SGTR: Learning Style-Generalizable Trajectory Reasoning for Zero-Shot Text Style Transfer

Anonymous ACL submission

Abstract

Text Style Transfer (TST) aims to rewrite text in a target style while preserving its semantics. However, existing methods rely on style-specific supervision and implicit generation, leading to poor generalization and semantic degradation given the inherent scarcity of parallel data. We propose Style-Generalizable Trajectory Reasoning (SGTR), a novel framework that formulates TST as the joint modeling of reasoning trajectories and outputs, decomposing the process into three interpretable stages: style understanding, deviation localization, and reflection refinement. To further enhance stylistic generalization, we introduce a latent contrastive learning objective, enabling the model to disentangle stylistic attributes from semantic content directly within its hidden representations. We distill these sophisticated analytical capabilities into a lightweight student model. Experiments demonstrate that SGTR consistently outperforms strong baselines in both style quality and semantic preservation, while exhibiting superior zero-shot generalization to unseen styles.

1 Introduction

Text Style Transfer (TST) aims to rewrite a given text into a target style while preserving its core semantic content. Due to its significant utility in applications such as toxic content neutralization (Ziegenbein et al., 2024), formality adjustment (Rao and Tetreault, 2018), and content personalization (Horvitz et al., 2024), TST remains a focal point of research in Natural Language Processing.

In recent years, the emergence of Large Language Models (LLMs) has demonstrated remarkable zero-shot capabilities across various instruction-driven generation tasks. TST research has transitioned from traditional disentangled representation learning (John et al., 2019) to frameworks anchored in LLMs and instruction tuning (Shu et al., 2024). Current state-of-the-art (SOTA) approaches typically employ an LLM as a teacher model to synthesize high-quality parallel data or generate self-explanatory rationales, which are then distilled into lightweight student models (Zhang et al., 2024; Han et al., 2024). Compared to early unsupervised methods that operated in the absence of parallel corpora (Shen et al., 2017), these advancements have significantly enhanced the quality of stylized generation.

Despite this progress, several critical bottlenecks persist. Prevailing methods adhere to a style-specific fine-tuning paradigm that necessitates training a distinct model for each transfer task. This unscalable approach requires large-scale parallel datasets for every newly introduced style, which is often untenable due to the inherent scarcity of supervised TST corpora. Furthermore, these approaches typically treat the transformation as an implicit black-box operation, frequently leading to semantic erosion or hallucinated rewrites, particularly in low-resource scenarios. We argue that these limitations stem from a fundamental reliance on surface-level textual imitation rather than deep stylistic comprehension. Thus, models lack the analytical capacity to localize specific stylistic mismatches and systematically optimize them without compromising the semantic integrity.

To address these issues, we propose a **Style-Generalizable Trajectory Reasoning (SGTR)** framework, shifting the TST paradigm from style-specific memorization to style-generalizable universal reasoning. As illustrated in Figure 1, SGTR reformulates TST as a joint distribution of Chain-of-Thought (CoT) (Wei et al., 2022) reasoning trajectories and execution outputs, decomposing the transformation into a structured, interpretable three-stage pipeline: Style Understanding, Deviation Localization, and Reflection Refinement. By establishing a standardized CoT reasoning paradigm, SGTR enables the model to apply a consistent analytical logic across unseen styles, thereby ensuring

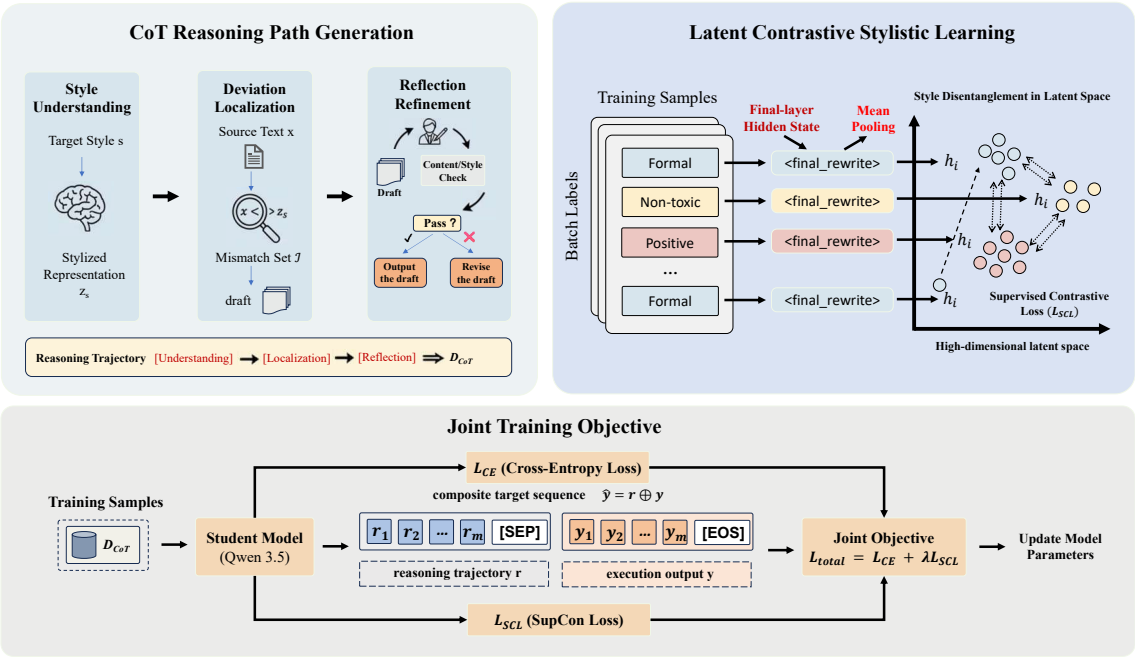

Figure 1: An overview of the SGTR framework. The top-left panel illustrates the CoT Reasoning Path Generation, decomposing the transfer process into style understanding, deviation localization, and reflection refinement. The top-right panel depicts the Latent Contrastive Stylistic Learning, which enforces fundamental stylistic disentanglement within the model's high-dimensional latent space. The bottom panel details the joint training process, where these sophisticated analytical capabilities are distilled into a student model optimized by a dual-objective function: the cross-entropy loss ($\mathcal{L}_{CE}$) for sequence generation and the latent contrastive loss ($\mathcal{L}_{SCL}$) for feature regularization.

robust generalizability. To empower the model with universal stylistic comprehension, SGTR introduces a latent contrastive stylistic learning objective. This representation-based constraint compels the model to decouple stylistic markers from semantic content directly within its hidden states. By pulling together representations of identical styles and pushing apart disparate ones, we transform the training objective from simple textual imitation to the fundamental mastery of stylistic features. We distill these sophisticated analytical capabilities into a lightweight student model Qwen3.5-9B (Qwen Team, 2026). Experimental results demonstrate that SGTR not only outperforms multiple strong baselines on seen styles but also exhibits exceptional zero-shot generalization to unseen styles with remarkably low data requirements.

