingchao.kong, wenji.mao}@ia.ac.cn
weipenghui2016@gmail.com

Abstract

In subjective natural language generation tasks, generating diverse perspectives is essential for fostering balanced discourse and mitigating bias. Argument generation with diverse perspectives plays a vital role in advancing the understanding of controversial claims. Despite the strong generative capabilities of large language models (LLMs), the diversity of perspectives remains insufficiently explored within argument generation task. Moreover, there remains a significant research gap in developing methods that explicitly generate multi-perspective arguments under the quality control of claim-stance alignment constraints. In this paper, we propose POEM, a **P**erspective-driven Preference **O**ptimization with **E**ntropy **M**aximization framework for diverse argument generation. It enhances perspective diversity through preference optimization based on the constructed preference dataset via perspective mining and diversity measuring. It further introduces entropy maximization to promote perspective diversity by encouraging dispersed semantic representations among the generated arguments. Experimental results on claim-stance argument generation benchmarks show that POEM is capable of generating diverse arguments while maintaining comparable performances in claim and stance controllability as well as text quality compared to the state-of-the-art baselines and human evaluation.

1 Introduction

Natural language generation (NLG) for subjective tasks such as argument generation requires not only fluency and coherence but also the ability to express a diverse range of perspectives. To mitigate potential bias, a well-designed NLG model should be capable of representing multiple perspectives rather than defaulting to a singular one or prevailing opinions (Hayati et al., 2024). Argument generation with diverse perspectives plays a vital role in advancing the understanding of controversial claims (Chen et al., 2019), and manifests foundational capabilities of human intelligence that are essential for modeling a wide range of human activities and common to human societies (Slonim et al., 2021). To ensure the quality of generated content, arguments should be consistent with the given stance and relevant to the corresponding claim. Moreover, expressing diverse perspectives is essential for avoiding potential bias and enhancing the persuasiveness of the arguments.

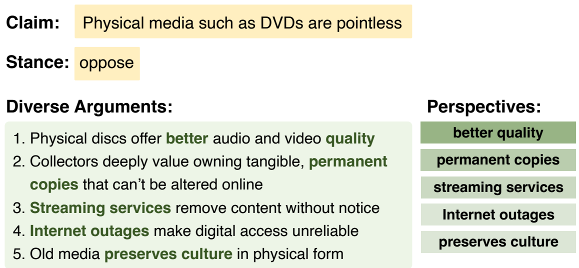

Figure 1: Illustrative example of arguments associated with diverse perspectives. In this example, for the claim "*Physical media such as DVDs are pointless*" and the *oppose* stance towards it, the diverse arguments reflect five distinct perspectives: better quality, permanent copies, streaming services, Internet outages, preserves culture.

Generating diverse arguments across multiple perspectives is an essential issue for online interactions. It can help the user articulate a wide spectrum of personal opinions, reflecting the richness of discussions in social media. For example, in Figure 1, there are five arguments to express *oppose* stance towards the claim "*Physical media such as DVDs are pointless*". Perspective provides a distinct pragmatic semantic focus that supports a given stance toward a claim. The perspectives in Figure 1 include noun phrases better quality, permanent copies, streaming services, Internet outages, preserves culture, which express topical diversity of opinions.

[*]Corresponding author

Findings of the Association for Computational Linguistics: EMNLP 2025, pages 22479–22496
November 4-9, 2025

To ensure the quality of argument generation, the generated arguments should be stance-alignment and maintain high relevance to the given claim. The majority of existing research has explored various control dimensions of argument generation to meet practical application requirements, including stance alignment (Sato et al., 2015; Gretz et al., 2020; Al Khatib et al., 2021), specific aspects (Schiller et al., 2021), user personalization (Alshomary et al., 2021, 2022), and factuality (Saha and Srihari, 2023). However, previous research has largely overlooked the important perspective diversity issue in argument generation.

With the rapid development of Large Language Models (LLMs), computational argumentation has become an essential tool across various domains, providing persuasive arguments aligned with a specific stance to enhance the understanding of controversial claims. Recently, there has been growing interest in leveraging their generative capabilities for argument generation. LLMs provide an important means to promoting positive online interactions and fostering active community communications among participants. However, the diversity of perspectives remains insufficiently explored in argument generation task.

To address this issue, recent studies have investigated prompting-based methods to elicit diverse arguments (Hayati et al., 2024). Due to the inherent limitations of auto-regressive decoding, LLMs are liable to produce repetitive contents or semantically different contents with similar perspectives. In particular, there remains a significant research gap in developing methods that explicitly generate multi-perspective arguments given a specific stance for a claim. Thus, from a computational account, there is still a lacking of LLM-enabled argument generation research encompassing diverse perspectives under the quality control of claim-stance alignment constraints. As a result, generating diverse arguments that effectively represent a broad and meaningful range of perspectives remains a fundamental challenge.

To address these challenges, in this paper, we propose a **P**erspective-driven Preference **O**ptimization with **E**ntropy **M**aximization (POEM) framework for diverse argument generation, comprised of instruction fine-tuning, perspective-driven preference optimization and entropy maximization modules. First, the instruction fine-tuning stage trains the model to equip initial ability of generating stance-consistent and claim-relevant arguments while implicitly capturing the multi-perspective nature of gold arguments. Then, we construct pairwise preference data via measuring perspective diversity among arguments to perform preference optimization, explicitly enhancing perspective diversity in generations. Finally, we devise an entropy maximization module that further encourages dispersed representations among generated arguments to reduce semantic redundancy and increase the coverage of perspectives.

The main contributions of our work are summarized as follows:

- To enhance perspective diversity, we design a perspective-driven preference optimization module to construct preference pairs and guide the model toward generating claim-relevant and stance-consistent arguments.
- We devise an entropy maximization module to further encourage perspective diversity through learning dispersed representations, effectively reducing semantic redundancy and increasing the coverage of perspectives.
- Experimental results on claim-stance argument generation benchmarks show that our proposed framework significantly enhances perspective diversity while maintaining competitive performances in claim-stance controllability and text quality compared to strong baselines as well as human evaluation.

2 Proposed Method

Given a controversial claim $\boldsymbol{c}$ and a stance $\boldsymbol{s}$, the stance-conditional argument generation task aims to generate an argument set $\boldsymbol{y} = \{y_1, \ldots, y_n\}$ containing n diverse arguments, where each argument $y_i \in \boldsymbol{y}$ reveals a perspective towards the claim. We define a **perspective** as a specific semantic focus or reasoning dimension that supports the given stance to the claim. As shown in Figure 1, the claim is approached from various perspectives including better quality, permanent copies, streaming services, Internet outages, preserves culture. These perspectives are grounded in noun phrases or topical phrases from the arguments.

Our proposed framework POEM consists of three modules: (1) **Instruction Fine-Tuning**, which enables the language model to produce arguments that are claim-relevant and stance-consistent, while promoting perspective diversity implicitly; (2) **Perspective-driven Preference Optimization**, where perspectives are formalized, measured, and explicitly optimized based on the constructed perspective-driven preference data; (3) **Entropy Maximization**, which further encourages perspective diversity through learning dispersed representations among arguments. Figure 2 illustrates an overview of the framework.

2.1 Instruction Fine-Tuning

To equip our model with the initial ability of generating arguments that are both claim-relevant and stance-consistent, we perform instruction fine-tuning on a dataset consisting of high-quality (claim $\boldsymbol{c}$, stance $\boldsymbol{s}$, argument set $\boldsymbol{y}$) triplets, where each argument set has a multi-perspective nature.

Formally, given a controversial claim $\boldsymbol{c}$ and a stance $\boldsymbol{s}$, our model $\pi_{\boldsymbol{\theta}}$ takes $(\boldsymbol{c}, \boldsymbol{s})$ as condition and generates an diverse argument set $\boldsymbol{y} = \{y_1, \ldots, y_n\}$, where each argument y_i is expected to reveal a perspective while maintaining alignment with $(\boldsymbol{c}, \boldsymbol{s})$. The training data for instruction fine-tuning $\mathcal{D}_{\text{instruct}}$ consists of examples where each instance contains a natural language instruction $\boldsymbol{x}$, a claim $\boldsymbol{c}$, a stance $\boldsymbol{s}$, and a multi-perspective argument set $\boldsymbol{y}$, and the learning objective is to minimize the negative log-likelihood:

$$\mathcal{L} = -\mathbb{E}_{(\boldsymbol{c},\boldsymbol{s},\boldsymbol{x},\boldsymbol{y})\sim\mathcal{D}_{\text{instruct}}} \left[\log \pi_{\boldsymbol{\theta}}(\boldsymbol{y} \mid \boldsymbol{c}, \boldsymbol{s}, \boldsymbol{x})\right], \tag{1}$$

where $\boldsymbol{\theta}$ refers to the trainable parameter set.

