

CEMT: Controllable Element-Oriented Machine Translation via Structured Linguistic Reasoning

Lingling Shi^{1,†} Haoyu Jin^{1,†} Ruiyu Fang¹ Shuangyong Song¹
Jinsong Su³ Yongxiang Li¹ Xuelong Li^{2,*}

¹Xingchen AGI Lab, China Telecom Artificial Intelligence Technology (Beijing) Co., Ltd

²Institute of Artificial Intelligence (TeleAI), China Telecom

³School of Informatics, Xiamen University

shill2@chinatelecom.cn, jinhaoyu@bupt.edu.cn, fangry@chinatelecom.cn, songshy@chinatelecom.cn,

jssu@xmu.edu.cn, liyx25@chinatelecom.cn, xuelong_li@ieee.org

Abstract

Large Language Models have shown strong performance in Machine Translation, yet they often suffer from paraphrasing errors, omissions, or hallucinations when the input contains translation-specific elements (e.g., URLs, slang, and idioms) that require strict preservation or controlled transformation, undermining the reliability of critical details. We propose CEMT¹, a **Controllable Element-Oriented Machine Translation** framework inspired by the analysis–strategy–generation paradigm in human translation. CEMT first employs an Element Detection Module to identify translation-specific elements, and then introduces a Translation Module that decomposes the translation process into linguistically grounded analysis, strategy formulation, and final generation, thereby guiding the reliable translation of these elements. We further introduce a CoT Judge model during training that provides step-wise supervision over the accuracy and consistency of the translation process. On the WMT23/24 Chinese↔English benchmarks, CEMT improves performance over existing Machine Translation models while significantly reducing element-level constraint violations.

1 Introduction

Machine Translation (MT) is a central task in Natural Language Processing (NLP), and recent years have seen remarkable progress driven by advances in Large Language Models (LLMs) (Brown et al., 2020; Ouyang et al., 2022; Hendy et al., 2023; Kocmi et al., 2023; Jiao et al., 2023; Gao et al., 2024; Chen et al., 2025; Zheng et al., 2025).

However, LLMs still face challenges regarding quality instability and decision transparency when handling translation-specific elements such

[†]Equal contribution.

^{*}Corresponding author.

¹Code is available at <https://github.com/ther7777/CEMT>.

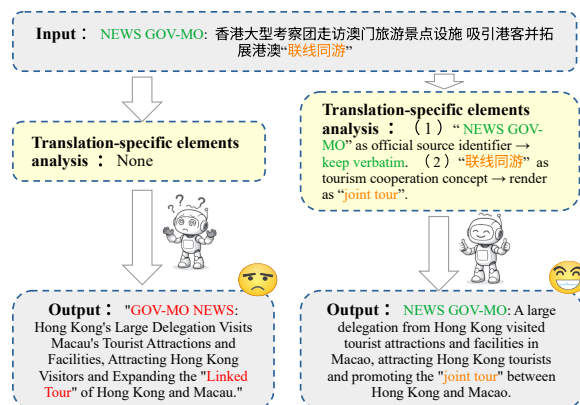


Figure 1: MT errors often arise from mishandling fine-grained expressions, undermining overall reliability.

as slang, URLs, or particular tokens. As shown in Figure 1, MT systems often struggle with such elements. Song and Xu (2024) indicate that such low-frequency components, due to their syntactic and semantic irregularities, are a primary bottleneck for current MT systems and a major source of hallucinations and errors (Guerreiro et al., 2023; Dale et al., 2023). Meanwhile, prior MT research has shown that explicit linguistic information can enhance translation quality (Sennrich and Haddow, 2016; Eriguchi et al., 2016; Chen et al., 2017; Marcheggiani et al., 2018). In contrast, the current LLM paradigm relies largely on implicit internal representations, which limits the interpretability of the translation process and leads to unpredictable handling of challenging linguistic elements. These observations motivate us to revisit whether explicit linguistic signals and structured translation reasoning can improve the reliability of LLM-based MT.

In addition, reasoning-enhanced methods based on Chain-of-Thought (CoT) and Reinforcement Learning (RL) (Wei et al., 2022; DeepSeek-AI et al., 2025; Li et al., 2025b; Wu et al., 2025b), such as DRT (Wang et al., 2025a), TEaR (Feng et al., 2025b), and ExTrans (Wang et al., 2025c),

have been widely applied to MT to improve analytical depth and reasoning transparency. However, the reward design of existing methods usually focuses on translation quality and output format, which introduces a credit assignment problem (Sutton and Barto, 2018), leaving it unclear whether the model genuinely engages in meaningful reasoning or merely exhibits formalistic thinking (Yue et al., 2025).

To address these challenges, inspired by the collaborative intelligence of AI Flow (An et al., 2025), we propose CEMT, which reframes MT from a purely generative task into a more controllable decision process by explicitly modeling key translation elements and introducing linguistically grounded reasoning. Specifically, CEMT first identifies translation-specific elements that require special handling, and performs translation via a staged reasoning process that integrates semantics & pragmatics, argument structure, syntactic structure, and strategy formulation, in which these elements are explicitly tracked and handled throughout the translation pipeline. In addition, to mitigate shallow or template-driven reasoning behaviors, CEMT incorporates process-level supervision during training, assessing the accuracy and fidelity of intermediate reasoning steps rather than relying solely on final translation quality.

The main contributions are as follows:

- **Translation-Specific Elements Modeling.** We introduce an Element Detection Module (EDM) that explicitly identifies Do-Not-Translate (DNT) elements and pragmatic cues, providing constraints to guide translation.
- **Dual-Track Linguistically Grounded CoT Reasoning.** We design a dual-track, linguistically grounded reasoning paradigm that integrates a basic linguistic analysis track with an element-oriented track, achieving a balance between overall semantic consistency and fine-grained element accuracy through a four-stage reasoning process.
- **Process-Level Reasoning Supervision.** We develop a CoT Judge that provides process-level supervision, including semantic-level, element-level, and global consistency rewards, encouraging that the generated reasoning traces are structurally complete, semantically faithful, and well aligned with the underlying translation decisions.

On the WMT23/24 (Kocmi et al., 2023, 2024) Chinese \leftrightarrow English benchmarks, CEMT, built upon Qwen2.5-7B (Yang et al., 2024) surpasses strong open-source baselines such as Hunyuan-MT-7B (Zheng et al., 2025) and substantially improves the handling of translation-specific elements. These results indicate that explicit constraints, linguistically structured reasoning, and process-level supervision together provide an effective pathway toward more reliable and interpretable MT systems.

2 The CEMT Framework

To achieve fine-grained control in LLM-based MT, we propose CEMT. As shown in Figure 2, CEMT consists of an EDM and a Translation Module. Given a source text, the EDM first performs an explicit analysis to identify translation-specific elements and produces a structured translation constraint set. Subsequently, this constraint set drives the Translation Module to execute a structured CoT analysis: first performs basic linguistic understanding and then formulates a translation strategy based on this information. To encourage adherence to this structured reasoning process, we employ SFT + RL training paradigm. During RL training, we further introduce a CoT Judge to construct a reward function based on the model’s reasoning process.

2.1 Element Detection Module

As the initial step of CEMT, the EDM performs an explicit analysis of the source text to identify translation-specific elements. We formalize this information as a translation constraint set, denoted as EA :

$$EA = (S_{comp}, E_{dnt}, E_{prag}) \quad (1)$$

where S_{comp} denotes the sentence complexity, categorized as {Complete Simple, Complete Complex, Incomplete}, and serves as a reasoning intensity signal that controls the depth of linguistic analysis. E_{dnt} denotes the set of DNT elements to be preserved verbatim (e.g., URLs, platform-specific tags, and emojis). E_{prag} denotes the set of pragmatic elements requiring contextual analysis prior to cross-lingual reconstruction (e.g., culture-specific expressions, idioms, and slang).

The EDM is initialized with the Qwen2.5-7B model, fine-tuned via SFT on approximately 4,000 augmented annotated samples. During the subsequent SFT and RL stages of the Translation Module, the EDM is pre-trained and completely frozen,

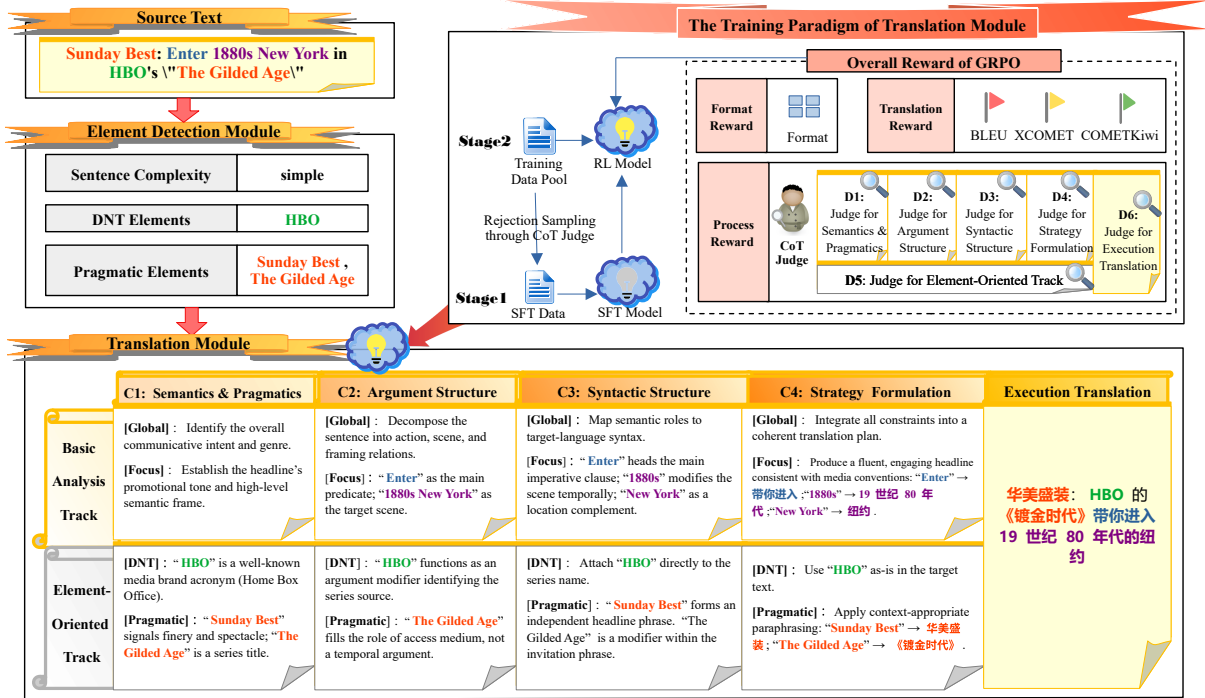


Figure 2: The overall architecture of the CEMT framework. The Inference Pipeline (Left) depicts the translation flow from Source Text \rightarrow EDM \rightarrow Translation Module. The Training Paradigm (Right) details the two-stage SFT and RL process of the Translation Module, highlighting the role of the CoT Judge.

providing fixed constraint signals without participating in gradient updates.

2.2 Translation Module

The Translation Module serves as the core component conditioned on the constraint set EA produced by the EDM. It performs translation by generating a linguistically grounded CoT under the specified constraints. To facilitate interpretability and process-level supervision, the structured reasoning process consists of four sequential stages: C1 (Semantics & Pragmatics) captures the overall meaning and pragmatic context. C2 (Argument Structure) clarifies the argument predicate structure. C3 (Syntactic Structure) provides necessary structural constraints. C4 (Strategy Formulation) integrates the information from the previous stages to guide generation.

To balance global linguistic coherence and fine-grained element control, we adopt a dual-track analysis structure that operates across all four reasoning stages: (1) Basic Analysis Track: performs general linguistic analysis and adaptively adjusts the depth of reasoning according to the sentence complexity label S_{comp} . (2) Element-Oriented Track: focuses on constraint-specific reasoning based on explicit immutability constraints from E_{dnt} and pragmatic

transformation requirements from E_{prag} .

2.3 CoT Judge

The CoT Judge is designed to audit the CoT generated by the Translation Module. It serves as a data filter during the SFT stage and as a process-level reward signal during the RL stage. The judge evaluates the reasoning process with respect to three global constraints that apply across all reasoning steps: (1) reasoning depth must be consistent with the complexity level indicated by S_{comp} ; (2) the dual-track analysis should be strictly followed, prohibiting cross-track drift; (3) the handling of E_{dnt} should follow the verbatim preservation rule, while the translation of E_{prag} should follow the context-function-equivalence logic.

Based on these global constraints, the judge employs a six-dimensional scoring scheme with clearly delineated roles: D1–D3 target the Basic Analysis Track at steps C1–C3, evaluating analytical precision to assign raw scores s_1 , s_2 , and s_3 . Subsequently, D4 evaluates the Basic Analysis Track at the C4 stage, determining whether the formulated strategy constitutes a logically consistent product of the initial analysis, yielding the score s_4 . In parallel, D5 focuses on the Element-Oriented Track throughout the process, evaluating the com-

pliance of handling plans for DNT or pragmatic elements, denoted as s_5 . Finally, D6 evaluates the coherence between the complete reasoning chain and the generated translation, assessing global logical and constraint-level consistency to assign s_6 .

All raw evaluation scores are discrete integers $s_i \in \{1, 2, 3, 4, 5\}$. The analysis quality reward is defined as the mean normalized score over the first five dimensions. To explicitly penalize violations in execution faithfulness, we define a discrete execution penalty based on s_6 :

$$R_{analysis} = \frac{1}{5} \sum_{i=1}^5 \hat{s}_i, \quad \text{where } \hat{s}_i = \frac{s_i - 1}{4} \quad (2)$$

$$P_{exec} = \begin{cases} 0.0, & \text{if } s_6 = 5 \\ -0.2, & \text{if } s_6 = 4 \\ -0.4, & \text{if } s_6 = 3 \\ -0.8, & \text{if } s_6 = 2 \\ -1.5, & \text{if } s_6 = 1 \end{cases} \quad (3)$$

$$R_{proc} = R_{analysis} + P_{exec} \quad (4)$$

For compactness, we denote the CoT Judge process reward as R_{proc} . The discrete mapping in P_{exec} encodes the severity levels of D6. Scores $s_6 \in \{5, 4, 3\}$ indicate valid execution with decreasing coherence, so we apply small step penalties. We assign a larger penalty to $s_6 = 2$ to mark a major mismatch between the reasoning chain and the final translation. When $s_6 = 1$, we apply the strongest penalty to discourage unfaithful execution. This shaping prevents high analysis scores from masking execution failures within R_{proc} .

