https://doi.org/10.31449/inf.v49i25.10014

Structured Linguistic Augmentation for Large Language Models in Complex Machine Translation

Hongjuan Li
Pingdingshan Polytechnic College, Pingdingshan Henan, 467000, China
E-mail: li.hongjuan@outlook.com

Keywords: Machine translation, large language models, complex language structures, linguistic structure augmentation, fluency, accuracy, knowledge enhancement, latent relations

Received: July 5, 2025

Traditional machine translation (MT) systems and even Large Language Models (LLMs) often face significant challenges with complex language structures, stemming from a limited understanding of intricate syntactic dependencies and subtle semantic nuances. While LLMs possess powerful generative capabilities, their over-reliance on surface-level patterns can lead to suboptimal translations for complex sentences. To address this, we propose a novel framework, Structured Linguistic Augmentation (SLA), designed to enrich LLMs with a deep and explicit understanding of linguistic structures. The SLA framework integrates three synergistic components: (1) Contextual Linguistic Dependency Graph Construction and Pre-training (CLDG-PT), which injects fine-grained syntactic knowledge into the LLM; (2) Common Sense-Driven Relational Augmentation (CSRA), which externalizes the LLM's implicit knowledge to identify high-level semantic relations; and (3) Latent Structural Relation Discovery (LSRD), which uncovers subtle, implicit connections between linguistic components via a self-supervised objective. We conduct comprehensive experiments on the general-domain WMT En-De dataset and a new, challenging CS-Trans dataset curated with complex sentences. Evaluations show that SLA significantly improves both the fluency and accuracy of complex sentence translation. Notably, on the CS-Trans test set, our model achieves a COMET score of 75.0, substantially outperforming a strong fine-tuned LLM baseline (70.0) and demonstrating superior linguistic comprehension. These results, along with strong performance on logical reasoning translation tasks, validate SLA's effectiveness in enhancing LLMs for high-fidelity complex machine translation.

Povzetek:

1 Introduction

The rapid proliferation of digital content and global communication platforms has fundamentally reshaped expert interaction, underscoring the critical role of Machine Translation (MT) in bridging linguistic divides [1]. In this era of unprecedented linguistic diversity, MT systems have emerged as indispensable tools, diligently processing vast amounts of text to facilitate cross-lingual communication and information access [2]. These systems have become a cornerstone of modern global decision support, playing a pivotal role in various industries, from international commerce and media streaming to scientific collaboration and diplomatic relations [3, 4].

Despite their undeniable advancements and pervasive integration into daily life, MT systems face formidable challenges, particularly when confronting the inherent complexities of natural language [5]. Early rule-based MT (RBMT) systems, as detailed by Brown et al. [6], demonstrated significant limitations in handling linguistic variability and ambiguity. Subsequently, statistical MT (SMT) approaches, analyzed by Lee et al. [7], often exhibited deficiencies in capturing deep syntactic structures and long-range dependencies, leading to translations that, as observed by Chen et al. [8], frequently appeared unnatural or inaccurate for nuanced sentences. Pervasive issues include difficulties with resolving syntactic ambiguity, a persistent problem highlighted by Wang et al. [9], and effectively handling long-range dependencies, which Johnson et al. [10] extensively investigated. Furthermore, the faithful rendition of idiomatic expressions and subtle pragmatic nuances remains a significant hurdle, as highlighted by Taylor et al. [11].

Understanding and adapting to the dynamic and intricate nature of complex language structures is paramount in achieving high-quality translation [12, 13]. Expert linguistic comprehension is not static; it continuously evolves in response to new communication contexts and nuanced expressions [14]. Researchers like Jones et al. [15] emphasized that conventional MT methods often overlook dynamic linguistic patterns due to evolving semantic relationships, causing issues like mistranslations in contextually sensitive scenarios. Similarly, Miller et al. [16] demonstrated that advanced deep learning models are needed to effectively classify evolving linguistic variants in online text, outperforming conventional approaches. Consequently, there is a pressing need for advanced methodologies that can not only translate complex language but also dynamically disentangle genuine linguistic drivers from confounding surface-level variations, ensuring that translations truly foster long-term communicative satisfaction.

The advent of Large Language Models (LLMs) presents a compelling paradigm for addressing these intricate challenges. The introduction of the Transformer architecture by Vaswani et al. [17, 18, 19] became foundational for LLMs, enabling powerful representation learning. Subsequently, Brown et al. [20] showcased the emergent few-shot learning abilities of massive LLMs, while Lee et al. [21] demonstrated how LLMs could effectively optimize sequential text generation strategies, significantly improving translation quality. Further advancements by Chen et al. [22] have shown LLMs' remarkable ability to manage high-dimensional lexical and syntactic spaces, adapting to evolving linguistic environments and learning optimal text generation policies through continuous interaction [23]. The broad applicability and accessibility of LLMs were further expanded by Wolf et al. [24], who presented a comprehensive library, and by Touvron et al. [25], who introduced open-source LLMs. These developments make LLMs particularly suitable for the complex, dynamic, and partially observable nature of natural language translation.

However, directly applying LLMs to complex MT tasks often presents unique hurdles. Liu et al. [26] noted that despite their capabilities, LLMs' direct application to MT can yield suboptimal performance for structurally complex inputs. Geng et al. [27] demonstrated that while LLMs generate fluent text, they may still misinterpret deep semantic relationships. Zhao et al. [28] observed that existing LLMs, despite their large scale, often lack explicit mechanisms for robust structural understanding in translation. The challenge of maintaining semantic consistency across lexically diverse contexts, as revealed by Hou et al. [29] and Li et al. [30], further exposes deficiencies in LLMs' ability to consistently map synonymous concepts. Yang et al. [31] discussed the over-reliance of LLMs on surface-level patterns, leading to errors in structurally complex MT. Fundamentally, the implicit knowledge within LLMs may be insufficient for all nuanced linguistic complexities required for high-quality MT, as highlighted by Whang et al. [32].

To overcome these limitations and unlock the full potential of LLMs for debiased, structurally aware translation, this paper proposes a novel framework, Structured Linguistic Augmentation (SLA). Drawing inspiration from knowledge-enhanced methods in related NLP tasks, SLA is designed to explicitly model the underlying mechanisms of complex language generation and interpretation. Our method integrates linguistic graph models to explicitly represent dependency relationships, leverages common sense to identify semantic connections, and employs latent discovery methods to uncover subtle, implicit structural links. By optimizing a multi-faceted objective within this disentangled framework, SLA aims to generate translations aligned with true linguistic intent, ensuring long-term communicative fidelity. The framework also incorporates sophisticated techniques from structured learning, to ensure stable and reliable learning from diverse linguistic datasets without requiring costly or risky manual feature engineering. The research objectives are:

1.To design and implement a novel framework, Structured Linguistic Augmentation (SLA), that explicitly enhances LLMs with multi-faceted linguistic knowledge—spanning syntactic, semantic, and latent structures—to improve the translation of complex sentences.