2.2 Perspective-driven Preference Optimization

To explicitly enhance perspective diversity, we construct a perspective-driven preference dataset to further tunes the model through the following steps.

Perspective Mining To extract salient perspectives, we perform a multi-step perspective mining process. After instruction fine-tuning, for each given claim-stance pair $(\boldsymbol{c}, \boldsymbol{s})$, we first generate m candidate argument sets $\mathcal{Y}_{\boldsymbol{c},\boldsymbol{s}} = \{\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(m)}\}$, where each argument set $\boldsymbol{y}^{(*)} = \{y_1^{(*)}, \ldots, y_n^{(*)}\}$ contains a multi-perspective argument set while remaining aligned with the same input condition $(\boldsymbol{c}, \boldsymbol{s})$. We then employ a noun phrase parser to extract concept-level candidate phrases. To extract representative perspectives, we choose Latent Dirichlet Allocation (LDA) (Blei et al., 2003) to cluster these noun phrases into K latent topics. From each topic, we select the most probable noun phrase to construct a perspective set $\mathcal{P}_{\boldsymbol{c},\boldsymbol{s}} = \{p_1, \ldots, p_K\}$, which serves as the perspective vocabulary for the claim-stance pair $(\boldsymbol{c}, \boldsymbol{s})$.

Diversity Measuring Given the extracted perspective set $\mathcal{P}_{\boldsymbol{c},\boldsymbol{s}} = \{p_1, \ldots, p_K\}$ for a given claim-stance pair $(\boldsymbol{c}, \boldsymbol{s})$, we propose to quantify the perspective diversity of m candidate argument sets $\mathcal{Y}_{\boldsymbol{c},\boldsymbol{s}} = \{\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(m)}\}$.

Consider an argument $y_i^{(*)}$ from the candidates $\mathcal{Y}_{\boldsymbol{c},\boldsymbol{s}}$ and a perspective $p_j \in \mathcal{P}_{\boldsymbol{c},\boldsymbol{s}}$, we compute their similarity using the cosine similarity between their embeddings:

$$s_{ij} = \text{CosSim}(y_i^{(*)}, p_j). \tag{2}$$

We then normalize the similarity scores into a probability distribution over the perspective set $\mathcal{P}_{\boldsymbol{c},\boldsymbol{s}} = \{p_1, \ldots, p_K\}$ for each argument:

$$v_{ij} = \frac{\exp(s_{ij})}{\sum_{j=1}^{K} \exp(s_{ij})}, \tag{3}$$

Here, the score $v_{ij} \in \boldsymbol{v}_i = [v_{i1}, v_{i2}, \ldots, v_{iK}]$ denotes the semantic similarity between an argument $y_i^{(*)}$ with a perspective $p_j \in \mathcal{P}_{\boldsymbol{c},\boldsymbol{s}}$. Therefore, for each argument set $\boldsymbol{y}^{(*)} = \{y_1^{(*)}, \ldots, y_n^{(*)}\} \in \mathcal{Y}_{\boldsymbol{c},\boldsymbol{s}}$, we obtain the measured diversity score matrix $V^{(*)} = \{\boldsymbol{v}_i, \ldots, \boldsymbol{v}_n\}$.

Then, to measure the perspective diversity of each argument set $\boldsymbol{y}^{(*)} = \{y_1^{(*)}, \ldots, y_n^{(*)}\} \in \mathcal{Y}_{\boldsymbol{c},\boldsymbol{s}}$, we compute the averaged forward and backward KL divergence between the similarity scores:

$$\mathbb{D}_{\text{div}}(\boldsymbol{y}^{(*)}) = \sum_{1\leq i<j\leq n} \Big(\text{KL}(\boldsymbol{v}_i \| \boldsymbol{v}_j) + \text{KL}(\boldsymbol{v}_j \| \boldsymbol{v}_i)\Big) \tag{4}$$

For each input condition $(\boldsymbol{c}, \boldsymbol{s})$, we employ the above diversity measuring method to rank the m generated candidate argument sets $\mathcal{Y}_{\boldsymbol{c},\boldsymbol{s}} = \{\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(m)}\}$. Specifically, we sample the top-d and bottom-d argument sets to form the highest-scoring subset $\mathcal{Y}_{\boldsymbol{c},\boldsymbol{s}}^{(+)} = \{\boldsymbol{y}^{(1+)}, \ldots, \boldsymbol{y}^{(d+)}\}$ and the lowest-scoring subset $\mathcal{Y}_{\boldsymbol{c},\boldsymbol{s}}^{(-)} = \{\boldsymbol{y}^{(1-)}, \ldots, \boldsymbol{y}^{(d-)}\}$, respectively. By randomly pairing samples from $\mathcal{Y}_{\boldsymbol{c},\boldsymbol{s}}^{(+)}$ and $\mathcal{Y}_{\boldsymbol{c},\boldsymbol{s}}^{(-)}$, we construct $d \times d$ preference pairs as perspective-driven preference data $\mathcal{D}_{\text{preference}} = \left\{\left(\boldsymbol{c}, \boldsymbol{s}, \boldsymbol{x}, \boldsymbol{y}^{(+)}, \boldsymbol{y}^{(-)}\right)\right\} |_{\boldsymbol{y}^{(+)}\sim\mathcal{Y}_{\boldsymbol{c},\boldsymbol{s}}^{(+)}, \boldsymbol{y}^{(-)}\sim\mathcal{Y}_{\boldsymbol{c},\boldsymbol{s}}^{(-)}}$.

Preference Optimization To explicitly enhance perspective diversity among generated arguments, we adopt direct preference optimization (DPO) (Rafailov et al., 2024) on the constructed

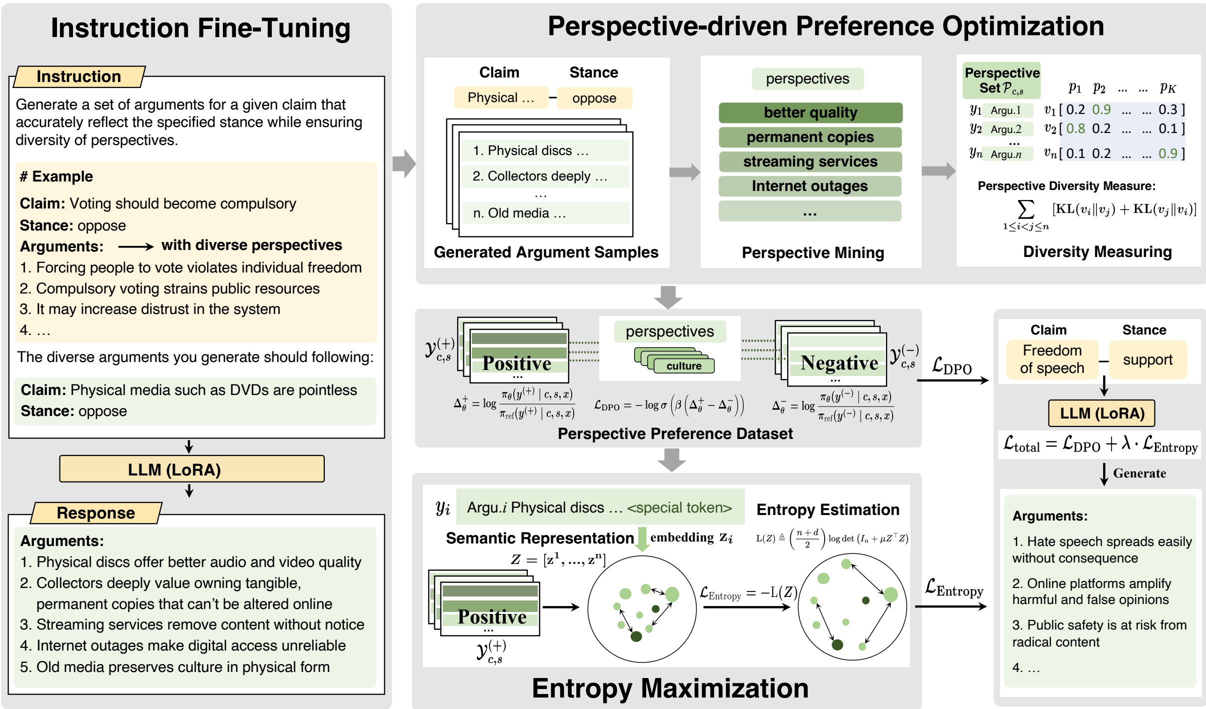

Figure 2: Overview of Perspective-driven Preference Optimization with Entropy Maximization for Diverse Argument Generation framework.