The reward design enforces a strict separation between analytical quality and execution faithfulness, reflecting the principle that correct reasoning is only valuable when faithfully executed.

2.4 Training paradigm

To enable the Translation Module to stably generate structured CoT and execute element-level controllable translation, we adopt a two-stage SFT+RL training paradigm.

Stage 1: SFT. This stage enables the model to learn the linguistic reasoning paradigm. The core lies in constructing a high-quality CoT dataset. We employ a generate-evaluate-filter pipeline: for each sample (x, EA) , we use Qwen3-235B to generate $n = 3$ candidate CoTs. Then, using the CoT Judge, we calculate the process reward score R_{proc} for

each candidate, retaining only the highest-scoring CoT that exceeds a threshold of 0.85.

Stage 2: RL. This stage uses GRPO (Shao et al., 2024) to refine CoT quality and translation consistency. Unlike the SFT stage, GRPO also utilizes low-quality samples that represent difficult scenarios, encouraging RL to learn broader complex patterns. We construct a composite reward integrating assessments of output format, reasoning process, and final translation quality.

- R_{format} (Format Reward): A binary validity flag for format compliance ($R_{format} \in \{0, 1\}$); invalid-format outputs receive $R_{total} = -1$.
- R_{proc} (Process Reward): Calculated only when the format is valid, derived from the CoT Judge defined in Section 2.3.
- R_{trans} (Translation Reward): Calculated only when the format is valid, based on automated translation evaluation metrics, including the scores from BLEU (Papineni et al., 2002) (S_{bleu}), COMETKiwi-23-XL (Rei et al., 2023) (S_{kiwi}), and XCOMET-XL (Guerreiro et al., 2024) (S_{xcm}).

We define the translation reward as a weighted combination of translation-level metrics:

$$R_{trans} = w_{bleu} S_{bleu} + w_{kiwi} S_{kiwi} + w_{xcm} S_{xcm} \quad (5)$$

$$R_{total} = \begin{cases} w_{proc} R_{proc} + R_{trans}, & \text{if } R_{format} = 1 \\ -1, & \text{if } R_{format} = 0 \end{cases} \quad (6)$$

Here, w_{bleu} , w_{kiwi} , and w_{xcm} are scalar weights that balance the contributions of different translation metrics, and w_{proc} controls the relative importance of the process reward. We report the concrete values in Section D.3.

3 Experiments

This section validates the effectiveness of CEMT.

3.1 Experimental Setup

Datasets and Metrics. We focus on ZH↔EN translation, using publicly released WMT test sets from 2017–2022 as additional parallel training data, comprising 11,560 ZH→EN and 9,860 EN→ZH sentence pairs. Evaluation is conducted on the WMT23 ZH→EN test set (1,976 samples) and

Model	ZH→EN				EN→ZH			
	BLEU	COMETKiwi	XCOMET	Avg	BLEU	COMETKiwi	XCOMET	Avg
Closed-Source General Purpose LLMs								
Claude 3.5 Sonnet [‡]	22.55	71.69	87.32	60.52	38.63	70.39	78.24	62.42
GPT-4o [‡]	22.57	71.63	87.22	60.47	41.13	69.01	75.43	61.86
Gemini 1.5 Pro [‡]	18.34	69.23	85.55	57.71	39.82	67.47	76.26	61.18
Open-Source General Purpose LLMs								
Qwen2.5-7B-Base [†]	21.35	69.42	85.00	58.59	32.67	61.75	73.40	55.94
TeleChat2-7B [†]	18.35	69.06	85.85	57.75	29.56	67.65	74.29	57.17
Qwen2.5-7B-Instruct [†]	21.36	68.44	84.69	58.16	37.11	66.16	72.74	58.67
Qwen2.5-72B-Instruct [‡]	21.96	70.95	87.07	59.99	39.29	69.04	76.97	61.77
LLaMA-3.1-70B-Instruct [‡]	25.19	70.43	86.21	60.61	39.82	68.05	75.17	61.01
Open-Source Translation Models (Non-Qwen2.5-7B Backbones)								
NLLB-200-3.3B [†]	22.29	65.55	81.33	56.39	28.27	53.76	63.27	48.43
X-ALMA-13B [†]	23.56	67.71	79.38	56.88	33.88	66.00	72.25	57.38
TowerInstruct-7B-v0.2 [†]	23.13	69.75	84.91	59.26	34.85	64.81	71.05	56.90
Hunyuan-MT-7B [†]	20.01	72.16	88.06	60.07	30.80	72.53	83.14	61.30
Open-Source Translation Models (Qwen2.5-7B Backbone)								
DRT-7B [†]	–	–	–	–	32.09	68.25	74.73	58.36
SSR-X-Zero-7B [†]	21.73	71.05	87.08	59.95	33.79	70.49	78.59	60.95
mExTrans-7B [†]	19.27	70.89	86.06	58.74	29.03	71.70	78.59	59.77
MT-R1-Zero-Mix-7B [‡]	23.98	70.81	86.17	60.32	40.97	69.43	76.36	62.25
TAT-R1-7B [†]	23.25	70.31	86.47	60.01	39.24	69.00	76.18	61.47
CEMT _{Qwen2.5-7B-Base}	24.42	72.21	87.56	61.39	39.51	69.80	78.63	62.65

Table 1: Main results on WMT23 (ZH→EN) and WMT24 (EN→ZH). Open-source translation models are grouped by whether they use the Qwen2.5-7B backbone. [†] Results are evaluated by us using the released checkpoints and the authors’ recommended decoding settings. [‡] Results are reported numbers taken from Feng et al. (2025a). DRT-7B only supports EN→ZH translation; ZH→EN results are therefore not reported (–). All models are evaluated using the same inference pipeline as Feng et al. (2025a). Within each block, the best score in each column is boldfaced; red shading marks the best Avg within a block, and green shading marks the best Avg overall.

the WMT24 EN→ZH test set (997 samples) using BLEU, COMETKiwi-23-XL, and XCOMET-XL. We additionally report results on two element-specific subsets (DNT and Pragmatic) filtered from each test set. For ZH→EN, the DNT and Pragmatic subsets contain 195 (9.86%) and 680 (34.41%) samples, respectively; for EN→ZH, they contain 96 (9.62%) and 456 (45.74%) samples. Dataset statistics and subset construction details are provided in Section A. For the DNT subset, we additionally report span level exact match accuracy to assess the verbatim preservation of detected spans.

Baselines. We consider a comprehensive set of strong baselines covering both Closed-Source and Open-Source models: (1) **Closed-Source LLMs** (GPT-4o (Hurst et al., 2024), Claude 3.5 Sonnet (Anthropic, 2024), and Gemini 1.5 Pro (Team et al., 2024)); and (2) **Open-Source Models**, including **General-Purpose LLMs** (Qwen2.5-7B-Instruct, Qwen2.5-7B-Base, TeleChat2-7B (Wang et al., 2025d, 2024; He et al., 2024; Liu et al., 2025; Li et al., 2024b,a), Qwen2.5-72B-Instruct, LLaMA-3.1-70B-Instruct (Grattafiori et al., 2024)), and **Translation Models** (NLLB (Team et al., 2022), TowerInstruct (Alves et al., 2024), Hunyuan-MT (Zheng et al., 2025), X-ALMA (Xu et al., 2025),

MT-R1-Zero-Mix (Feng et al., 2025a), TAT-R1 (Li et al., 2025a), DRT (Wang et al., 2025a), SSR-X-Zero-7B (Yang et al., 2025), and mExTrans-7B (Wang et al., 2025c)). Details are provided in Section C.

Model Configuration. CEMT utilizes Qwen2.5-7B-Base as the backbone for both the EDM and the Structured Translation Module. Qwen3-235B-A22B is used as an offline CoT-based judge to score candidates in GRPO. Detailed prompts are provided in Section E. Following common practice for large language model inference, we report results from a single run under a fixed parameter setting.

3.2 Main Results

Table 1 presents the performance of CEMT. We analyze the performance from the following aspects:

CEMT vs. Open-Source Models. CEMT obtains the highest Avg for both directions across open-source systems, including general-purpose LLMs and translation-specific models. Under the same Qwen2.5-7B backbone, CEMT consistently leads among variants. Specifically, it achieves the best group results for ZH→EN and

Model	DNT BLEU	DNT Acc	Prag Kiwi
Qwen2.5-7B-Base	25.16	68.85%	63.19
X-ALMA-13B	35.70	70.08%	62.62
TowerInstruct-7B	33.13	73.33%	64.91
TAT-R1-7B	32.06	60.18%	67.94
CEMT (Ours)	35.59	80.92%	69.17

Table 2: Performance on element-focused subsets from WMT test sets for DNT and pragmatic translation. All columns report bidirectional averages; DNT Acc denotes span level exact match accuracy (micro-average). Detailed directional results are provided in Appendix Table 9.

leads in semantic evaluation and Avg for EN→ZH (XCOMET 78.63, Avg 62.65), with competitive BLEU and COMETKiwi, outperforming DRT, TAT-R1, mExTrans, SSR-X-Zero, and MT-R1-Zero-Mix. This advantage stems from explicit modeling of translation-specific elements, linguistically grounded structured reasoning, and process-level supervision.

CEMT vs. Closed-Source Models. CEMT is competitive with leading closed-source models on the reported automatic metrics under the same evaluation protocol. In Table 1, it achieves higher average scores than GPT-4o (Avg 60.47/61.86) and Claude 3.5 Sonnet (Avg 60.52/62.42) on both ZH→EN and EN→ZH. Overall, these results indicate that CEMT outperforms several closed-source models on these benchmarks, despite having only 7B parameters. This consistent advantage suggests the effectiveness of the design of CEMT.

3.3 Element-Specific Subsets Evaluation

CEMT focuses on translation-specific elements including DNT elements and pragmatic elements. In order to evaluate the effectiveness of CEMT on these elements, we evaluate on two element-specific subsets filtered from the WMT test sets (Table 2), details of subset construction are provided in Section A. BLEU and COMETKiwi-23-XL are used for the DNT and Pragmatic subsets respectively.

Analysis on DNT subset. CEMT substantially improves BLEU on the DNT subset, achieving a global average of 35.59. Furthermore, our model achieves a DNT accuracy (DNT Acc) of 80.92%, outperforming competitive baseline models (directional results are provided in Table 9). This indicates that explicitly identifying and enforcing DNT

Model Variant	ZH→EN Avg	EN→ZH Avg	Global Avg	Δ (Drop)
CEMT (Full)	61.39	62.65	62.02	-
Direct Distillation [†]	60.49	61.83	61.16	0.86 ↓
<i>Element Detection Ablation</i>				
w/o DNT & Prag	60.37	62.12	61.25	0.77 ↓ *
w/o Pragmatic	60.73	61.82	61.27	0.75 ↓
w/o DNT	60.76	61.98	61.37	0.65 ↓
<i>Structured CoT Ablation</i>				
w/o Sem. & Pra.	60.67	61.72	61.20	0.82 ↓ *
w/o Argument	60.37	62.45	61.41	0.61 ↓
w/o Syntactic	60.48	62.35	61.41	0.61 ↓
w/o Strategy	60.55	62.55	61.55	0.47 ↓

Table 3: Ablation results of CEMT. Δ denotes the drop in Global Avg relative to the Full model. * marks the most critical component in each ablation category. † denotes the system baseline supervised entirely on target side final translations without CoT constraints.

elements effectively reduces corruption, omission, and unintended modification of non-translatable spans, leading to more reliable preservation of critical surface forms. X-ALMA-13B achieves a slightly higher global average DNT BLEU (35.70 vs. 35.59), primarily driven by its EN→ZH performance, which likely reflects the benefits of larger model capacity and BLEU’s bias toward lexical overlap.

Analysis on Pragmatic subset. CEMT improves the global average of COMETKiwi-23-XL to 69.17 on the Pragmatic subset. Compared with TAT-R1 that focuses on the fidelity of key terms, CEMT yields larger gains on the Pragmatic subset. This suggests that explicitly modeling pragmatic elements helps reduce inappropriate paraphrasing and supports context-appropriate, controlled rewrites.

3.4 Ablation Study

To dissect the sources of CEMT’s performance, we conduct ablations along two axes: the presence of Element Detection signals for translation-specific elements and the contribution of structured CoT steps together with their aligned process-level supervision. All variants are trained from scratch with the same SFT+RL pipeline, while the direct distillation baseline utilizes identical training stages but focuses on distilling final translations directly from Qwen3-235B without intermediate structured CoT inputs or process level rewards. Table 3 reports the results (we discuss Global Avg unless noted).

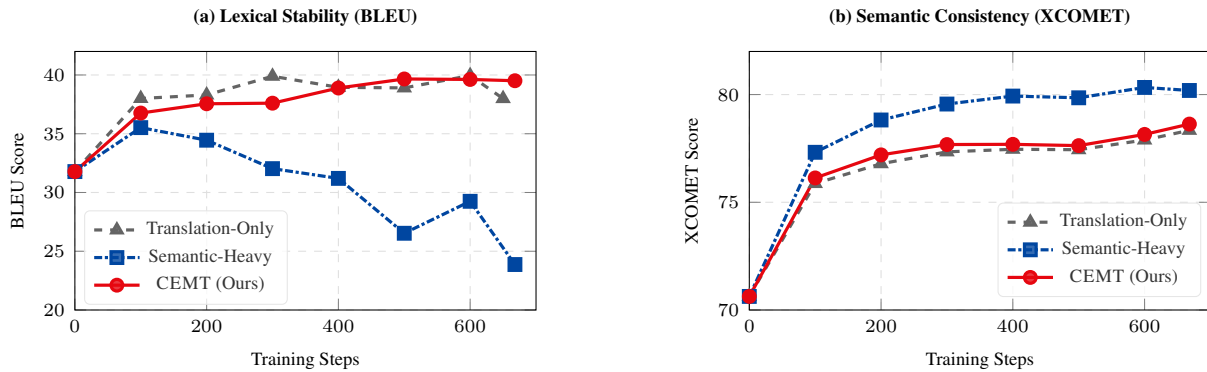


Figure 3: Training dynamics under different reward configurations on the WMT24 EN→ZH test set. CEMT demonstrates more stable optimization than Semantic-Heavy and Translation-Only variants.