Our contributions are summarized as follows:

(1) We propose SGTR, a novel framework that reformulates TST as the joint modeling of CoT reasoning trajectories and execution outputs, shifting the paradigm from style-specific memorization to style-generalizable universal reasoning.

(2) We introduce a latent contrastive stylistic learning objective. This representation-based constraint empowers the model to decouple stylistic markers from semantic content in the hidden space, enabling it to master underlying stylistic features rather than merely imitating surface-level textual patterns.

(3) Extensive experiments demonstrate that SGTR outperforms multiple strong baselines with high data efficiency and exhibits exceptional zero-shot generalization to unseen stylistic domains, offering a scalable solution for universal style transfer in low-resource scenarios. We **open-source our code and the dataset** in the supplementary material to facilitate future research in this domain.

2 Related Work

Traditional Style Transfer. TST research has long been hampered by the inherent scarcity of parallel corpora (Toshevska and Gievska, 2025), leading to a historical emphasis on unsupervised methodologies (Shen et al., 2017; Prabhumoye et al., 2018; Scalercio and Paes, 2024). Early studies primarily revolved around disentangled representation learning (John et al., 2019) or the Delete-

Retrieve-Generate framework (Li et al., 2018), aiming to decouple stylistic attributes from semantic content within the latent space. However, these methods frequently suffer from semantic erosion, generating sentences that lack authenticity and logical fluency (Lample et al., 2019; Mir et al., 2019). To address these limitations, recent advancements have proposed a method that ensures contextual coherence through hierarchical structural mapping at both sentence and paragraph levels (Wu and Deng, 2025), while DisenTrans (Han et al., 2024) proposed a disentangled data synthesis paradigm to construct high-quality pseudo-parallel data by identifying stylistic attribute components. Collectively, these developments signify a paradigm shift from surface-level word manipulation toward a more holistic understanding of discourse structure and stylistic attributes.

Style Transfer with LLMs. The emergence of LLMs has significantly broadened the paradigm of TST, enabling models to perform more general rewriting tasks through instruction following and knowledge distillation (Sanh et al., 2021). Recent work has demonstrated the efficacy of instruction tuning for TST rewriting (Shu et al., 2024), while the subsequent work (Ziegenbein et al., 2024) extends the application scope to the document-level normalization of inappropriate argumentation, or leveraged decoupled reinforcement learning to enhance generic rewriting performance (Li et al., 2025). Furthermore, inspired by the reasoning capabilities elicited by Chain-of-Thought (CoT) prompting (Wei et al., 2022), prior work distills self-explanatory reasoning processes to synthesize high-quality parallel data (Zhang et al., 2024), thereby improving performance in low-resource scenarios. Despite their strong performance, these approaches often rely on style-specific tuning and exhibit limited generalization to unseen styles, with constrained interpretability of the underlying generation process. To address these limitations, we propose the SGTR framework, shifting the TST paradigm toward a style-generalizable, trajectory-based reasoning approach for superior zero-shot generalizability.

3 Methodology

3.1 Probabilistic Formulation

We formulate the Text Style Transfer (TST) task as a conditional generation problem. Traditional approaches typically aim to learn the conditional distribution $p(y \mid x, s)$, where $x \in \mathcal{X}$ is the source sequence, $s \in \mathcal{S}$ is the target style, and $y \in \mathcal{Y}$ is the stylized output. However, the direct modeling of this distribution is often hindered by the scarcity of high-quality parallel corpora and the inherent difficulty in achieving precise stylistic transformation without eroding the core semantic integrity of the source text.

To address this, SGTR introduces a reasoning variable $r \in \mathcal{R}$ to represent the analytical and decision-making process, decomposing the implicit transformation into structured intermediate inference steps. The learning objective is extended to the joint modeling of the reasoning trajectory and the execution output:

$$p(y, r \mid x, s) = p(r \mid x, s) \cdot p(y \mid r, x, s) \quad (1)$$

where $p(r \mid x, s)$ denotes the generation of the reasoning trajectory and $p(y \mid r, x, s)$ executes the textual reconstruction conditioned on the reasoning trajectory. Based on this formulation, we curate a reasoning-augmented dataset $\mathcal{D}_{\text{cot}}$ comprising chain-of-thought inference paths. This dataset not only provides explicit textual supervision for the multi-stage reasoning process but also serves as the foundational data structure for applying our latent contrastive regularization in the subsequent representation learning phase.

3.2 CoT Reasoning Path Generation.

As illustrated in Figure 2, SGTR employs structured instructions to guide the LLM in generating reasoning trajectories in a standardized format and defines the reasoning variable r through a three-stage pipeline: Style Understanding, Deviation Localization, and Reflection Refinement.

Style Understanding. We enable the model to understand the specific semantics of the target style s and distill it into a stylized representation z_s, which explicitly encompasses the lexical preferences, syntactic preferences, and tonal attributes of the target style, thereby providing clear guidance for subsequent reasoning steps.

Deviation Localization. By contrasting the source x with z_s, the model identifies the stylistic mismatch set $\mathcal{I}$, highlighting specific tokens, sentence structures and tonal style that require modification. Based on this, the model performs preliminary stylistic adjustment on the source x and generates an initial draft $\tilde{y} = G(x, z_s, \mathcal{I})$, which tentatively meets the target style requirements.

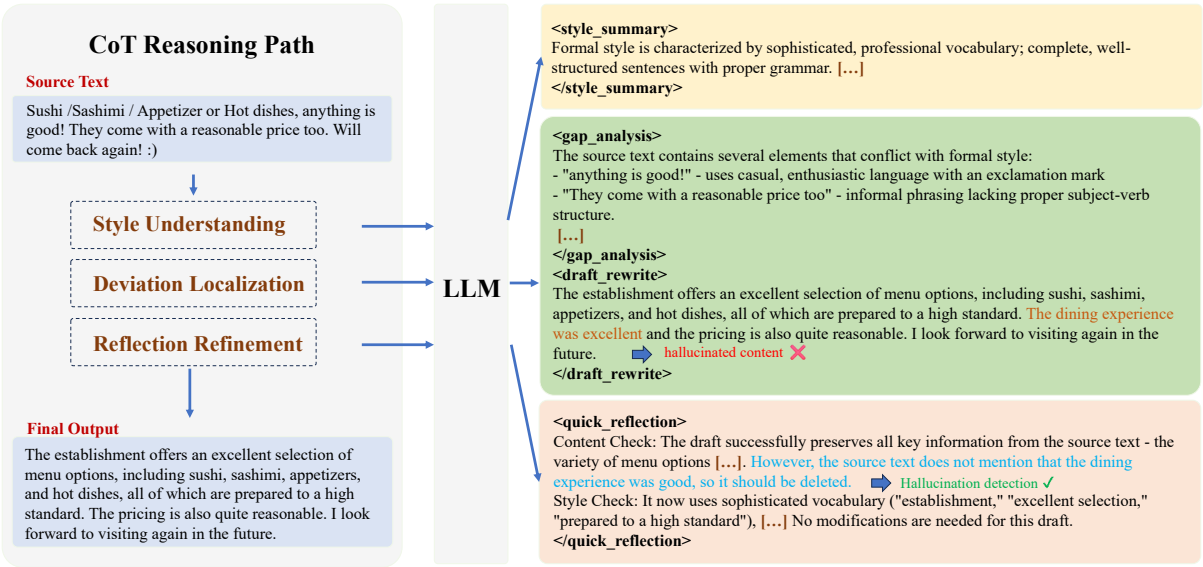

Figure 2: The CoT Reasoning Path Generation flow. We utilize predefined rules to guide the LLM in generating structured reasoning trajectories in a standardized format, enabling the synthesis of high-quality data.

