Thinker-DDM: Modeling Deliberation for Machine Translation with a Drift-Diffusion Process

Hongbin Na[♠,*], **Zimu Wang**[♡,*], **Mieradilijiang Maimaiti**[♣,†], **Tong Chen**[♡],
Wei Wang[♡], **Tao Shen**[♠], **Ling Chen**[♠]
[♠]Australian AI Institute, University of Technology Sydney
[♡]Xi'an Jiaotong-Liverpool University [♣]Chinese Academy of Sciences
Hongbin.Na@student.uts.edu.au, {Zimu.Wang19,Tong.Chen19}@student.xjtlu.edu.cn
miradel_51@hotmail.com, Wei.Wang03@xjtlu.edu.cn
{Tao.Shen,Ling.Chen}@uts.edu.au

Abstract

Large language models (LLMs) have demonstrated promising potential in various downstream tasks, including machine translation. However, prior work on LLM-based machine translation has mainly focused on better utilizing training data, demonstrations, or predefined and universal knowledge to improve performance, with a lack of consideration of decision-making like human translators. In this paper, we incorporate **Thinker** with the **D**rift-**D**iffusion **M**odel (Thinker-DDM) to address this issue. We then redefine the Drift-Diffusion process to emulate human translators' dynamic decision-making under constrained resources. We conduct extensive experiments under the high-resource, low-resource, and commonsense translation settings using the WMT22 and CommonMT datasets, in which Thinker-DDM outperforms baselines in the first two scenarios. We also perform additional analysis and evaluation on commonsense translation to illustrate the high effectiveness and efficacy of the proposed method.

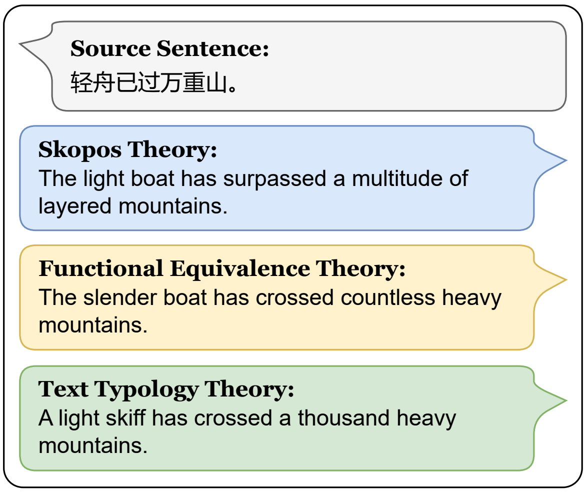

Figure 1: An example of translating a Chinese poetry to English using different translation theories. **Skopos theory** emphasizes the objectives (mood and emotion); **functional equivalence theory** underscores target culture; **text typology theory** concentrates on the characteristics of different text types (poetic and rhythmic).

1 Introduction

Large language models (LLMs), such as GPT-4 (OpenAI et al., 2023), GLM-130B (Zeng et al., 2023), and LLaMA (Touvron et al., 2023a), have recently achieved the state-of-the-art performance in a variety of downstream tasks, such as information extraction (Peng et al., 2023), text summarization (Wang et al., 2023b), and adversarial attacks (Wang et al., 2023d). One particular area where LLMs have demonstrated promising potential is machine translation (Zhang et al., 2022; Vilar et al., 2023), and they also excel under low-resource and zero-resource scenarios (Moslem et al., 2023a; Zhu et al., 2023a).

Recent research on LLM-based machine translation can be broadly categorized into two paradigms: in-context learning (ICL, Brown et al., 2020) that conditions on natural language instructions or demonstrations, and fine-tuning that updates the model parameters based on the availability of usually limited amount of labelled data. In the context of ICL, researchers concentrate on leveraging optimal in-context examples (Agrawal et al., 2023; Sarti et al., 2023; Iyer et al., 2023), dictionary knowledge (Ghazvininejad et al., 2023; Lu et al., 2023), adaptive learning (Moslem et al., 2023a; Reinauer et al., 2023), and translation memories (Reheman et al., 2023). Numerous studies have also fine-tuned LLMs to augment their capacity in translating unseen languages (Yang et al., 2023; Mao and Yu, 2024) and domains (Moslem et al., 2022, 2023b) and building multilingual LLMs (Zhang et al., 2023; Zhu et al., 2023b).

Despite this success, prior work on LLM-based machine translation, whether through in-context

[*]Equal contribution.
[†]Corresponding author.

Proceedings of The 23rd Annual Workshop of the Australasian Language Technology Association, pages 45–63
November 26-28, 2025 ©2025 Association for Computational Linguistics

learning or fine-tuning, shares a common limitation: they typically apply a uniform translation strategy (He et al., 2023) to all sentences within a given task. This static approach contrasts sharply with the methodology of human translators. A human expert engages in a dynamic and adaptive process, deliberating on the unique challenges posed by each sentence. For instance, they may consider various established translation strategies, such as the Skopos theory (Reiß and Vermeer, 2010), functional equivalence theory (Nida, 1964), and text typology theory (Reiss, 1989) (see Figure 1), and apply them interchangeably based on the source text's nuances. The crucial challenge, therefore, is not merely providing the model with access to different strategies, but equipping it with a mechanism to decide when and how to apply them, mimicking the resource-constrained deliberation of a human. This cognitive aspect of translation—the sentence-level, dynamic decision-making—remains largely unexplored.

We introduce a new machine translation approach that incorporates **Thinker** with **D**rift-**D**iffusion **M**odel (Thinker-DDM) to address the aforementioned challenges. Thinker-DDM incorporates a decision-making process characterized by drift and diffusion to simulate human cognitive behavior. Specifically, we first design relevant translation strategy prompts in line with the three representative theories in translation. Then, we redefine the Drift-Diffusion process, incorporating the processes of initial bias (drift), evidence gathering (diffusion), and boundary-driven decision-making to emulate how human translators behave with resource constraints. We conduct extensive experiments under the high-resource, low-resource, and commonsense translation settings using the WMT22 (Kocmi et al., 2022) and CommonMT (He et al., 2020) datasets. Evaluation results show that Thinker-DDM outperforms Microsoft Translator, GPT-3.5, and Hybrid Max-Routing (Hendy et al., 2023) baselines in the first two scenarios. We also carry out additional analysis and evaluation on the translation strategies and DDM, which highlight the effectiveness and efficacy of the proposed method.

The main contributions are summarized as follows: (1) We propose Thinker-DDM, a novel machine translation approach to simulate human translators' dynamic decision-making process. (2) We define relevant translation strategy prompts in line with representative theories in translation and redefine the Drift-Diffusion process to adapt to the machine translation task. (3) We conduct comprehensive experiments across multiple languages to demonstrate the high effectiveness and efficacy of the proposed method.

2 Preliminaries

Translation Studies. Translation studies is a multidisciplinary field that centers on translation as a means of cross-cultural and cross-linguistic communication (Munday et al., 2022). Theorists have developed a variety of theories to guide and interpret the translation process, encompassing significant schools of thought and perspectives. As shown in Figure 1, *Skopos theory* (Reiß and Vermeer, 2010) posits that translation activities should revolve around their pre-determined objectives, emphasizing the guiding role of these purposes in shaping translation strategies. Complementing this, *functional equivalence theory* (Nida, 1964) underscores the function and impact of translations within the target culture, focusing on how translations bridge meaning across diverse cultures. Additionally, *text typology theory* (Reiss, 1989) concentrates on the unique characteristics of different text types and their influence on translation requirements, advocating for adaptive translation strategies tailored to specific text categories, such as informative, expressive, or operative. Together, these theories construct a profound perspective that transcends literal translation, highlighting the importance of cultural adaptability, target audience needs, and effective communication in the translation process (Saroukhil et al., 2018).