Effect of Element Detection Signals. We ablate element detection by selectively removing either DNT or Pragmatic signals, or both, from both the constraint construction and the structured CoT input. Starting from the w/o DNT & Prag baseline (61.25), adding only DNT (+0.02) or only Pragmatic elements (+0.12) yields marginal gains, indicating that modeling a single class of translation-specific elements is insufficient. In contrast, enabling both signals improves Global Avg to 62.02 (+0.77), suggesting that jointly modeling preservation-oriented and rewriting-oriented elements is critical for reducing element-level errors and constraint violations.

Effect of Structured CoT and Process-level Supervision. To assess the contribution of structured linguistic reasoning and its associated supervision, we ablate individual CoT steps by removing both the corresponding reasoning component from the CoT template and its process-level reward during GRPO. Ablating key steps in the structured CoT consistently degrades performance, highlighting the importance of maintaining end-to-end reasoning integrity. Removing the Sem. & Pra. results in the largest drop (61.20), often leading to failures where global element constraints are violated. Removing Strategy (61.55) weakens the alignment between analysis and execution, increasing the risk of element-handling errors despite locally correct analysis.

Overall, these ablations show that CEMT’s gains arise from the combination of explicit element modeling and structured CoT with aligned process-level supervision, which together reduce translation errors.

Setting	DNT Acc	XCOMET
CEMT (Full)	89.53	81.93
Interference (Mask DNT)	68.60	81.38
Baseline (Instruct)	70.93	77.08

Table 4: Mask DNT Interference (EN-ZH). Span level exact match accuracy for DNT elements under an interference setting that masks detection signals at inference time.

S_{comp} Level	Count	Total Tokens
0 (Simple)	240	~552
1 (Standard)	1032	~712
2 (Complex)	704	~971

Table 5: End to end token usage by sentence complexity level.

3.5 Inference Robustness and Cost

To balance inference cost and translation quality, and evaluate the necessity of detector signals, we present results for the DNT masking interference study and token usage statistics in Table 4 and Table 5.

Mask DNT Interference. In Table 4, we simulate detector failure by masking DNT signals during inference. While DNT accuracy drops to 68.60%, semantic quality (XCOMET) remains stable at 81.38, well above the Instruct baseline (77.08). This demonstrates robustness against incomplete constraint signals.

Inference Efficiency. Table 5 breaks down inference costs across source complexities (S_{comp}). CEMT adaptively scales reasoning effort, using fewer tokens for simple sentences (~552) and allocating more steps to complex ones (~971), effec-

tively balancing cost and quality.

3.6 Analysis of Reward Design

We analyze training dynamics across reward configurations in Figure 3. Besides CEMT, we design two reward variants to isolate the roles of semantic emphasis and process-level supervision. Semantic-Heavy excludes BLEU and emphasizes semantic rewards (COMETKiwi and XCOMET), while retaining a lightly weighted process term. Translation-Only removes the process reward and evenly weights BLEU, COMETKiwi, and XCOMET. Detailed reward formulations are provided in Section D.3.

Semantic-Heavy variant maximizes XCOMET at the cost of lexical stability, with BLEU dropping to 23.87. This suggests that without n-gram anchors, the model may be incentivized to over-optimize semantic variant rewards via excessive paraphrasing. Translation-Only initially performs well but exhibits late-stage instability; the degradation is correlated with weaker reasoning behavior and a looser link between analysis and execution (BLEU regresses from 39.95 to 37.99). In contrast, CEMT introduces R_{proc} as an effective regularizer. It encourages the model to maintain a coherent reasoning chain, helping generation stay aligned with the analysis and leading to more stable improvements on the tracked metrics.

3.7 Case Study

To qualitatively assess CEMT, we compare it with Qwen2.5-7B-Instruct on two types of cases: DNT handling and pragmatic paraphrasing (Table 6).

DNT Handling. On inputs containing non-translatable spans, the baseline frequently violates DNT requirements by altering or translating protected segments (Table 6, Cases 1–3). In contrast, CEMT explicitly identifies DNT elements (E_{dnt}) and adopts a verbatim-copy strategy in its reasoning, preserving surface forms and avoiding these errors in all shown cases.

Pragmatic Paraphrasing. For pragmatically marked expressions, the baseline often defaults to literal translation, resulting in semantically inappropriate outputs (Table 6, Cases 4–6). CEMT instead activates a pragmatic reasoning track and favors functionally equivalent paraphrases (e.g., “*ski bums*” → “滑雪族” (ski enthusiasts)), indicating that explicit pragmatic reasoning supports meaning-preserving rewrites.

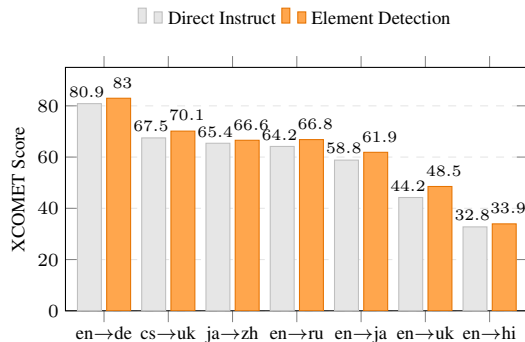


Figure 4: Effect of Element Detection on Qwen2.5-7B-Instruct across diverse language pairs in zero-shot setting, evaluated with XCOMET.

3.8 Zero-Shot Multilingual Analysis of Element Detection

To assess the generality of CEMT beyond ZH↔EN, we apply the Element Detection mechanism to multiple language pairs in a zero-shot setting, without task-specific training. Figure 4 reports XCOMET results on 7 selected language pairs, and Section D provides metrics (BLEU, XCOMET-XL, and COMETKiwi-23-XL) for 10 pairs.

Across all tested pairs, Element Detection yields consistent gains on semantic metrics. XCOMET improves for each of the 7 pairs in Figure 4, and Section D further shows consistent improvements on XCOMET-XL and COMETKiwi-23-XL across all 10 pairs. In a few cases, we observe a mild trade-off in lexical overlap: while semantic metrics improve, BLEU dips slightly (e.g., EN→HI and EN→RU in Table 10), suggesting a shift toward meaning preservation rather than strict n-gram matching.

4 Related Work

CoT for MT. CoT has proven effective in eliciting reasoning in LLMs (Wei et al., 2022; Zhang et al., 2025; Chang et al., 2026), with related structured language understanding surveyed in (Wu et al., 2025a). In MT, CoT has been used to enhance semantic analysis and refinement, e.g., DRT (Wang et al., 2025a) and TEaR (Feng et al., 2025b), but most systems still produce free-form traces with limited structural constraints, making the reasoning process difficult to quantify and supervise.

RL for Translation Optimization. Recent RL-based MT methods optimize COMET or mixed-metric rewards, such as R1-T1 (He et al., 2025), MT-R1-Zero (Feng et al., 2025a), and LLM-

Type	Source Text	Qwen2.5-7B-Instruct	CEMT (Ours)
Scenario A: DNT Handling	(1) ...https://fgc.network/objects/0f1b42c6-cbb1-49bb-91f5-db81bc71ea14...	...https://fgc.network/对象/0f1b42c6-cbb1-49bb-91f5-db81bc71ea14...	...https://fgc.network/objects/0f1b42c6-cbb1-49bb-91f5-db81bc71ea14...
	(2) ...follow Skin Potion on Instagram @kpotionus that's S K I N O T I O N Skin potion...	...账号是 @kpotionus (K 皮质感)账号是 @kpotionus, 即S K I N...
	(3) 1. start of HTML document up to <body>	1. HTML文档的开始, 直到</body>	1. HTML文档的开始, 直到<body>
Scenario B: Pragmatic Paraphrasing	(4) @user25 ooo <i>thats a shout</i> , defo one once I've got some skills in	...哦哦, 这是在喊吧, 等我练一些技能...	...哦, 这真是个好主意, 等我掌握了...
	(5) Skiing is a major travelling activity... occasionally known as "ski bums," planning entire vacations...	...有时被称为“滑雪bum”, 会专门...	...有时被称为“滑雪族”, 他们会...
	(6) ...不断开新店, 看似能够循环往复, 实则真吃卯粮, 无法形成...	...actually eating next month's grain...	...actually living beyond one's means...

Table 6: Case study of translation outputs, contrasting typical baseline failures with CEMT’s behavior. (1) URL component mistranslated into a lexical item; (2) Hallucinated explanation injected for an identifier; (3) Wrong HTML tag/completion; (4) Literal misinterpretation of slang; (5) Mixed-language or overly literal rendering of a culture-specific term; (6) Literal translation of an idiom.

feedback approaches like DeepTrans (Wang et al., 2025b). However, these methods remain largely outcome-driven and suffer from credit-assignment limitations (Sutton and Barto, 2018), while dense process-level supervision is still underexplored (Lightman et al., 2024; Uesato et al., 2022; Li et al., 2025b).

Fine-grained Control in Translation. Fine-grained translation elements have disproportionate impact on translation quality (Song and Xu, 2024). Earlier work explored constraint-based decoding (Post and Vilar, 2018; Hu et al., 2019), and recent LLM-based approaches (Li et al., 2025a) improve terminology consistency via alignment signals. However, existing approaches typically focus on either hard lexical constraints or pragmatic control, while integrating heterogeneous constraints in MT remains challenging and underexplored (Huang et al., 2023).

5 Conclusion

This paper proposes CEMT, which combines an EDM with a dual-track structured CoT to support constraint-guided reasoning in MT. By integrating a CoT Judge with SFT+RL, we place more emphasis on supervising the reasoning process, offering a complementary perspective on credit assignment for constraint-sensitive translation. Experiments show that CEMT achieves the best average performance among the 7B open-source baselines on WMT benchmarks under our evaluation protocol. Ablations further suggest that element-level constraints and pragmatic reasoning are complementary, and that preserving the integrity of the struc-

tured CoT improves analysis-to-generation alignment.

6 Limitations

While our results provide evidence that CEMT is effective for MT, several limitations are worth noting. First, we mainly evaluate on ZH↔EN benchmarks under established protocols; future work should extend the full pipeline to more language pairs, especially low-resource directions. Second, we instantiate CEMT with a 7B backbone to keep comparisons fair under a fixed compute budget; future work should study scaling behaviors across model sizes and model families. Third, our pipeline relies on automatic element detection, which may introduce occasional false positives/negatives; future work could improve reliability with human-annotated supervision and stricter evaluation (Section B). Fourth, our process-level supervision uses an LLM-based CoT Judge (Qwen3-235B-A22B in this work) during training/data construction, which adds extra compute overhead; future work could distill the judge or adopt cheaper verification signals. Fifth, inference requires element detection and structured reasoning outputs, which may increase token usage; future work could shorten the reasoning steps, merge stages, or distill multi-step reasoning into a single-pass translator.

References

Duarte M. Alves, José Pombal, Nuno M. Guerreiro, Pedro H. Martins, João Alves, Amin Farajian, Ben Peters, Ricardo Rei, Patrick Fernandes, Sweta Agrawal, Pierre Colombo, José G. C. de Souza, and André F. T. Martins. 2024. *Tower: An open multilingual*

- large language model for translation-related tasks. *Preprint*, arXiv:2402.17733.
- Hongjun An, Wenhan Hu, Sida Huang, Siqi Huang, Runjun Li, Yuanzhi Liang, Jiawei Shao, Yiliang Song, Zihan Wang, Cheng Yuan, Chi Zhang, Hongyuan Zhang, Wenhao Zhuang, and Xuelong Li. 2025. [Ai flow: Perspectives, scenarios, and approaches](#). *Preprint*, arXiv:2506.12479.
- Anthropic. 2024. [Claude 3.5 sonnet model card addendum](#).
- Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel Ziegler, Jeffrey Wu, Clemens Winter, and 12 others. 2020. [Language models are few-shot learners](#). In *Advances in Neural Information Processing Systems*, volume 33, pages 1877–1901. Curran Associates, Inc.
- Wenhan Chang, Tianqing Zhu, Yu Zhao, Shuangyong Song, Ping Xiong, and Wanlei Zhou. 2026. [Chain-of-lure: A universal jailbreak attack framework using unconstrained synthetic narratives](#). *Preprint*, arXiv:2505.17519.
- Huadong Chen, Shujian Huang, David Chiang, and Jijun Chen. 2017. [Improved neural machine translation with a syntax-aware encoder and decoder](#). In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1936–1945, Vancouver, Canada. Association for Computational Linguistics.
- Jinyang Chen, Haolun Wu, Jianhong Pang, Yihua Wang, Dell Zhang, and Changzhi Sun. 2025. [Tool learning with language models: a comprehensive survey of methods, pipelines, and benchmarks](#). *Vicinagearth*, 2(1):16.
- David Dale, Elena Voita, Loic Barrault, and Marta R. Costa-jussà. 2023. [Detecting and mitigating hallucinations in machine translation: Model internal workings alone do well, sentence similarity Even better](#). In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 36–50, Toronto, Canada. Association for Computational Linguistics.
- DeepSeek-AI, Daya Guo, Dejian Yang, Haowei Zhang, Junxiao Song, Ruoyu Zhang, Runxin Xu, Qihao Zhu, Shirong Ma, Peiyi Wang, Xiao Bi, Xiaokang Zhang, Xingkai Yu, Yu Wu, Z. F. Wu, Zhibin Gou, Zhihong Shao, Zhuoshu Li, Ziyi Gao, and 181 others. 2025. [Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning](#). *Preprint*, arXiv:2501.12948.
- Akiko Eriguchi, Kazuma Hashimoto, and Yoshimasa Tsuruoka. 2016. [Tree-to-sequence attentional neural machine translation](#). In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 823–833, Berlin, Germany. Association for Computational Linguistics.
- Zhaopeng Feng, Shaosheng Cao, Jiahao Ren, Jiayuan Su, Ruizhe Chen, Yan Zhang, Jian Wu, and ZuoZhu Liu. 2025a. [MT-r1-zero: Advancing LLM-based machine translation via r1-zero-like reinforcement learning](#). In *Findings of the Association for Computational Linguistics: EMNLP 2025*, pages 18685–18702, Suzhou, China. Association for Computational Linguistics.
- Zhaopeng Feng, Yan Zhang, Hao Li, Bei Wu, Jiayu Liao, Wenqiang Liu, Jun Lang, Yang Feng, Jian Wu, and ZuoZhu Liu. 2025b. [TEaR: Improving LLM-based machine translation with systematic self-refinement](#). In *Findings of the Association for Computational Linguistics: NAACL 2025*, pages 3922–3938, Albuquerque, New Mexico. Association for Computational Linguistics.
- Yuan Gao, Ruili Wang, and Feng Hou. 2024. [How to design translation prompts for chatgpt: An empirical study](#). In *Proceedings of the 6th ACM International Conference on Multimedia in Asia Workshops, MMA-sia '24 Workshops*, New York, NY, USA. Association for Computing Machinery.
- Aaron Grattafiori, Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha Letman, Akhil Mathur, Alan Schelten, Alex Vaughan, Amy Yang, Angela Fan, Anirudh Goyal, Anthony Hartshorn, Aobo Yang, Archi Mitra, Archie Sravankumar, Artem Korenev, Arthur Hinsvark, and 542 others. 2024. [The llama 3 herd of models](#). *Preprint*, arXiv:2407.21783.
- Nuno M. Guerreiro, Duarte M. Alves, Jonas Waldendorf, Barry Haddow, Alexandra Birch, Pierre Colombo, and André F. T. Martins. 2023. [Hallucinations in large multilingual translation models](#). *Transactions of the Association for Computational Linguistics*, 11:1500–1517.
- Nuno M. Guerreiro, Ricardo Rei, Daan van Stigt, Luisa Coheur, Pierre Colombo, and André F. T. Martins. 2024. [xCOMET: Transparent machine translation evaluation through fine-grained error detection](#). *Transactions of the Association for Computational Linguistics*, 12:979–995.
- Mingui He, Yilun Liu, Shimin Tao, Yuanchang Luo, Hongyong Zeng, Chang Su, Li Zhang, Hongxia Ma, Daimeng Wei, Weibin Meng, Hao Yang, Boxing Chen, and Osamu Yoshie. 2025. [R1-t1: Fully incentivizing translation capability in llms via reasoning learning](#). *Preprint*, arXiv:2502.19735.
- Zhongjiang He, Zihan Wang, Xinzhang Liu, Shixuan Liu, Yitong Yao, Yuyao Huang, Xuelong Li, Yongxiang Li, Zhonghao Che, Zhaoxi Zhang, Yan Wang, Xin Wang, Luwen Pu, Huinan Xu, Ruiyu Fang, Yu Zhao, Jie Zhang, Xiaomeng Huang, Zhilong Lu, and 17 others. 2024. [Telechat technical report](#). *Preprint*, arXiv:2401.03804.