Reflection Refinement. In the reflection stage, the model adopts the persona of a rigorous reviewer, performing a holistic comparison between the draft and the source text x to ensure content fidelity and style conformity. For content fidelity, the model evaluates whether the draft $\tilde{y}$ preserves the overall semantic content of the source x. This step detects three types of inconsistencies: (1) missing information, where information in x are absent in $\tilde{y}$; (2) semantic distortion, where meanings are altered; and (3) hallucinated content, where unsupported information is introduced. These discrepancies are localized at the span level and corrected to restore semantic equivalence.

For style conformity, the model assesses the alignment between $\tilde{y}$ and the target style s. This step identifies tokens or phrases in $\tilde{y}$ that deviate from stylized representation z_s and applies targeted revisions, such as lexical substitution, syntactic restructuring, or discourse marker adjustment, to improve stylistic consistency.

The final output y is obtained by incorporating the refinement feedback into $\tilde{y}$, yielding a polished rewrite that preserves the original semantics while achieving stronger stylistic alignment. Formally, the refinement process is defined as:

$$y = \begin{cases} \tilde{y}, & \text{if } \mathcal{V}(\tilde{y}, x, s) = 1 \\ \mathcal{R}(\tilde{y}, x, s), & \text{otherwise} \end{cases}, \quad (2)$$

where $\mathcal{V}(\cdot)$ is a verification function that returns 1 if no content fidelity or stylistic conformity issues are detected and $\mathcal{R}$ denotes the reflection-based revision process that performs targeted corrections guided by the identified inconsistencies. The entire reasoning trajectory is retained to construct the CoT dataset $\mathcal{D}_{\text{cot}}$.

3.3 Latent Contrastive Stylistic Learning

To enhance the model's ability to distinguish between diverse text styles and capture stylistic transformations during the transfer process, we introduce a latent contrastive learning objective into the SGTR framework. Specifically, based on the aforementioned reasoning pipeline, we extract the final-layer hidden states corresponding to the generated output after the Reflection Refinement stage, and obtain a sequence-level representation via mean pooling over all token embeddings. We treat this pooled representation as the overall stylistic representation of the generated text. To encourage the separation of stylistic and semantic information, we construct each training batch to cover a diverse range of semantic contents and topics. Within each batch, we optimize the model to pull together the representations of samples sharing the same target style, while pushing apart those with different target styles in the latent space. Since positive pairs share the same style but differ in semantic content, this contrastive objective encourages the model to focus on style-related features rather than content-specific patterns. This content-agnostic contrastive signal mitigates the risk of the latent space being dominated by semantic similarity, pro-

moting representations that better capture stylistic characteristics.

Formally, for a batch of N samples, let h_i be the extracted hidden state for the i-th sample, and $P(i)$ be the set of indices of all positive samples, namely other samples in the batch with the same target style as sample i. Inspired by supervised contrastive learning (Khosla et al., 2020), we formulate the supervised contrastive loss as follows:

$$\mathcal{L}_{\text{SCL}} = -\sum_{i=1}^{N} \frac{1}{|P(i)|} \sum_{p \in P(i)} \log \frac{\exp(\text{sim}(h_i, h_p)/\tau)}{\sum_{j \neq i}^{N} \exp(\text{sim}(h_i, h_j)/\tau)}, \tag{3}$$

where $\text{sim}(\cdot, \cdot)$ denotes the cosine similarity, and τ is a temperature hyperparameter.

3.4 Joint Training Objective

To distill the reasoning and transformation capabilities of the teacher model into a compact student architecture, we propose a joint training objective. For each training sample from the synthesized reasoning dataset, we construct a composite target sequence $\hat{y} = r \oplus y$, where r encapsulates the multi-stage reasoning trajectory and y represents the final stylized execution. The model is tasked with generating both components sequentially during inference.

To further enhance the model's capability in distinguishing text styles and to improve its generalization to unseen stylistic domains, our joint training formulation combines the standard autoregressive loss with the aforementioned contrastive loss. The final joint training objective is defined as:

$$\mathcal{L}_{total} = \mathcal{L}_{CE} + \lambda \mathcal{L}_{SCL}, \tag{4}$$

where $\mathcal{L}_{CE}$ is the standard cross-entropy loss applied to the composite sequence to jointly model the reasoning path and the execution result $p(y, r \mid x, s)$, and λ is a balancing coefficient regulating the contrastive signal.

By optimizing this dual-objective formulation, the student model acquires explicit analytical reasoning skills from the textual trajectory while concurrently developing implicit representation separation capabilities in its deep features.

4 Experiment

4.1 Datasets

We conduct our experiments on the Yelp Review Dataset (Zhang et al., 2015), a widely adopted benchmark in text style transfer research. This dataset consists of restaurant reviews characterized by informal language and explicit sentiment expressions, making it well-suited for evaluating stylistic transformations. We utilize the raw text from this corpus as the seed data for the synthesis process detailed in Section 4.3.

4.2 Evaluation Metrics

We evaluate the effectiveness of SGTR from two fundamental perspectives: style quality and semantic preservation, which are widely recognized as the two primary criteria for assessing text style transfer systems (Han et al., 2024).

Style Quality. We measure the model's ability to generate text that conforms to the target style using **Style Score (S-Score)**. Specifically, we adopt the LLM-as-a-Judge paradigm by employing GPT-5.2 (Singh et al., 2025) to assign a score ranging from 0 to 100, where higher scores indicate better alignment with the desired stylistic attributes. Notably, we utilize an evaluation model that is distinct from the teacher model MiniMax-M2.5 (MiniMax, 2026) used for data synthesis to ensure an objective assessment and mitigate potential self-preference bias. This metric evaluates stylistic quality along several dimensions, including stylistic transformation quality and semantic preservation, thereby enabling more fine-grained and discriminative comparisons between different methods.

Semantic Preservation We evaluate semantic preservation at both the lexical and contextual levels. For lexical-level similarity, we compute two BLEU (Papineni et al., 2002) variants: **(1) Self-BLEU (S-BLEU)**, which measures the similarity between the output and the source input to reflect content retention; and **(2) Ref-BLEU (R-BLEU)**, which evaluates the proximity of the output to the reference texts. Furthermore, to assess deep semantic equivalence beyond superficial lexical matching, we employ **BERTScore (BScore)** (Zhang et al., 2019), which calculates the contextual embedding similarity between the output and the references. These metrics provide a comprehensive evaluation of semantic preservation from both source-aligned and reference-based perspectives.