perspective-driven preference data. The learning objective compares the log-likelihood ratios between the language model π_θ to be learned and the reference model π_{ref} from previous instruction fine-tuning stage:

$$\mathcal{L}_{\text{DPO}} = -\,\mathbb{E}\left[\log\sigma\left(\beta\log\frac{\pi_\theta(\boldsymbol{y}^{(+)}\mid \boldsymbol{c},\boldsymbol{s},\boldsymbol{x})}{\pi_{\text{ref}}(\boldsymbol{y}^{(+)}\mid \boldsymbol{c},\boldsymbol{s},\boldsymbol{x})} - \beta\log\frac{\pi_\theta(\boldsymbol{y}^{(-)}\mid \boldsymbol{c},\boldsymbol{s},\boldsymbol{x})}{\pi_{\text{ref}}(\boldsymbol{y}^{(-)}\mid \boldsymbol{c},\boldsymbol{s},\boldsymbol{x})}\right)\right],$$
$$\text{where}\ \left(\boldsymbol{c},\boldsymbol{s},\boldsymbol{x},\boldsymbol{y}^{(+)},\boldsymbol{y}^{(-)}\right)\sim\mathcal{D}_{\text{preference}} \tag{5}$$

here $\sigma(\cdot)$ denotes logistic function and β is a hyperparameter.

2.3 Entropy Maximization

Similar perspectives often lead to semantic similarity, which may result in semantic collapse. To address this issue and encourage richer diversity across perspectives, our semantic regularization loss leverages an entropy-based regularization mechanism. The diversity of generated texts is closely linked to the probability distribution of output tokens. In particular, diversity can be understood as the similarity between the output distribution and a uniform distribution. In information theory, this relationship is quantified using Shannon Entropy (Shannon, 1948).

Based on this idea, we aim to incorporate entropy maximization into the training process. However, directly maximizing the entropy of the output probabilities may lead to semantic collapse, where the outputs become meaningless sequences that simply maximize the entropy of the token set. To avoid this, we propose maximizing entropy in the model's embedding space. Specifically, we devise a semantic regularization loss $\mathcal{L}_{\text{Entropy}}$, which encourages the argument embeddings to be more uniformly distributed across the hypersphere. This objective reduces semantic redundancy by increasing the separation between representations, while preserving alignment with the claim and stance, so as to further promote the targeted perspective diversity.

Semantic Representation To obtain semantic representations during training, we design a model-based approach to extract semantic representations from the positive subset $\mathcal{Y}^{(+)}_{\boldsymbol{c},\boldsymbol{s}} = \{\boldsymbol{y}^{(1+)},\ldots,\boldsymbol{y}^{(d+)}\}$. Specifically, a special token is appended to each argument $y_i \in \boldsymbol{y}$, and the hidden state corresponding to this token in the final decoder layer is used as the semantic representation of the argument y_i:

$$\mathbf{z_i} = \text{HiddenState}_{\text{last layer}}(y_i[\texttt{<special>}]). \tag{6}$$

Entropy Estimation By restricting the distribution to a semantically meaningful support set, the issue of semantic collapse is naturally reduced. However, calculating the entropy of learned representations remains challenging because estimating the distribution of high-dimensional vectors is difficult. To address this, we adopt the Coding-Length Function (CLF) (Ma et al., 2007), which is defined as follows:

$$\mathrm{L}(Z) \triangleq \left(\frac{n+d}{2}\right) \log\det\left(I_n + \mu Z^\top Z\right), \quad (7)$$

where $Z = [\mathbf{z^1}, ..., \mathbf{z^n}] \in \mathbb{R}^{n\times d}$ represents the semantic embeddings extracted in Eq. 6, n and d denote the number of representations and their dimensionality, respectively, and I_n is an n-by-n identity matrix. In the original CLF, μ is determined by n, d, and the expected decoding error. For simplicity, we set μ to 1.

Intuitively, the CLF measures the volume of the representation space. Maximizing this estimator leads to an expansion of the representation space, which in turn results in more diverse output contexts. To further encourage a uniform spread of semantic embeddings and mitigate redundancy, we define the semantic loss $\mathcal{L}_{\text{Entropy}}$ as follows:

$$\mathcal{L}_{\text{Entropy}} = -\mathrm{L}(Z). \quad (8)$$

Joint Objective The total training objective is a combination of the perspective-driven preference loss and the entropy maximization loss:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{DPO}} + \lambda \cdot \mathcal{L}_{\text{Entropy}}, \quad (9)$$

where λ is a hyperparameter that balances entropy maximization with preference optimization.

3 Experiments

3.1 Datasets

Following (Park et al., 2019), we evaluate our method on the PERSPECTRUM dataset (Chen et al., 2019), which contains claims paired with diverse stance-conditioned arguments. To assess generalization, we further introduce the CHANGE MY VIEW (CMV) (Hayati et al., 2024) and MIXED datasets for out-of-domain evaluation.

PERSPECTRUM Building upon the work of (Park et al., 2019), which focuses on generating sentential arguments from multiple perspectives, we extend the task by incorporating multi-scale stance control to better align generated arguments with user preferences. Following (Park et al., 2019), we evaluate our method on PERSPECTRUM dataset (Chen et al., 2019), which contains claims and their associated sentential arguments spanning diverse perspectives. In our setting, we use the claim along with its stance as input, and treat the set of arguments representing different perspectives as the target output. The dataset consists of 907 claims and 11,164 related arguments. We adopt the official split from (Park et al., 2019), which partitions the data into 541 claims for training, 139 for validation, and 227 for testing. The split ensures that claims on the same subject are grouped in the same partition, helping to prevent overfitting to claim-specific patterns. Additionally, to simulate real-world communication scenarios where users express varying degrees of stance, we leverage the dataset with stance control labels: *support* and *oppose*. In the PERSPECTRUM dataset, arguments labeled as *neutral* are randomly sampled from either *support* or *oppose* instances. Therefore, following prior work (Li et al., 2021; Vamvas and Sennrich, 2020), we exclude the *neutral* class from our setting.

CHANGE MY VIEW (CMV) After training and evaluating on the PERSPECTRUM dataset, we further assess the generalization capability of our model on unseen and varied claims. Inspired by (Hayati et al., 2024), we evaluate our method on the CMV dataset. CHANGE MY VIEW (CMV) comprises debate threads collected from the Change My View subreddit (Hidey et al., 2017). We extract only the discussion titles, which typically serve as the claims, resulting in 67 unique claims. This dataset allows us to test whether LLMs can generate diverse perspectives on highly subjective and potentially controversial claims (van Eemeren et al., 2015).

MIXED To evaluate performance on complex and politically charged claims, we construct a mixed-domain evaluation dataset by sampling claims from four widely used out-of-domain stance detection datasets: WT-WT (Conforti et al., 2020), P-Stance (Li et al., 2021), SemEval-2016 (Mohammad et al., 2016), and COVID19-Stance (Glandt et al., 2021). These datasets span a variety of domains, including politics, social issues, and public health, and collectively provide a comprehensive benchmark consisting of 18 politically claims for evaluating the generalization ability of argument generation beyond the scope of PERSPECTRUM.