- Amr Hendy, Mohamed Abdelrehim, Amr Sharaf, Vikas Raunak, Mohamed Gabr, Hitokazu Matsushita, Young Jin Kim, Mohamed Afify, and Hany Hassan Awadalla. 2023. [How good are gpt models at machine translation? a comprehensive evaluation](#). *ArXiv*, abs/2302.09210.
- J. Edward Hu, Huda Khayrallah, Ryan Culkin, Patrick Xia, Tongfei Chen, Matt Post, and Benjamin Van Durme. 2019. [Improved lexically constrained decoding for translation and monolingual rewriting](#). 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 839–850, Minneapolis, Minnesota. Association for Computational Linguistics.
- Xuancheng Huang, Zijun Liu, Peng Li, Tao Li, Maosong Sun, and Yang Liu. 2023. [An extensible plug-and-play method for multi-aspect controllable text generation](#). *Preprint*, arXiv:2212.09387.
- Aaron Hurst, Adam Lerer, Adam P. Goucher, Adam Perelman, Aditya Ramesh, Aidan Clark, AJ Ostrow, Akila Welihinda, Alan Hayes, Alec Radford, Aleksander Madry, Alex Baker-Whitcomb, Alex Beutel, Alex Borzunov, Alex Carney, Alex Chow, Alex Kirillov, Alex Nichol, Alex Paino, and 79 others. 2024. [Gpt-4o system card](#). *CoRR*, abs/2410.21276.
- Wenxiang Jiao, Wenxuan Wang, Jen tse Huang, Xing Wang, Shuming Shi, and Zhaopeng Tu. 2023. [Is chatgpt a good translator? yes with gpt-4 as the engine](#).
- Tom Kocmi, Eleftherios Avramidis, Rachel Bawden, Ondřej Bojar, Anton Dvorkovich, Christian Federmann, Mark Fishel, Markus Freitag, Thamme Gowda, Roman Grundkiewicz, Barry Haddow, Marzena Karpinska, Philipp Koehn, Benjamin Marie, Christof Monz, Kenton Murray, Masaaki Nagata, Martin Popel, Maja Popović, and 3 others. 2024. [Findings of the WMT24 general machine translation shared task: The LLM era is here but MT is not solved yet](#). In *Proceedings of the Ninth Conference on Machine Translation*, pages 1–46, Miami, Florida, USA. Association for Computational Linguistics.
- Tom Kocmi, Eleftherios Avramidis, Rachel Bawden, Ondřej Bojar, Anton Dvorkovich, Christian Federmann, Mark Fishel, Markus Freitag, Thamme Gowda, Roman Grundkiewicz, Barry Haddow, Philipp Koehn, Benjamin Marie, Christof Monz, Makoto Morishita, Kenton Murray, Masaaki Nagata, Toshiaki Nakazawa, Martin Popel, and 3 others. 2023. [Findings of the 2023 conference on machine translation \(WMT23\): LLMs are here but not quite there yet](#). In *Proceedings of the Eighth Conference on Machine Translation*, pages 1–42, Singapore. Association for Computational Linguistics.
- Xiang Li, Yiqun Yao, Xin Jiang, Xuezhi Fang, Chao Wang, Xinzhang Liu, Zihan Wang, Yu Zhao, Xin Wang, Yuyao Huang, Shuangyong Song, Yongxiang Li, Zheng Zhang, Bo Zhao, Aixin Sun, Yequan Wang, Zhongjiang He, Zhongyuan Wang, Xuelong Li, and Tiejun Huang. 2024a. [52b to 1t: Lessons learned via tele-film series](#). *Preprint*, arXiv:2407.02783.
- Xiang Li, Yiqun Yao, Xin Jiang, Xuezhi Fang, Chao Wang, Xinzhang Liu, Zihan Wang, Yu Zhao, Xin Wang, Yuyao Huang, Shuangyong Song, Yongxiang Li, Zheng Zhang, Bo Zhao, Aixin Sun, Yequan Wang, Zhongjiang He, Zhongyuan Wang, Xuelong Li, and Tiejun Huang. 2024b. [Tele-film technical report](#). *Preprint*, arXiv:2404.16645.
- Zheng Li, Mao Zheng, Mingyang Song, and Wenjie Yang. 2025a. [Tat-r1: Terminology-aware translation with reinforcement learning and word alignment](#). *Preprint*, arXiv:2505.21172.
- Zhongqiu Li, Shiquan Wang, Ruiyu Fang, Mengjiao Bao, Zhenhe Wu, Shuangyong Song, Yongxiang Li, and Zhongjiang He. 2025b. [MR-UIE: multi-perspective reasoning with reinforcement learning for universal information extraction](#). *Vicinityearth*, 2(1):17.
- Hunter Lightman, Vineet Kosaraju, Yuri Burda, Harrison Edwards, Bowen Baker, Teddy Lee, Jan Leike, John Schulman, Ilya Sutskever, and Karl Cobbe. 2024. [Let’s verify step by step](#). In *The Twelfth International Conference on Learning Representations, ICLR 2024, Vienna, Austria, May 7-11, 2024*. OpenReview.net.
- Xinzhang Liu, Chao Wang, Zhihao Yang, Zhuo Jiang, Xuncheng Zhao, Haoran Wang, Lei Li, Dongdong He, Luobin Liu, Kaizhe Yuan, Han Gao, Zihan Wang, Yitong Yao, Sishi Xiong, Wenmin Deng, Haowei He, Kaidong Yu, Yu Zhao, Ruiyu Fang, and 35 others. 2025. [Training report of telechat3-moe](#). *Preprint*, arXiv:2512.24157.
- Diego Marcheggiani, Jasmijn Bastings, and Ivan Titov. 2018. [Exploiting semantics in neural machine translation with graph convolutional networks](#). In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*, pages 486–492, New Orleans, Louisiana. Association for Computational Linguistics.
- Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul F Christiano, Jan Leike, and Ryan Lowe. 2022. [Training language models to follow instructions with human feedback](#). In *Advances in Neural Information Processing Systems*, volume 35, pages 27730–27744. Curran Associates, Inc.
- 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, Philadelphia,

- Pennsylvania, USA. Association for Computational Linguistics.
- Matt Post and David Vilar. 2018. [Fast lexically constrained decoding with dynamic beam allocation for neural machine translation](#). 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 1314–1324, New Orleans, Louisiana. Association for Computational Linguistics.
- Ricardo Rei, Nuno M. Guerreiro, José Pombal, Daan van Stigt, Marcos Treviso, Luisa Coheur, José G. C. de Souza, and André F. T. Martins. 2023. [Scaling up CometKiwi: Unbabel-IST 2023 submission for the quality estimation shared task](#). In *Proceedings of the Eighth Conference on Machine Translation*, pages 841–848, Singapore. Association for Computational Linguistics.
- Rico Sennrich and Barry Haddow. 2016. [Linguistic input features improve neural machine translation](#). In *Proceedings of the First Conference on Machine Translation: Volume 1, Research Papers*, pages 83–91, Berlin, Germany. Association for Computational Linguistics.
- Zhihong Shao, Peiyi Wang, Qihao Zhu, Runxin Xu, Junxiao Song, Xiao Bi, Haowei Zhang, Mingchuan Zhang, Y. K. Li, Y. Wu, and Daya Guo. 2024. [Deepseekmath: Pushing the limits of mathematical reasoning in open language models](#). *Preprint*, arXiv:2402.03300.
- Huacheng Song and Hongzhi Xu. 2024. [A deep analysis of the impact of multiword expressions and named entities on Chinese-English machine translations](#). In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 6154–6165, Miami, Florida, USA. Association for Computational Linguistics.
- Richard S. Sutton and Andrew G. Barto. 2018. *Reinforcement Learning: An Introduction*, 2 edition. MIT Press, Cambridge, MA.
- Gemini Team, Petko Georgiev, Ving Ian Lei, Ryan Burnell, Libin Bai, Anmol Gulati, Garrett Tanzer, Damien Vincent, Zhufeng Pan, Shibo Wang, Soroosh Mariooryad, Yifan Ding, Xinyang Geng, Fred Alcober, Roy Frostig, Mark Omernick, Lexi Walker, Cosmin Paduraru, Christina Sorokin, and 1118 others. 2024. [Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context](#). *Preprint*, arXiv:2403.05530.
- NLLB Team, Marta R. Costa-jussà, James Cross, Onur Çelebi, Maha Elbayad, Kenneth Heafield, Kevin Hefernan, Elahe Kalbassi, Janice Lam, Daniel Licht, Jean Maillard, Anna Sun, Skyler Wang, Guillaume Wenzek, Al Youngblood, Bapi Akula, Loic Barrault, Gabriel Mejia Gonzalez, Prangthip Hansanti, and 20 others. 2022. [No language left behind: Scaling human-centered machine translation](#). *Preprint*, arXiv:2207.04672.
- Jonathan Uesato, Nate Kushman, Ramana Kumar, Francis Song, Noah Siegel, Lisa Wang, Antonia Creswell, Geoffrey Irving, and Irina Higgins. 2022. [Solving math word problems with process-and outcome-based feedback](#). *arXiv preprint arXiv:2211.14275*.
- Jiaan Wang, Fandong Meng, Yunlong Liang, and Jie Zhou. 2025a. [DRT: Deep reasoning translation via long chain-of-thought](#). In *Findings of the Association for Computational Linguistics: ACL 2025*, pages 6770–6782, Vienna, Austria. Association for Computational Linguistics.
- Jiaan Wang, Fandong Meng, and Jie Zhou. 2025b. [Deeptrans: Deep reasoning translation via reinforcement learning](#). *Preprint*, arXiv:2504.10187.
- Jiaan Wang, Fandong Meng, and Jie Zhou. 2025c. [Extrans: Multilingual deep reasoning translation via exemplar-enhanced reinforcement learning](#). *Preprint*, arXiv:2505.12996.
- Zihan Wang, XinZhang Liu, Shixuan Liu, Yitong Yao, Yunyao Huang, Mengxiang Li, Zhongjiang He, Yongxian Li, Luwen Pu, Huinan Xu, Chao Wang, and Shuangyong Song. 2024. [TeleChat: An open-source bilingual large language model](#). In *Proceedings of the 10th SIGHAN Workshop on Chinese Language Processing (SIGHAN-10)*, pages 10–20, Bangkok, Thailand. Association for Computational Linguistics.
- Zihan Wang, Xinzhang Liu, Yitong Yao, Chao Wang, Yu Zhao, Zhihao Yang, Wenmin Deng, Kaipeng Jia, Jiaxin Peng, Yuyao Huang, Sishi Xiong, Zhuo Jiang, Kaidong Yu, Xiaohui Hu, Fubei Yao, Ruiyu Fang, Zhuoru Jiang, Ruiting Song, Qiyi Xie, and 19 others. 2025d. [Technical report of telechat2, telechat2.5 and t1](#). *Preprint*, arXiv:2507.18013.
- Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed H. Chi, Quoc V. Le, and Denny Zhou. 2022. [Chain-of-thought prompting elicits reasoning in large language models](#). In *Advances in Neural Information Processing Systems 35: Annual Conference on Neural Information Processing Systems 2022, NeurIPS 2022, New Orleans, LA, USA, November 28 - December 9, 2022*.
- Di Wu, Ruiyu Fang, Liting Jiang, Shuangyong Song, Xiaomeng Huang, Shiquan Wang, Zhongqiu Li, Lingling Shi, Mengjiao Bao, Yongxiang Li, and Hao Huang. 2025a. [Multi-intent spoken language understanding: a survey of methods, trends, and challenges](#). *Viciniagearth*, 2(1):20.
- Zhenhe Wu, Jian Yang, Jiaheng Liu, Xianjie Wu, Changzai Pan, Jie Zhang, Yu Zhao, Shuangyong Song, Yongxiang Li, and Zhoujun Li. 2025b. [Table-rl: Region-based reinforcement learning for table understanding](#). *Preprint*, arXiv:2505.12415.
- Haoran Xu, Kenton Murray, Philipp Koehn, Hieu Hoang, Akiko Eriguchi, and Huda Khayrallah. 2025.