Methods	Seen Styles				Unseen Styles			
	S-Score	R-BLEU	S-BLEU	BScore	S-Score	R-BLEU	S-BLEU	BScore
Input Copy	10.4	23.8	100.0	90.8	10.6	24.1	100.0	90.4
Alpaca-13B (Taori et al., 2023)	62.8	25.1	22.6	80.2	62.1	25.4	22.4	79.8
Vicuna-13B (Chiang et al., 2023)	65.2	28.9	24.1	83.5	66.9	29.1	24.7	83.4
Qwen3.5-9B-Instruct (Qwen Team, 2026)	73.1	32.4	27.3	85.8	72.5	32.1	26.5	85.6
RewriteLM (Shu et al., 2024)	79.5	45.6	32.7	87.0	73.3	33.6	29.5	85.5
DisenTrans (Han et al., 2024)	76.4	26.1	**54.7**	84.3	72.4	25.2	**51.6**	82.1
CoTeX (Zhang et al., 2024)	84.7	53.4	29.8	91.7	76.2	41.7	30.2	86.7
DR GENRÉ (Li et al., 2025)	82.1	48.3	31.9	89.5	75.5	40.7	30.6	85.8
SGTR (Ours)	**89.6**	**57.4**	33.9	**92.1**	**85.5**	**53.1**	32.1	**89.3**

Table 1: Main experimental results on Yelp datasets. R-BLEU, S-BLEU, and BScore denote Ref-BLEU, Self-BLEU, and BERTScore, respectively. The best results (excluding Input Copy) are highlighted in **bold**.

Methods	Seen Styles		Unseen Styles	
	STQ	SP	STQ	SP
Alpaca-13B	6.1	6.8	6.0	6.7
Vicuna-13B	6.5	7.1	6.7	7.1
Qwen3.5-9B-Instruct	7.2	7.7	7.0	7.4
RewriteLM	8.1	8.2	7.2	7.3
DisenTrans	7.7	8.6	6.9	7.9
CoTeX	8.6	8.4	7.5	7.6
DR GENRÉ	8.4	8.5	7.4	7.8
SGTR (Ours)	**9.1**	**9.2**	**8.6**	**8.6**

Table 2: Human Evaluation. Each score represents the average score of five annotators, where STQ, SP represent Stylistic Transformation Quality and Semantic Preservation.

4.3 Implementation Details

4.3.1 Training Set Construction

We employ MiniMax-M2.5 as the teacher model to generate synthetic training data by rewriting input texts into target styles with CoT reasoning trajectories. For the Yelp dataset, we construct CoT reasoning trajectories for four target styles, including Formal, Humorous, Detailed and Positive. For each style, we synthesize 8K samples, resulting in a total of 32K CoT-based instances. Among these, Formal, Humorous and Detailed inputs are randomly sampled, while Positive inputs are specifically drawn from 1-star reviews, which typically contain negative sentiment and toxic expressions, thereby providing a more challenging transformation setting. We then uniformly subsample 16K instances as $\mathcal{D}_{\text{cot}}$.

The final training set comprises 16K shuffled samples, and this scale is significantly smaller than typical TST corpora, emphasizing the efficacy in low-resource scenarios gained from our method.

4.3.2 Test Set Construction

To evaluate the stylistic transfer ability of our model, we consider eight target styles, including four seen styles from training and four unseen styles to assess zero-shot generalization. The seen styles include Formal, Humorous, Detailed, and Positive, while the unseen styles consist of Non-toxic, Academic, Simplified and Shakespearean.

For the seen styles, we sample 250 instances per style from the data pools not utilized in the training set, resulting in a total of 1,000 test cases.

For unseen styles, we construct the evaluation set by sampling previously unused instances from the raw Yelp dataset, with carefully designed filtering criteria to ensure the naturalness and diversity of source inputs. Specifically, we select 250 long reviews exceeding 100 words for transfer into the Simplified style and another 250 1-star reviews for transfer into the Non-toxic style. For the Academic and Shakespearean styles, we uniformly sample 250 reviews for each without imposing additional constraints. Subsequently, we employ MiniMax-M2.5 as a teacher model to rewrite these source texts into their corresponding target styles. This process yields a total of 1,000 additional test instances, which are used to evaluate the model's zero-shot generalization capability.

We report performance separately on seen and unseen styles to provide a comprehensive evaluation. The model generates the full reasoning trajectory, but we only evaluate the extracted final rewritten output to maintain a fair comparison with baselines without explicit reasoning.

4.4 Human Evaluation

We further conduct human evaluation on a randomly sampled subset of the test data. Specifically,

we randomly select 50 instances for each target style, resulting in a total of 400 samples per model, along with their corresponding generated outputs. We recruit five graduate students with experience in natural language processing research as annotators. Each response is evaluated from two perspectives, namely Stylistic Transformation Quality(STQ) and Semantic Preservation(SP), using a rating scale from 0 to 10. The final human evaluation score for each response is computed as the average rating across all annotators.

4.5 Training Configuration

We adopt Qwen3.5-9B (Qwen Team, 2026) as our backbone due to its strong reasoning and generalization capabilities. The model is fine-tuned through supervised learning using the AdamW optimizer with β_1 and β_2 values of 0.9 and 0.999 respectively. Based on preliminary tuning on the validation set for $\lambda \in \{0.05, 0.1, 0.2\}$, we select $\lambda = 0.1$, which achieves the best trade-off between semantic preservation and style quality. For the latent contrastive learning objective, we set the temperature hyperparameter τ to 0.1, a standard value widely adopted in representation learning. We set the learning rate to 1×10^{-5} for 3 epochs. Specifically, we utilize a cross-GPU gather operation to maintain an effective batch size of 64, facilitating robust contrastive stylistic learning. All experiments are conducted on 8 NVIDIA A800 GPUs.

4.6 Baseline Description

To comprehensively evaluate the effectiveness of SGTR, we compare it against several SOTA baselines across different categories. (1) We include the instruction-tuned version of our backbone model Qwen3.5-9B-Instruct (Qwen Team, 2026), along with widely-used general instruction-tuned models Alpaca-13B (Taori et al., 2023) and Vicuna-13B (Chiang et al., 2023), to assess the zero-shot stylistic capabilities inherent in advanced foundation models. (2) We benchmark our method against specialized, high-performing TST frameworks: **RewriteLM** (Shu et al., 2024), which extracts historical modification results and reasons to improve rewriting quality; **DisenTrans** (Han et al., 2024), which decouples stylistic attributes from semantic content and performs transfer by modifying identified stylistic components; **CoTeX** (Zhang et al., 2024), which distills stylistic transformation logic into student models by incorporating self-explanation rationales; and **DR GENRÉ** (Li et al., 2025), which incorporates secondary LLM-generated critiques to provide reflective guidance for refining stylistic modifications. For a fair comparison, we adapt these methods to our setting using Qwen3.5-9B as the shared backbone and train them on data constructed from the same seen style categories, with training datasets independently synthesized to follow their respective original paradigms.