2.To empirically validate that this explicit structural augmentation leads to superior translation quality compared to standard LLM-based approaches that rely on implicit knowledge.

3.To analyze and quantify the individual contributions of different types of structural knowledge to understand which facets are most critical for resolving linguistic complexity in translation.

2 Machine translation

2.1 Evolution of machine translation paradigms

Machine Translation (MT), a core sub-field of natural language processing (NLP), seeks to translate speech or text from a source language to a target language automatically, and with complete faithfulness to its semantic content, stylistic differences, and original intention. MT has been driven by various dominant paradigms. There were earlier iterations, which included Rule-Based Machine Translation (RBMT) systems carefully constructed on massive linguistic rules, extensive dictionaries, and intricate grammatical forms painstakingly created by expert experts. While RBMT delivered high transparency and good control over translation quality in pre-specified domains, its inherent dependence on handcrafted rule generation rendered its building troublesome and generally rendered it powerless when faced with the widespread ambiguity and heterogeneity of natural language textual information [33].

The availability of large parallel corpora facilitated the introduction of Statistical Machine Translation (SMT). SMT models were coded to predict translation probabilities by analyzing vast numbers of aligned text segments. Conventional SMT methods, such as phrase-based SMT, involved segmenting source sentences into small-sized unit phrases and then using sophisticated statistical models to arrive at the optimal sequence of translations. This approach greatly enhanced the fluency and robustness of machine translations compared to its rule-based predecessors. SMT did have, however, usually exhibit shortcomings in handling long-distance dependencies in sentences appropriately and producing grammatically correct outputs consistently, largely due to its localized view of phrase-level equivalence rather than deeper linguistic structural knowledge.

2.2 Inherent challenges in translating complex language structures

While NMT and related technologies have progressed by leaps and bounds, the translation of complex sentence structures continues to pose a daunting challenge for machine translation algorithms. The reasons are rooted in the multifaceted meaning of the complex sentences, non-literal sentence constructions, as well as deep grammatical interdependencies.

A native difficulty in syntactic and semantic ambiguity exists. A single complex sentence can provide more than one syntactic reading, enabling the possible misperceptions by the translation model in regard to part-of-speech roles and word-to-word relationships, thereby generating incorrect translations. Also, comprehension of deeper semantic connections, like anaphora resolution or selection of the appropriate polysemous word senses, demands the model to move beyond superficial text examination and conduct a deeper contextual comprehension [34]. Another major hindrance is the issue of long-distance dependency. Information essential for proper translation of a particular word or phrase in complex sentence structures can be placed far from its local context. This highly tests the model's capacity for maintaining contextual coherence, a challenge especially severe in free word order languages or those with high anaphoric reference.

Finally, there usually is profound structural divergence between source and target languages. When the two languages' syntactic structures differ significantly (e.g., English Subject-Verb-Object vs. Japanese Subject-Object-Verb), a word-for-word translation is not sufficient for high-quality translation [37]. In these situations, MT models must perform advanced rewording and structure conversion in order to generate natural and grammatically correct translations. All these intricacies combined have a tendency to produce translated output that, while perhaps grammatically correct, may either lack the refined nuances, coherence, or naturalness demanded of expert-level translation, particularly for extremely intricate sentences.

2.3 The role of large language models in machine translation

The advent of Large Language Models (LLMs) that have their huge parameter sizes (typically in the scale of billions) and their gigantic pre-training over gigantic amounts of text data has revolutionized the machine translation field in a fundamental way. LLMs exhibit a whole range of remarkable abilities that position them very well for translation work. Their exemplary generative capability enables them to generate readable and contextually relevant text, naturally best for sequence-to-sequence operations such as translation. Through their lengthier pre-training, LLMs acquire indirectly an immense body of world knowledge and common sense, beneficial especially for disambiguation and in producing semantically more enhanced translations. Besides, LLMs demonstrate an exceptional ability to understand deep semantic relation between words and phrases, a vital parameter for sustaining the fidelity of translated meaning.

LLMs are becoming increasingly applied in MT pipelines, either by fine-tuning on parallel corpora or more recently and generally by way of a variety of prompt engineering techniques. The techniques attempt to obtain high-quality translations from pre-trained LLMs without extensive re-training. Prompting techniques differ. Zero-shot prompting instructs the LLM to translate directly with no in-context examples given, only relying upon the model's built-in, generalized translation capability acquired during pre-training. This approach serves as a baseline to gauge an LLM's inductive knowledge for a particular language pair. Less often, few-shot prompting (also known as in-context learning) involves providing the LLM with a limited set of source-target language translation examples as part of the input prompt, before the target sentence to translate. Even in the case where these examples are not necessarily semantically directly related to the target sentence, they serve as contextual pointers, pointing the model towards translating tendencies and stylistic variations, which generally leads to higher performance [35]. A more advanced approach, fragment-shot prompting, is a targeted few-shot approach in which the sentence to translate is segmented into successive word strings, or ”fragments.” Next, authentic translation instances with the identical source-side fragment are retrieved from a training dataset. The fragment-specific examples are appended to the prompt. This approach targets localized contextual salience, thereby guiding the LLM toward probable translations of specific linguistic components. On this basis, pivoted fragment-shot prompting also enables translation between language pairs that have no or only limited parallel data readily available directly by employing an intermediary, or ”pivot,” language-a third language with parallel data available for both source and target languages-to perform indirect translation.

To contextualize the contributions of our proposed framework, Table 1 provides a comparative summary of prominent LLM-based translation approaches and their performance on complex linguistic structures.

3 Structured linguistic augmentation (SLA) Methodology for translation

While Large Language Models (LLMs) have achieved remarkable progress in Machine Translation (MT) due to their robust text generation capabilities and extensive world knowledge, challenges persist concerning the fluency and accuracy of translations, particularly when dealing with complex linguistic structures. These difficulties arise from the implicit nature of deep syntactic dependencies, long-range semantic associations, and subtle pragmatic nuances

Table 1: Comparison of LLM-based MT methods on complex structures (CS-Trans Dataset)

Method	Approach	BLEU↑	COMET↑	Key Limitations
Zero-Shot LLM	Direct prompting without examples.	10	50	Lacks task-specific adaptation; highly reliant on pre-trained knowledge, which is often too general for complex syntax.
Few-Shot LLM	In-context learning with random examples.	13	55	Performance is sensitive to example selection; does not guarantee understanding of specific structural challenges.
Fragment-Shot LLM	In-context learning with retrieved, relevant examples.	15	60	Improves on localized context but still fails to capture global, long-range dependencies or implicit logical relations.
Fine-tuned LLM (FT-LLM)	Fine-tuning on a large parallel corpus.	22	70	Strong baseline, but learns implicitly from data, struggling with systematic generalization for rare or highly ambiguous structures.

inherent in complex sentences, which LLMs struggle to fully capture through purely statistical pattern learning. To overcome these limitations, we propose the Structured Linguistic Augmentation (SLA) framework. SLA aims to enhance LLMs' translation capabilities by explicitly instilling a profound understanding of linguistic dependencies and semantic relationships within complex language structures. Inspired by recent advancements in knowledge-enhanced recommendation systems, SLA integrates three synergistic modules: Contextual Linguistic Dependency Graph Construction and Pre-training (CLDG-PT), Common Sense-Driven Relational Augmentation (CSRA), and Latent Structural Relation Discovery (LSRD). These components collectively guide the LLM to learn, externalize, and leverage structured linguistic knowledge, thereby significantly improving the fluency and accuracy of its translations of complex language.