X-ALMA: plug & play modules and adaptive rejection for quality translation at scale. In *The Thirteenth International Conference on Learning Representations, ICLR 2025, Singapore, April 24-28, 2025*. OpenReview.net.

An Yang, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chengyuan Li, Dayiheng Liu, Fei Huang, Haoran Wei, Huan Lin, Jian Yang, Jianhong Tu, Jianwei Zhang, Jianxin Yang, Jiaxi Yang, Jingren Zhou, Junyang Lin, Kai Dang, and 22 others. 2024. [Qwen2.5 technical report](#). *CoRR*, abs/2412.15115.

Wenjie Yang, Mao Zheng, Mingyang Song, Zheng Li, and Sitong Wang. 2025. [Ssr-zero: Simple self-rewarding reinforcement learning for machine translation](#). *Preprint*, arXiv:2505.16637.

Yang Yue, Zhiqi Chen, Rui Lu, Andrew Zhao, Zhaokai Wang, Yang Yue, Shiji Song, and Gao Huang. 2025. [Does reinforcement learning really incentivize reasoning capacity in llms beyond the base model?](#) *Preprint*, arXiv:2504.13837.

Wei Zhang, Xiangyuan Guan, Lu Yunhong, Jie Zhang, Shuangyong Song, Xianfu Cheng, Zhenhe Wu, and Zhoujun Li. 2025. [Lemur: Log parsing with entropy sampling and chain-of-thought merging](#). *Preprint*, arXiv:2402.18205.

Mao Zheng, Zheng Li, Bingxin Qu, Mingyang Song, Yang Du, Mingrui Sun, and Di Wang. 2025. [Hunyuan-mt technical report](#). *Preprint*, arXiv:2509.05209.

A Dataset Statistics and Element-Specific Subsets Construction

Data sources and cleaning. We collect EN↔ZH parallel sentence pairs from publicly released WMT test sets in 2017–2022 as training data, following common practice. Our evaluation is conducted on the more recent WMT23 ZH→EN and WMT24 EN→ZH test sets. Unless otherwise stated, all data cleaning and CoT generation are performed by Qwen3-235B-A22B using the prompts in Section E.2 and Section E.1. We further construct two element-specific subsets from the held-out WMT23/24 test sets using element annotations produced by the same Qwen3-235B-A22B annotator with the inference prompt in Section E.1. At inference time, our translation pipeline relies on a 7B element detector to produce the element report consumed by the translator (Section B).

We apply the Data Cleaning & Alignment Prompt (Section E.2) to filter out low-quality pairs, including information-asymmetry cases (falsifiable facts present in the target but unsupported by the source) and strict DNT violations. After cleaning, we obtain 21,420 aligned pairs in total, consisting of 11,560 ZH→EN pairs and 9,860 EN→ZH pairs.

SFT and GRPO data construction. For SFT, we use the SFT Sample Generation with CoT prompt (Section E.2) to generate $n=3$ candidate CoTs for each cleaned example. We then score all candidates with the CoT Judge (Section E.3) and retain the top-scoring CoT subject to a threshold of 0.85, resulting in 16,925 SFT instances. For GRPO, we train on a larger pool of 21,420 instances, which includes all cleaned pairs, following the same data sources and cleaning pipeline.

Test sets and element-specific subsets. We evaluate on the held-out WMT23 ZH→EN test set (1,976 test examples) and WMT24 EN→ZH test set (997 test examples). We further construct two element-specific subsets from these held-out test sets using element detection annotations produced by the same Qwen3-235B-A22B annotator with the inference prompt in Section E.1. Specifically, we take sentences whose detection report contains only DNT elements as the DNT subset, and sentences whose report contains only Pragmatic elements as the Pragmatic subset. Table 7 summarizes the dataset statistics.

Note on model usage. Qwen3-235B-A22B is used only during data construction for prompted

Split	Direction	#Pairs
Cleaned training corpus	ZH→EN	11,560
	EN→ZH	9,860
	Total	21,420
SFT training set	ZH↔EN	16,925
GRPO training pool	ZH↔EN	21,420
Main test set	WMT23 ZH→EN	1,976
Main test set	WMT24 EN→ZH	997
DNT subset	ZH→EN	195
DNT subset	EN→ZH	96
Pragmatic subset	ZH→EN	680
Pragmatic subset	EN→ZH	456

Table 7: Dataset statistics used in our experiments. All counts are sentence pairs. Targeted subsets are extracted from the corresponding held-out test sets.

cleaning, CoT generation, and element annotation. At inference time, the translation pipeline uses a 7B element detector and the translation model follows the setup described in Section 3.1.

B Element Detection Consistency with Annotator Models

Since translation-specific elements lack a unified human gold standard, we report the agreement between the same 7B EDM used in our main experiments and the annotations produced by Qwen3-235B-A22B. This agreement study is intended to assess the reliability of the detector signals used in our pipeline, rather than to claim absolute accuracy.

Setup. We stratify-sample 100 instances from the annotated pool as a held-out test set. We compare the predicted element report at the level that is directly consumed by the translation pipeline, including (i) the sentence-complexity label S_{comp} and (ii) whether the report contains any Do-Not-Translate elements ($E_{\text{dnt}} \neq \emptyset$) and/or any pragmatic elements ($E_{\text{prag}} \neq \emptyset$).

Results. On this 100-instance test set, we observe perfect agreement (100%) between the 7B detector predictions and the model annotations on all the above coarse-grained fields. This indicates that, under our protocol, the detector outputs are consistent with the model labels at the interface level consumed by the translation pipeline. We emphasize that this agreement is not an absolute measure of extraction correctness: we do not evaluate span level overlap for extracted elements in this study. Constructing a stricter human-annotated benchmark with span level metrics is left for future work. We

Category	Hyperparameter	Value
Stage 1: SFT	Base Model	Qwen2.5-7B-Base
	Scheduler Configuration	2 Epochs (Cosine Decay)
	Global Batch Size	256
	Learning Rate	1e-5
	Max Sequence Length	6000
Stage 2: RL	Algorithm	GRPO
	Epochs	2
	Total Training Steps	668
	Train Batch Size	64
	Group Size (G)	6
	Learning Rate	1e-6 (Constant)
	KL Coefficient (β)	0.005
	Max Response Length	2000
Reward Function	Mixed (BLEU : Kiwi : XCOMET : CoT = 3:3:3:1)	
Infrastructure	Training Compute	1 Node \times 8 A100 GPUs
	Reward Judge (Qwen3-235B)*	2 Nodes \times 8 A100 GPUs
	Reward Metrics (COMETs)	2 \times A100 (1 GPU per model)

Table 8: Hyperparameters for SFT and GRPO stages. *Decoupled inference services used for reward computation, not counted towards the gradient training compute budget.

expect that additional high-quality human supervision could provide cleaner targets and potentially improve detection robustness in edge cases.

Robustness to imperfect element reports. We qualitatively observe that the translation module is not overly sensitive to occasional noise in the detector’s element report. One possible reason is that our training exposes the model to imperfect detections in multiple stages: (i) the SFT data are constructed from generated annotations that can be noisy, and the model learns to translate correctly under such imperfect signals; (ii) GRPO further penalizes downstream failures when the element report is uninformative or partially incorrect; and (iii) the CoT Judge used during training scores the consistency between the reasoning, detected elements, and the final translation, which may discourage the model from blindly following spurious detections. We emphasize that, beyond the controlled Mask DNT interference study in Section 3.5, this paragraph provides qualitative observations rather than a controlled robustness evaluation. A dedicated study (e.g., injecting synthetic perturbations into element reports and measuring constraint violations

and translation quality) is left for future work.

C Experimental Setup & Hyperparameters

This section details the training configurations used to produce the reported results. The core training process (Actor/Rollout) was conducted on a single node with 8 NVIDIA A100 GPUs using the ver1 framework.

C.1 Hyperparameters

We employ a two-stage training pipeline. Table 8 lists the specific hyperparameters. Conceptually, the SFT stage primarily teaches the model to follow the structured output format (i.e., element report and structured CoT), while the GRPO stage enforces reward-aligned behavioral constraints and stabilizes the learned format under on-policy sampling. Importantly, the judge model (Qwen3-235B-A22B) is only used to generate CoT supervision in SFT and to score the intermediate reasoning traces in GRPO; it does not generate the final translations, and the final translation rewards are computed by metric models rather than the 235B judge.

C.2 Evaluation Details

In the main results (Table 1), we report numbers from three sources.

Results taken from MT-R1-Zero. For models whose results are marked with ‡ in Table 1, we directly use the reported numbers from Feng et al. (2025a). These results are not re-evaluated by us; decoding settings and evaluation protocols follow the original paper.

Baselines evaluated by us. For models whose results are marked with † (or otherwise not marked as ‡) in Table 1, we evaluate the released checkpoints ourselves. We use the official checkpoints and the authors’ recommended inference parameters (and recommended inference framework when provided) to ensure a fair comparison. Our pre-processing and scoring pipeline follows the public MT-R1-Zero evaluation protocol, and we use the official implementations for BLEU, COMETKiwi-23-XL, and XCOMET-XL to keep metric computation consistent across systems.

For Qwen2.5-7B-Base, we apply a minimal post-processing step to enable reliable automatic evaluation. Since the base model is not instruction-tuned, it often appends extra free-form text after

Model	DNT (BLEU)		DNT (Acc)		Prag (Kiwi)	
	ZH→EN	EN→ZH	ZH→EN	EN→ZH	ZH→EN	EN→ZH
Qwen2.5-7B-Base	16.92	33.39	66.8%	70.9%	66.08	60.29
X-ALMA-13B	24.67	46.72	67.8%	72.3%	60.82	64.42
TowerInstruct-7B	22.67	43.58	71.7%	75.0%	66.25	63.56
TAT-R1-7B	23.75	40.36	65.2%	55.2%	67.16	68.71
CEMT (Ours)	25.63	45.55	72.3%	89.5%	68.54	69.79

Table 9: Directional results on element-focused subsets.

producing the translation, which occurs in over 60% of our samples. In most cases, the translation and the trailing text are separated by a newline character. Therefore, when computing automatic metrics for Qwen2.5-7B-Base, we keep only the substring before the first newline as the model translation. This behavior reflects the model’s tendency to prepend unsolicited explanations under instruction-following prompts, rather than deficiencies in translation quality.

Our model evaluation. We decode our model with vLLM using the translation prompt in Section E.1 and the model’s chat template. Unless otherwise stated, we use greedy decoding (temperature= 0) for deterministic evaluation; for analyses that require mild stochasticity, we set temperature= 0.2 with the same top- $p = 0.95$. We set max_tokens to 2000 and cap the context length with max_model_len= 6000, matching the training setup. All other decoding parameters follow vLLM defaults. For evaluation, we extract the final translation as the text after the </think> tag.

D Additional Experimental Results

This section presents additional experimental results, including training dynamics, directional subset evaluation, and multilingual zero-shot analysis.

D.1 Element-Focused Subset Directional Breakdown

Table 9 provides the directional breakdown of the bidirectional averages reported in Table 2. Notably, DNT Acc exhibits significant directional asymmetry: on the EN→ZH direction, accuracy improves from 70.9% (baseline) to 89.5%, whereas on the ZH→EN direction, the gain is more modest, from 66.8% to 72.3%. This disparity stems from the inherent distributional differences of non-translatable elements across translation directions—EN→ZH involves abundant English letters, proper nouns, and abbreviations that must be preserved

(e.g., “IBM”, “AI”), while Chinese elements requiring retention in ZH→EN (e.g., Chinese characters, sinicized proper nouns) are extremely rare. Consequently, the model has greater room for improvement in identifying and preserving DNT elements in the EN→ZH direction.

D.2 Full Multilingual Zero-Shot Results

Table 10 presents the complete zero-shot performance across 10 language pairs spanning diverse language families. We compare the direct inference of Qwen2.5-7B-Instruct (Base) against the Element Detection inference mode (+Element Detection), where the same backbone is used and the only change is injecting an Element Detection Report at inference time. This evaluation is conducted in a strictly zero-shot setting without any task-specific training, serving as a stress test of the generality of the proposed mechanism.

Overall trends. Across all 10 directions, +Report yields consistent improvements on semantic metrics, with minor trade-offs on lexical overlap in a small number of directions. In particular, XCOMET-XL and COMETKiwi-23-XL improve for every language pair, with average gains of +2.17 XCOMET-XL (57.50→59.67) and +2.00 COMETKiwi-23-XL (47.95→49.95). BLEU also improves on average (+0.65; 16.49→17.14), indicating that the injected report is generally compatible with lexical overlap as well.

Where the gains concentrate. The largest semantic gains are observed on directions that are either typologically distant or involve morphologically rich targets. For example, EN→UK shows the most pronounced improvements (+4.30 XCOMET-XL and +4.65 COMETKiwi-23-XL), suggesting that explicit element-centric signals can substantially reduce adequacy errors when surface realizations are flexible. Similarly, EN→JA exhibits strong gains (+3.05 XCOMET-XL and +2.83 COMETKiwi-23-XL), consistent with the intuition

Lang Pair	BLEU		XCOMET-XL		COMETKiwi-23-XL	
	Base	+Element Detection	Base	+Element Detection	Base	+Element Detection
cs→uk	13.39	15.11	67.48	70.13	46.14	48.07
en→cs	12.40	13.56	52.41	54.01	42.77	44.50
en→de	21.73	22.64	80.86	82.96	57.48	59.25
en→es	32.93	34.13	76.91	78.12	64.26	64.42
en→hi	9.49	9.43	32.75	33.93	38.65	41.68
en→is	5.43	5.50	32.05	33.88	20.08	22.27
en→ja	16.50	17.10	58.81	61.86	60.97	63.80
en→ru	15.67	15.20	64.15	66.82	56.57	58.14
en→uk	11.13	12.01	44.22	48.52	35.98	40.63
ja→zh	26.27	26.76	65.39	66.56	56.56	56.78
Avg	16.49	17.14	57.50	59.67	47.95	49.95

Table 10: Detailed zero-shot performance on 10 language pairs. **Base**: Qwen2.5-7B-Instruct (Direct Translation). **+Element Detection**: Same model with Element Detection Report injection. Best results are highlighted in **bold**.

that the report helps stabilize translation decisions under higher structural divergence.