Individual Decision-Making Processes. In exploring the nuances of individual decision-making, dual-system theory offers a key framework, differentiating two modes of thinking: intuitive System 1 and analytical System 2. System 1 governs low-risk, familiar decisions with its fast, automated processes, whereas System 2 is engaged in complex, logical decision-making scenarios (Kahneman, 2011). Despite its insights, this theory has limitations in fully capturing decision-making dynamics. Here, DDM presents a more nuanced understanding by portraying decision-making as a continuous accumulation of evidence, leading to a choice (Ratcliff, 1978). It highlights the process's stability and speed, especially in decisions grounded in clear value orientations, aligning with value-based decision theory (Rowland, 1946), elucidating how internal values and external information guide rapid, precise choices (Ratcliff et al., 2016).

3 Methodology

Following the overall framework of our proposed Thinker-DDM method illustrated in Figure 2, in this section, we introduce each part of the framework in detail.

3.1 Translation Strategy Prompts

In the pursuit of optimizing translation quality, we first develop a series of translation strategy prompts that draw on representative theories in translation studies. These prompts facilitate a deeper analysis of source sentences, taking into account various aspects during human translation, such as target audiences, key information, information equivalence, and culture equivalence. This ensures a satisfactory level of linguistic accuracy as well as cultural and contextual relevance. Figure 3 shows our suggested translation strategies, which are categorized according to their underlying theoretical framework: Skopos theory, functional equivalence theory, and text type theory. With each prompt category targeting a specific facet of the translation process, it is possible to conduct a thorough analysis that will guide the translation strategy effectively. The detailed content and examples of each of the seven prompts are organized in Appendix A.

This structured approach, grounded in well-established theoretical frameworks, offers a comprehensive methodology to address the multi-faceted nature of translation. Systematically applying these prompts to source sentences enriches the translation process and guarantees that translations are accurate linguistically as well as appropriately attuned to contextual and cultural nuances. While the translation strategy prompts provide valuable insights for source sentences, they may also introduce extraneous noises, which lowers the overall quality and accuracy of the translation.

3.2 Drift-Diffusion Model

DDM is a framework for quantifying the process of rapid decision-making under uncertainty. It decomposes the decision-making process into several essential components: **drift**, **diffusion**, and **boundary conditions**, as shown in Figure 2. This model is particularly prevalent in studies involving reaction time and decision accuracy, offering a mathematical approach to understanding cognitive processes (Bogacz et al., 2006).

Drift represents the average rate of information accumulation in favor of a particular decision. It is defined as the mean of the stochastic process and is denoted by μ. Mathematically, it can be expressed as:

$$\mu = \frac{dE[X(t)]}{dt}, \tag{1}$$

where $E[X(t)]$ is the expected value of the decision variable X at time t. Drift rate is a critical parameter as it reflects the strength of evidence and influences the speed-accuracy tradeoff in the decision-making process.

Diffusion refers to the stochastic or random component of the decision-making process, which accounts for the inherent variability and unpredictability in accumulating evidence. This random fluctuation is often modeled mathematically as a process with a variance parameter σ^2, encapsulating the element of uncertainty in evidence accumulation. The diffusion component of the decision process can be expressed as:

$$dX(t) = \mu dt + \sigma dB(t), \tag{2}$$

where $dX(t)$ represents the instantaneous change in the decision variable at time t, quantifying the accumulated evidence and $dB(t)$ denotes the incremental, stochastic change, analogous to the random noise, in the evidence accumulation over time. This term captures the essence of uncertainty and randomness in the decision-making process, reflecting how real-world decisions are influenced by factors that cannot always be predicted or quantified precisely.

Boundary Conditions define the thresholds for making a decision. When the accumulated evidence, represented by the decision variable $X(t)$, reaches one of these boundaries, a decision is made. The boundaries are often set symmetrically at $\pm A$, with A being a positive constant. These boundaries not only determine the decision outcome but also play a crucial role in modeling reaction times. The time at which $X(t)$ first hits either boundary is the predicted reaction time for the decision.

3.3 Translation Thinker with DDM

In this study, we redefine and apply DDM to translation evaluation, aiming to develop an innovative model for choosing the optimal translation of a

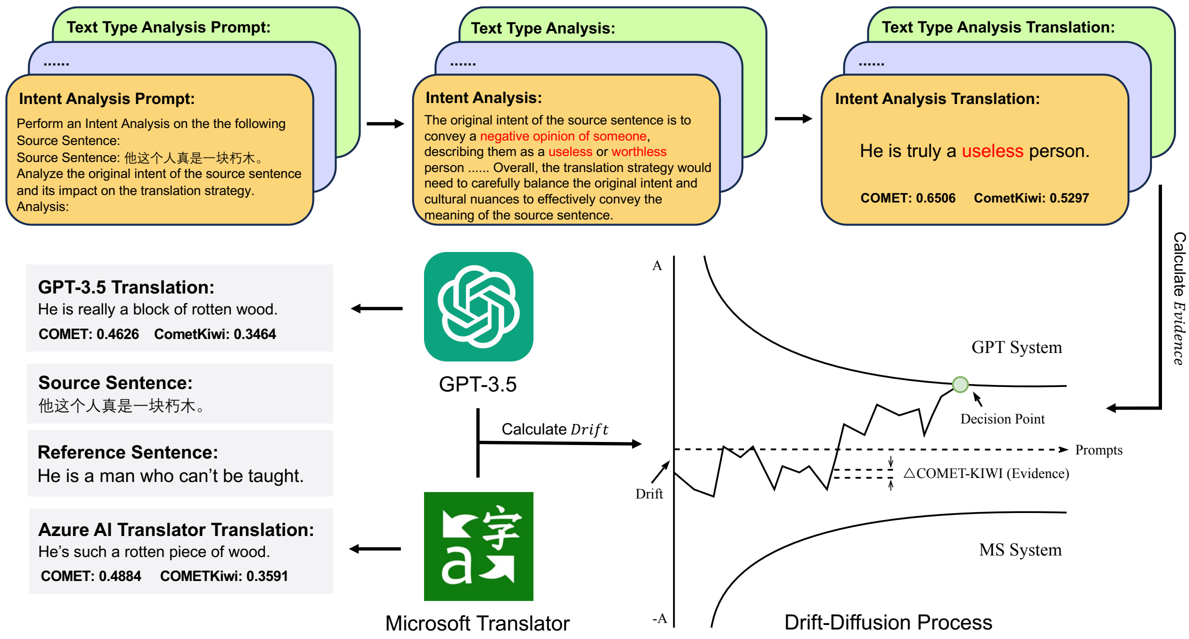

Figure 2: An overview of our proposed Thinker-DDM method. It first utilizes relevant translation strategy prompts in line with the three representative theories in translation, and then redefine the Drift-Diffusion process, incorporating the processes of initial bias (drift), evidence gathering (diffusion), and boundary-driven decision-making.

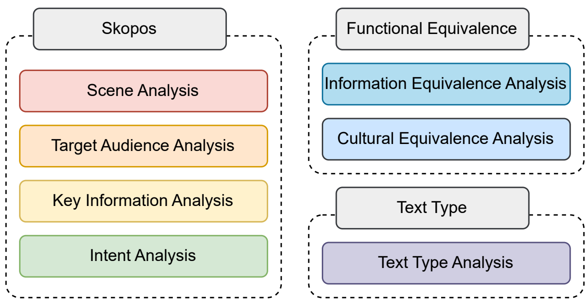

Figure 3: Our suggested translation strategies for Thinker-DDM and their underlying theoritical frameworks.

given source sentence. It involves a series of iterative steps to determine the optimal translation for a specific sentence, as the overall procedure illustrated in Algorithm 1.