Method	PERSPECTRUM					CMV				
	Diversity		Controllability		Quality	Diversity		Controllability		Quality
	Perspective↑	Semantic↑	Claim↑	Stance↑	PPL↓	Perspective↑	Semantic↑	Claim↑	Stance↑	PPL↓
ArgU (Saha and Srihari, 2023)	7.70	0.41	8.90	9.20	55.77	7.51	0.43	8.45	8.58	62.23
Crit.1 (Hayati et al., 2024)	7.98	0.37	9.16	8.86	32.47	7.75	0.42	8.44	8.16	32.62
Crit.2 (Hayati et al., 2024)	7.71	0.46	8.83	8.80	12.64	7.58	0.48	8.12	7.90	13.80
Meta-Llama-3-70B-Instruct	7.57	0.35	8.94	9.09	**10.09**	7.38	0.35	8.54	8.35	**11.44**
Qwen2.5-72B-Instruct	8.06	0.41	**9.62**	**9.58**	12.45	7.87	0.42	**8.98**	<u>8.70</u>	13.51
GPT-4-turbo	8.32	0.37	9.40	9.21	16.44	8.01	0.40	8.66	8.37	18.27
Deepseek R1	7.85	0.47	9.11	9.25	21.38	7.75	0.49	8.66	8.63	27.68
POEM (Meta-Llama-3-8B-Instruct)	<u>8.46</u>	<u>0.64</u>	<u>9.56</u>	9.16	<u>10.68</u>	<u>8.31</u>	<u>0.68</u>	**8.98**	**8.83**	<u>12.15</u>
POEM (Qwen2.5-7B-Instruct)	**8.72**	**0.70**	<u>9.56</u>	<u>9.36</u>	11.67	**8.44**	**0.72**	<u>8.79</u>	8.67	12.54

Table 1: Comparison of our method and baselines on the PERSPECTRUM (left) and CMV (right) datasets. The best results are in bold; the second-best are underlined. Here, Crit.1 represents the Criteria-based Prompting based on Meta-Llama-3-8B-Instruct and Crit.2 represents the Criteria-based Prompting based on Qwen2.5-7B-Instruct. **Perspective** refers to Perspective Diversity, **Semantic** to Semantic Diversity. **Claim** indicates Claim Relevance, **Stance** denotes Stance Alignment, and **PPL** measures the Perplexity of generated text.

Variant	PERSPECTRUM					CMV				
	Diversity		Controllability		Quality	Diversity		Controllability		Quality
	Perspective↑	Semantic↑	Claim↑	Stance↑	PPL↓	Perspective↑	Semantic↑	Claim↑	Stance↑	PPL↓
POEM (Meta-Llama-3-8B-Instruct)	8.46	0.64	9.56	9.16	10.68	8.31	0.68	8.98	8.83	12.15
– Entropy Maximization	8.18	0.52	9.51	9.02	10.13	8.01	0.54	8.85	8.61	12.24
– Perspective-driven Preference Optimization	7.93	0.41	9.43	8.96	10.24	7.77	0.44	8.73	8.46	11.89
– Instruction Fine-Tuning	7.79	0.33	9.33	8.79	11.18	7.64	0.36	8.67	8.42	13.22
POEM (Qwen2.5-7B-Instruct)	8.72	0.70	9.56	9.36	11.67	8.44	0.72	8.79	8.67	12.54
– Entropy Maximization	8.44	0.58	9.43	9.21	11.31	8.13	0.62	8.71	8.44	12.50
– Perspective-driven Preference Optimization	8.17	0.46	9.36	9.11	11.26	7.89	0.51	8.50	8.27	12.30
– Instruction Fine-Tuning	7.97	0.40	9.31	8.89	12.37	7.75	0.46	8.43	8.14	13.38

Table 2: Ablation results of our proposed method and its variants on the **PERSPECTRUM** (left) and **CMV** (right) datasets. Here, **Perspective** refers to Perspective Diversity, **Semantic** to Semantic Diversity. **Claim** indicates Claim Relevance, **Stance** denotes Stance Alignment, and **PPL** measures the Perplexity of generated text.

3.2 Evaluation Metrics

We evaluate the generated arguments from three dimensions: **diversity**, **controllability**, and **text quality**. **Diversity** measures whether the model generates a broad range of distinct perspectives and semantic representations, including perspective diversity and semantic diversity. **Controllability** assesses whether generated arguments are relevant to the given claim and consistent with the specified stance. **Text Quality** assesses the fluency of generated arguments, measured by perplexity.

Perspective Diversity We employ GPT-4o-mini as an evaluator to assess the diversity of perspectives expressed within each generated argument set. Each set is rated on a 0.0–10.0 scale based on the number and distinctness of reasoning dimensions it covers. Evaluation details and prompts are provided in Appendix B.1.

Semantic Diversity Following Hayati et al. (2024), we evaluate semantic diversity by first encoding each generated argument using Sentence-BERT (Reimers and Gurevych, 2019) with the DistilRoBERTa encoder (Sanh et al., 2019). We then compute the pairwise cosine distances between all arguments in a set and average them to obtain a semantic diversity score for that set. The final metric is the average score across all evaluation samples.

Controllability To evaluate whether the generated arguments are aligned with the intended stance and relevant to the given claim, we again use GPT-4o-mini as an evaluator. Each sample is rated from 0.0 to 10.0 on *Stance Alignment* and *Claim Relevance*. Details and rating prompts are provided in Appendix B.1.

Text Quality We assesses the fluency of generated arguments, measured by perplexity (PPL) computed using a pretrained GPT-2_{medium} language model (Radford et al., 2019), where lower values indicate more fluent outputs.

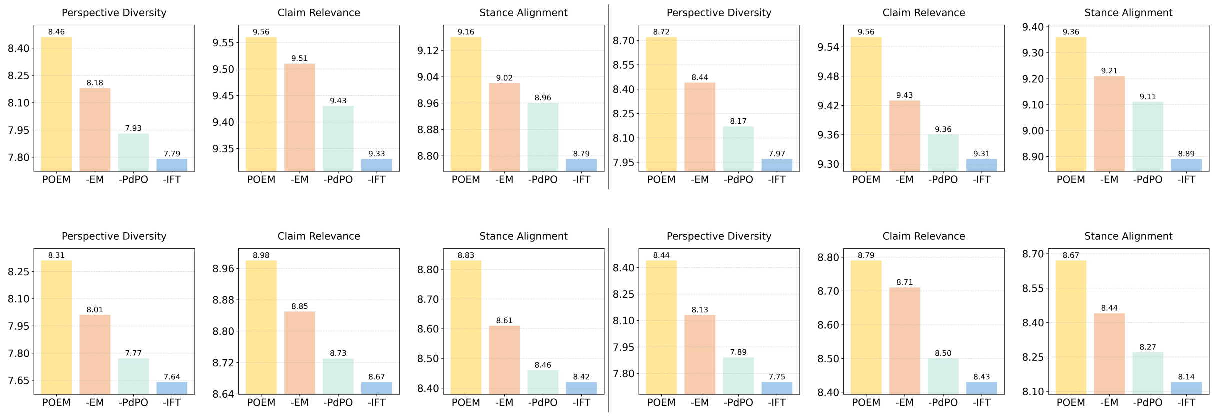

Figure 3: Ablation study results on: (**TOP**) PERSPECTRUM dataset and (**BOTTOM**) CMV dataset. For each dataset, the three bars on the left represent POEM (Meta-Llama-3-8B-Instruct) results, while the three right bars show POEM (Qwen2.5-7B-Instruct) results. **EM**: Entropy Maximization; **PdPO**: Perspective-driven Preference Optimization; **IFT**: Instruction Fine-Tuning.

3.3 Baselines

We compare our method against several representative baselines for diverse argument generation, including both traditional approaches and recent LLM-based models: (1) **ArgU** (Saha and Srihari, 2023) introduces a BART-based (Lewis et al., 2020) neural argument generator that incorporates argument schemes using control codes; (2) **Criteria-based Prompting** (Hayati et al., 2024) elicits diverse outputs from LLMs through criteria-guided prompting strategy; (3) **GPT Variants** are strong instruction-following decoder-only LLMs; we experiment with GPT-4-turbo, Qwen2.5-72B-Instruct and Meta-Llama-3-70B-Instruct; (4) **DeepSeek R1** is a recently released model emphasizing multi-step reasoning, included to assess performance against models equipped with reasoning capabilities.