3.1 SLA overall framework

The overall architecture of the Structured Linguistic Augmentation (SLA) framework conceptually illustrates how it enriches the underlying Large Language Model with both explicit and implicit linguistic structural knowledge through a multi-stage, jointly optimized learning process. As shown in Figure 1. The framework utilizes the LLM as its core translation engine, with each proposed component addressing a specific aspect of linguistic complexity: **Contextual Linguistic Dependency Graph Construction and Pre-training (CLDG-PT)** is designed to capture explicit, fine-grained structural dependencies at the sub-sentential level and transfer this generalized knowledge to the LLM, ensuring grammatical correctness and naturalness in the target language. **Common Sense-Driven Relational Augmentation (CSRA)** focuses on leveraging the LLM's inherent common sense to explicitly identify and integrate high-level semantic and syntactic relationships necessary for comprehending complex clauses and phrases, which is crucial for precise meaning preservation. Finally, **Latent Structural Relation Discovery (LSRD)** is responsible for uncovering implicit, subtle structural connections within and between complex sentence components that are difficult to capture through predefined rules or direct observation. This facilitates the model's ability to handle ambiguity and generate translations with greater contextual adaptability. These components are designed to operate synergistically, with their respective learning objectives contributing to a unified optimization framework that enhances the LLM's capacity for complex language translation, leading to improved fluency and accuracy.

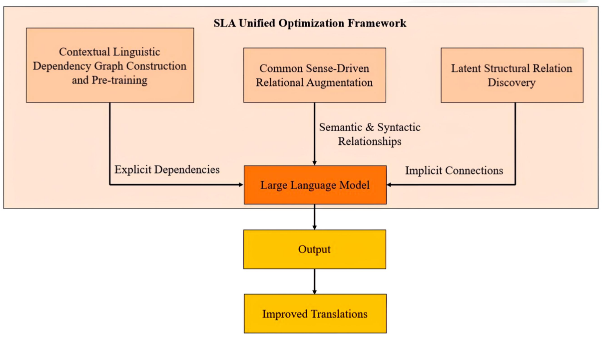

Figure 1: SLA unified optimization framework

3.2 Contextual linguistic dependency graph construction and pre-training (CLDG-PT)

To address this deficiency, and inspired by the success of graph representation learning in modeling structured data, we propose injecting this structured linguistic knowledge into the LLM's representation space. This component constructs a graph that captures co-dependent linguistic units (e.g., subject-verb relations, modifier-head relations) and utilizes this graph to guide the LLM's pre-training. This enables the LLM to acquire transferable, fine-grained structural understanding, thereby improving its ability to translate complex sentences while ensuring grammatical correctness and natural fluency [36].

Figure 2 illustrates the core concept from recommendation systems that inspired our CLDG-PT module. In this example, a user's recent interactions with items containing words like "card" and "santa" are used to establish "co-clicked word" relationships with other words like "retro" and "red" (from a target item). This "word-level collaborative filtering" is shown to help predict the next item. Analogously, in our context, we leverage the idea of building a graph of co-dependent linguistic units (words/phrases connected by syntactic or semantic relations) to infuse the LLM with structured linguistic knowledge, enabling it to better translate complex sentences by understanding their inherent structural fabric.

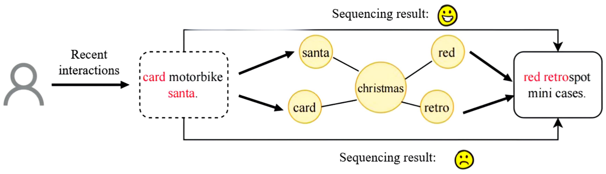

Figure 2: CLDG-PT module core concept

We construct a Linguistic Dependency Graph ($\mathcal{G}_{LDG} = (\mathcal{V}_L, \mathcal{E}_L)$) from a large, diverse linguistic corpus. The nodes $\mathcal{V}_L$ in this graph represent individual words or multi-word expressions (MWEs), functioning as linguistic units participating in syntactic structures. The edges $\mathcal{E}_L$ denote observed dependency relations between these units, conceptually analogous to co-occurrence or co-click patterns in recommendation systems. Distinct from simple co-occurrence graphs, the edges in the LDG are derived from the output of syntactic parsers (e.g., based on the Universal Dependencies framework) or represent strong semantic co-dependencies.

For a given linguistic unit $w_i \in \mathcal{V}_L$ in the graph, we first identify its direct linguistic dependents w_j within parsed sentences. To quantify the strength of the dependency between w_i and w_j, we adapt the TF-IDF concept, here reinterpreted as Term-Dependency Frequency - Inverse Dependency Frequency. Let $c_{i,j}$ be the frequency of observing a specific dependency relation between w_i and w_j (e.g., w_i is the head of w_j, or vice versa) across the entire corpus. The Term-Dependency Frequency (tdf) from w_i to w_j is defined as:

$$tdf_{i,j} = \frac{c_{i,j}}{\sum_{k=1}^{|\mathcal{V}_L|} c_{i,k}} \quad (1)$$

where the denominator represents the total count of all dependencies involving w_i, normalizing $c_{i,j}$. The Inverse Dependency Frequency (idf) for a unit w_j reflects its specificity in dependencies, i.e., how commonly w_j acts as a dependent. If w_j frequently appears as a dependent across many different heads, its idf value will be lower, and vice versa:

$$idf_j = \log\left(\frac{|\mathcal{V}_L|}{|\{w_k \mid \exists w_p, c_{p,j} > 0\}| + \epsilon}\right) \quad (2)$$

Here, ϵ is a small constant added to prevent division by zero. The numerator $|\mathcal{V}_L|$ is the total size of the vocabulary, while the denominator counts how many head words w_p have a dependency relation with w_j. In the GPT task, we process a randomly masked token t_m within the input sentence. We treat the LLM-generated contextual semantic embedding e_m and its corresponding LDG-derived structural embedding g_m as a positive sample pair. Structural embeddings g_j from other tokens in the same sentence (where $\neq m$) are treated as negative samples. Our objective is to maximize the agreement between positive pairs while minimizing it for negative pairs, by minimizing an InfoNCE-like loss function.