We redefine ***drift*** as an initial preference, implying that there is a bias towards a particular translation system prior to gathering evidence. We utilize the CometKiwi model $\mathcal{C}$ (Rei et al., 2022b) to conduct a reference-free quality assessment to imitate the human translation process. We assess the translation by both Microsoft Translator[1] $T_{MS}(S)$ and GPT-3.5 $T_{GPT}(S)$, where S is the original sentence. Mathematically, drift is defined as the difference in the CometKiwi scores between these two methods of translation:

$$Drift = \mathcal{C}(T_{GPT}(S)) - \mathcal{C}(T_{MS}(S)). \quad (3)$$

Subsequently, a series of translation strategy prompts $\mathcal{P} = \{P_1, \dots P_i, \dots, P_n\}$ are proposed to generate new translation versions $T_{P_i}(S)$. The

Algorithm 1 Translation Thinker with DDM

Input: Original sentence S
Output: Best translation T_{best}
Initialize $T_{MS}(S), T_{GPT}(S), \mathcal{P}$;
Initialize A_{up}, A_{low}, decay;
$Drift \leftarrow \mathcal{C}(T_{GPT}(S)) - \mathcal{C}(T_{MS}(S))$;
for each prompt P_i in $\mathcal{P}$ **do**
  Generate new translation $T_{P_i}(S)$;
  $Diffusion_i \leftarrow \mathcal{C}(T_{P_i}(S)) - \mathcal{C}(T_{MS}(S))$;
  Update $Drift \leftarrow Drift + Diffusion_i$;
  Update boundaries:
  $A_{up/low} \leftarrow A_{up/low} \times e^{-\text{decay}}$;
  if $Drift \geq A_{up}$ or $Drift \leq A_{low}$ **then**
    break;
if $Drift \geq A_{up}$ **then**
  $T_{best} \leftarrow$ translation with highest $\mathcal{C}$;
else if $A_{up} > Drift > A_{low}$ **then**
  $T_{best} \leftarrow$ translation with highest $\mathcal{C}$;
else if $Drift \leq A_{low}$ **then**
  $T_{best} \leftarrow T_{MS}(S)$;
return T_{best}

[1] https://www.microsoft.com/en-us/translator/

order of cues is randomized to simulate the stochastic process of diffusion. For each prompt P_i, its corresponding ***diffusion*** is calculated as:

$$Diffusion_i = \mathcal{C}(T_{P_i}(S)) - \mathcal{C}(T_{MS}(S)). \quad (4)$$

in which $Diffusion_i$ represents the evidence collected in each iteration. In each iteration, the drift value *Drift* is updated to $Drift + Diffusion_i$, while checking whether the pre-determined decision boundaries, the upper boundary A_{up} and the lower boundary A_{low}, have been reached. The boundary values are dynamically adjusted and decreases in an exponential decay manner, formally,

$$A^{new}_{up/low} = A^{old}_{up/low} \times e^{-\text{decay}}. \quad (5)$$

Such exponential decay of the boundary values allows the model to maintain a wider decision range in the initial stages, allowing for ample exploration; as time progresses, the boundary values decrease, enabling the model to make decisions faster, thereby enhancing efficiency.

Based on the final value of the *Drift*, we select the optimal translation. If $Drift \geq A_{up}$, we choose the translation with the highest CometKiwi score from all generated translations; otherwise, if $Drift \leq A_{low}$, we select Microsoft Translator's result $T_{MS}(S)$. For those that do not reach the boundary until the end of the iteration, the approach is consistent with reaching the upper boundary, and we select the translation with the highest CometKiwi score since all translation candidates have been generated.

4 Experiments

4.1 Datasets

We conducted experiments on the WMT22 dataset (Kocmi et al., 2022), which is widely utilized in machine translation research. It encompasses both high-resource and low-resource languages, predominantly featuring English (EN) ⇔ Chinese (ZH), English (EN) ⇔ German (DE), English (EN) ⇔ Japanese (JA), and German (DE) ⇔ French (FR) language pairs. Additionally, to explore translation performance in low-resource languages, particular emphasis of the dataset was placed on Czech (CS) ⇔ Ukrainian (UK) and English (EN) ⇔ Ukrainian (UK). We sampled a subset of 500 sentences for each language pair from the test set of the dataset for experiments.

To further investigate the capability of translation models in handling semantic ambiguities, we incorporated Chinese-to-English translation samples from the CommonMT dataset (He et al., 2020).

4.2 Evaluation Metrics

We adopted two mainstream reference-based evaluation metrics in machine translation, COMET (Rei et al., 2022a) and BLEURT (Sellam et al., 2020), to assess Thinker-DDM's performance. These model-based metrics have shown to be superior to conventional string-based metrics, such as BLEU (Papineni et al., 2002), and have been widely employed in LLM-based machine translation literature (Hendy et al., 2023; Moslem et al., 2023a). In accordance with the previous work, we utilized `wmt22-comet-da`[2] and `BLEURT-20`[3] checkpoints for the selected metrics, respectively.

4.3 Baselines

We examined Thinker-DDM against a variety of representative baselines, encompassing both the single-candidate and multiple-candidate approaches:

- **Microsoft Translator (MS-Translator)** is an off-the-shelf commercial machine translation system provided by Microsoft that is able to translate text instantly or in batches across more than 100 languages. We accessed the translator through the public API provided by Microsoft Azure[4].
- **GPT-3.5** is the standard zero-shot translator using the GPT-3.5 model. We employed the prompts as shown in Table 12 to ensure precise and cohesive translations.
- **Hybrid Max-Routing (Max-Routing)** determines the upper limit by choosing the superior translation from the two aforementioned systems, i.e. MS-Translator and GPT-3.5, based on CometKiwi evaluations. We referred to it as the "Hybrid Max-Routing" approach as illustrated in Hendy et al. (2023).

4.4 Experimental Settings

Thinker-DDM iteratively selected the optimal translation by calculating the new evidence provided by new translation candidates, for which the results were obtained by the GPT-3.5 model.

[2] https://github.com/Unbabel/COMET
[3] https://github.com/google-research/bleurt
[4] https://azure.microsoft.com/en-us/products/ai-services/ai-translator/

We accessed GPT-3.5 by calling the official API released by OpenAI[5], and we selected the `gpt-3.5-turbo-instruct` checkpoint, a refined version of GPT-3 designed to perform natural language tasks with heightened accuracy and reduced toxicity. We set the temperature as 0.5 to simulate the randomness in human reasoning. Given that both systems already possess high translation capabilities, we set the initial boundaries A_{up} and A_{low} as 0.05 and -0.05, respectively, and we set the decay as 0.2.

4.5 Main Results

High-resource Translation Results The high-resource translation results of Thinker-DDM against baselines are reported in Table 1. The Thinker-ALL method took all the translation results across different translation strategies as well as MS-Translator and GPT-3.5 into account. It selected the optimal translation results under CometKiwi for the nine candidates. From the table, it is evident that Thinker-DDM and Thinker-ALL exhibited superior translation quality across most evaluated language pairs, outperforming all baselines with both COMET22 and BLEURT metrics. This indicates its effectiveness and generalizability, which is not limited to specific language pairs but is consistent across various linguistic combinations. In addition, there are only marginal differences in scores between Thinker-DDM and Thinker-ALL. Such close performance suggested that Thinker-DDM is highly efficient and stable. Since Thinker-DDM acheived the optimal or second-best results in most cases, it is even more stable than Thinker-ALL. In addition, it potentially reduces computational resources and processing requests.