Lexical-semantic trade-off in a few cases.

While semantic metrics improve universally, we observe a mild trade-off in lexical overlap for a small subset of directions. Specifically, BLEU slightly decreases on EN→HI and EN→RU, even though both XCOMET-XL and COMETKiwi-23-XL still increase. This pattern suggests that +Report may encourage paraphrases that better preserve meaning and constraints, which can be rewarded by semantic metrics but not necessarily by n-gram matching.

Takeaway. Overall, the results indicate that Element Detection Report injection is a lightweight yet effective mechanism to improve semantic adequacy and robustness across languages in a zero-shot manner. The consistent improvements on XCOMET-XL and COMETKiwi-23-XL across all 10 pairs suggest that the mechanism can generalize beyond ZH↔EN under a zero-shot setting, while the few BLEU dips indicate that lexical-overlap metrics may under-estimate gains when the model shifts toward meaning-preserving rewrites.

D.3 Reward configurations

All variants differ only in how translation-level metrics and process supervision are weighted. We first define the translation-level reward as a weighted combination of automatic translation metrics:

$$R_{\text{trans}} = w_{\text{bleu}}S_{\text{bleu}} + w_{\text{kiwi}}S_{\text{kiwi}} + w_{\text{xcm}}S_{\text{xcm}} \quad (7)$$

Here, S_{bleu} , S_{kiwi} , and S_{xcm} correspond to BLEU, COMETKiwi, and XCOMET, respectively, and

weights w_{bleu} , w_{kiwi} , w_{xcm} control their contributions.

Following Eq. (6) in the main text, the overall GRPO reward applies a hard format gate using a binary validity flag $R_{\text{format}} \in \{0, 1\}$:

$$R_{\text{total}} = \begin{cases} w_{\text{proc}}R_{\text{proc}} + R_{\text{trans}}, & \text{if } R_{\text{format}} = 1 \\ -1, & \text{if } R_{\text{format}} = 0 \end{cases} \quad (8)$$

Here, R_{proc} supervises intermediate reasoning quality via the CoT Judge, and R_{trans} measures translation quality. Under invalid-format outputs, we directly set $R_{\text{total}} = -1$; under valid-format outputs, we optimize the weighted sum of process and translation rewards.

Under this formulation, the three reward configurations differ only in the (unnormalized) weights assigned to each term. **CEMT (Ours)** uses $(w_{\text{bleu}}, w_{\text{kiwi}}, w_{\text{xcm}}, w_{\text{proc}}) = (3, 3, 3, 1)$; **Translation-Only** sets $(3, 3, 3, 0)$, removing the process term; and **Semantic-Heavy** sets $(0, 4.5, 4.5, 1)$, discarding BLEU while upweighting semantic metrics. All weights are relative and are not required to sum to one.

Semantic-Heavy: early gains with late-stage lexical degradation.

Semantic-Heavy initially improves all metrics (e.g., BLEU increases from 31.78 at Step 0 to 35.52 at Step 100), suggesting that the policy can still benefit from on-policy optimization under predominantly semantic signals. However, as training proceeds, BLEU gradually deteriorates and ends substantially lower (23.87 at the final checkpoint), whereas XCOMET and Kiwi remain high (80.19 and 71.53, respectively). We do not interpret this as conclusive causal evi-

Model Variant	Step	BLEU	XCOMET	Kiwi	Avg
<i>Baseline (Starting Point)</i>					
Qwen2.5-7B (SFT)	0	31.78	70.63	62.56	54.99
Semantic-Heavy	100	35.52	77.32	69.86	60.90
	200	34.45	78.82	71.14	61.47
	300	32.03	79.56	71.27	60.95
	400	31.20	79.93	71.56	60.90
	500	26.53	79.85	72.01	59.46
	600	29.25	80.33	71.98	60.52
	Final (668)	23.87	80.19	71.53	58.53
Translation-Only	100	38.00	75.87	68.71	60.86
	200	38.30	76.79	68.69	61.26
	300	39.89	77.34	69.27	62.17
	400	39.95	77.46	69.26	62.22
	500	38.90	77.44	69.35	61.90
	600	38.95	77.89	69.51	62.12
	Final (668)	37.99	78.34	69.77	62.03
CEMT (Ours)	100	36.76	76.13	68.23	60.37
	200	37.55	77.20	69.08	61.28
	300	37.60	77.68	69.24	61.51
	400	38.89	77.69	69.35	61.98
	500	39.66	77.63	69.53	62.27
	600	39.62	78.15	69.52	62.43
	Final (668)	39.51	78.63	69.80	62.65

Table 11: Step-wise training dynamics on WMT24 EN→ZH. The **Semantic-Heavy** variant shows a collapse in BLEU scores (highlighted in red) despite high XCOMET scores. The **Translation-Only** variant shows instability in the final steps. CEMT achieves the most balanced convergence.

dence, but the pattern is consistent with a potential reward-exploitation tendency, in which the absence of lexical constraints may allow the policy to favor outputs that score well under neural semantic metrics while being less aligned with surface realizations, as reflected in the BLEU decline despite strong XCOMET.

Translation-Only: strong peak but mild end-stage regression. Translation-Only achieves a strong peak in BLEU (39.95 at Step 400) and maintains steady improvements in XCOMET/Kiwi throughout training. Nevertheless, the final phase shows a mild regression in BLEU from Step 600 to the end (38.95→37.99), even though semantic metrics continue to increase (XCOMET 77.89→78.34; Kiwi 69.51→69.77).

In qualitative inspection, we observe that, in the absence of process supervision, the generated CoT gradually becomes less informative and is sometimes only loosely related to the final translation output. While this does not directly affect reward computation (which depends only on the final translation), it may reduce the amount of useful guidance provided during exploration in late-stage up-

dates. We view this as one possible contributing factor to the observed fluctuations, where continued gains in semantic metrics do not consistently translate into stable, cross-metric improvements.

CEMT: the most balanced and stable trajectory. In contrast, CEMT shows the most stable training trajectory: it improves steadily across steps and retains strong final performance without the BLEU degradation observed in Semantic-Heavy or the mild end-stage regression seen in Translation-Only. At the end of training, CEMT attains the highest Avg score (62.65) together with a strong BLEU of 39.51, as well as competitive XCOMET (78.63) and Kiwi (69.80). Rather than indicating a direct causal mechanism, we view these results as suggestive evidence that R_{proc} may play a regularizing role during optimization, helping to avoid updates that disproportionately favor a subset of reward components and supporting more consistent progress across metrics.

Why the SFT starting point may not maximize translation metrics. The Step-0 model (the SFT checkpoint used to initialize GRPO) is designed primarily as a format-learning cold start rather than

a translation-quality optimizer. In our setup, SFT mainly encourages structured output compliance (element report + structured CoT), so that RL training begins from a policy that reliably produces the required format. In addition, the CoT supervision is generated by Qwen3-235B-A22B conditioned on reference translations; since reference translations are not always of perfect quality, the resulting CoT (and the SFT objective built upon it) may not always provide a strictly stronger translation target than the base model’s own preferences. It is therefore unsurprising that the SFT checkpoint sometimes yields limited gains or mild degradation on automatic metrics, while GRPO is primarily responsible for recovering and improving final translation behavior under explicit reward signals.

E Prompt Templates

This appendix presents the exact prompt templates used in the CEMT framework. They are categorized by their lifecycle stage: Inference, Data Construction, and Evaluation.

Terminology and Coding Scheme Alignment. For faithfulness to the training and inference pipelines, we keep the original field names used in the data schema and prompt templates (e.g., `feature_code`, `feature_report`). Conceptually, these fields correspond to the element-detection formulation in Section 2, where

$$EA = (S_{comp}, E_{dnt}, E_{prag}) \quad (9)$$

feature_code. The variable `feature_code` is a 3-dimensional discrete code (c_0, c_1, c_2) , where $c_0 \in \{0, 1, 2\}$ denotes the sentence-complexity level, and $c_1, c_2 \in \{0, 1\}$ are binary indicators of whether the corresponding element types are present (1) or absent (0).

where each position has the following semantics:

- **First position** c_0 — sentence complexity (maps to S_{comp}):

$$c_0 \in \{0, 1, 2\}$$

- $c_0 = 0$: Incomplete / fragmented sentence
- $c_0 = 1$: Complete but simple sentence
- $c_0 = 2$: Complete and complex sentence

- **Second position** c_1 — binary indicator of DNT elements (maps to E_{dnt}):

$$c_1 \in \{0, 1\}$$

- $c_1 = 0$: No DNT elements present
- $c_1 = 1$: At least one DNT element present

- **Third position** c_2 — binary indicator of pragmatic elements (maps to E_{prag}):

$$c_2 \in \{0, 1\}$$

- $c_2 = 0$: No pragmatic/culture-specific elements
- $c_2 = 1$: Pragmatic elements present (idioms, slang, emotion-bearing expressions)

Intuitively, c_0 controls the depth of reasoning, while c_1 and c_2 indicate whether DNT/pragmatic handling is required.

feature_report / feature_analysis_json. Both refer to the structured JSON report produced by the EDM, containing concrete instances of E_{dnt} and E_{prag} (not only their binary flags). We keep these names unchanged because they are directly used in the training codebase and inference pipelines.

E.1 Inference Prompts

E.1.1 EDM Prompt (Feature Extraction)

DIMENSION 1: COGNITIVE COMPLEXITY
(Ternary Scale: 2, 1, 0)

- Code 2 (*Deep Analysis*): Text containing multiple clauses, complex logic (e.g., “Although... but...”), or deep nesting.
- Code 1 (*Base Analysis*): Grammatically complete simple sentences or clear imperatives (e.g., “Send me the file”).
- Code 0 (*Identification*): Structurally incomplete fragments (e.g., single nouns, phrases without predicates).

DIMENSION 2: ABSOLUTE DNT ELEMENTS
(Binary Scale: 1, 0)

- Code 1: Strictly contains elements from the *Hard DNT List*:
 - *Identifiers*: URLs, Emails, API endpoints.
 - *Technical*: Code snippets, Error codes, Version/Serial numbers.
 - *Visual*: Emojis, Kaomoji, Trademarks.

Negative Constraint: Do NOT extract common terms, translatable proper nouns (e.g.,

“Belt and Road”), Organization names, Time/Dates, or Units (e.g., kg, 220V).

- Code 0: Contains no such elements.

DIMENSION 3: PRAGMATIC ELEMENTS (Binary Scale: 1, 0)

- Code 1: Contains slang, idioms, buzzwords, or cultural expressions requiring paraphrasing. Includes abbreviations with pragmatic functions.
- Code 0: Contains none.

TASK INPUT: [source_text](#)

E.1.2 CEMT Translation Prompt (Module 2)

Role Setting

You are a world-class [src_lang](#)→[tgt_lang](#) translation expert. Your primary goal is to produce translations that are **natural and idiomatic**, avoiding “translationese.” You adhere to a precise “Analysis-Strategy-Execution” workflow.

Core Task

Your task is to generate a structured Chain of Thought (CoT) analysis based on the provided “Element Detection Report,” and then produce the final translation. Your CoT and final translation must be logically consistent.

Instructions

1. Your output must begin with a `<think>` block.
2. Your reasoning must follow a “**Base + Incremental**” logic, guided by the `feature_code` in the Element Detection Report.
3. You must fill the following four sub-tags in order. Your **analysis depth must be proportional to the complexity level** (`feature_code[0]`):

- `<holistic_semantics_pragmatics_analysis>`
 - **Base Analysis:** Analyze overall context/tone. **Comp 2** must be comprehensive; **Comp 1** must be concise; **Comp 0** should be identification-only.
 - **Incremental Response:** If `feature_code[2] == 1` (Pragmatic element present), add an analysis

section specifically identifying this pragmatic element.

- `<argument_predicate_analysis>`
 - **Base Analysis:** Analyze the core predicate-argument structure. **If Comp 2**, this *must* identify key logical/semantic difficulties (even untagged ones). **If Comp 0**, this may be “N/A” (e.g., for fragments).
 - **Incremental Response:** If `feature_code[1] == 1` (DNT element present), add an analysis section to: (1) **Classify** the DNT. (2) **Recommend** a handling strategy.
- `<syntactic_structure_analysis>`
 - **Base Analysis:** Analyze the syntactic structure/sentence type. **If Comp 2**, this must cover key structural challenges. **If Comp 0**, this may be “N/A”.
 - **Incremental Response:** (N/A for this tag)
- `<translation_strategy_formulation>`
 - **Base Strategy:** Formulate the overall translation approach based on **all difficulties** identified in the Base and Incremental analyses above.
 - **Incremental Strategy:**
 - * **DNT:** Formulate specific “**directives**” for any DNT elements.
 - * **Pragmatics:** Pragmatic strategies must follow a ‘**Context-Function-Equivalence**’ reasoning loop to find a natural, functionally equivalent expression.
 - **Concluding Commitment:** [CRITICAL] You must end this tag with the fixed phrase: “**In summary, the final translation is determined to be:**” followed immediately by the complete translation.

4. After the `</think>` block, output the final translation. This translation must be **identical** to the one stated in your concluding commitment.

Task Inputs

-Source Text:
[source_text](#)

-Element Detection Report:
feature_analysis_json
Mandatory Output Format

```
<think>
.....
</think>
...Your Final Translation...
```

E.1.3 Baseline Prompts

1. Direct Translation Prompt

Translate the following **src_lang**→**tgt_lang** into the target language directly. Do not output any explanations or analysis.

Source: **source_text**

Translation:

2. Zero-shot Element Detection Prompt

You are a professional **src_lang**→**tgt_lang** translator. Below is an “Element Detection Report” regarding the linguistic features of the source text. Please refer to it to improve your translation accuracy.

[Element Detection Report]

feature_analysis_json

[Task] Translate the source text into the target language.

Source: **source_text**

Translation:

E.2 Data Construction Prompts

E.2.1 Data Cleaning & Alignment Prompt

ROLE: Senior Data Quality Audit Expert.
TASK: Identify translation pairs containing “Information Asymmetry” or “DNT Violations.”