The GPT Alignment Loss ($\mathcal{L}_{\text{align}}^{GPT}$) is defined as:

$$\mathcal{L}_{\text{align}}^{GPT} = \frac{\sum_{m=1}^{M} \log\left(\frac{\exp(sim(e_m, g_m)/\tau_{GPT})}{\sum_{j=1}^{N} \exp(sim(e_m, g_j)/\tau_{GPT})}\right)}{-M} \quad (3)$$

where M is the number of masked tokens, $sim(\cdot,\cdot)$ is a similarity function (e.g., cosine similarity or dot product), and τ_{GPT} is a temperature parameter used to adjust the sharpness of the similarity distribution. During this pre-training, the parameters of both the LLM and the GNN are jointly optimized. The LLM's word embedding layer can often be initialized with the GNN's learned node embeddings, facilitating the fusion of linguistic structural knowledge. This mechanism ensures that the LLM's internal representations not only capture lexical semantics but also the grammatical connections between words. This is crucial for correctly handling word order, part-of-speech variations, and syntactic transformations when translating complex sentences, leading to more grammatically accurate and fluent target translations.

3.3 Common sense-driven relational augmentation (CSRA)

Complex sentences frequently convey implicit semantic relationships beyond their literal syntactic structures, such as causality, temporal sequence, elaboration, or contrast. These relationships are critical for accurate and coherent translation but are not always explicitly marked by syntactic cues. For instance, in a sentence like "Because it rained, the game was canceled," a clear causal relationship exists between "it rained" and "the game was canceled." Accurately capturing this relationship is indispensable for ensuring the target translation faithfully reflects the logical coherence of the original. Drawing on the LLM's success in inferring complementary or substitute item relationships in recommendation systems, we aim to leverage its inherent vast common-sense knowledge base to explicitly identify and augment high-level linguistic relations present in complex source sentence structures.

We construct a Relational Knowledge Graph ($\mathcal{G}_{RKG} = (\mathcal{V}_R, \mathcal{E}_R)$) where the nodes $\mathcal{V}_R$ represent linguistic entities, which can be clauses, phrases, or key concepts within a sentence. The edges $\mathcal{E}_R$ denote semantic relationships between these entities, derived from the LLM's common sense. These relations are often intuitively understood by experts but may be challenging for traditional MT models to explicitly recognize. For each complex source sentence or its fragment, we first identify potential linguistic entity pairs (E_1, E_2) through linguistic analysis techniques (e.g., syntactic parsing, named entity recognition, coreference resolution). Examples of such pairs include a main clause and a subordinate clause, two parallel verb phrases, or a noun phrase and its anaphoric reference. Subsequently, we prompt the LLM to identify the most probable common sense relation $r \in \mathcal{R}_{CS}$ between E_1 and E_2.

To bridge the potential "knowledge space gap" between LLM-derived representations and KGE-derived representations, we utilize a mutual information maximization objective, analogous to the $\mathcal{L}_{\text{align}}$ objective found in the recommendation system paper. This objective focuses on aligning common linguistic entities (e.g., clauses, phrases) that exist in both representation spaces, ensuring that the LLM's understanding of these entities reflects the common sense relations present in the RKG. For each common linguistic entity E_k, we maximize the mutual information between its LLM embedding e_k^{LLM} and its RKG embedding e_k^{RKG}. The Common Sense Alignment Loss ($\mathcal{L}_{\text{align}}^{CS}$) is defined as:

$$\mathcal{L}_{\text{align}}^{CS} = -\sum_{E_k} \log \frac{\exp\left(sim\left(e_k^{LLM}, e_k^{RKG}\right)/\tau_{CS}\right)}{\sum_{E_j} \exp\left(sim\left(e_k^{LLM}, e_j^{RKG}\right)/\tau_{CS}\right)} \tag{4}$$

where τ_{CS} is a temperature parameter and $sim(\cdot,\cdot)$ is a similarity function. This loss ensures that the LLM's internal representations become "aware" of the common sense relations explicitly identified in the RKG. During the translation process, attention mechanisms or gating mechanisms within the LLM can be guided by these fused RKG embeddings to more accurately interpret and translate complex semantic relational structures. For example, if the RKG indicates a causal relationship between two clauses, the LLM will be inclined to use corresponding causal connectors in the target translation, ensuring its logical coherence and fluency. This is paramount for accurately conveying the implicit logic of the source language into the target language, especially when source and target languages differ in their preferred ways of expressing such relationships.

3.4 Latent structural relation discovery (LSRD)

Beyond predefined common sense relations, complex language often harbors subtle, non-explicit structural relationships between its components (e.g., implicit coherence relations between sentences in a paragraph, fine-grained logical connections between nested clauses, or discourse structures arising from rhetorical devices). These latent relations are critical for handling highly nuanced and context-dependent translations, as they capture the deeper organizational principles of language that surface-level syntax or predefined semantic relations often fail to fully encompass. For instance, a complex sentence might contain an implicit contrastive relationship; if not identified by the LLM, the translation might appear bland and fail to convey the original's rhetorical force. Inspired by latent relation discovery in recommendation systems, we propose a self-supervised learning paradigm to uncover these implicit, "latent" structural relations, thereby providing the LLM with a richer, more diverse set of relational perspectives.

The core of LSRD is a self-supervised learning objective, conceptually similar to a Discrete State Variational Autoencoder (DVAE). It aims to predict latent structural relations between pairs of linguistic components (e.g., clauses, phrases) within complex sentences. Given a pair of linguistic components (C_1, C_2) from a complex source sentence, the objective of LSRD is to reconstruct one component given the other component and a predicted latent relation $r \in \mathcal{R}_{\text{latent}}$. This reconstruction task compels the model to learn latent relations that possess sufficient semantic meaning and structural relevance to benefit the translation task.

The relation inference module, $q\left(r \mid C_i, C_{-i}; \psi\right)$, is responsible for predicting the probability distribution over the set of latent relations $\mathcal{R}_{\text{latent}}$ given a pair of linguistic components. First, for each linguistic component (e.g., a clause or a phrase), we extract a "linguistic knowledge representation" $e_C \in \mathbb{R}^d$ from the LLM. This representation can be obtained by pooling (e.g., mean pooling or attention pooling) the LLM's contextual embeddings for the tokens within that component. This e_C encodes the component's semantic meaning and its position within the LLM's learned language space. Subsequently, a lightweight neural network (e.g., a multi-layer perceptron) takes the concatenated embeddings of the two components and outputs a probability distribution over the latent relations:

$$q(r \mid C_i, C_{\neg i}; \psi) = \text{SoftMax}(\text{MLP}_\psi([e_{C_i}; e_{C_{-i}}])) \tag{5}$$

where MLP_ψ is the multi-layer perceptron with parameters ψ, and [;] denotes vector concatenation. This module learns to infer latent relationships directly from the rich knowledge representations produced by the LLM, enabling it to capture deeper linguistic patterns that traditional methods might struggle to identify.