Low-resource Translation Results Machine translation under the low-resource scenario presents a challenge, primarily due to the limited data available for these languages. To investigate the performance of Thinker-DDM under such settings, we conducted experiments and reported the experimental results of translation between low-resource languages (i.e. CS ⇔ UK) and between high-resource and low-resource languages (i.e. EN ⇔ UK) under the COMET evaluation metrics in Table 2. The results were also consistent with those in high-resource translation. Hybrid Max-Routing, Thinker-DDM, and Thinker-ALL exhibited commendable translation quality, emphasizing the significance of customized and optimized algorithms in the translation of low-resource languages.

[5] https://platform.openai.com/

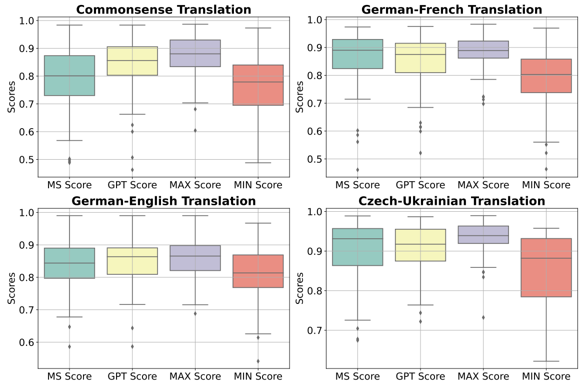

Figure 4: Experimental results of the COMET scores of MS-Translator's and GPT-3.5's translations and the maximum and minimum scores post-prompt.

Commonsense Translation Results In the field of machine translation, data-driven approaches have become dominant, leading to a notable decline in the explicit exploration of world knowledge(He et al., 2020). Our method attempted to bridge this gap to handle translation tasks requiring a comprehensive grasp of commonsense knowledge and cultural background. However, as shown in Table 3, the performance of Thinker-DDM in commonsense translation did not meet expectations. We hypothesized that this might be related to the inability of the CometKiwi evaluation metric to assess translation quality accurately.

We conducted supplementary experiments to test our hypothesis, wherein the translation quality evaluator, CometKiwi, was substituted with COMET. As shown in Table 3, the experimental results witnessed a substantial enhancement with both COMET and BLEURT metrics, implying that incorporating translation theories can endow models with a more profound comprehension of context and cultural nuances. Contrarily, CometKiwi lacks the capability to discriminate between varying levels of translation quality in commonsense machine translation; this also implies the need in designing a better reference-free evaluation metric in the future.

4.6 Additional Analysis

Effectiveness of Translation Strategies In exploring the impact of translation strategies, our initial focus was on their role in expanding the candidate space. To understand their roles more profoundly, we conducted an additional experiment

Method	DE-EN	EN-DE	ZH-EN	EN-ZH	JA-EN	EN-JA	FR-DE	DE-FR	Average
					COMET22				
MS-Translator	0.8532	0.8716	0.8118	0.8712	0.8234	0.8904	0.8655	0.8429	0.8538
GPT-3.5	0.8527	0.8688	0.8290	0.8663	0.8298	0.8850	0.8689	0.8386	0.8549
Max-Routing	0.8585	0.8762	<u>0.8299</u>	<u>0.8809</u>	0.8362	0.8969	**0.8772**	<u>0.8478</u>	0.8630
Thinker-DDM	<u>0.8588</u>	**0.8765**	<u>0.8299</u>	**0.8810**	<u>0.8369</u>	<u>0.8989</u>	<u>0.8757</u>	**0.8500**	<u>0.8635</u>
Thinker-ALL	**0.8597**	<u>0.8764</u>	**0.8304**	0.8806	**0.8398**	**0.9012**	0.8745	0.8466	**0.8637**
					BLEURT				
MS-Translator	0.7430	0.7767	0.6898	0.7270	0.6903	0.6959	0.7842	<u>0.7317</u>	0.7298
GPT-3.5	0.7422	0.7726	0.7163	0.7199	0.6959	0.6717	0.7888	0.7119	0.7274
Max-Routing	0.7495	<u>0.7814</u>	0.7159	**0.7365**	0.7049	0.6997	<u>0.7979</u>	0.7301	0.7395
Thinker-DDM	**0.7518**	0.7810	<u>0.7173</u>	<u>0.7339</u>	<u>0.7058</u>	<u>0.7029</u>	**0.7981**	**0.7373**	**0.7410**
Thinker-ALL	<u>0.7513</u>	**0.7821**	**0.7211**	0.7307	**0.7088**	**0.7034**	0.7978	0.7250	<u>0.7400</u>

Table 1: High-resource translation results of Thinker-DDM against baselines on the WMT22 dataset, in which the optimal and second-best results are highlighted in **bold** and <u>underlined</u>, respectively.

Method	CS-UK	UK-CS	EN-UK	UK-EN
MS-Translator	0.9096	0.9149	0.8842	0.8566
GPT-3.5	0.9014	0.9056	0.8746	0.8529
Max-Routing	0.9156	0.9211	0.8898	**0.8634**
Thinker-DDM	**0.9195**	<u>0.9243</u>	<u>0.8950</u>	<u>0.8630</u>
Thinker-ALL	<u>0.9176</u>	**0.9245**	**0.8980**	0.8574

Table 2: Low-resource translation results of Thinker-DDM against baselines under the COMET evaluation metric.

Method	COMET	BLEURT
MS-Translator	0.8156	0.7010
GPT-3.5	0.8405	0.7326
Max-Routing	0.8383	0.7278
Thinker-DDM	0.8340	0.7244
Thinker-ALL	0.8364	0.7253
Thinker-DDM (COMET)	<u>0.8739</u>	<u>0.7713</u>
Thinker-ALL (COMET)	**0.8818**	**0.7853**

Table 3: Commonsense translation results of Thinker-DDM against baselines.

as follows: we first randomly sampled 50 sentence pairs from commonsense translation, DE-FR, DE-EN, and CS-UK. Then we calculated the scores of MS-Translator and GPT-3.5, as well as the maximum and minimum scores of post-prompt under the COMET metric.

As shown in Figure 4, the translation strategies across all translation settings provided both higher upper and lower bottom limits, indicating the positive role of translation strategies in enhancing translation performance. More importantly, the risks of performance degradation due to the lower limit can be efficiently eliminated through the Drift-Diffusion process. This ensures the quality and stability of translation while expanding the range of candidate translations.

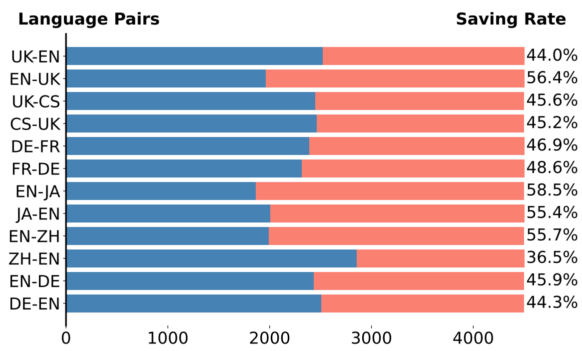

Figure 5: Query saving rate across different language pairs.

Effectiveness of Drift-Diffusion Model In the main experiments, we concluded the effectiveness of DDM on translation accuracy; however, the effectiveness of DDM in translation efficiency remains explored. Following the experimental results shown in Figure 5, Thinker-DDM achieved an average reduction of 48% in query volume compared to the Thinker-ALL, underscoring our approach's enhanced efficiency and underlining the significance of decision theory optimization in settings with limited resources while achieving satisfactory translation outcomes.