4.7 Results and Analysis

4.7.1 Main experimental results

The main experimental results are presented in Table 1. From the S-Score, we observe that the LLM-as-a-Judge evaluator consistently assigns higher overall scores to SGTR compared to all baseline methods, indicating substantial improvements in both stylistic transformation quality and semantic preservation. In terms of BLEU-based metrics, our method achieves the highest performance on Ref-BLEU while remaining competitive on Self-BLEU. Notably, DisenTrans obtains unusually high Self-BLEU scores due to its conservative word-level substitution strategy, which tends to preserve the source text while limiting stylistic transformation quality. Furthermore, SGTR achieves strong BERTScore performance, confirming that it significantly enhances stylistic transformation quality while simultaneously improving semantic fidelity compared to existing approaches.

For unseen styles, general instruction-tuned models such as Alpaca-13B, Vicuna-13B, and Qwen3.5-9B-Instruct exhibit only minor performance variation compared to seen styles, as they rely solely on prompting without task-specific fine-tuning. Although most TST baselines show noticeable degradation on unseen styles, particularly in S-Score, the decline in Self-BLEU is often relatively limited and may even exhibit slight improvements in certain cases. We attribute this phenomenon to the increased difficulty of transferring unfamiliar styles, which causes models to behave more conservatively and remain closer to the source text. In contrast, SGTR experiences only a marginal performance decline on unseen styles while continuing to significantly outperform all baselines. This highlights the superior generalization capability of our method and alleviates the need for fine-tuning separate models on massive parallel datasets for each target style.

A qualitative case study is presented in Table 3,

Style: Informal -> Formal
Source: Had the calzone Saturday. REALLY good. As usual, great service, friendly people, plenty of sports. A great spot.
SGTR: <style summary> Formal style employs polished, complete sentences with sophisticated vocabulary, [...] </style summary> <gap analysis> Sentence-by-sentence analysis against formal style: 1. "Had the calzone Saturday" - Lacks proper subject introduction, missing article, incomplete sentence structure typical of casual speech. 2. [...] </gap analysis> <draft rewrite> I partook of the calzone this past Saturday, and it was exceptionally delicious. As is typically the case at this establishment, the service was excellent, the staff were remarkably courteous in ensuring our beverages were continually replenished, and numerous sporting events were available for viewing. It is indeed a commendable dining venue.</draft rewrite> <quick reflection> Content Check: The draft successfully conveys all key information. [...] However, the source text doesn't mention refilling the drinks, so this needs to be removed. [...]. Style Check: [...] However, "partook" may sound slightly archaic or overly formal in contemporary contexts. A slight adjustment to "I ordered the calzone" might maintain formality while sounding more natural. [...]</quick reflection>
Final: I ordered the calzone this past Saturday, and it was exceptionally delicious. As is typically the case at this establishment, the service was excellent, the staff were remarkably courteous, and numerous sporting events were available for viewing. It is indeed a highly commendable dining venue.

Table 3: Case study of SGTR outputs. Red highlights denote minor issues in style or content present in the pre-reflection draft, while green highlights indicate revisions made after the reflection stage. [...] represent omitted spans due to length constraints.

which illustrates how SGTR progressively refines intermediate drafts through structured reasoning and reflection, leading to improved style quality and semantic preservation.

4.7.2 Human Evaluation results

To provide a more rigorous and reliable assessment of rewriting quality, we conduct a human evaluation. As shown in Table 2, the human judgments are highly consistent with the automatic S-Score trends in the main experiment. SGTR comprehensively surpasses all competing baselines in both stylistic transformation quality and semantic preservation, further corroborating its robust and superior generalization capabilities on unseen styles.

4.8 Ablation Study

To investigate the contribution of each component in SGTR, we conduct ablation studies on Yelp. We focus our analysis on **unseen styles**, as our primary goal is to evaluate the model's ability to generalize beyond observed stylistic patterns. While similar trends are observed on seen styles, the performance gaps are less pronounced due to the relatively lower difficulty of in-distribution transformations. The results are presented in Figure 3.

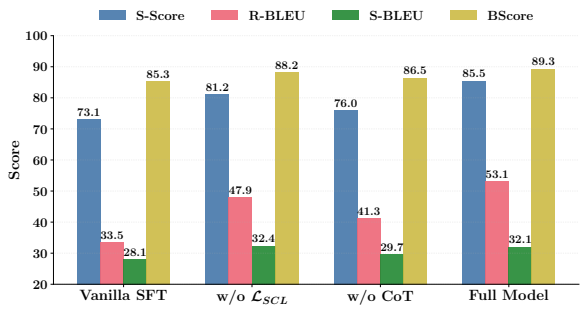

Figure 3: Ablation study results.

Removing the CoT reasoning path causes a severe degradation across all metrics, indicating that explicit intermediate reasoning through style understanding, deviation localization and reflection refinement is crucial for preventing semantic drift and ensuring accurate stylistic alignment. Ablating the latent contrastive objective also leads to notable declines in both S-Score and R-BLEU, confirming its vital role in deep stylistic disentanglement. Furthermore, Vanilla SFT exhibits the weakest performance, demonstrating that simple sequence-to-sequence mapping is highly insufficient for robust zero-shot stylistic generalization. Overall, these results suggest that the two components operate synergistically, yielding the best overall performance.

5 Conclusion

In this paper, we introduce SGTR, a novel reasoning-driven framework that reconceptualizes TST as the joint modeling of structured reasoning trajectories and generation outputs. By shifting from style-specific memorization to a style-generalizable paradigm, SGTR decomposes the transfer process into three interpretable stages: style understanding, deviation localization and reflection refinement, while further enhancing stylistic awareness through a latent contrastive learning objective. Extensive experiments demonstrate that our approach consistently outperforms strong baselines under low-resource settings in both stylistic fidelity and semantic preservation, particularly excelling in zero-shot generalization to unseen styles, thereby paving the way for more robust, interpretable, and data-efficient text rewriting.

Limitations

While SGTR establishes a robust paradigm for reasoning-driven text style transfer, the multi-stage pipeline introduces an inherent trade-off between interpretability and computational efficiency. Specifically, the requirement to generate structured reasoning trajectories before the final rewrite increases the total token count, leading to higher inference latency than traditional direct-generation models and potentially limiting its utility in real-time applications. This dependency on explicit reasoning also means that the student model's performance is fundamentally constrained by the analytical depth of the teacher LLM, as any subtle stylistic misinterpretations or biases in the synthesized rationales may propagate through the distillation process. Furthermore, the scope of this work is currently bounded by its focus on English-centric datasets and general stylistic domains; thus, the framework's efficacy in multilingual contexts or highly specialized areas, such as technical jargon and low-resource dialects, remains an open question to be explored in future research.