We adopt a unified instruction format to guide the generation process. The prompt directs the model to generate a set of arguments based on a given claim and stance, with an emphasis on perspective diversity. The specific prompt instruction used is as follows:

> **Instruction:** Generate a set of arguments for a given claim that accurately reflect the specified stance while ensuring diversity of perspectives.

3.4 Main Results and Analysis

From the experimental results shown in Table 1 and Table 5, we observe that our proposed method consistently achieves superior perspective and semantic diversity across all three datasets. At the same time, it maintains competitive performance in stance alignment, claim relevance, and text quality, demonstrating the effectiveness of our entropy maximization module. Our method performs robustly on the CMV and MIXED datasets, which contain more open-ended and politically sensitive claims. In these settings, it consistently surpasses baselines in both diversity and controllability, showing strong generalization ability beyond the in-domain PERSPECTRUM dataset. Compared to strong LLM baselines such as GPT-4 and DeepSeek R1, our method also produces lower perplexity while preserving high output quality. These findings highlight the robustness and adaptability of our approach across different model architectures and data domains.

In addition to the main results, we assess the robustness of our method across multiple runs. Tables 7, 8, and 9 in Appendix A.3 report the standard deviations for all evaluation metrics on three datasets. The small standard deviations show that the results of our method remain stable across different runs, which are not caused by randomness in sampling or optimization. These results demonstrate the consistency and robustness of our method across multiple runs.

3.5 Ablation Study

We conduct ablation studies to evaluate the contribution of each component in the POEM framework. Figure 3 provides an intuitive illustration of the ablation results on PERSPECTRUM and CMV datasets. Table 2 and Table 6 report quantitative results across the three datasets. We could

Claim: America is a better place because of the 55 million abortions it's had. **Stance**: support
1. The reduction in population has allowed for better allocation of healthcare and education resources. (Environmental benefits of reduced population) 2. Abortion gives women autonomy to pursue education and careers, leading to greater gender equality. (Women's autonomy and empowerment through abortion) 3. It is more ethical to prevent unwanted pregnancies than to force unprepared parents to raise children. (Ethical considerations of preventing unwanted pregnancies) 4. Legal access to abortion has contributed to broader cultural acceptance of diverse family structures. (Cultural acceptance of diverse family structures) 5. The shift from coercive population control policies to reproductive rights reflects societal progress. (Learning from historical mistakes regarding reproductive rights)

Table 3: Representative case from the CMV dataset. Text in parentheses denotes the sentence-level perspectives extracted by GPT-4, illustrating the underlying reasoning behind each argument.

Method	PERSPECTRUM				CMV			
	Perspective↑	Claim↑	Stance↑	Fluency↑	Perspective↑	Claim↑	Stance↑	Fluency↑
GPT-4	8.39	8.92	9.10	9.31	8.48	9.02	9.10	9.24
POEM (Meta-Llama-3-8B-Instruct)	8.54	9.15	9.18	9.29	8.57	9.13	9.23	9.17
POEM (Qwen2.5-7B-Instruct)	8.43	8.99	9.09	9.26	8.50	9.05	9.13	9.32

Table 4: Human evaluation results on the PERSPECTRUM and CMV datasets, averaged over two independent annotators. Each dimension is rated on a 0.0–10.0 scale. The average Cohen's kappa coefficients (Cohen, 1960) κ of the inter-annotator agreement for human evaluation on perspective, claim, stance and fluency are 0.79, 0.62, 0.58 and 0.60, respectively (note that $0.6 \leq \kappa \leq 0.8$ means substantial agreement and $\kappa \geq 0.8$ means almost perfect agreement). **Perspective**: Perspective Diversity, **Stance**: Stance Alignment, **Claim**: Claim Relevance, **Fluency**: Linguistic Fluency. All scores: higher is better (↑).

see that removing the entropy maximization module (i.e., using only perspective-driven preference optimization) leads to a clear decline in semantic diversity, highlighting its importance in promoting deeper semantic separation among generated arguments. Omitting the perspective-driven preference optimization module (i.e., using only instruction fine-tuning) results in further drops in diversity, particularly in perspective diversity, indicating that explicit preference signals are crucial for encouraging varied reasoning. Lastly, excluding the instruction fine-tuning stage (i.e., using 0-shot prompting only) significantly degrades stance alignment and claim relevance, suggesting that fine-tuning is essential for improving controllability. These findings demonstrate that each component plays a complementary role in enhancing both diversity and controllability in stance-aware argument generation.

3.6 Case Study

To qualitatively assess the effectiveness of our method, we present a representative case from the CMV dataset. Table 3 illustrates diverse arguments generated by POEM (Qwen2.5-7B-Instruct), each reflecting distinct and nuanced perspectives aligned with the given stance. For instance, arguments span from practical considerations, such as better resource allocation due to reduced population, to deeper ethical implications like women's autonomy and prevention of unwanted parenthood. These examples highlight POEM's strength in explicitly capturing diverse and semantically distinct perspectives, thereby effectively minimizing redundancy and enhancing the comprehensiveness of generated arguments. More case studies and comparisons with baseline models are provided in Appendix D.

3.7 Human Evaluation

To complement automatic metrics, we conduct a human evaluation on the PERSPECTRUM and CMV datasets to assess the quality of the generated argument sets, using two independent annotators and report the average score. We also provide the average kappa coefficients on four evaluation metrics that reflect the inter-rater agreement on the knowledge evaluation. GPT-4-turbo exhibits strong perspective diversity among baseline methods in Table 1, we select it as the comparison model for human evaluation. As shown in Table 4, our method POEM significantly improves perspective diversity while maintaining comparable performance in stance consistency, claim relevance, and fluency. POEM with Meta-Llama-3-8B-Instruct achieves top performance in perspective diversity and claim relevance, while also slightly outperforming GPT-4 in stance alignment. These results validate the effectiveness of our stance-conditional framework. More details are provided in Appendix B.2.

4 Related Work

4.1 Argument Generation

Argument generation aims to automatically produce persuasive and coherent arguments (Chen et al., 2024). Research has explored various control dimensions of argument generation to meet practical application requirements, including stance alignment (Sato et al., 2015; Gretz et al., 2020; Al Khatib et al., 2021), specific aspects (Schiller et al., 2021), user personalization (Alshomary et al., 2021, 2022), and factuality (Saha and Srihari, 2023). Recently, research on exploiting large language models (LLMs) for argument generation has gained more attention. The related research focuses on multi-round debates (Li et al., 2024), logical fallacy (Mouchel et al., 2025) and diversity (Hayati et al., 2024). Due to the inherent limitations of auto-regressive generation in producing diverse outputs, recent research has explored alternative prompting strategies to enhance diversity in argument generation (Hayati et al., 2024). Specifically, to extract maximum diversity from LLMs, (Hayati et al., 2024) proposes a criteria-based prompting technique to ground diverse opinions. However, current methods rarely address the generation of diverse perspectives under stance control through computational means, which is essential for fostering a more comprehensive understanding of controversial claims.

4.2 Diversity in Natural Language Generation

Diversity in Natural Language Generation (NLG) has been explored across multiple levels, including lexical (Dušek and Kasner, 2020; Tevet and Berant, 2021), syntactic (Shen et al., 2019; Wen et al., 2023; Holtzman et al., 2020; Jinnai et al., 2024), semantic (Diao et al., 2021; Zhang et al., 2024), and perspective (Hayati et al., 2021; Santy et al., 2023). To solve the problem that implementations of GANs tend to be lack semantic diversity, TILGAN (Diao et al., 2021) combines a Transformer auto-encoder and a GAN in the latent space based on the Kullback-Leibler (KL) divergence. To summarize informative opinions, Wei et al. (2021) design a two-stage graph-to-sequence learning framework to promote salience and non-redundancy. Perspective diversity is constrained by training data and model design biases. (Hayati et al., 2021) show that BERT's interpretation of stylistic cues often diverges from human perception, while (Santy et al., 2023) reveal cultural biases favoring Western, educated, English-speaking views. Despite progress in diversity, perspective diversity remains underexplored in the context of argument generation. Existing approaches still struggle to generate diverse perspectives with sufficient breadth and depth, particularly on polarizing and controversial topics.