Building on the aforementioned insights, we fi-

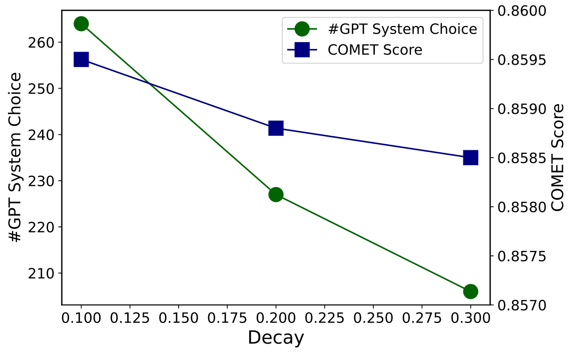

Figure 6: Correlations between the decay factor and the number of GPT system choice and translation performance.

nally turned our attention to the influence of the decay factor on decision-making in translation tasks. To this end, we focused on the DE-EN translation, set the boundary at 0.05 and experimented with decay values of $\{0.1, 0.2, 0.3\}$. As shown in Figure 6, the decay value had a negative correlation with the number of selecting the GPT system, indicating the significance of selecting an appropriate decay value for balancing translation effectiveness and efficiency.

5 Related Work

Decision Theory in LLMs LLMs have recently made significant strides in emulating human cognitive processes, particularly focusing on enhancing decision-making and problem-solving capabilities. System 2 Attention (Weston and Sukhbaatar, 2023) and Tree of Thoughts (ToT, Yao et al., 2023) are exemplary in this regard. Both initiatives mark a shift towards mimicking human "System 2" cognitive processes, involving conscious, analytical thought, and logical reasoning. In a similar vein, developments in agent-based systems have also garnered attention (Liu et al., 2023; Tian et al., 2023; Wang et al., 2023c). These systems collectively emphasize the integration of rapid responsiveness, robust reasoning, and emotional in LLMs, showcasing a trend towards creating more interactive, intuitive, and human-like LLMs. Additionally, DOMINO (Wang et al., 2023a) presents a dual-system approach for multi-step visual language reasoning, adeptly managing complex questions and information extraction. Together, these advancements illustrate the broadening scope of decision theory in LLMs, moving towards systems that not only mimic human cognitive processes but also interact and reason in dynamically complex environments.

Machine Translation Machine translation has been extensively studied over the past few years by transforming multiple paradigms. Early research utilizes rule-based methods (Forcada et al., 2011) to handle the machine translation task; some other methods exploit statistical (Koehn et al., 2003, 2007) and deep learning (Zheng et al., 2021; Maimaiti et al., 2022) techniques to enhance translation performance.

Recently, the rapid advancement of LLMs catalyzes the research on LLM-based machine translation, consisting of two paradigms: ICL and finetuning. In the context of ICL, researchers concentrate on leveraging optimal in-context examples (Agrawal et al., 2023; Sarti et al., 2023; Iyer et al., 2023), dictionary knowledge (Ghazvininejad et al., 2023; Lu et al., 2023), adaptive learning (Moslem et al., 2023a; Reinauer et al., 2023), and translation memories (Reheman et al., 2023). Additionally, numerous studies have also fine-tuned LLMs to augment their capacity in translating unseen languages (Yang et al., 2023; Mao and Yu, 2024) and domains (Moslem et al., 2022, 2023b) and building multilingual LLMs (Zhang et al., 2023; Zhu et al., 2023b). Furthermore, some research also concentrates on post-editing the translation outcomes (Moslem et al., 2023c; Raunak et al., 2023) and leveraging LLMs for machine translation evaluation (Fu et al., 2023; Fernandes et al., 2023).

6 Conclusion and Future Work

We introduced Thinker-DDM, a novel machine translation approach that simulates the dynamic decision-making process of human translators. We designed relevant translation strategy prompts in line with the three representative theories in translation and then redefined the Drift-Diffusion process to emulate human translators' decision-making under constrained resources. We conducted extensive experiments to validate the effectiveness and efficacy of our proposed method. In the future, we will incorporate Thinker-DDM into more natural language generation tasks, such as text summarization and question answering.

Limitations

We organized the limitations of Thinker-DDM twofold: (1) Thinker-DDM adopted CometKiwi to evaluate the translation qualities; however, since it is a reference-free evaluation model, there exists some gaps between the evaluation results obtained from CometKiwi and COMET, leaving a potential line in developing a more effective reference-free metric. (2) We only tested the performance of Thinker-DDM on one LLM, GPT-3.5, in our experiments. There remain investigations of Thinker-DDM on other commonly used LLMs, e.g. LLaMA 2 (Touvron et al., 2023b) and Mistral (Jiang et al., 2023), in the future.

References

Sweta Agrawal, Chunting Zhou, Mike Lewis, Luke Zettlemoyer, and Marjan Ghazvininejad. 2023. In-context examples selection for machine translation. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 8857–8873, Toronto, Canada. Association for Computational Linguistics.

Rafal Bogacz, Eric Brown, Jeff Moehlis, Philip Holmes, and Jonathan D Cohen. 2006. The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological review*, 113(4):700.

Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared 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 M. Ziegler, Jeffrey Wu, Clemens Winter, and 12 others. 2020. Language models are few-shot learners. *Preprint*, arXiv:2005.14165.

Patrick Fernandes, Daniel Deutsch, Mara Finkelstein, Parker Riley, André Martins, Graham Neubig, Ankush Garg, Jonathan Clark, Markus Freitag, and Orhan Firat. 2023. The devil is in the errors: Leveraging large language models for fine-grained machine translation evaluation. In *Proceedings of the Eighth Conference on Machine Translation*, pages 1066–1083, Singapore. Association for Computational Linguistics.

Mikel L. Forcada, Mireia Ginestí-Rosell, Jacob Nordfalk, Jim O'Regan, Sergio Ortiz-Rojas, Juan Antonio Pérez-Ortiz, Felipe Sánchez-Martínez, Gema Ramírez-Sánchez, and Francis M. Tyers. 2011. Apertium: A free/open-source platform for rule-based machine translation. *Machine Translation*, 25(2):127–144.

Jinlan Fu, See-Kiong Ng, Zhengbao Jiang, and Pengfei Liu. 2023. Gptscore: Evaluate as you desire. *Preprint*, arXiv:2302.04166.

Marjan Ghazvininejad, Hila Gonen, and Luke Zettlemoyer. 2023. Dictionary-based phrase-level prompting of large language models for machine translation. *Preprint*, arXiv:2302.07856.

Jie He, Tao Wang, Deyi Xiong, and Qun Liu. 2020. The box is in the pen: Evaluating commonsense reasoning in neural machine translation. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 3662–3672, Online. Association for Computational Linguistics.

Zhiwei He, Tian Liang, Wenxiang Jiao, Zhuosheng Zhang, Yujiu Yang, Rui Wang, Zhaopeng Tu, Shuming Shi, and Xing Wang. 2023. Exploring human-like translation strategy with large language models. *Preprint*, arXiv:2305.04118.

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. *Preprint*, arXiv:2302.09210.

Vivek Iyer, Pinzhen Chen, and Alexandra Birch. 2023. Towards effective disambiguation for machine translation with large language models. In *Proceedings of the Eighth Conference on Machine Translation*, pages 482–495, Singapore. Association for Computational Linguistics.

Albert Q. Jiang, Alexandre Sablayrolles, Arthur Mensch, Chris Bamford, Devendra Singh Chaplot, Diego de las Casas, Florian Bressand, Gianna Lengyel, Guillaume Lample, Lucile Saulnier, Lélio Renard Lavaud, Marie-Anne Lachaux, Pierre Stock, Teven Le Scao, Thibaut Lavril, Thomas Wang, Timothée Lacroix, and William El Sayed. 2023. Mistral 7b. *Preprint*, arXiv:2310.06825.

Daniel Kahneman. 2011. *Thinking, fast and slow*. macmillan.