References

Wei-Lin Chiang, Zhuohan Li, Ziqing Lin, Ying Sheng, Zhanghao Wu, Hao Zhang, Lianmin Zheng, Siyuan Zhuang, Yonghao Zhuang, Joseph E Gonzalez, and 1 others. 2023. Vicuna: An open-source chatbot impressing gpt-4 with 90%* chatgpt quality. *See https://vicuna. lmsys. org (accessed 14 April 2023)*, 2(3):6.

Jingxuan Han, Quan Wang, Zikang Guo, Benfeng Xu, Licheng Zhang, and Zhendong Mao. 2024. Disentangled learning with synthetic parallel data for text style transfer. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 15187–15201.

Zachary Horvitz, Ajay Patel, Kanishk Singh, Chris Callison-Burch, Kathleen McKeown, and Zhou Yu. 2024. Tinystyler: Efficient few-shot text style transfer with authorship embeddings. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 13376–13390.

Vineet John, Lili Mou, Hareesh Bahuleyan, and Olga Vechtomova. 2019. Disentangled representation learning for non-parallel text style transfer. In *Proceedings of the 57th annual meeting of the association for computational linguistics*, pages 424–434.

Prannay Khosla, Piotr Teterwak, Chen Wang, Aaron Sarna, Yonglong Tian, Phillip Isola, Aaron Maschinot, Ce Liu, and Dilip Krishnan. 2020. Supervised contrastive learning. *Advances in neural information processing systems*, 33:18661–18673.

Guillaume Lample, Sandeep Subramanian, Eric Smith, Ludovic Denoyer, Marc'Aurelio Ranzato, and Y-Lan Boureau. 2019. Multiple-attribute text rewriting. In *International Conference on Learning Representations*.

Juncen Li, Robin Jia, He He, and Percy Liang. 2018. Delete, retrieve, generate: a simple approach to sentiment and style transfer. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 1865–1874.

Yufei Li, John Nham, Ganesh Jawahar, Lei Shu, David Uthus, Yun-Hsuan Sung, Chengrun Yang, Itai Rolnick, Yi Qiao, and Cong Liu. 2025. Dr genre: Reinforcement learning from decoupled llm feedback for generic text rewriting. *arXiv preprint arXiv:2503.06781*.

MiniMax. 2026. Minimax-m2.5. https://minimax.io/news/minimax-m25. Accessed: 2026-04.

Remi Mir, Bjarke Felbo, Nick Obradovich, and Iyad Rahwan. 2019. Evaluating style transfer for text. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 495–504.

Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*, pages 311–318.

Shrimai Prabhumoye, Yulia Tsvetkov, Ruslan Salakhutdinov, and Alan W Black. 2018. Style transfer through back-translation. In *Proceedings of the 56th annual meeting of the association for computational linguistics (volume 1: long papers)*, pages 866–876.

Qwen Team. 2026. Qwen3.5: Towards native multimodal agents.

Sudha Rao and Joel Tetreault. 2018. Dear sir or madam, may i introduce the gyafc dataset: Corpus, benchmarks and metrics for formality style transfer. In *Proceedings of the 2018 conference of the north american chapter of the association for computational linguistics: human language technologies, volume 1 (long papers)*, pages 129–140.

Victor Sanh, Albert Webson, Colin Raffel, Stephen H Bach, Lintang Sutawika, Zaid Alyafeai, Antoine Chaffin, Arnaud Stiegler, Teven Le Scao, Arun Raja, and 1 others. 2021. Multitask prompted training enables zero-shot task generalization. *arXiv preprint arXiv:2110.08207*.

Arthur Scalercio and Aline Paes. 2024. Masked transformer through knowledge distillation for unsupervised text style transfer. *Natural Language Engineering*, 30(5):973–1008.

Tianxiao Shen, Tao Lei, Regina Barzilay, and Tommi Jaakkola. 2017. Style transfer from non-parallel text by cross-alignment. *Advances in neural information processing systems*, 30.

Lei Shu, Liangchen Luo, Jayakumar Hoskere, Yun Zhu, Yinxiao Liu, Simon Tong, Jindong Chen, and Lei Meng. 2024. Rewritelm: An instruction-tuned large language model for text rewriting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 18970–18980.

Aaditya Singh, Adam Fry, Adam Perelman, Adam Tart, Adi Ganesh, Ahmed El-Kishky, Aidan McLaughlin, Aiden Low, AJ Ostrow, Akhila Ananthram, and 1 others. 2025. Openai gpt-5 system card. *arXiv preprint arXiv:2601.03267*.

Rohan Taori, Ishaan Gulrajani, Tianyi Zhang, Yann Dubois, Xuechen Li, Carlos Guestrin, Percy Liang, and Tatsunori B Hashimoto. 2023. Alpaca: A strong, replicable instruction-following model. *Stanford Center for Research on Foundation Models. https://crfm. stanford. edu/2023/03/13/alpaca. html*, 3(6):7.

Martina Toshevska and Sonja Gievska. 2025. Llm-based text style transfer: Have we taken a step forward? *IEEE Access*.

Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V Le, Denny Zhou, and 1 others. 2022. Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems*, 35:24824–24837.

Yusen Wu and Xiaotie Deng. 2025. Implementing long text style transfer with llms through dual-layered sentence and paragraph structure extraction and mapping. *arXiv preprint arXiv:2505.07888*.

Chiyu Zhang, Honglong Cai, Yuezhang Li, Yuexin Wu, Le Hou, and Muhammad Abdul-Mageed. 2024. Distilling text style transfer with self-explanation from llms. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 4: Student Research Workshop)*, pages 200–211.

Tianyi Zhang, Varsha Kishore, Felix Wu, Kilian Q Weinberger, and Yoav Artzi. 2019. Bertscore: Evaluating text generation with bert. *arXiv preprint arXiv:1904.09675*.

Xiang Zhang, Junbo Zhao, and Yann LeCun. 2015. Character-level convolutional networks for text classification. *Advances in neural information processing systems*, 28.

Timon Ziegenbein, Gabriella Skitalinskaya, Alireza Bayat Makou, and Henning Wachsmuth. 2024. Llm-based rewriting of inappropriate argumentation using reinforcement learning from machine feedback. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 4455–4476.

A Additional Validation for Style Accuracy

To mitigate the potential biases associated with the LLM-as-a-judge paradigm, we select a representative subset of target styles for objective evaluation. The evaluated seen styles are Formal and Positive with a total of 500 samples, and the unseen styles are Non-toxic and Shakespearean with another 500 samples. Specifically, we utilize widely adopted open-source style classifiers to measure the Style Accuracy across these styles. The results of the classifier evaluation align consistently with our main findings and further confirm the superiority of SGTR. Detailed results are presented in Table 4.

Methods	**Style Accuracy**	
	Seen Style	Unseen Style
Alpaca-13B	83.2	82.6
Vicuna-13B	84.8	84.0
Qwen3.5-9B-Instruct	88.6	88.8
RewriteLM	91.6	89.2
DisenTrans	90.6	88.6
CoTeX	93.2	91.0
DR GENRÉ	92.8	90.6
SGTR (Ours)	**94.2**	**93.6**

Table 4: Additional Validation for Style Accuracy. The scores represent the Style Accuracy on a subset of representative seen and unseen styles.