5 Conclusion

In this paper, we propose POEM, a **P**erspective-driven Preference **O**ptimization with **E**ntropy **M**aximization framework for diverse argument generation. Our approach integrates instruction fine-tuning, perspective-driven preference optimization, and an entropy maximization module to jointly promote perspective diversity while preserving stance consistency and claim relevance. Our perspective-driven preference optimization guides the model to generate claim-relevant, stance-consistent arguments by measuring and enhancing perspective diversity through perspective-driven preference data. Meanwhile, the entropy maximization module promotes diversity by learning dispersed representations, reducing redundancy and increasing perspective coverage. Experimental results on claim-stance benchmarks demonstrate that POEM significantly enhances perspective diversity while maintaining strong performance in stance alignment, claim relevance and text quality compared to state-of-the-art baselines also human evaluation.

Limitations

Our method is primarily designed for paragraph-level argument generation and may face challenges when applied to longer or discourse-level argumentative writing that requires more complex logical structures and extended coherence. Moreover, our work does not address the case where the stance is neutral or none, which is common in real-world communication. Existing research lacks a clear definition of the neutral stance, and such cases are often represented in datasets by a mixture of supportive and opposing arguments. Modeling this type of stance ambiguity remains an open challenge and a promising direction for future research.

Ethics Statement

This work is conducted purely for academic research and does not aim to promote or endorse any specific political or social stance. While our framework generates arguments over potentially sensitive and controversial claims, the model is not explicitly trained to produce biased, offensive, or harmful content. The perspectives expressed in generated arguments do not reflect the views of the authors. Our experiments are based on publicly available datasets, and human evaluations were conducted in accordance with ethical guidelines and with informed consent. We acknowledge that large language models may carry inherent social biases. Therefore, caution should be exercised when deploying such systems in real-world applications. We also encourage users to critically assess generated outputs, especially when applying the model to sensitive domains. We recommend that additional safeguards be considered to prevent potential harms in downstream use.

Acknowledgments

This work is supported in part by the National Natural Science Foundation of China under Grant #72293575, and the Joint Research Project on Culture-Science & Technology Integration between Chinese Academy of Sciences and Hunan Province under Grant #2024JK4003.

References

Khalid Al Khatib, Lukas Trautner, Henning Wachsmuth, Yufang Hou, and Benno Stein. 2021. Employing argumentation knowledge graphs for neural argument generation. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing*, pages 4744–4754.

Milad Alshomary, Wei-Fan Chen, Timon Gurcke, and Henning Wachsmuth. 2021. Belief-based generation of argumentative claims. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics*, pages 224–233.

Milad Alshomary, Roxanne El Baff, Timon Gurcke, and Henning Wachsmuth. 2022. The moral debater: A study on the computational generation of morally framed arguments. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics*, pages 8782–8797.

David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *Journal of machine Learning research*, pages 993–1022.

Guizhen Chen, Liying Cheng, Anh Tuan Luu, and Lidong Bing. 2024. Exploring the potential of large language models in computational argumentation. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics*, pages 2309–2330.

Sihao Chen, Daniel Khashabi, Wenpeng Yin, Chris Callison-Burch, and Dan Roth. 2019. Seeing things from a different angle:discovering diverse perspectives about claims. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics*, pages 542–557.

Jacob Cohen. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1):37–46.

Costanza Conforti, Jakob Berndt, Mohammad Taher Pilehvar, Chryssi Giannitsarou, Flavio Toxvaerd, and Nigel Collier. 2020. Will-they-won't-they: A very large dataset for stance detection on Twitter. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 1715–1724.

Shizhe Diao, Xinwei Shen, Kashun Shum, Yan Song, and Tong Zhang. 2021. TILGAN: Transformer-based implicit latent GAN for diverse and coherent text generation. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 4844–4858.

Ondřej Dušek and Zdeněk Kasner. 2020. Evaluating semantic accuracy of data-to-text generation with natural language inference. In *Proceedings of the 13th International Conference on Natural Language Generation*, pages 131–137.

Kyle Glandt, Sarthak Khanal, Yingjie Li, Doina Caragea, and Cornelia Caragea. 2021. Stance detection in COVID-19 tweets. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing*, pages 1596–1611.

Shai Gretz, Yonatan Bilu, Edo Cohen-Karlik, and Noam Slonim. 2020. The workweek is the best time to start a family – a study of GPT-2 based claim generation. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 528–544.

Shirley Anugrah Hayati, Dongyeop Kang, and Lyle Ungar. 2021. Does BERT learn as humans perceive?

understanding linguistic styles through lexica. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 6323–6331.

Shirley Anugrah Hayati, Minhwa Lee, Dheeraj Rajagopal, and Dongyeop Kang. 2024. How far can we extract diverse perspectives from large language models? In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 5336–5366.

Christopher Hidey, Elena Musi, Alyssa Hwang, Smaranda Muresan, and Kathy McKeown. 2017. Analyzing the semantic types of claims and premises in an online persuasive forum. In *Proceedings of the 4th Workshop on Argument Mining*, pages 11–21.

Ari Holtzman, Jan Buys, Li Du, Maxwell Forbes, and Yejin Choi. 2020. The curious case of neural text degeneration. In *International Conference on Learning Representations*.

Yuu Jinnai, Ukyo Honda, Tetsuro Morimura, and Peinan Zhang. 2024. Generating diverse and high-quality texts by minimum Bayes risk decoding. In *Findings of the Association for Computational Linguistics: ACL 2024*.

Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer. 2020. BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7871–7880.

Ming Li, Jiuhai Chen, Lichang Chen, and Tianyi Zhou. 2024. Can LLMs speak for diverse people? tuning LLMs via debate to generate controllable controversial statements. In *Findings of the Association for Computational Linguistics: ACL 2024*, pages 16160–16176.

Yingjie Li, Tiberiu Sosea, Aditya Sawant, Ajith Jayaraman Nair, Diana Inkpen, and Cornelia Caragea. 2021. P-stance: A large dataset for stance detection in political domain. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 2355–2365.

Yi Ma, Harm Derksen, Wei Hong, and John Wright. 2007. Segmentation of multivariate mixed data via lossy data coding and compression. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(9):1546–1562.

Saif Mohammad, Svetlana Kiritchenko, Parinaz Sobhani, Xiaodan Zhu, and Colin Cherry. 2016. SemEval-2016 task 6: Detecting stance in tweets. In *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, pages 31–41.

Luca Mouchel, Debjit Paul, Shaobo Cui, Robert West, Antoine Bosselut, and Boi Faltings. 2025. A logical fallacy-informed framework for argument generation. In *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 7296–7314.

ChaeHun Park, Wonsuk Yang, and Jong C Park. 2019. Generating sentential arguments from diverse perspectives on controversial topic. In *Proceedings of the Second Workshop on Natural Language Processing for Internet Freedom: Censorship, Disinformation, and Propaganda*, pages 56–65.

Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever, et al. 2019. Language models are unsupervised multitask learners. *OpenAI blog*.

Rafael Rafailov, Archit Sharma, Eric Mitchell, Christopher D Manning, Stefano Ermon, and Chelsea Finn. 2024. Direct preference optimization: Your language model is secretly a reward model. *Advances in Neural Information Processing Systems*.

Nils Reimers and Iryna Gurevych. 2019. Sentence-BERT: Sentence embeddings using Siamese BERT-networks. 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 3982–3992.

Sougata Saha and Rohini Srihari. 2023. ArgU: A controllable factual argument generator. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*, pages 8373–8388.

Victor Sanh, Lysandre Debut, Julien Chaumond, and Thomas Wolf. 2019. Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter. *arXiv preprint arXiv:1910.01108*.

Sebastin Santy, Jenny Liang, Ronan Le Bras, Katharina Reinecke, and Maarten Sap. 2023. NLPositionality: Characterizing design biases of datasets and models. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*, pages 9080–9102.

Misa Sato, Kohsuke Yanai, Toshinori Miyoshi, Toshihiko Yanase, Makoto Iwayama, Qinghua Sun, and Yoshiki Niwa. 2015. End-to-end argument generation system in debating. In *Proceedings of ACL-IJCNLP 2015 System Demonstrations*, pages 109–114.

Benjamin Schiller, Johannes Daxenberger, and Iryna Gurevych. 2021. Aspect-controlled neural argument generation. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 380–396.

Claude E Shannon. 1948. A mathematical theory of communication. *The Bell system technical journal*, 27(3):379–423.