Tom Kocmi, Rachel Bawden, Ondřej Bojar, Anton Dvorkovich, Christian Federmann, Mark Fishel, Thamme Gowda, Yvette Graham, Roman Grundkiewicz, Barry Haddow, Rebecca Knowles, Philipp Koehn, Christof Monz, Makoto Morishita, Masaaki Nagata, Toshiaki Nakazawa, Michal Novák, Martin Popel, and Maja Popović. 2022. Findings of the 2022 conference on machine translation (WMT22). In *Proceedings of the Seventh Conference on Machine Translation (WMT)*, pages 1–45, Abu Dhabi, United Arab Emirates (Hybrid). Association for Computational Linguistics.

Philipp Koehn, Hieu Hoang, Alexandra Birch, Chris Callison-Burch, Marcello Federico, Nicola Bertoldi, Brooke Cowan, Wade Shen, Christine Moran, Richard Zens, Chris Dyer, Ondřej Bojar, Alexandra Constantin, and Evan Herbst. 2007. Moses: Open source toolkit for statistical machine translation. In *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics Companion Volume Proceedings of the Demo and Poster Sessions*, pages 177–180, Prague, Czech Republic. Association for Computational Linguistics.

Philipp Koehn, Franz J. Och, and Daniel Marcu. 2003. Statistical phrase-based translation. In *Proceedings* *of the 2003 Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics*, pages 127–133.

Jijia Liu, Chao Yu, Jiaxuan Gao, Yuqing Xie, Qingmin Liao, Yi Wu, and Yu Wang. 2023. Llm-powered hierarchical language agent for real-time human-ai coordination. *Preprint*, arXiv:2312.15224.

Hongyuan Lu, Haoyang Huang, Dongdong Zhang, Haoran Yang, Wai Lam, and Furu Wei. 2023. Chain-of-dictionary prompting elicits translation in large language models. *Preprint*, arXiv:2305.06575.

Mieradilijiang Maimaiti, Yang Liu, Huanbo Luan, and Maosong Sun. 2022. Enriching the transfer learning with pre-trained lexicon embedding for low-resource neural machine translation. *Tsinghua Science and Technology*, 27(1):150–163.

Zhuoyuan Mao and Yen Yu. 2024. Tuning llms with contrastive alignment instructions for machine translation in unseen, low-resource languages. *Preprint*, arXiv:2401.05811.

Yasmin Moslem, Rejwanul Haque, John Kelleher, and Andy Way. 2022. Domain-specific text generation for machine translation. In *Proceedings of the 15th biennial conference of the Association for Machine Translation in the Americas (Volume 1: Research Track)*, pages 14–30, Orlando, USA. Association for Machine Translation in the Americas.

Yasmin Moslem, Rejwanul Haque, John D. Kelleher, and Andy Way. 2023a. Adaptive machine translation with large language models. In *Proceedings of the 24th Annual Conference of the European Association for Machine Translation*, pages 227–237, Tampere, Finland. European Association for Machine Translation.

Yasmin Moslem, Rejwanul Haque, and Andy Way. 2023b. Fine-tuning large language models for adaptive machine translation. *Preprint*, arXiv:2312.12740.

Yasmin Moslem, Gianfranco Romani, Mahdi Molaei, John D. Kelleher, Rejwanul Haque, and Andy Way. 2023c. Domain terminology integration into machine translation: Leveraging large language models. In *Proceedings of the Eighth Conference on Machine Translation*, pages 902–911, Singapore. Association for Computational Linguistics.

Jeremy Munday, Sara Ramos Pinto, and Jacob Blakesley. 2022. *Introducing translation studies: Theories and applications*. Routledge.

Eugene Albert Nida. 1964. *Toward a science of translating: with special reference to principles and procedures involved in Bible translating*. Brill Archive.

OpenAI, :, Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, Red Avila, Igor Babuschkin, Suchir Balaji, Valerie Balcom, Paul Baltescu, Haiming Bao, Mo Bavarian, and 263 others. 2023. Gpt-4 technical report. *Preprint*, arXiv:2303.08774.

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.

Hao Peng, Xiaozhi Wang, Jianhui Chen, Weikai Li, Yunjia Qi, Zimu Wang, Zhili Wu, Kaisheng Zeng, Bin Xu, Lei Hou, and Juanzi Li. 2023. When does in-context learning fall short and why? a study on specification-heavy tasks. *Preprint*, arXiv:2311.08993.

Roger Ratcliff. 1978. A theory of memory retrieval. *Psychological Review*, 85:59–108.

Roger Ratcliff, Philip L. Smith, Scott D. Brown, and Gail McKoon. 2016. Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, 20:260–281.

Vikas Raunak, Amr Sharaf, Yiren Wang, Hany Awadalla, and Arul Menezes. 2023. Leveraging GPT-4 for automatic translation post-editing. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 12009–12024, Singapore. Association for Computational Linguistics.

Abudurexiti Reheman, Tao Zhou, Yingfeng Luo, Di Yang, Tong Xiao, and Jingbo Zhu. 2023. Prompting neural machine translation with translation memories. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37(11):13519–13527.

Ricardo Rei, José G. C. de Souza, Duarte Alves, Chrysoula Zerva, Ana C Farinha, Taisiya Glushkova, Alon Lavie, Luisa Coheur, and André F. T. Martins. 2022a. COMET-22: Unbabel-IST 2022 submission for the metrics shared task. In *Proceedings of the Seventh Conference on Machine Translation (WMT)*, pages 578–585, Abu Dhabi, United Arab Emirates (Hybrid). Association for Computational Linguistics.

Ricardo Rei, Marcos Treviso, Nuno M. Guerreiro, Chrysoula Zerva, Ana C Farinha, Christine Maroti, José G. C. de Souza, Taisiya Glushkova, Duarte Alves, Luisa Coheur, Alon Lavie, and André F. T. Martins. 2022b. CometKiwi: IST-unbabel 2022 submission for the quality estimation shared task. In *Proceedings of the Seventh Conference on Machine Translation (WMT)*, pages 634–645, Abu Dhabi, United Arab Emirates (Hybrid). Association for Computational Linguistics.

Raphael Reinauer, Patrick Simianer, Kaden Uhlig, Johannes E. M. Mosig, and Joern Wuebker. 2023. Neural machine translation models can learn to be few-shot learners. *Preprint*, arXiv:2309.08590.

Katharina Reiss. 1989. Text types, translation types and translation assessment. *Readings in translation theory*, 19771989.

Katharina Reiß and Hans J Vermeer. 2010. *Grundlegung einer allgemeinen Translationstheorie*, volume 147. Walter de Gruyter.

E. Rowland. 1946. Theory of games and economic behavior. *Nature*, 157:172–173.

Morteza Abdi Saroukhil, Omid Ghalkhani, and Ali Hashemi. 2018. A critical review of translation: A look forward. *International Journal of Education and Literacy Studies*, 6(2):101–110.

Gabriele Sarti, Phu Mon Htut, Xing Niu, Benjamin Hsu, Anna Currey, Georgiana Dinu, and Maria Nadejde. 2023. RAMP: Retrieval and attribute-marking enhanced prompting for attribute-controlled translation. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 1476–1490, Toronto, Canada. Association for Computational Linguistics.

Thibault Sellam, Dipanjan Das, and Ankur Parikh. 2020. BLEURT: Learning robust metrics for text generation. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7881–7892, Online. Association for Computational Linguistics.

Xiaoyu Tian, Liangyu Chen, Na Liu, Yaxuan Liu, Wei Zou, Kaijiang Chen, and Ming Cui. 2023. Duma: a dual-mind conversational agent with fast and slow thinking. *Preprint*, arXiv:2310.18075.

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

Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shruti Bhosale, Dan Bikel, Lukas Blecher, Cristian Canton Ferrer, Moya Chen, Guillem Cucurull, David Esiobu, Jude Fernandes, Jeremy Fu, Wenyin Fu, and 49 others. 2023b. Llama 2: Open foundation and fine-tuned chat models. *Preprint*, arXiv:2307.09288.