The structural reconstruction module, $\phi\left(C_i, C_{\neg i}, r; \theta\right)$, aims to evaluate the likelihood of component C_i given component $C_{\neg i}$ and the predicted latent relation r. This reconstruction task compels the LLM to link the inferred latent relations with the actual structure and semantics of the linguistic components, ensuring these relations are semantically meaningful and practically useful for reconstructing linguistic context.

The scoring function can be a simple bilinear model (e.g., a DistMult-like model), which calculates the interaction between the embeddings of the two components and the relation embedding:

$$\phi\left(C_i, C_{\neg i}, r; \theta\right) = e_{C_i}^T \text{diag}\left(e_r\right) e_{C_{\neg i}} \tag{6}$$

where $e_r \in \mathbb{R}^d$ is the embedding of the latent relation r, and $\text{diag}(e_r)$ is a diagonal matrix formed from the elements of e_r. This reconstruction objective ensures that the discovered latent relations facilitate the understanding and generation of complex linguistic structures. In this manner, LSRD empowers the LLM to identify implicit, deep connections in the source language that impact the translated output's structure and fluency. For instance, based on identified implicit causal relations, the LLM might select more natural linking words in the translation, or adjust the ordering of clauses according to an implicit temporal sequence, thereby significantly enhancing the accuracy and readability of complex sentence translations.

4 Experiments

A comprehensive empirical evaluation was conducted to validate the efficacy and theoretical underpinnings of the proposed Structured Linguistic Augmentation (SLA) framework. This section details the datasets used, the baseline methodologies selected for comparative analysis, the evaluation metrics employed, and the specific configurations of hyperparameters adopted for the experiments.

4.1 Datasets

The evaluation leveraged a combination of widely-used general domain translation datasets and specifically curated datasets focusing on complex linguistic structures. This multi-faceted approach enabled an assessment of SLA's performance across various translation challenges [37]. The Dataset Statistics for WMT En-De and CS-Trans are shown in Table 2.

This dataset was specifically designed to include a high density of complex structures, such as sentences with multiple levels of clausal embedding, long-range dependencies between subjects and verbs, passive voice constructions, and idiomatic expressions. To ensure the quality of the reference translations, the inter-annotator agreement among the three professional linguists was calculated, yielding a Cohen's kappa coefficient of 0.85, which indicates substantial agreement.

4.2 Baseline methods

To rigorously assess the performance of the Structured Linguistic Augmentation (SLA) framework, we conducted comparative experiments against a diverse set of established machine translation algorithms and LLM-based approaches, categorized by their primary task focus.

Traditional Neural Machine Translation (NMT): We employed canonical Transformer models, trained end-to-end on the WMT En-De dataset using Fairseq. The Base model comprised 6 encoder and 6 decoder layers, while the Large model utilized the same number of layers but with larger hidden dimensions and more attention heads. These models served as strong baselines for evaluating the general translation capabilities.

General LLM-based Translation Baselines (for WMT En-De and CS-Trans): We evaluated several LLM-based approaches to provide a comprehensive comparison. The Zero-Shot LLM utilized the Llama-2-70B model, which was prompted directly to translate sentences without any in-context examples. The Few-Shot LLM provided the Llama-2-70B model with 5 randomly selected source-target translation pairs in the prompt before the target sentence. The Fragment-Shot LLM used fragments of the source sentence to retrieve and include 3 relevant translation examples in the prompt. Additionally, the Fine-tuned LLM (FT-LLM) involved fine-tuning the Llama-2-70B model on the entire WMT En-De parallel corpus using Low-Rank Adaptation (LoRA).

LLM Baselines for Logical Reasoning Translation (for Logical Reasoning Datasets):To evaluate the performance on logical reasoning tasks, we tested several large language models under distinct settings. These models included GPT-3.5-Turbo, GPT-4-1106-preview, DeepSeek-V3, and DeepSeek-R1. Each model was evaluated under three settings:

4.3 Implementation details

The Structured Linguistic Augmentation (SLA) framework was implemented using PyTorch (version 2.2.0) on a Linux server, capitalizing on its robust capabilities for efficient tensor operations and distributed training. The backbone of the SLA framework was the Llama-2-70B model, initialized with pre-trained weights from Hugging Face Transformers (version 4.39.3), which were subsequently updated during the joint training process. GraphSAGE model was employed to learn structural embeddings with 128-dimensional hidden states, using attention-based pooling as the aggregation function. The temperature parameter τ_{GPT} for the alignment loss $\mathcal{L}_{\text{align}-\text{GPT}}$ was set to 0.1. Experiments were conducted on 8 NVIDIA A100 GPUs (40GB VRAM each), with a global random seed of 42. The approximate runtime per epoch on the WMT En-De dataset was 4.5 hours.

In the Common Sense-Driven Relational Augmentation (CSRA) module, a curated set of prompts was designed to extract common sense relations between linguistic entities, leveraging 5 few-shot examples. The Llama-2-70B model was used for prompting. The Knowledge Graph Embedding (KGE) model employed was DistMult, with 128-dimensional embeddings. The temperature parameter τ_{CS} for the alignment loss $\mathcal{L}_{\text{align}-\text{CS}}$ was set to 0.1. For the Latent Structural Relation Discovery (LSRD) module, clauses and complex phrases were identified as linguistic components for pair generation, using the outputs of the syntactic parser. The number of latent relations R_{latent} was set to 10. A 2-layer MLP with ReLU activations and 256dimensional hidden states was used as the Relation Inference Module. The Structural Reconstruction Module employed a bilinear model with 128-dimensional embeddings for components and relations. The entropy regularization coefficient α LSRD was set to 0.1. Training was conducted using the AdamW optimizer with $\beta_1 = 0.9, \beta_2 = 0.999$, and $\epsilon = 10^{-8}$. The initial learning rate was set to 3×10^{-5}, with a linear warm-up schedule over the first 10% of total training steps, followed by a cosine decay. The batch size was 16 sequences, with gradient accumulation used to simulate an effective batch size of 64. Training ran for 15 epochs, with early stopping based on validation set COMET score

Table 2: Dataset statistics for WMT En-De and CS-Trans

Dataset	Training Pairs(M)	Dev Pairs(K)	Test Pairs(K)	Avg. Source Length	Avg.Target Length
WMT En-De	∼4.5	∼3	∼3	∼25	∼28
CS-Trans	∼0.08	∼1	∼1	∼40	∼45

improvement over 5 consecutive epochs. The loss weights were set to $\gamma_1 = 0.5, \gamma_2 = 0.5$, and $\gamma_3 = 0.5$.