Tianxiao Shen, Myle Ott, Michael Auli, and Marc'Aurelio Ranzato. 2019. Mixture models for diverse machine translation: Tricks of the trade. In *International conference on machine learning*, pages 5719–5728.

Noam Slonim, Yonatan Bilu, Carlos Alzate, Roy Bar-Haim, Ben Bogin, Francesca Bonin, Leshem Choshen, Edo Cohen-Karlik, Lena Dankin, Lilach Edelstein, et al. 2021. An autonomous debating system. *Nature*, 591(7850):379–384.

Guy Tevet and Jonathan Berant. 2021. Evaluating the evaluation of diversity in natural language generation. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, pages 326–346.

Jannis Vamvas and Rico Sennrich. 2020. X-stance: A multilingual multi-target dataset for stance detection. *arXiv preprint arXiv:2003.08385*.

Frans H. van Eemeren, Sally Jackson, and Scott Jacobs. 2015. *Argumentation*, pages 3–25.

Penghui Wei, Jiahao Zhao, and Wenji Mao. 2021. A graph-to-sequence learning framework for summarizing opinionated texts. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 29:1650–1660.

Yuqiao Wen, Yongchang Hao, Yanshuai Cao, and Lili Mou. 2023. An equal-size hard EM algorithm for diverse dialogue generation. In *The Eleventh International Conference on Learning Representations*.

Tianhui Zhang, Bei Peng, and Danushka Bollegala. 2024. Improving diversity of commonsense generation by large language models via in-context learning. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 9226–9242.

A Additional Results and Robustness

A.1 Main Results on MIXED Dataset

We provide the main experimental results on the MIXED dataset in Table 5. As shown in the table, our proposed method POEM consistently outperforms all baselines across most evaluation dimensions. In particular, POEM with Qwen2.5-7B achieves the strongest overall performance, demonstrating the effectiveness of integrating all three components of our framework in enhancing diversity and controllability. POEM with Meta-Llama-3-8B also performs competitively, ranking closely behind in core metrics while producing text with notably higher fluency. Compared to strong baselines such as GPT-4-turbo and Criteria-based prompting, POEM shows clear advantages in generating arguments that are both diverse and well-aligned with stance and claim. The consistent improvements across models and datasets reflect the transferability and flexibility of our framework.

A.2 Ablation Study on MIXED Dataset

Table 6 provides the detailed ablation results on the MIXED dataset. We report the performance of POEM and its three ablated variants across multiple dimensions, including perspective diversity, semantic diversity, stance alignment, claim relevance, and text quality.

A.3 Robustness Across Multiple Runs

To further examine the robustness of our method, we report the means and standard deviations over three independent runs on the PERSPECTRUM, CMV, and MIXED datasets in Tables 7, 8, and 9, respectively. These results demonstrate the consistency and robustness of our method across multiple runs.

B Evaluation

B.1 Evaluation Prompts for GPT4o

To enable more accurate evaluation of *Perspective Diversity*, we design a two-stage instruction for GPT-4o-mini. The model is first prompted to identify the distinct perspective behind each generated comment, and then rate the overall diversity based on the extracted reasoning types. The actual prompt is as follows:

Task: Evaluate the generated comments based on the given claim and stance using the following criteria. Assign scores from 0.0 to 10.0 (rounded to one decimal place):

1. **Claim Relevance (0.0–10.0)** Assess how well the generated comments relate to the given claim. A higher score indicates stronger relevance and logical connection.
    - Focus only on topic relevance, regardless of stance or diversity.
    - 0.0 means completely irrelevant; 10.0 means highly relevant.
2. **Stance Alignment (0.0–10.0)** Evaluate how accurately the comments reflect the specified stance (support, oppose).
    - Assign 0.0 if the comment is unrelated to the claim.
    - If relevant, score based on stance consistency.
    - This metric ignores topic diversity.
3. **Perspective Diversity (0.0–10.0)** First, label each comment with a short `Perspective: <summary>` describing its distinct reasoning.
    - Do not rephrase the comments; only extract their underlying reasoning.
    - Then evaluate the overall diversity of perspectives.
    - 0.0 indicates near-duplicate perspectives; 10.0 indicates rich and clearly distinct viewpoints.

B.2 Human Evaluation

To complement the automatic metrics, we conduct human evaluation to assess the quality of the generated argument sets. Each set is evaluated along four dimensions: **Perspective Diversity**, **Stance Alignment**, **Claim Relevance**, and **Fluency**.

We randomly sample 90 argument sets from the test data and recruit two annotators with NLP backgrounds to rate each set on a scale from 0.0 to 10.0 for each dimension, with one decimal point of precision. Higher scores indicate better performance. Final scores are averaged across annotators.

Detailed scoring criteria for each evaluation dimension are provided in Tables 10–13.

C Experimenmtal Details

C.1 Implementation Details

We extract concept-level noun phrases from generated argument sets using the `en_core_web_sm` model of spacy for perspective mining, which performs syntactic parsing and noun chunk identification. We retain only multi-word noun phrases

Method	MIXED Dataset				
	Diversity		Controllability		Quality
	Perspective↑	Semantic↑	Claim↑	Stance↑	PPL↓
ArgU (Saha and Srihari, 2023)	7.57	0.41	8.64	9.42	55.28
Crit.1 (Hayati et al., 2024)	7.88	0.46	9.01	9.23	43.56
Crit.2 (Hayati et al., 2024)	7.72	0.49	8.75	9.40	<u>11.67</u>
Meta-Llama-3-70B-Instruct	7.68	0.40	8.71	9.35	11.15
Qwen2.5-72B-Instruct	7.89	0.46	9.26	<u>9.64</u>	12.67
GPT-4-turbo	8.12	0.45	9.04	9.51	16.18
Deepseek R1	7.75	0.54	8.90	9.55	24.92
POEM (Meta-Llama-3-8B)	<u>8.21</u>	<u>0.71</u>	<u>9.47</u>	9.62	**10.96**
POEM (Qwen2.5-7B)	**8.52**	**0.76**	**9.53**	**9.83**	11.94

Table 5: Comparison of our method and baselines on the **MIXED** dataset. The best results are in bold; the second-best are underlined. Here, Crit.1 represents the Criteria-based Prompting based on Meta-Llama-3-8B-Instruct and Crit.2 represents the Criteria-based Prompting based on Qwen2.5-7B-Instruct. **Perspective** refers to Perspective Diversity, **Semantic** to Semantic Diversity. **Claim** indicates Claim Relevance, **Stance** denotes Stance Alignment, and **PPL** measures the Perplexity of generated text.

Variant	Diversity		Controllability		Text Quality
	Perspective↑	Semantic↑	Claim↑	Stance↑	PPL↓
POEM (Meta-Llama-3-8B-Instruct)	8.21	0.71	9.47	9.62	10.96
– Entropy Maximization	7.99	0.58	9.43	9.47	10.95
– Perspective-driven Preference Optimization	7.81	0.48	9.41	9.40	10.72
– Instruction Fine-Tuning	7.67	0.40	9.31	9.30	11.79
POEM (Qwen2.5-7B-Instruct)	8.52	0.76	9.53	9.83	11.94
– Entropy Maximization	8.29	0.64	9.48	9.69	12.29
– Perspective-driven Preference Optimization	8.07	0.53	9.41	9.67	11.43
– Instruction Fine-Tuning	7.90	0.44	9.38	9.31	12.69

Table 6: Ablation results of the variants of our proposed method on the **MIXED** dataset. Here, **Perspective** refers to Perspective Diversity, **Semantic** to Semantic Diversity. **Stance** denotes Stance Alignment, **Claim** indicates Claim Relevance, and **PPL** measures the Perplexity of generated text.

Model	PERSPECTRUM Dataset				
	Diversity		Controllability		Quality
	Perspective↑	Semantic↑	Claim↑	Stance↑	PPL↓
POEM (Meta-Llama-3-8B-Instruct)	8.46 ± 0.04	0.64 ± 0.02	9.56 ± 0.03	9.16 ± 0.02	10.68 ± 0.05
POEM (Qwen2.5-7B-Instruct)	8.72 ± 0.05	0.70 ± 0.03	9.56 ± 0.04	9.36 ± 0.01	11.67 ± 0.06

Table 7: Evaluation results with standard deviations over three runs on the **PERSPECTRUM** dataset.

to ensure the extracted concepts are semantically meaningful. These phrases are then used as input units for LDA-based topic modeling. The topic model LDA is implemented with gensim, where the number of latent topics is set to 5. The number of candidate argument sets n for each claim stance pair is set to 6. When sample the top-d and bottom-d outputs to form the highest-scoring subset and lowest-scoring subset, d is set to 2.