David Vilar, Markus Freitag, Colin Cherry, Jiaming Luo, Viresh Ratnakar, and George Foster. 2023. Prompting PaLM for translation: Assessing strategies and performance. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 15406–15427, Toronto, Canada. Association for Computational Linguistics.

Peifang Wang, Olga Golovneva, Armen Aghajanyan, Xiang Ren, Muhao Chen, Asli Celikyilmaz, and Maryam Fazel-Zarandi. 2023a. Domino: A dual-system for multi-step visual language reasoning. *Preprint*, arXiv:2310.02804.

Yiming Wang, Zhuosheng Zhang, and Rui Wang. 2023b. Element-aware summarization with large language models: Expert-aligned evaluation and chain-of-thought method. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 8640–8665, Toronto, Canada. Association for Computational Linguistics.

Zhilin Wang, Yu Ying Chiu, and Yu Cheung Chiu. 2023c. Humanoid agents: Platform for simulating human-like generative agents. *Preprint*, arXiv:2310.05418.

Zimu Wang, Wei Wang, Qi Chen, Qiufeng Wang, and Anh Nguyen. 2023d. Generating valid and natural adversarial examples with large language models. *Preprint*, arXiv:2311.11861.

Jason Weston and Sainbayar Sukhbaatar. 2023. System 2 attention (is something you might need too). *Preprint*, arXiv:2311.11829.

Wen Yang, Chong Li, Jiajun Zhang, and Chengqing Zong. 2023. Bigtranslate: Augmenting large language models with multilingual translation capability over 100 languages. *Preprint*, arXiv:2305.18098.

Shunyu Yao, Dian Yu, Jeffrey Zhao, Izhak Shafran, Thomas L. Griffiths, Yuan Cao, and Karthik Narasimhan. 2023. Tree of Thoughts: Deliberate problem solving with large language models. *Preprint*, arXiv:2305.10601.

Aohan Zeng, Xiao Liu, Zhengxiao Du, Zihan Wang, Hanyu Lai, Ming Ding, Zhuoyi Yang, Yifan Xu, Wendi Zheng, Xiao Xia, Weng Lam Tam, Zixuan Ma, Yufei Xue, Jidong Zhai, Wenguang Chen, Zhiyuan Liu, Peng Zhang, Yuxiao Dong, and Jie Tang. 2023. GLM-130b: An open bilingual pre-trained model. In *The Eleventh International Conference on Learning Representations*.

Shaolei Zhang, Qingkai Fang, Zhuocheng Zhang, Zhengrui Ma, Yan Zhou, Langlin Huang, Mengyu Bu, Shangtong Gui, Yunji Chen, Xilin Chen, and Yang Feng. 2023. Bayling: Bridging cross-lingual alignment and instruction following through interactive translation for large language models. *Preprint*, arXiv:2306.10968.

Susan Zhang, Stephen Roller, Naman Goyal, Mikel Artetxe, Moya Chen, Shuohui Chen, Christopher Dewan, Mona Diab, Xian Li, Xi Victoria Lin, Todor Mihaylov, Myle Ott, Sam Shleifer, Kurt Shuster, Daniel Simig, Punit Singh Koura, Anjali Sridhar, Tianlu Wang, and Luke Zettlemoyer. 2022. Opt: Open pre-trained transformer language models. *Preprint*, arXiv:2205.01068.

Yuanhang Zheng, Zhixing Tan, Meng Zhang, Mieradilijiang Maimaiti, Huanbo Luan, Maosong Sun, Qun Liu, and Yang Liu. 2021. Self-supervised quality estimation for machine translation. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 3322–3334, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.

Wenhao Zhu, Hongyi Liu, Qingxiu Dong, Jingjing Xu, Shujian Huang, Lingpeng Kong, Jiajun Chen, and Lei Li. 2023a. Multilingual machine translation with large language models: Empirical results and analysis. *Preprint*, arXiv:2304.04675.

Wenhao Zhu, Yunzhe Lv, Qingxiu Dong, Fei Yuan, Jingjing Xu, Shujian Huang, Lingpeng Kong, Jiajun Chen, and Lei Li. 2023b. Extrapolating large language models to non-english by aligning languages. *Preprint*, arXiv:2308.04948.

A Translation Strategy Prompts

INSTRUCTION: Perform a Scene Analysis on the the following Source Sentence: Source Sentence: `{source_sentence}` Analyze the context on the following source sentence and consider how this context impacts the translation strategy. Analysis:
EXAMPLE: I will just arrange a returns label for you now.
ANSWER: The source sentence is spoken by a customer service representative to a customer who needs to return an item. The context of this sentence is a professional and polite interaction between the representative and the customer. The representative is offering to help the customer by arranging a returns label, which indicates that the company has a specific process in place for returns. **The use of the word "just" suggests that this is a quick and easy task for the representative to complete.** Overall, the context of this sentence requires a translation strategy that is clear, concise, and professional. The translation should accurately convey the representative's willingness to assist the customer and the company's established process for returns. It may also be important to maintain the polite tone and use of the word "just" to convey the ease of the task. Additionally, the translation should be culturally appropriate and considerate of any potential language barriers between the representative and the customer.
MS-TRANSLATOR TRANSLATION: 我现在只为您安排一个退货标签。
ANALYSIS-BASED TRANSLATION: 我现在会为您安排一个退货标签。

Table 4: Instruction and example for conducting scene analysis on a given source sentence.

Instruction: Perform a Target Audience Analysis on the the following Source Sentence: Source Sentence: `{source_sentence}` Describe and analyze the characteristics of the target audience, discussing how these characteristics influence the translation. Analysis:
Example: For quality purposes this chat interaction will be closed, please remember you can always come back and we will be happy to continue assisting you.
Answer: The target audience for this source sentence appears to be customers or clients who are engaging in a chat interaction with a customer service representative. These individuals are likely seeking assistance or support with a product or service. **The first characteristic of the target audience is that they are seeking quality service.** This suggests that they are likely to be detail-oriented and have high expectations for the assistance they receive. As a result, the translation should accurately convey the message of the source sentence and maintain a professional tone. ... Overall, the target audience for this source sentence appears to be professional, detail-oriented, tech-savvy, time-conscious, and potentially long-term customers. These characteristics should be considered when translating the sentence to ensure that the message is accurately conveyed and the tone is appropriate for the audience.
MS-Translator Translation: 出于质量目的，此聊天互动将关闭，请记住您可以随时回来，我们将很乐意继续为您提供帮助。
Analysis-based Translation: 为了保证服务质量，此次聊天交互将会关闭，请记住您随时可以回来，我们将很乐意继续为您提供帮助。

Table 5: Instruction and example for conducting target audience analysis on a given source sentence.

INSTRUCTION: Perform a Key Information Analysis on the the following Source Sentence: Source Sentence: `{source_sentence}` Identify and analyze the key information in the source sentence, discussing how to maintain this information's integrity in the translation. Analysis:
EXAMPLE: Joint Aid for Dogs can be fed to all dogs from any age at the 'General Support' level to maintain freedom of movement and muscle condition throughout their lives.
ANSWER: **In this source sentence, the key information is related to the product "Joint Aid for Dogs" and its usage.** The product is specifically designed for dogs and can be given to all dogs, regardless of their age. The recommended level for usage is "General Support," which suggests that the product is suitable for regular maintenance rather than treating a specific condition. The purpose of the product is to maintain the dog's freedom of movement and muscle condition throughout their lives. ... In conclusion, to maintain the integrity of the key information in the translation, it is essential to accurately convey the product name, target audience, recommended level, and purpose. This can be achieved by using appropriate terminology and language, and ensuring that the translation conveys the same level of effectiveness and suitability as the source sentence.
MS-TRANSLATOR TRANSLATION: 狗关节援助可以喂给任何年龄的"一般支持"级别的所有狗，以保持其一生的活动自由和肌肉状况。
ANALYSIS-BASED TRANSLATION: Joint Aid for Dogs可以在"一般支持"级别下喂养所有年龄段的狗，以保持它们一生中的运动自由和肌肉状况。

Table 6: Instruction and example for conducting key information analysis on a given source sentence.