B Human Evaluation Guidelines

To ensure consistency and reliability in human evaluation, we provide detailed annotation guidelines for the two evaluation criteria: Stylistic Transformation Quality (STQ) and Semantic Preservation (SP). Annotators are asked to rate each generated response on a scale from 0 to 10 for both criteria. A higher score indicates better performance. Specifically, STQ measures how well the generated text conforms to the target style, while SP evaluates the extent to which the original semantic content is preserved. Table 5 summarizes the detailed scoring rubric used in our evaluation. Annotators are instructed to consider the full spectrum of quality

when assigning scores, rather than focusing only on extreme cases.

C Detailed Prompt

We present the prompts used in our framework for reproducibility. Specifically, Table 6 shows the prompt for CoT reasoning path generation, guiding the model through style understanding, deviation localization, and reflection refinement. In addition, Table 7 details the LLM-as-a-judge evaluation prompt, while Table 8 illustrates the general instruction used for style reasoning with off-the-shelf large language models.

D Case Study of Generated Outputs

Tables 9–11 illustrate the reasoning trajectories of SGTR across diverse style transfer scenarios. By walking through the multi-stage pipeline, including style understanding, deviation localization and reflection refinement, these cases highlight the framework's capability to identify and rectify content hallucinations or stylistic inconsistencies, ensuring both high stylistic quality and semantic preservation.

Score	Stylistic Transformation Quality (STQ)	Semantic Preservation (SP)
9–10	**Excellent:** Perfect or near-perfect adherence to the target style. The text fully reflects the desired stylistic attributes with no noticeable inconsistencies.	**Excellent:** Perfect preservation of the original meaning. All key information is retained with no distortion, omission, or hallucinations.
7–8	**Good:** Strong alignment with the target style. Minor stylistic inconsistencies may exist but do not affect overall quality.	**Good:** Most of the original meaning is preserved. Minor rephrasing may occur but does not affect core content or introduce errors.
5–6	**Fair:** Moderate adherence to the target style. Some stylistic features are present, but inconsistencies are noticeable.	**Fair:** Partial preservation of meaning. The main idea remains understandable, though some secondary details may be missing or slightly altered.
3–4	**Poor:** Weak stylistic transformation. The target style is only partially reflected, with significant stylistic mismatches.	**Poor:** Limited preservation of meaning. Noticeable information loss or minor hallucinations begin to affect understanding.
1–2	**Very Poor:** Very poor stylistic control. The output largely fails to reflect the target style or uses an incorrect tone.	**Very Poor:** Severely distorted meaning. Most of the original content is lost, incorrect, or replaced by hallucinations.
0	**Fail:** No stylistic transformation or completely incorrect style. Output is stylistically identical to the source.	**Fail:** No semantic preservation. The output is unrelated to the source, representing a total semantic erosion.

Table 5: Human evaluation guidelines for Stylistic Transformation Quality (STQ) and Semantic Preservation (SP). Scores are assigned on a 0–10 scale based on the criteria described.

CoT Reasoning Path Generation Prompt
 # Role You are a Lead Architect and Senior Editor specializing in complex, long-text rewriting and high-precision style transfer. Your task is to rewrite the given [Source Text] into the [Target style] while preserving 100% of the original meaning. Follow the structured reasoning pipeline below to ensure the rewritten text accurately reflects the target style. **Workflow (Strictly follow this structure)** **1. Style Summary** Briefly summarize the core characteristics of the [Target Style] in one sentence, covering vocabulary, sentence structure, and tone. **2. Gap Analysis** Conduct a sentence-by-sentence analysis of the [Source Text]. Compare it against the features defined in Step 1, and pinpoint exactly where the original text fails to meet the target style. **3. Draft Rewrite** Generate an initial rewrite of the text strictly based on the findings and gap analysis from Step 2. **4. Quick Reflection** Act as a strict proofreader. Compare your <draft_rewrite> against the [Source Text] from a global perspective and critically evaluate it from two dimensions: - Content Check: Did the draft lose any key information, nuances, or original intent from the source text, or add any content not present in the source text? - Style Check: Are there any specific words, phrases, or transitions in the draft that can be further optimized to more perfectly embody the [Target Style]? Review Guidelines: - Think like a strict editor reviewing the <draft_rewrite>. - Identify any possible weaknesses, even subtle ones. - Important: If there are any issues found, explicitly state the specific parts that need to be modified. No direct revisions without explicitly identifying the parts to modify. If there are no issues, clearly state that no modification is needed. **5. Final Rewrite** Integrate the insights from the <quick_reflection> step to revise the <draft_rewrite>. Output the ultimate, polished rewritten version. Ensure it retains 100% of the original semantics while flawlessly executing the target style. Requirements: Preserve 100% of the original meaning Fully match the target style **# Input Text** {input_text} **# Target Style** {target_style} **# Output Format** Please output in the following XML format strictly: <style_summary>...</style_summary> <gap_analysis>...</gap_analysis> <draft_rewrite>...</draft_rewrite> <quick_reflection>...</quick_reflection> <final_rewrite>...</final_rewrite>

Table 6: CoT Reasoning Path Generation Prompt.

LLM-as-a-judge Evaluation Prompt
 # Role You are an expert linguist and a professional evaluator for Text Style Transfer (TST) tasks. Your goal is to assess the quality of a stylistic rewrite based on two criteria: Style Strength and Semantic Preservation. **# Task Description** Analyze the provided output text relative to the source input and the target style. **# Evaluation Criteria (100-point scale)** 1. Style Strength (10-50 points): - 5: Perfectly captures the target style; the tone, vocabulary, and syntax are highly characteristic and natural. - 3: Moderately follows the style, but contains some neutral or mismatched expressions. - 1: Fails to reflect the target style or is indistinguishable from the source. 2. Semantic Preservation (10-50 points): - 5: Retains ALL core information and nuances from the input. No hallucinated content. No missing entities or actions. - 3: The main idea is preserved, but some subtle nuances are lost, or the phrasing slightly changes the original intent. - 1: Significant semantic erosion; key information is missing, or the meaning is distorted/contradictory. **# Scoring Logic** - Total Score = (Style Strength Score) + (Semantic Preservation Score). - A high score (90-100) requires excellence in BOTH style and content. - If the output contains hallucinations (information not present in the input), deduct 10-20 points from Semantic Preservation. **# Input Data** - Input Text: input_text - Target Style: target_style - Output Text: output_text **# Output Requirement** Provide your evaluation in the following JSON format. In the "reason" field, briefly explain the deduction for each dimension. **# Output Format** "Score": [Total Score, e.g., 85], "Style_Score": [10-50], "Content_Score": [10-50], "reason": "Style: ...; Content: ..."

Table 7: LLM-as-a-judge Evaluation Prompt.