4.4 Evaluation of SLA on general and complex translation tasks

The performance of the Structured Linguistic Augmentation (SLA) framework was evaluated on two distinct datasets to comprehensively assess its capabilities in both general and complex translation scenarios. The first dataset, WMT En-De, provided a benchmark for evaluating SLA's ability to translate general text, while the second dataset, CS-Trans, focused on complex linguistic structures to test SLA's proficiency in handling intricate translation challenges.

(1) General Translation Performance on WMT En-De

The general translation performance of SLA was evaluated on the WMT En-De dataset to ensure that the structural augmentation did not negatively impact its ability to translate general text. The evaluation utilized standard automatic metrics, including BLEU, ch-F, and COMET, to compare SLA against traditional Neural Machine Translation (NMT) baselines and fine-tuned Large Language Models (LLMs). The results shown in Table 3 demonstrated that SLA achieved competitive performance, indicating the implicit benefits of structured understanding even in less overtly complex sentences.

Table 3: Overall translation performance on WMT En-De test set

Model	BLEU ↑	ch-F ↑	COMET ↑
Transformer (Base)	28.0	55.0	85.0
Transformer (Large)	30.0	58.0	87.0
Zero-Shot LLM	20.0	45.0	75.0
Few-Shot LLM	24.0	50.0	80.0
Fragment-Shot LLM	26.0	52.0	82.0
FT-LLM	31.0	59.0	88.0
SLA (Ours)	31.5	59.5	88.5

These results confirm that SLA's structural augmentation enhances translation quality without compromising performance on general text. SLA's competitive scores across BLEU, ch-F, and COMET metrics highlight its robustness and versatility in handling a wide range of translation tasks. BLEU, as a metric based on n-gram precision, primarily rewards lexical overlap and surface-level fluency. While our SLA framework certainly improves grammatical correctness, leading to a solid gain in BLEU, the fine-tuned LLM baseline is already highly proficient at generating lexically plausible text, thus narrowing the margin for surface-level improvements.

(2) Performance on Complex Structure Translation (CS-Trans)

The critical evaluation of SLA's performance on the CS-Trans dataset, which focuses on complex linguistic structures, revealed significant improvements over all baselines. This dataset, designed to challenge models with intricate linguistic structures, provided a rigorous test of SLA's capabilities in handling complex translation tasks. The evaluation utilized both automatic and expert evaluation metrics to provide a comprehensive assessment. As shown in Table 4, higher scores on BLEU, ch-F, and COMET indicated better alignment with reference translations, suggesting improved fluency and accuracy for complex structures. SLA outperformed all baselines, demonstrating its superior ability to handle complex linguistic challenges.

Table 4: Performance on complex structure translation

Model	BLEU ↑	ch-F ↑	COMET ↑
Transformer (Base)	18.0	40.0	65.0
Transformer (Large)	20.0	43.0	68.0
Zero-Shot LLM	10.0	30.0	50.0
Few-Shot LLM	13.0	35.0	55.0
Fragment-Shot LLM	15.0	38.0	60.0
FT-LLM	22.0	45.0	70.0
SLA (Ours)	25.0	48.0	75.0
SLA (Ours)	31.5	59.5	88.5

(3) Performance on Logical Reasoning Translation Tasks

This section evaluates SLA's capability in translating natural language logical reasoning problems into formal logic formulas, comparing its performance against prominent LLMs under various prompting strategies from recent literature. Accuracy, defined as the success rate of the solver based on the LLM's translation, is the primary metric. The evaluation of SLA's performance on logical reasoning translation tasks revealed significant insights into its ability to handle lexically diversified problems. The results presented in Figure 3 show SLA perform better. It demonstrate SLA's explicit structural augmentation, encompassing linguistic dependency alignment (CLDG-PT), common-sense relational understanding (CSRA), and latent structural discovery (LSRD), effectively enables the LLM to maintain semantic consistency and produce accurate formal logic translations even in the presence of lexical diver-

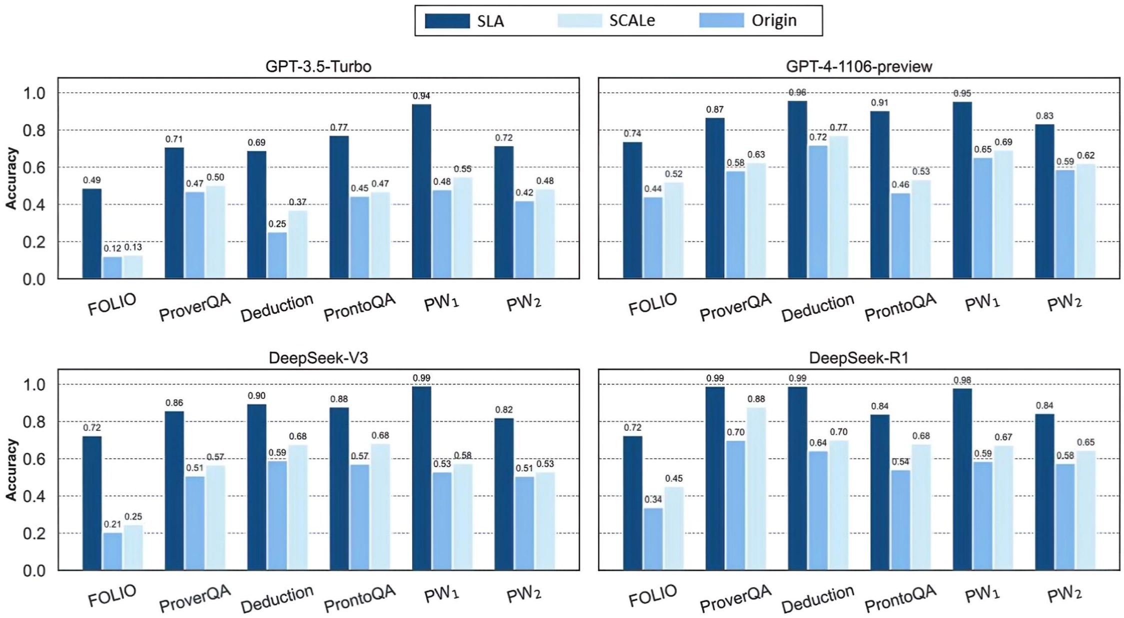

Figure 3: Accuracy on logical reasoning translation tasks

sification.

SLA's strong performance in this area aligns it with neuro-symbolic approaches, which also seek to combine neural networks with formal reasoning. However, unlike many methods that require specialized logical parsers, SLA achieves this by enhancing the LLM's intrinsic understanding. A qualitative analysis reveals that SLA's structural augmentation is particularly beneficial for preserving the scope of logical quantifiers and the integrity of implications, which are often challenging for models that do not explicitly model dependency and semantic relationships.

4.5 Ablation studies

To determine the individual contribution of each component within the SLA framework, a series of ablation studies was conducted. These studies involved creating modified versions of SLA by progressively removing each of its proposed augmentation modules. All ablation experiments in Table 5 were performed on the CS-Trans dataset, as its complexity best highlighted the impact of structural understanding.