We fine-tune Meta-LLaMA-3-8B-Instruct and Qwen2.5-7B-Instruct models using the LLaMA-Factory framework. During both the supervised fine-tuning (SFT) and Direct Preference Optimization (DPO) stages, we employ the AdamW opti-

Model	CMV Dataset				
	Diversity		Controllability		Quality
	Perspective↑	Semantic↑	Claim↑	Stance↑	PPL↓
POEM (Meta-Llama-3-8B-Instruct)	8.31 ± 0.03	0.68 ± 0.03	8.98 ± 0.04	8.83 ± 0.02	12.15 ± 0.04
POEM (Qwen2.5-7B-Instruct)	8.44 ± 0.02	0.72 ± 0.02	8.79 ± 0.03	8.67 ± 0.04	12.54 ± 0.05

Table 8: Evaluation results with standard deviations over three runs on the **CMV** dataset.

Model	MIXED Dataset				
	Diversity		Controllability		Quality
	Perspective↑	Semantic↑	Claim↑	Stance↑	PPL↓
POEM (Meta-Llama-3-8B-Instruct)	8.21 ± 0.04	0.71 ± 0.02	9.47 ± 0.02	9.62 ± 0.02	10.96 ± 0.03
POEM (Qwen2.5-7B-Instruct)	8.52 ± 0.02	0.76 ± 0.03	9.53 ± 0.02	9.83 ± 0.04	11.94 ± 0.06

Table 9: Evaluation results with standard deviations over three runs on the **MIXED** dataset.

Score Range	Description
0.0–2.0	All arguments express nearly the same idea or use repeated reasoning. No noticeable perspective variation.
3.0–5.0	Minor variation across arguments, but significant semantic overlap exists. Perspectives are not clearly distinguishable.
6.0–8.0	Most arguments introduce distinct angles or aspects. Reasoning is moderately diverse and topic coverage is broadened.
9.0–10.0	Arguments cover clearly different and meaningful perspectives. Each one contributes a unique viewpoint to the claim.

Table 10: Human evaluation scoring criteria for Perspective Diversity.

Score Range	Description
0.0–2.0	Arguments contradict the given stance or support the opposite side.
3.0–5.0	Arguments are loosely related to the stance, with some inconsistencies.
6.0–8.0	Arguments generally support the given stance, but may contain ambiguous or mixed signals.
9.0–10.0	Arguments strongly and clearly support the specified stance with consistent reasoning.

Table 11: Human evaluation scoring criteria for Stance Alignment.

Score Range	Description
0.0–2.0	Arguments are irrelevant or completely off-topic from the given claim.
3.0–5.0	Arguments mention related ideas but fail to directly address the claim.
6.0–8.0	Arguments address the claim with moderate clarity and relevance.
9.0–10.0	Arguments are highly relevant, directly engaging with the claim in a focused and coherent way.

Table 12: Human evaluation scoring criteria for Claim Relevance.

mizer with a weight decay of 0.1. The learning rates are set to 7e-6 for SFT and 1e-6 for DPO, with a cosine learning rate schedule and a warm-up ratio of 0.01. We train all models with a batch size of 16 and a gradient accumulation step of 8, resulting in an effective batch size of 128. The models

Score Range	Description
0.0–2.0	Arguments are grammatically incorrect or incoherent, with major fluency issues.
3.0–5.0	Texts contain noticeable grammatical errors or awkward phrasing that affect readability.
6.0–8.0	Generally fluent, with minor issues in grammar or flow.
9.0–10.0	Highly fluent, well-structured, and natural-sounding language.

Table 13: Human evaluation scoring criteria for Fluency.

are trained for 3 epochs in the SFT stage and 10 epochs in the DPO stage. For entropy maximization, the λ is fixed at 0.1 across all models and datasets.

D Case Study

As shown in Tables 14–16, we present three representative cases from the CMV dataset to qualitatively evaluate the effectiveness of our method. Each case includes the input claim and stance, along with the generated standpoints from POEM (Meta-Llama-3-8B-Instruct) and POEM (Qwen2.5-7B-Instruct). Text in parentheses denotes the sentence-level perspectives extracted by GPT-4, illustrating the underlying reasoning behind each argument.

Claim: I don't think anyone should have to pay child support. **Stance**: oppose
POEM (Meta-Llama-3-8B-Instruct) 1. The government should take responsibility for supporting children, not individuals. (Child support is essential for children's welfare.) 2. Paying child support is a form of patriarchal oppression. (Parents have a responsibility to support their children financially.) 3. Child support laws are often unfair and arbitrary. (Eliminating child support could harm children's development.) 4. Raising children is a personal choice, not a financial obligation. (Child support ensures fairness in parenting responsibilities.)
POEM (Qwen2.5-7B-Instruct) 1. Non-custodial parent's contributions are essential for family stability. (Financial hardship for custodial parents and children) 2. Child support ensures basic needs like food and education are met. (Child support ensures children's basic needs are met) 3. Ability to pay varies, but abolishing support is unfair. (Factors affecting ability to pay child support) 4. Non-payment has legal consequences like imprisonment. (Legal consequences of non-payment of child support) 5. Payments help cover healthcare and development needs. (Child support covers essential expenses for child development)

Table 14: Case Study 1: Child Support Debate

Claim: Instead of "bands" for income tax there should be a continuous formula for tax rates. **Stance**: support
POEM (Meta-Llama-3-8B-Instruct) 1. Ensures equitable contributions based on actual income. (Equity in tax contributions) 2. Simplifies the tax system by eliminating brackets. (Efficiency through reduced complexity) 3. Enables progressive tax changes with income. (Progressive tax rate increases) 4. Promotes transparency and predictability in taxation. (Transparency in tax structure) 5. Adapts easily to economic fluctuations. (Responsiveness to economic changes) 6. Supports fairness and justice under law. (Fairness and justice in taxation)
POEM (Qwen2.5-7B-Instruct) 1. Efficient resource allocation based on income. (Fairness in tax burden distribution) 2. Prevents unnecessary over-taxation. (Fairness in tax burden distribution) 3. Reduces confusion in the tax code. (Simplification of tax code) 4. Encourages higher earnings without bracket penalty. (Encouragement of higher earnings) 5. Reduces cost of tax administration. (Reduction of administrative costs) 6. Reflects actual income more accurately. (Accuracy in reflecting income)

Table 15: Case Study 2: Continuous Tax Formula

Claim: Mobile websites are a relic of the early mobile internet era and are inferior to actual websites when using your smartphone.
Stance: support

POEM (Meta-Llama-3-8B-Instruct)
1. Mobile sites are clunky and harder to navigate. (Clunky mobile websites hinder navigation)
2. Often lack full content and features. (Limited content frustrates mobile users)
3. Design is often cramped or unappealing. (Poor design choices overwhelm users)
4. Poor SEO reduces search visibility. (Mobile sites struggle with search optimization)
5. Trust and security are lacking. (Security risks make mobile sites less trustworthy)

POEM (Qwen2.5-7B-Instruct)
1. Mobile sites offer slower performance. (Mobile websites lack engaging multimedia content)
2. Less media-rich and less engaging. (Mobile web has security vulnerabilities and performance issues)
3. Prone to vulnerabilities via plugins. (Mobile web has security vulnerabilities and performance issues)
4. Actual websites optimized for mobile. (Actual websites are optimized for a better mobile experience)
5. Lower maintenance and better SEO. (Developing actual websites is more cost-effective long-term)
6. Broader access and consistency. (Actual websites have better SEO and device accessibility)
7. Better data efficiency. (Mobile web can lead to higher data costs for users)
8. Cross-platform compatibility. (Actual websites ensure consistent user experience across platforms)
9. Enhanced user engagement tools. (Mobile web limits access to advanced features for engagement)
10. More interactive design elements. (Actual websites can include more interactive elements)

Table 16: Case Study 3: Mobile Websites vs. Actual Websites