PRINCIPLES OF ACCEPTANCE (IS_ALIGNED: TRUE)

1. **Liberal Translation & Inference:** Permitted if faithful to the core message and based on general world knowledge (e.g., adding “US President” to “Obama” is valid).
2. **Register Shift:** Permitted (e.g., changing “Let’s go” to “Scheduled departure”) unless it introduces falsifiable specific facts.
3. **Entity Exemption:** Proper nouns (People, Places, Orgs) *may* be translated, transliter-

ated, or kept in the original language.
STRICT DNT (DO-NOT-TRANSLATE) RULES The following elements must remain **verbatim**. Translating them is a strict violation:

- *Internet:* URLs, Emails, Hashtags (e.g., #DIY), Handles (@user).
- *Technical:* Code variables, Placeholders (name), Paths, Regex.
- *Visual:* Emojis, Trademarks (iPhone, Tesla).
- *Codes:* Promo codes (SAVE20), Serial numbers.

REJECTION CRITERIA (IS_ALIGNED: FALSE)

1. **Information Asymmetry:** Target contains specific, falsifiable facts (names, dates, numbers) missing from the Source (Hallucination).
2. **DNT Violation:** Any modification to the form of DNT elements (even if semantically correct).

INPUT: Source: **source_text** | Target: **target_text**

E.2.2 SFT Sample Generation with CoT

ROLE: Expert in translation methodology and CoT generation. TASK: Generate an ideal, gold-standard CoT that logically deduces the provided **target_text** from the **source_text**, adhering to the “Analysis-Strategy-Execution” workflow.

SOP: STANDARD OPERATING PROCEDURE

1. **Base Analysis** (Dependent on feature_code[0])

- Comp 2: Comprehensive analysis of logical/semantic difficulties.
- Comp 1: Concise confirmation of core elements.
- Comp 0: Identification only (or “N/A” for inapplicable tags).

2. **Incremental Response**

- **DNT Handling** (inside <argument_predicate_analysis>):
 - *Diagnosis:* Classify as Hard DNT (Strict Replication, e.g., URLs) or Soft DNT

(Discretionary).

- *Strategy*: Propose specific handling (Strict vs. Discretionary).

- **Pragmatic Handling** (inside <translation_strategy_formulation>):

- Must reflect a **Context-Function-Equivalence Loop**: Identify Context → Analyze Function (e.g., venting vs. playfulness) → Select Functional Equivalent.

3. Concluding Commitment [CRITICAL]

- End the strategy tag with: “**In summary, the final translation is determined to be:**” followed immediately by the text identical to **target_text**.

INPUT: Source: **source_text** | Target: **target_text** | Report: **feature_analysis_json**

E.3 Evaluation Prompts

E.3.1 CoT Judge Prompt

ROLE: Exacting Senior Translation Evaluation Chair. TASK: Audit the Chain of Thought (CoT) and final translation based on the *Element Detection Report*. You must ensure the reasoning is logical, proportional, and consistent with the *Concluding Commitment*.

CORE AUDIT PRINCIPLES

- **Principle A: Proportional Reasoning.** Analysis depth must match feature_code[0]: *Comp 2* (Comprehensive/Deep), *Comp 1* (Concise), *Comp 0* (ID only).
- **Principle B: Structural Location.** Incremental analysis must be in correct tags: Pragmatics → <holistic_...> tag; DNT → <argument_...> tag.
- **Principle C: DNT Handling.** Hard DNTs (URLs, Code, IDs) must be replicated verbatim.
- **Principle D: Pragmatic Equivalence.** Execute the *Context-Function-Equivalence (C-F-E) Loop*; no literal translation of idioms/slang.

EVALUATION RUBRICS (S1–S6)

- **S1: Holistic Analysis Quality (1–5):** Evaluates the *Base Analysis* in the

<holistic_...> tag according to Principle A.

- **S2: Semantic Analysis Quality (1–5):** Evaluates the *Base Analysis* in the <argument_...> tag according to Principle A.
- **S3: Syntactic Analysis Quality (1–5):** Evaluates the *Base Analysis* in the <syntactic_...> tag according to Principle A.
- **S4: Strategy & Grounding (1–5):** Evaluates the logical derivation of the translation guidelines from S1–S3.
- **S5: Incremental Response Quality (1–5):**
 - If code[1, 2]==0: Score 5 (“Not applicable”).
 - If code[1, 2]==1: *5 (Perfect)*: Correct location (Principle B) and high-quality content (Principles C/D). *3 (Acceptable)*: Correct translation but **Structural Location error** (Principle B violation). *1 (Fail)*: No response or literal translation of pragmatics.
- **S6: Execution & Fidelity Audit (1–5):**
 - *Step 1 (Form)*: Does the final translation **exactly match** the text after the phrase “*In summary, the final translation is determined to be:*”? If no, **Score 1**.
 - *Step 2 (Logic)*: Is the committed translation a logical product of the CoT?
 - *Penalty*: Hard DNT violation or factual error results in **Score 1**.

INPUTS: Source Text: **source_text** | CoT: **cot_analysis** | Final Translation: **final_translation** | Report: **feature_report**

F Qualitative Case Studies: Full Reasoning Traces

In this section, we present the **unabridged** inference traces of CEMT across six distinct scenarios. These examples illustrate how the model handles DNTs and Pragmatic Elements through its structured “Analysis-Strategy-Execution” workflow.

The following case studies visualize the model’s structured reasoning traces, which facilitate post-hoc analysis of element detection and translation

decisions. We do not claim these traces to be faithful explanations, but they offer useful insights into the model's decision process.

Case Study 1: URL Preservation

Source Text:

<https://fgc.network/objects/0f1b42c6-cbb1-49bb-91f5-db81bc71ea14>

Generated CoT & Translation:

<think>

<holistic_semantics_pragmatics_analysis>

[BASE ANALYSIS] This input is a URL string with a feature_code of (0, 1, 0), indicating no semantic or pragmatic complexity. The analysis is limited to identifying and confirming a Hard Do-Not-Translate (DNT) element.

<argument_predicate_analysis>

[BASE ANALYSIS] The string is a URL that serves as a technical identifier for a specific digital object. The substring 0f1b42c6-cbb1-49bb-91f5-db81bc71ea14 is a unique object identifier (UUID) on the fgc.network platform.

[INCREMENTAL RESPONSE – DNT ELEMENT]

DNT Directive: The URL <https://fgc.network/objects/0f1b42c6-cbb1-49bb-91f5-db81bc71ea14> is a Hard DNT. It is a fixed, technical identifier that is not subject to translation or transformation, and therefore must be preserved exactly as it is in the target text.

<syntactic_structure_analysis>

[BASE ANALYSIS] The structure follows a standard URL pattern: a scheme (https://), a domain (fgc.network), a path (/objects/), and a unique identifier (0f1b42c6-cbb1-49bb-91f5-db81bc71ea14). The syntax is simple and conforms to conventional web URL formatting.

<translation_strategy_formulation>

1. **Base Strategy:** Given the feature_code of (0, 1, 0), the analysis is limited to confirming the Hard DNT nature of the URL. As a technical identifier, it must be preserved exactly as it is to maintain its functionality and uniqueness.

2. **Incremental Strategy (DNT & Pragmatics):**

- **DNT Directive:** The URL is a Hard DNT and must be preserved verbatim. It is a fixed technical string that is not subject to translation or transformation.

- **Pragmatic Directive:** No pragmatic elements (e.g., idioms, slang) are present. The URL is a factual technical string and requires no additional functional-equivalence transformation.

3. **Concluding Commitment:**

Based on the analysis above, the strategy is to preserve the Hard DNT nature of the URL. The final translation must be identical to the source URL.

In summary, the final translation is determined to be: <https://fgc.network/objects/0f1b42c6-cbb1-49bb-91f5-db81bc71ea14>

</think>

Final Output:

<https://fgc.network/objects/0f1b42c6-cbb1-49bb-91f5-db81bc71ea14>

Case Study 2: Social Media Handles & Mixed Constraints

Source Text:

Now I need to let y'all know I have, I've never done this before, but I'm going to do it now. Skin potion is handcrafted with organic and all natural ingredients that deeply moisturize and illuminate the skin... And you can follow Skin Potion on Instagram @kpotionus that's S K I N O T I O N Skin potion or Skin potion us and visit their website, www.shopskinpotion.

Generated CoT & Translation:

<think>

<holistic_semantics_pragmatics_analysis>

[BASE ANALYSIS] This is a sentence with cognitive complexity of 2, requiring a comprehensive and in-depth analysis due to its promotional and instructional nature. The sentence provides a narrative structure where the speaker is introducing a new product ("Veil of Venus") and its brand ("Skin Potion") while emphasizing the organic and natural ingredients used in the product. The key challenge lies in accurately capturing the tone, the sequence of information, and the functional equivalence of the pragmatic elements.

[INCREMENTAL RESPONSE - PRAGMATIC ELEMENT]

- **Pragmatic Element:** "y'all", "I've never done this before", "but I'm going to do it now", "handcrafted", "deeply moisturize and illuminate", "newest addition", "available right now for pre order", "that's S K I N O T I O N", "visit their website"

- **Function:** These phrases are used to convey a sense of personalization, humility, and promotional enthusiasm. "y'all" is a colloquial term for "you all" used in informal speech to address a group, adding a friendly and inclusive tone. The phrases like "I've never done this before" and "but I'm going to do it now" are used to build a narrative arc and emphasize the speaker's commitment to a new endeavor. "Handcrafted" and "deeply moisturize and illuminate" are evaluative terms with positive connotations, used to describe the quality and benefits of the product. "Newest addition" and "available right now for pre order" are promotional phrases indicating the product's status and availability. "That's S K I N O T I O N" is a phonetic transcription of the brand name, which is a key identifier. "Visit their website" is a directive phrase used to guide the audience to a specific online resource.

- **Equivalence Loop:** To achieve functional equivalence, these phrases must be translated into Chinese equivalents that preserve the informal tone, promotional nature, and the brand name.

- "y'all" is best rendered as "你们" or "各位" to maintain the inclusive and friendly tone.

- "I've never done this before" is best translated as "我以前从未做过" to preserve the humility.

- "but I'm going to do it now" is best rendered as "但我会现在做" to maintain the contrastive and proactive tone.

- "Handcrafted" is best translated as "手工制作" to preserve the artisanal quality.

- "Deeply moisturize and illuminate" is best rendered as "深层保湿和提亮" to maintain the descriptive and functional benefit.

- "Newest addition" is best translated as "最新产品" to preserve the product's status.

- "Available right now for pre order" is best rendered as "现在可以预订" to maintain the immediacy and availability.
- "That's S K I N O T I O N" is best transliterated as "那就是Skin Potion" to preserve the brand name.
- "Visit their website" is best translated as "访问他们的网站" to maintain the directive function.

<argument_predicate_analysis>

[BASE ANALYSIS] The core predicate-argument structure consists of two main clauses connected by a causal or sequential relationship:

1. "I need to let y'all know I have... but I'm going to do it now" — a general statement about the speaker's intention to inform the audience.
2. "Skin potion is handcrafted with organic and all natural ingredients that deeply moisturize and illuminate the skin. Veil of Venus is the newest addition to the Skin Potion family. It is available right now for pre order. And you can follow Skin Potion on Instagram @kpotionus that's S K I N O T I O N Skin potion or Skin potion us and visit their website, www.shopskinpotion" — a detailed description of the product and its availability.

The key predicate in the second clause is "is handcrafted with organic and all natural ingredients that deeply moisturize and illuminate the skin," which describes the product's composition and benefits. The phrase "Veil of Venus is the newest addition to the Skin Potion family" establishes the product's position within a family of products. The phrase "It is available right now for pre order" indicates the current status and availability. The phrase "you can follow Skin Potion on Instagram @kpotionus" provides a social media handle for engagement. The phrase "visit their website, www.shopskinpotion" is a directive to access the official website.

[INCREMENTAL RESPONSE - DNT ELEMENT]

- **DNT Directive:** The phrases "www.shopskinpotion" and "@kpotionus" are **URL and Handle DNTs**. They are fixed strings that must be preserved exactly as they are in translation to maintain the functional link and brand identity.

- **Recommendation:** These DNTs should be directly replicated in Chinese as they are part of the brand's online presence and are used for navigation and social media engagement.

<syntactic_structure_analysis>

[BASE ANALYSIS] The sentence is a complex narrative sentence with a main clause and a subordinate clause. The structure is: Subject + "I need to let y'all know" + Clause 1 + "but I'm going to do it now" + Clause 2. The second clause contains a compound sentence with two parallel clauses: "Skin potion is handcrafted with organic and all natural ingredients that deeply moisturize and illuminate the skin" and "Veil of Venus is the newest addition to the Skin Potion family. It is available right now for pre order. And you can follow Skin Potion on Instagram @kpotionus that's S K I N O T I O N Skin potion or Skin potion us and visit their website, www.shopskinpotion." The structure is complex due to the multiple clauses and the sequence of information. The phrase "that's S K I N O T I O N" is a phonetic transcription of the brand name, which must be preserved as a direct translation of the brand name.

<translation_strategy_formulation>

1. **Base Strategy:** Given the cognitive complexity of 2, the translation must be comprehensive and semantically accurate, preserving the promotional tone, brand names, and the sequence of information. The key challenge

is to maintain the causal structure and the functional equivalence of the pragmatic elements.

2. Incremental Strategy (DNT & Pragmatics):

- **DNT Directive:** The URLs and handles must be directly replicated: "www.shopskinpotion" → "www.shopskinpotion" and "@kpotionus" → "@kpotionus".
- **Pragmatic Directive:** The phrases must be analyzed through the context-function-equivalence loop:
 - "y'all" → "你们" (informal and inclusive)
 - "I've never done this before" → "我以前从未做过" (humility)
 - "but I'm going to do it now" → "但我会现在做" (contrastive and proactive)
 - "handcrafted" → "手工制作" (preserves the artisanal quality)
 - "deeply moisturize and illuminate" → "深层保湿和提亮" (functional benefit)
 - "newest addition" → "最新产品" (product status)
 - "available right now for pre order" → "现在可以预订" (availability)
 - "that's S K I N O T I O N" → "那就是Skin Potion" (brand name)
 - "visit their website" → "访问他们的网站" (directive)

3. Concluding Commitment:

Based on the analysis above, the strategy involves a comprehensive translation that preserves the promotional tone, brand names, and the sequence of information. The DNT elements are directly replicated. The pragmatic elements are translated to maintain the informal and functional tone.