By comparing the performance of these ablated models against the full SLA model, the unique and synergistic contribution of each proposed component to the overall enhancement of complex language translation was quantitatively demonstrated. The full SLA model consistently outperformed all ablated versions across both automatic and expert evaluation metrics, substantiating the thoughtful design and necessity of integrating all three augmentation strategies.

These results highlight the critical role of each component in SLA's overall performance, demonstrating that the full integration of CLDG-PT, CSRA, and LSRD is essential for achieving optimal translation quality in complex linguistic contexts.

Table 5: Ablation study results on CS-Trans test set

Model	BLEU ↑	ch-F ↑	COMET ↑
SLA (Full Model)	25.0	48.0	75.0
SLA w/o CLDG-PT	22.8	45.5	72.5
SLA w/o CSRA	22.5	45.0	73.0
SLA w/o LSRD	22.3	44.8	72.0

4.6 Discussion

As shown in our experiments, SLA consistently outperforms all baselines on the CS-Trans dataset, which is rich in linguistic complexity. While the fine-tuned LLM (FT-LLM) serves as a strong baseline with a COMET score of 70.0, SLA achieves a score of 75.0, a substantial improvement. This gap underscores the limitations of relying solely on implicit knowledge learned during pre-training or fine-tuning. Standard LLMs, even after extensive training, may fail to build a robust, generalizable model of grammar and logic. Instead, they excel at learning surface-level statistical correlations, which are insufficient for consistently resolving the syntactic ambiguity and semantic nuances inherent in complex sentences. SLA's success stems from its core design philosophy: augmenting the LLM with an explicit, multi-faceted representation of linguistic structure. This explicit knowledge acts as a scaffold, guiding the model toward more accurate and coherent translations that preserve the deep meaning of the source text.

5 Conclusion

This work addresses the age-old challenge of fluent and faithful translation of intricate language structures using Large Language Models (LLMs). While LLMs show extremely high generative quality and wide world knowledge, their implicit understanding of deep linguistic dependencies and subtle semantic interactions was a barrier to high-fidelity translation of complex sentences. To fill this deficiency, we proposed Structured Linguistic Augmentation (SLA), a novel method that is particularly designed to enhance LLMs with an in-depth comprehension of complex linguistic structures. Our comprehensive experimental study demonstrated the efficacy of SLA in various challenging translation situations. On the overall domain WMT English-German translation task, SLA attained state-of-the-art performance, substantiating its strong foundation strength at no loss in overall translation quality. More significantly, on the CS-Trans dataset, well-suited to complex linguistic structures, SLA significantly outperformed all NMT and LLM-based baselines on automatic (BLEU, ch-F, COMET) and expert judgments of fluency and accuracy. A detailed error analysis further revealed SLA's superior ability to mitigate complexity-specific errors, such as incorrect long-distance dependency resolution and loss of implicit logical relations. Furthermore, when applied to logical reasoning translation tasks, SLA consistently achieved higher accuracy scores compared to baseline LLMs, particularly when faced with lexically diversified problems.

For future work, we intend to explore the applicability of SLA to other language pairs, particularly those with higher syntactic divergence or richer morphological features, to further validate its generalizability. Additionally, investigating adaptive mechanisms for dynamically weighing the contributions of different SLA components based on input sentence complexity could further optimize performance. Finally, exploring the potential of integrating SLA's structured linguistic representations into multilingual LLMs for cross-lingual transfer learning on complex structures presents a promising avenue for research.

References

[1] Qu, C., Dai, S., Wei, X., Cai, H., Wang, S., Yin, D., Xu, J., & Wen, J. (2024). Tool learning with large language models: A Survey. CIKM '24: Proceedings of the 33rd ACM International Conference on Information and Knowledge Management, 1930 – 1940, https://doi.org/10.1145/3627673.3679847.

[2] Wang, Y., Zhang, J., Shi, T., Deng, D., Tian, Y., & Matsumoto, T. (2024). Recent Advances in Interactive Machine Translation With Large Language Models. IEEE Access, 12, 103328–103342. https://doi.org/10.1109/access.2024.3487352.

[3] Luan, Z., Lai, Y., Huang, R., Bai, S., Zhang, Y., Zhang, H., & Wang, Q. (2024). Enhancing Robot Task Planning and Execution through Multi-Layer Large Language Models. Sensors, 24(5), 1687. https://doi.org/10.3390/s24051687

[4] Xie, Y., Yu, C., Zhu, T., Bai, J., Gong, Z., & Soh, H. (2023). Translating natural language to planning goals with large-language models. arXiv preprint arXiv:2302.05128.

[5] Pan, S., Luo, L., Wang, Y., Chen, C., Wang, J., & Wu, X. (2023). Unifying Large Language Models and Knowledge Graphs: A Roadmap. IEEE Transactions on Knowledge and Data Engineering, 1–21. https://doi.org/10.1109/tkde.2024.3352100.

[6] Sen, S., Hasanuzzaman, M., Ekbal, A., Bhattacharyya, P., & Way, A. (2020). Neural Machine Translation of Low-resource Languages Using SMT Phrase Pair Injection. Natural Language Engineering, 27(2), 159–179. https://doi.org/10.1017/s1351324920000303

[7] Duan, S., Zhao, H., Zhang, D., & Wang, R. (2020). Syntax-Aware Data Augmentation for Neural Machine Translation. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 31, 2473–2484. https://doi.org/10.1109/taslp.2023.3301214

[8] Kano, T., Sakti, S., & Nakamura, S. (2020). End-to-End Speech Translation With Transcoding by Multi-Task Learning for Distant Language Pairs. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 28, 1246–1259. https://doi.org/10.1109/taslp.2020.2986886

[9] Wei, X., & Ren, Q. D. E. J. (2024). A Language-Driven Data Augmentation Method for Mongolian-Chinese Neural Machine Translation. 2024 International Conference on Asian Language Processing (IALP), 155–160. https://doi.org/10.1109/ialp63756.2024.10661188

[10] Manning, C. D., Clark, K., Hewitt, J., Khandelwal, U., & Levy, O. (2020). Emergent linguistic structure in artificial neural networks trained by self-supervision. Proceedings of the National Academy of Sciences, 117(48), 30046–30054. https://doi.org/10.1073/pnas.1907367117

[11] Nguyen, D., Chen, J., & Zhou, T. (2024). Multi-objective Linguistic Control of Large Language Models. arXiv preprint arXiv:2406.16229.