INSTRUCTION: Perform an Intent Analysis on the the following Source Sentence: Source Sentence: `{source_sentence}` Analyze the original intent of the source sentence and its impact on the translation strategy. Analysis:
EXAMPLE: Your online life and real life are merging into one seamless experience, and you need security that can keep viruses, identity theft, and other digital threats from becoming real-life headaches.
ANSWER: The original intent of the source sentence is to emphasize the increasing integration of online and real-life experiences and the need for security measures to protect against digital threats. **This sentence is likely promoting a product or service that offers such security.** The impact on the translation strategy would be to accurately convey the message of the sentence and effectively promote the product or service. This would involve using clear and concise language to highlight the merging of online and real-life experiences, as well as the importance of security in preventing potential threats. Additionally, the translation should effectively convey the sense of urgency and potential consequences of not having proper security measures in place. It may also be beneficial to use persuasive language to encourage the audience to take action and invest in the promoted product or service.
MS-TRANSLATOR TRANSLATION: 您的在线生活和现实生活正在融合为一种无缝体验，您需要能够防止病毒、身份盗用和其他数字威胁成为现实生活中令人头疼的问题的安全性。
ANALYSIS-BASED TRANSLATION: 您的在线生活和现实生活正在融合为一个无缝的体验，您需要安全保障来防止病毒、身份盗窃和其他数码威胁成为现实生活的头痛。

Table 7: Instruction and example for conducting intent analysis on a given source sentence.

INSTRUCTION: Perform a Information Equivalence Analysis on the the following Source Sentence: Source Sentence: `{source_sentence}` Analyze how to convey the same information and emotions of the source sentence in `{target}`. Analysis:
EXAMPLE: Instanced Player Housing - From tents to palaces and everything in between, players can furnish and customize their own homes and can hire NPCs to keep your place clean or purchase things from the marketplace.
ANSWER: The Chinese equivalent conveys the same information and emotions as the source sentence. “Instanced Player Housing” is translated as “实例化玩家住房”, which captures the concept of players having their own individual and unique homes. **The phrase “From tents to palaces and everything in between” is translated as “从帐篷到宫殿，应有尽有”, which maintains the idea of a wide range of options for players to choose from.** The phrase "furnish and customize" is translated as “装饰和定制”, which conveys the idea of personalizing one’s home. The phrase “hire NPCs to keep your place clean” is translated as “雇佣NPC来保持清洁”, which captures the practical aspect of hiring NPCs to maintain the player’s home. Lastly, the phrase “purchase things from the marketplace” is translated as “从市场购买物品”, which conveys the idea of being able to buy items for the player’s home from a marketplace. Overall, the Chinese equivalent effectively conveys the same information and emotions as the source sentence.
MS-TRANSLATOR TRANSLATION: 实例化玩家住房 - 从帐篷到宫殿以及介于两者之间的一切，玩家可以布置和定制自己的房屋，并可以雇用NPC来保持您的地方清洁或从市场上购买东西。
ANALYSIS-BASED TRANSLATION: 实例化玩家住房 - 从帐篷到宫殿，玩家可以装饰和定制自己的家，还可雇佣NPC保持清洁或从市场购买物品。

Table 8: Instruction and example for conducting intent analysis on a given source sentence, in which “`{target}`” refers to the target translation language.

INSTRUCTION: Perform a Cultural Equivalence Analysis on the the following Source Sentence: Source Sentence: `{source_sentence}` Discuss the differences between the source and `{target}` cultures, and analyze how these differences impact the translation. Analysis:
EXAMPLE: Alfie Barbeary's all-action display capped by a captivating try had given Wasps hope but he was unexpectedly replaced at half-time and it became one-way traffic from the moment Campbell showed his class to finish in the 43rd minute.
ANSWER: The source sentence is from a sports article about a rugby match, specifically discussing the performance of a player named Alfie Barbeary. The sentence mentions his "all-action display" and "captivating try" that gave his team, Wasps, hope. However, he was unexpectedly replaced at half-time, and the match turned in favor of the opposing team, with the moment of a player named Campbell showing his "class" and scoring a try in the 43rd minute. **Lastly, the use of the phrase "one-way traffic" may not have a direct equivalent in ZH culture. This phrase is used to describe a situation in which one team is dominating the match and controlling the flow of play.** A literal translation may not convey the same meaning and impact as it does in the source culture. In conclusion, the cultural differences between the source and ZH cultures present several challenges for the translation of this sentence. The translator must carefully consider the cultural context and find ways to convey the same meaning and impact in the target language. This may require adapting certain phrases or finding cultural equivalents that accurately convey the intended message.
MS-TRANSLATOR TRANSLATION: 阿尔菲·巴比里（Alfie Barbeary）的全能表现加上一次迷人的尝试，这给了黄蜂希望，但他在中场休息时出人意料地被换下，从坎贝尔在第43分钟展示他的班级的那一刻起，它就变成了单向交通。
ANALYSIS-BASED TRANSLATION: 阿尔菲·巴伯里（Alfie Barbeary）的全力表现和引人注目的进球给了瓦斯普斯（Wasps）希望，但他在半场结束时意外被换下，从坎贝尔（Campbell）在第43分钟展现出他的实力后，比赛局势就一边倒了。

Table 9: Instruction and example for conducting cultural equivalence analysis on a given source sentence, in which "`{target}`" refers to the target translation language.

INSTRUCTION: Perform a Text Type Analysis on the following Source Sentence: Source Sentence: `{source_sentence}` Determine the text type of the source sentence (informative, expressive, operative) and analyze how this classification guides the translation strategy. Analysis:
EXAMPLE: Please keep tapping where the images are showing, are you able to see the images and follow where to tap?
ANSWER: The text type of the source sentence is operative. This type of text is used to give instructions or commands to the reader, with the goal of eliciting a specific action or response. In this case, the sentence is giving instructions on how to interact with the images being displayed. ... **Secondly, the use of imperative verbs, such as "keep tapping" and "follow", is common in operative texts and should be maintained in the translation.** These verbs convey a sense of urgency and give a clear indication of what the reader should do. ... Overall, the operative text type guides the translation strategy to focus on clear and direct language, maintain the tone and formality of the source sentence, and consider any cultural or contextual differences that may affect the understanding of the instructions.
MS-TRANSLATOR TRANSLATION: 请继续点击图像显示的位置，您是否能够看到图像并点击位置?
ANALYSIS-BASED TRANSLATION: 请在图像显示的位置持续点击，你能看到图像并按照指示进行点击吗?

Table 10: Instruction and example for conducting text type analysis on a given source sentence.

INSTRUCTION: Source Sentence: `{source_sentence}` Analysis: `{analysis}` Please give the best translation from `{source}` to `{target}` based on the above analysis without explanation. Best Translation:

Table 11: Instruction for analysis-based translation, in which "`{source}`", "`{target}`", and "`{analysis}`" refer to the source language, target translation language, and analysis, respectively.

INSTRUCTION: Source Sentence: `{source_sentence}` Please give the best translation from `{source}` to `{target}` without explanation. Best Translation:

Table 12: Instruction for zero-shot translation, in which "`{source}`" and "`{target}`" refer to the source and target translation languages, respectively.