In summary, the final translation is determined to be: 现在, 我想告诉你们, 我以前从未做过这件事, 但我现在要开始了。Skin Potion 是由有机和天然成分手工制作的, 能够深层保湿和提亮肌肤。Veil of Venus 是Skin Potion 家族的最新产品。目前可以提前预订。您还可以在Instagram 上关注Skin Potion, 账号是@kpotionus (即Skin Potion), 或者访问他们的网站www.shopskinpotion。

</think>

Final Output:

现在, 我想告诉你们, 我以前从未做过这件事, 但我现在要开始了。Skin Potion 是由有机和天然成分手工制作的, 能够深层保湿和提亮肌肤。Veil of Venus 是Skin Potion 家族的最新产品。目前可以提前预订。您还可以在Instagram 上关注Skin Potion, 账号是@kpotionus (即Skin Potion), 或者访问他们的网站www.shopskinpotion。

Case Study 3: Technical Syntax Preservation

Source Text:

1. start of HTML document up to <body>

Generated CoT & Translation:

<think>

<holistic_semantics_pragmatics_analysis>

[BASE ANALYSIS] This is a sentence with cognitive complexity of 1, indicating a factual description of the structure of an HTML document. The analysis should be concise, confirming the core semantic content and the presence of a DNT element.

<argument_predicate_analysis>

[BASE ANALYSIS] The core predicate-argument struc-

ture consists of two parts:

1. "start of HTML document" — a general term referring to the beginning of an HTML file.
2. "<body>" — a specific HTML tag that marks the start of the visible content of a web page.

The phrase "up to <body>" indicates the boundary or limit of the initial part of the document, which includes everything from the start of the document up to the opening of the <body> tag.

[INCREMENTAL RESPONSE - DNT ELEMENT]

- **DNT Directive:** The phrase "<body>" is a Hard DNT. It is a fixed HTML tag that is directly replicated in translation to maintain its structural and functional role. Hard DNTs are preserved exactly as-is because they are part of the markup language and are not subject to semantic or functional change.

<syntactic_structure_analysis>

[BASE ANALYSIS] The structure is a simple phrase: "start of HTML document up to <body>". The phrase is a compound phrase where "start of HTML document" is a general term and "<body>" is a specific tag. The phrase "up to <body>" acts as a boundary marker, indicating the end of the initial part of the document.

<translation_strategy_formulation>

1. **Base Strategy:** Given the cognitive complexity level of 1, the translation should be concise and factually accurate, preserving the structural and functional nature of the <body> tag. The phrase "start of HTML document" should be translated as "HTML文档的开始" to maintain the general context.

2. **Incremental Strategy (DNT & Pragmatics):**

- **DNT Directive:** The Hard DNT for "<body>" must be preserved exactly as-is. It is a fixed tag that is directly replicated in Chinese to maintain its structural role. Therefore, "<body>" → "<body>".

- **Pragmatic Directive:** No pragmatic elements (e.g., idioms, slang) are present. The phrase is factual and technical, requiring a direct and accurate translation.

3. **Concluding Commitment:**

Based on the analysis above, the strategy involves a natural and idiomatic translation that preserves the technical nature and the Hard DNT for the <body> tag.

In summary, the final translation is determined to be: 1. HTML文档的开始, 直到<body>

</think>

Final Output:

1. HTML文档的开始, 直到<body>

asm, and affirmation.

- "ooo" — interjection expressing agreement or excitement.
- "thats a shout" — British slang meaning "that's a good idea" or "great suggestion."
- "defo" — abbreviation of "definitely," reinforcing a strong affirmative tone.
- "one once I've got some skills in" — conveys a conditional commitment: "I'll do it once I'm skilled enough."

• **Equivalence Loop:** Functional equivalents in Chinese should maintain informality and positive tone:

- "ooo" → 哦
- "thats a shout" → 这真是个好主意 / 这太棒了
- "defo" → 当然 / 肯定
- "one once I've got some skills in" → 等我掌握了些技能后 / 等我熟练了之后

<argument_predicate_analysis>

[BASE ANALYSIS] The core structure contains two clauses connected by a sequential relation: the interjectional clause expressing enthusiasm and the conditional clause expressing intent.

• **Predicate:** "is a shout" (means "is a good idea").

• **Argument:** "defo one once I've got some skills in" (means "I'll definitely do it when I'm skilled enough").

[INCREMENTAL RESPONSE - DNT ELEMENT]

DNT Directive: The username "@user25" is a Hard DNT. It must remain unchanged in translation to preserve identity and context.

<syntactic_structure_analysis>

[BASE ANALYSIS] The utterance forms a compound structure: [@user25] + [ooo] + [thats a shout] + [defo one once I've got some skills in]. The final phrase acts as a dependent clause expressing a future-oriented condition (等我掌握了技能之后).

<translation_strategy_formulation>

1. **Base Strategy:** Given its low cognitive complexity, translation should remain concise, colloquial, and natural. Preserve tone and rhythm by using brief, spoken-style equivalents.

2. **Incremental Strategy (DNT & Pragmatics):**

- **DNT Directive:** Preserve "@user25" exactly.

- **Pragmatic Directive:** Apply context-function-equivalence mapping:

- "ooo" → 哦
- "thats a shout" → 这真是个好主意
- "defo" → 当然
- "one once I've got some skills in" → 等我掌握了些技能后

3. **Concluding Commitment:** The translation must maintain the username and the light, enthusiastic tone. Pragmatic elements are replaced with fluent, idiomatic Chinese equivalents.

In summary, the final translation is determined to be: @user25 哦, 这真是个好主意, 等我掌握了些技能后就会做到了。

</think>

Final Output:

@user25 哦, 这真是个好主意, 等我掌握了些技能后就会做到了。

Case Study 4: Slang & Pragmatic Equivalence

Source Text:

@user25 ooo thats a shout, defo one once I've got some skills in

Generated CoT & Translation:

<think>

<holistic_semantics_pragmatics_analysis>

[BASE ANALYSIS] This sentence has a cognitive complexity of 1, indicating a factual yet conversational utterance. It features casual, colloquial markers that convey enthusiasm and agreement.

[INCREMENTAL RESPONSE - PRAGMATIC ELEMENT]

- **Pragmatic Elements:** "ooo", "thats a shout", "defo", "one once I've got some skills in"

- **Function:** These phrases express informality, enthusi-

Case Study 5: Cultural Pragmatics

Source Text:

Skiing is a major travelling activity with many enthusiasts, occasionally known as "ski bums," planning entire vacations around skiing at a particular location.

Generated CoT & Translation:

<think>

<holistic_semantics_pragmatics_analysis>

[BASE ANALYSIS] This is a sentence with cognitive complexity of 1, describing a social phenomenon where skiing is a popular travel activity with many enthusiasts, who are sometimes referred to as "ski bums" and plan their entire vacations around skiing at a specific location.

[INCREMENTAL RESPONSE - PRAGMATIC ELEMENT]

- **Pragmatic Element:** "ski bums"
- **Function:** This term is a **pejorative slang term** used to describe individuals who are passionate about skiing and spend a significant portion of their lives traveling to different ski resorts. It is often used in a playful or humorous context to refer to dedicated skiers who live a nomadic lifestyle, often working part-time or not at all, to afford their passion.
- **Equivalence Loop:** To achieve functional equivalence, a natural Chinese equivalent must preserve the playful and informal nature of the term. While there is no direct equivalent in Chinese, terms like "滑雪狂热者" or "滑雪爱好者" are close in meaning. However, "ski bums" is best rendered as "滑雪族" or "滑雪爱好者" in Chinese, as it is a commonly used term in Chinese to refer to dedicated skiers. The term "ski bums" is best preserved in translation to maintain the playful and informal tone.

<argument_predicate_analysis>

[BASE ANALYSIS] The core predicate-argument structure consists of two main clauses connected by a causal or sequential relationship:

1. "Skiing is a major travelling activity with many enthusiasts"
2. "occasionally known as 'ski bums', planning entire vacations around skiing at a particular location"

The key predicate in the second clause is "planning entire vacations around skiing at a particular location," which is a complex phrase meaning that these enthusiasts dedicate their entire vacations to skiing at a specific destination.

[INCREMENTAL RESPONSE - DNT ELEMENT]

- **DNT Directive:** No DNT elements are present (feature_code[1] == 0). Therefore, no DNT handling is required.

<syntactic_structure_analysis>

[BASE ANALYSIS] The sentence is a complex English sentence with a main clause and a subordinate clause introduced by the phrase "occasionally known as 'ski bums'." The structure is: Subject ("Skiing") + Predicate ("is a major travelling activity with many enthusiasts") + Clause ("occasionally known as 'ski bums', planning entire vacations around skiing at a particular location"). The phrase "at a particular location" is a prepositional phrase modifying "skiing," indicating the specific destination.

<translation_strategy_formulation>

1. **Base Strategy:** Given the cognitive complexity level of 1, the translation should be concise and focused on preserving the main idea and the playful tone of the term "ski bums." The key challenge is accurately rendering the phrase "occasionally known as 'ski bums'" into Chinese, which should be translated as "有时被称为'滑雪族'" to

maintain the informal and descriptive nature of the term.

2. Incremental Strategy (Pragmatics):

- **Pragmatic Directive:** The term "ski bums" must be analyzed through the context-function-equivalence loop:
 - **Context:** The context is about dedicated skiers who live a nomadic lifestyle and spend their vacations skiing.
 - **Function:** The function of "ski bums" is to describe a subculture of passionate skiers who prioritize skiing over other activities and often live a lifestyle centered around their hobby.
 - **Equivalence:** A direct translation like "滑雪bum" is not idiomatic in Chinese. A natural and functionally equivalent term is "滑雪族" or "滑雪爱好者" which are commonly used in Chinese to refer to dedicated skiers. Therefore, "ski bums" is best rendered as "滑雪族" to preserve the playful and descriptive nature of the term.

3. Concluding Commitment:

Based on the analysis above, the strategy involves a natural and idiomatic translation that preserves the main idea, the playful tone, and the functional equivalence of the term "ski bums." The phrase "planning entire vacations around skiing at a particular location" is best rendered as "计划整个假期在某个地点滑雪" to maintain the idea of dedicated vacation planning.

In summary, the final translation is determined to be: 滑雪是一项深受许多爱好者喜爱的旅行活动, 有时被称为“滑雪族”, 他们会计划整个假期在某个特定地点进行滑雪。

</think>

Final Output:

滑雪是一项深受许多爱好者喜爱的旅行活动, 有时被称为“滑雪族”, 他们会计划整个假期在某个特定地点进行滑雪。

Case Study 6: Complex Logic & Idiom Translation

Source Text:

通过预售会员卡获取现金流, 不断开新店, 看似能够循环往复, 实则寅吃卯粮, 无法形成稳定的盈利模式。

Generated CoT & Translation:

<think>

<holistic_semantics_pragmatics_analysis>

[BASE ANALYSIS] This is a sentence with cognitive complexity of 2, requiring a comprehensive and in-depth analysis due to its strategic and structural complexity. The sentence describes a business strategy involving pre-selling membership cards to generate cash flow and opening new stores, which appears sustainable but is actually unsustainable and leads to financial depletion.

[INCREMENTAL RESPONSE - PRAGMATIC ELEMENT]

- **Pragmatic Element:** “寅吃卯粮”
- **Function:** This idiom means “living beyond one’s means” or “eating the seed corn,” referring to spending money that should be saved for the future, often used in economic or financial contexts to describe unsustainable practices.
- **Equivalence Loop:** To achieve functional equivalence, a natural English equivalent must preserve the metaphorical and cautionary tone. While there are several possible translations (e.g., “living beyond one’s

means,” “eating the seed corn,” “spending tomorrow’s money today”), the most idiomatic and contextually accurate equivalent in English is “living beyond one’s means” or a close variant like “spending money that should be saved for the future.”

<argument_predicate_analysis>

[BASE ANALYSIS] The core predicate-argument structure consists of two main clauses connected by a causal or contrastive relationship:

1. “通过预售会员卡获取现金流” (by pre-selling membership cards to generate cash flow)
2. “不断开新店” (continuously opening new stores)

These two actions are presented as a seemingly sustainable strategy, but the final clause “实则寅吃卯粮，无法形成稳定的盈利模式” (in reality, it is living beyond one’s means, and cannot form a stable profit model) provides a critical evaluation.

[INCREMENTAL RESPONSE - DNT ELEMENT]

- **DNT Directive:** No DNT elements are present (feature_code[1] == 0). Therefore, no DNT handling is required.

<syntactic_structure_analysis>

[BASE ANALYSIS] The sentence is a complex compound sentence with a main clause and a subordinate clause introduced by the phrase “实则...” (in reality...). The structure is: [Through pre-selling membership cards, we can obtain cash flow, and continuously open new stores, which seems to be a sustainable cycle, but in reality, it is living beyond one’s means, and cannot form a stable profit model.] The phrase “实则寅吃卯粮” is a rhetorical contrastive clause that introduces the true nature of the strategy, which is unsustainable.

<translation_strategy_formulation>

1. **Base Strategy:** Given the cognitive complexity of 2, the translation must be comprehensive and semantically accurate, preserving the causal structure and critical tone. The key challenge is accurately rendering the idiom “寅吃卯粮” into an English equivalent that maintains the metaphorical and cautionary nature.

2. Incremental Strategy (Pragmatics):

- **Pragmatic Directive:** The idiom “寅吃卯粮” must be analyzed through the context-function-equivalence loop:

- **Context:** The context is financial and strategic, where a business model is being critiqued for its unsustainable nature.
- **Function:** The function is to convey the idea of short-term gain at the expense of long-term viability.
- **Equivalence:** A direct translation like “eating the seed corn” or “living beyond one’s means” is functionally equivalent in English and preserves the metaphorical and critical tone. “Spending money that should be saved for the future” is also a close equivalent.

- Therefore, “实则寅吃卯粮” is best rendered as “in reality, it is living beyond one’s means” or a similar phrase that conveys the idea of unsustainable spending.

3. Concluding Commitment:

Based on the analysis above, the strategy involves a comprehensive translation that preserves the causal structure, critical tone, and the idiom “寅吃卯粮” is rendered as “living beyond one’s means” for functional equivalence. In summary, the final translation is determined to be: By pre-selling membership cards, we can obtain cash flow, and continuously open new stores, which seems to be a

sustainable cycle, but in reality, it is living beyond one’s means, and cannot form a stable profit model.

</think>

Final Output:

By pre-selling membership cards, we can obtain cash flow, and continuously open new stores, which seems to be a sustainable cycle, but in reality, it is living beyond one’s means, and cannot form a stable profit model.