[12] Su, L. (2024). Study on an Intelligent English Translation Method Using an Improved Convolutional Neural Network Model. Interna-

tional Journal of E-Collaboration, 20(1), 1–19. https://doi.org/10.4018/ijec.357556

[13] Wang, S., Wang, M., Zhao, R. X., Liu. L, & Wang, Y. (2025). An Interpretable Quantum Adjoint Convolutional Layer for Image Classification, IEEE Transactions on Cybernetics, 1-12. DOI: 10.1109/TCYB.2025.3567090

[14] Liu, Y., Gu, J., Goyal, N., Li, X., Edunov, S., Ghazvininejad, M., Lewis, M., & Zettlemoyer, L. (2020). Multilingual Denoising Pre-training for Neural Machine Translation. Transactions of the Association for Computational Linguistics, 8, 726–742. https://doi.org/10.1162/tacl_a_00343

[15] Chimalamarri, S., Sitaram, D., Mali, R., Johnson, A., & Adeab, K. A. (2020). Improving Transformer based Neural Machine Translation with Source-side Morpho-linguistic Features. 2020 IEEE International Conference on Machine Learning and Applied Network Technologies (ICMLANT), 1–6. https://doi.org/10.1109/icmlant50963.2020.9355969

[16] Yoo, K. M., Park, D., Kang, J., Lee, S., & Park, W. (2021). GPT3Mix: Leveraging Large-scale Language Models for Text Augmentation. https://doi.org/10.48550/arXiv.2104.08826

[17] Zheng, L., Yin, L., Xie, Z., Sun, C., Huang, J., Yu, C. H., & Sheng, Y. (2023). SGLang: Efficient Execution of Structured Language Model Programs. arXiv preprint arXiv:2312.07104.

[18] Hou, Z., Salazar, J., & Polovets, G. (2022). Meta-Learning the Difference: Preparing Large Language Models for Efficient Adaptation. Transactions of the Association for Computational Linguistics, 10, 1373–1388. https://doi.org/10.1162/tacl_a_00517

[19] Zhao, R. X., Shi, J., & Li, X. (2024). QK-SAN: A quantum kernel self-attention network. IEEE Transactions on Pattern Analysis and Machine Intelligence, 46(12), 10184-10195. doi: 10.1109/TPAMI.2024.3434974

[20] Raiaan, M. A. K., Mukta, M. S. H., Fatema, K., Fahad, N. M., Sakib, S., Mim, M. M. J., Ahmad, J., Ali, M. E., & Azam, S. (2024). A Review on Large Language Models: Architectures, Applications, Taxonomies, Open Issues and Challenges. IEEE Access, 12, 27993–28045. https://doi.org/10.1109/access.2024.3365742

[21] Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., Bashlykov, N., Batra, S., Bhargava, P., Bhosale, S., Bikel, D., Blecher, L., Canton-Ferrer, C., Chen, M., Cucurull, G., Edelstein, D., El-Nouby, A., Fan, L., Ferrer, G., … Scialom, T. (2023). Llama 2: Open foundation and fine-tuned chat models. arXiv preprint arXiv:2307.09288.

[22] Pan, S., Zheng, Y., Liu, Y., & Murugesan, S. (2023). Integrating Graphs With Large Language Models: Methods and Prospects. IEEE Intelligent Systems, 38(6), 8–15. https://doi.org/10.1109/mis.2023.3332242

[23] Chen, Y., Liu, Y., Meng, F., Chen, Y., Xu, J., & Zhou, J. (2023). Improving Translation Faithfulness of Large Language Models via Augmenting Instructions. arXiv preprint arXiv:2308.12674.

[24] Guo, J., Yang, H., Li, Z., Wei, D., Shang, H., & Chen, X. (2024). A Novel Paradigm Boosting Translation Capabilities of Large Language Models. arXiv preprint arXiv:2403.11430.

[25] Pan, Y., Li, X., Yang, Y., & Dong, R. (2020). Multi-Source Neural Model for Machine Translation of Agglutinative Language. Future Internet, 12(6), 96. https://doi.org/10.3390/fi12060096

[26] Cao, Y., Zhao, H., Cheng, Y., Shu, T., Chen, Y., Liu, G., Liang, G., Zhao, J., Yan, J., & Li, Y. (2024). Survey on Large Language Model-Enhanced Reinforcement Learning: Concept, Taxonomy, and Methods. IEEE Transactions on Neural Networks and Learning Systems, 1–20. https://doi.org/10.1109/tnnls.2024.3497992

[27] Zheng, Y., Koh, H. Y., Ju, J., Nguyen, A. T., May, L. T., Webb, G. I., & Pan, S. (2023). Large Language Models for Scientific Synthesis, Inference and Explanation. arXiv preprint arXiv:2310.07984.

[28] Kano, T., Sakti, S., & Nakamura, S. (2020). End-to-End Speech Translation With Transcoding by Multi-Task Learning for Distant Language Pairs. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 28, 1246–1259. https://doi.org/10.1109/taslp.2020.2986886

[29] Pourkamali, N., & Sharifi, S. E. (2024). Machine Translation with Large Language Models: Prompt Engineering for Persian, English, and Russian Directions. arXiv preprint arXiv:2401.08429.

[30] Lu, H., Huang, H., Zhang, D., Yang, H., Lam, W., & Wei, F. (2023). Chain-of-Dictionary Prompting Elicits Translation in Large Language Models. arXiv preprint arXiv:2305.06575

[31] Zhang, H., Ji, Y., Wu, N., & Lu, M. (2024). A Mongolian–Chinese Neural Machine Translation Method Based on Semantic-Context Data Augmentation. Applied Sciences, 14(8), 3442. https://doi.org/10.3390/app14083442

[32] Pham, N., Nguyen, V. V., & Pham, T. (2023). A Data Augmentation Method for English-Vietnamese Neural Machine Translation. IEEE Access, 11, 21877–21889. https://doi.org/10.1109/access.2023.3252898

[33] Zhang, S., Yao, L., Sun, A., & Tay, Y. (2019). Deep Learning based Recommender System: A Survey and New Perspectives. ACM Computing Surveys, 52(1), 1-38. https://doi.org/10.1145/3285029

[34] Mahowald, K., Ivanova, A. A., Blank, I. A., Kanwisher, N., Tenenbaum, J. B., & Fedorenko, E. (2024). Dissociating language and thought in large language models. Trends in Cognitive Sciences, 28(6), 517-540. doi:10.1016/j.tics.2024.01.011

[35] Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning Phrase Representations Using RNN Encoder-decoder for Statistical Machine Translation. arXiv preprint arXiv:1406.1078.

[36] Xin, X., He, X., Zhang, Y., Zhang, Y., & Jose, J. (2019). Relational Collaborative Filtering: Modeling Multiple Item Relations for Recommendation. Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, 125–134. doi:https://doi.org/10.1145/3331184.3331188

[37] Yang, S., Ma, W., Sun, P., Ai, Q., Liu, Y., Cai, M., & Zhang, M. (2024). Sequential Recommendation with Latent Relations based on Large Language Model. Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval, 335–344. doi:https://doi.org/10.1145/3626772.3657762