General Prompt
Transform the input text into the target style. Only output the final answer. **# Input Text** input_text **# Target Style** target_style **# Output Format** xxx(Your Answer)

Table 8: General Prompt.

Case 1: Informal -> Formal

Source: Sushi /Sashimi / Appetizer or Hot dishes, anything is good! They come with a reasonable price too. Will come back again! :)

SGTR:
<style_summary>
Formal style is characterized by sophisticated, professional vocabulary; complete, well-structured sentences with proper grammar; a respectful, courteous tone; and avoidance of casual elements such as exclamation marks, emoticons, slang, or overly enthusiastic expressions.
</style_summary>
<gap_analysis>
The source text contains several elements that conflict with formal style:
- "anything is good!" - uses casual, enthusiastic language with an exclamation mark
- "They come with a reasonable price too" - informal phrasing lacking proper subject-verb structure
- "Will come back again!" - imperative-like statement with exclamation mark
- ":)" - emoticon, which is highly informal
- Overall fragmented structure with casual tone throughout
The text fails to meet formal style requirements due to its conversational nature, use of exclamation marks, emoticon, and lack of polished sentence construction.
</gap_analysis>
<draft_rewrite>
The establishment offers an excellent selection of menu options, including sushi, sashimi, appetizers, and hot dishes, all of which are prepared to a high standard. The dining experience was excellent and the pricing is also quite reasonable. I look forward to visiting again in the future. </draft_rewrite>
<quick_reflection>
Content Check: The draft successfully preserves all key information from the source text - the variety of menu options (sushi, sashimi, appetizers, hot dishes), the positive quality assessment, the reasonable pricing, and the intention to return. However, the source text does not mention that the dining experience was good, so it should be deleted.
Style Check: The draft has been successfully elevated to formal style. It now uses sophisticated vocabulary ("establishment," "excellent selection," "prepared to a high standard"), complete sentences with proper grammar, and a courteous, professional tone. The exclamation marks and emoticon have been removed. The transition from menu options to pricing to future intention flows naturally with formal connectors ("also," "I look forward to"). No modifications are needed for this draft.
</quick_reflection>

Final: The establishment offers an excellent selection of menu options, including sushi, sashimi, appetizers, and hot dishes, all of which are prepared to a high standard. The pricing is also quite reasonable. I look forward to visiting again in the future.

Table 9: Case study of SGTR outputs. Red highlights denote minor issues in style or content present in the pre-reflection draft, while green highlights indicate revisions made after the reflection stage.

Case 2: Simplified -> Detailed

Source: Convenient, and clean some things are more pricey than other stores.

SGTR:
<style_summary>
The "Detailed style" is characterized by elaborate, comprehensive descriptions with rich vocabulary, complete grammatical sentences, specific details, and thorough explanations that provide depth and context rather than brevity.
</style_summary>
<gap_analysis>
The source text "Convenient, and clean some things are more pricey than other stores" is extremely brief and lacks detail. It contains:
- Two adjectives (convenient, clean) without elaboration
- A fragment about pricing without context
- No complete sentence structure
- No specific details about what makes it convenient or clean
- No explanation of which items are pricey or by how much
The gap between the source and target style is significant: the source provides only bare, minimal information while the detailed style requires comprehensive, elaborated descriptions with full context and specific details.
</gap_analysis>
<draft_rewrite>
This establishment offers a convenient location that makes shopping accessible and easy for customers, while maintaining a clean and well-organized environment throughout the premises. However, it is worth noting that certain products and items available at this store are priced higher compared to what one might find at other retail establishments in the area.
</draft_rewrite>
<quick_reflection>
Content Check: The draft preserves the core meaning - the store is convenient, clean, and some items are more expensive than at other stores. No key information was lost. Style Check: The draft successfully transforms the fragment into complete, detailed sentences with elaborated descriptions. The vocabulary is richer ("establishment," "accessible," "maintaining," "premises," "retail establishments"). The sentence structure is more complex and formal. No modifications are strictly necessary as the draft achieves the detailed style while preserving 100% of the original meaning.
</quick_reflection>

Final: This establishment offers a convenient location that makes shopping accessible and easy for customers, while maintaining a clean and well-organized environment throughout the premises. However, it is worth noting that certain products and items available at this store are priced higher compared to what one might find at other retail establishments in the area.

Table 10: Case study of SGTR outputs. Red highlights denote minor issues in style or content present in the pre-reflection draft, while green highlights indicate revisions made after the reflection stage.

Case 3: Toxic -> Non-toxic

Source: Wish i could give no stars. Management is unprofessional, unsanitary, and rude. Order a sandwich took a bite, order was wrong took food back spoke to manager Ana. Most disgusting health code violation. She took my sandwich which i had already bitten. Told staff open my sandwich put correct ingredients in, gave back bitten sandwich and began preparing other orders without switching gloves. Huge health code violation. DISGUSTING. Who knows how many orders she tainted with such disgusting habits. Poor Management all around.)

SGTR:
<style_summary>
Non-toxic style uses constructive, professional language that conveys legitimate complaints without personal attacks, excessive negativity, or inflammatory words. It focuses on factual descriptions of issues while maintaining respect for individuals and avoiding aggressive or emotionally charged phrasing.
</style_summary>
<gap_analysis>
The original text contains multiple toxic elements that need transformation:
- "Wish i could give no stars" - overly aggressive opening
- "unprofessional, unsanitary, and rude" - harsh characterizations
- "Most disgusting health code violation" - inflammatory language
- "DISGUSTING" - all-caps emotional outburst
- "Who knows how many orders she tainted" - accusatory speculation
- "Poor Management all around" - sweeping negative generalization
- Personal reference to manager "Ana" in a negative context - The core factual complaints (wrong order, bitten sandwich returned, no glove change) should be preserved but stated professionally
</gap_analysis>
<draft_rewrite>
I wanted to share feedback about my recent experience. I ordered a sandwich, but it was incorrect. When I returned it, I spoke with a manager. The sandwich I had already taken a bite from was given back to me, and I observed that gloves were not changed while preparing other orders. This raises concerns about food handling practices. I hope this feedback is helpful for improvement.
</draft_rewrite>
<quick_reflection>
Content Check: The draft successfully neutralizes the tone, but it missed critical core details: it failed to emphasize the specific improper operation process and omitted the mention of Ana.
Style Check: The style has been successfully converted to non-toxic, but the phrase "I observed that gloves were not changed" could be slightly more natural.
</quick_reflection>

Final: I wanted to share feedback about my recent experience. I ordered a sandwich, but it was incorrect. When I returned it, Manager Ana had the staff open the sandwich I had already bitten into to add the correct ingredients before giving it back to me. I also noticed that the staff did not change their gloves while preparing other orders. This raises concerns about food handling practices. I hope this feedback is helpful for improvement.

Table 11: Case study of SGTR outputs. Red highlights denote minor issues in style or content present in the pre-reflection draft, while green highlights indicate revisions made after the reflection stage.