
Exploring the Advantages and Challenges of a Concept-Guided Approach in Large Language Model Aided Machine Translation: Integrating Generative AI And Human-like Cognition

Ming Qian

Human-Centered AI, Charles River Analytics, Cambridge, MA 02138, USA

mqian@cra.com

Chuiqing Kong

Linguist, San Mateo, CA 94403, USA

chuiqingkong@gmail.com

Abstract

Humans outperform large language models (LLMs) on sophisticated tasks because human cognition involves a range of cognitive functions and their dynamic interactions. This study explores how integrating human cognition through concept-guided instruction and few-shot teaching in the prompt can guide LLMs to improve translation outcomes. We first demonstrate that for simple and widely used concepts, concept-guided prompting approaches offer significant benefits. We then test prompt engineering with Chinese-to-English translation examples, using hypothetical spaces—generated by GPT4—to estimate the complexity of various concepts and Likert scores—generated by human experts—to evaluate the translation performance. Our findings show that LLM translation performance declines as concept complexity increases. We also identify additional challenges: LLMs struggle with continuity in explaining and practicing sophisticated concepts due to the lack of human-like cognitive functions, such as cognitive dissonance. Additionally, LLMs lack a graceful speed-accuracy tradeoff because they do not possess the dynamic information processing, response strategies, and performance assessment that humans do. However, LLMs can mitigate some of these challenges by using Chain-of-Thought (CoT) reasoning, which is especially effective for problems requiring consistent, well-structured reasoning steps. Despite this, LLMs can only represent the effects of complex human cognitive functions through (often) fragmented linguistic descriptions, whereas humans excel at understanding critical and broader contexts and the interconnections between cognitive aspects.

1 Introduction

Humans possess superior capabilities compared to large language models (LLMs) because human cognition encompasses a range of mental functions: language processing, perception, attention, memory, reasoning, problem-solving, planning, decision-making, and executing actions. According to triadic cognition theory (Cook 2018), dual process theory (Chaiken 1999; Daniel 2017), embodied cognition (Shapiro 2014), and cognitive psychology (Neisser 2014), human cognition arises from the dynamic interaction of these functions, yielding a more nuanced, multifaceted understanding of real-world contexts than a

simplistic and static computational model.

Recent advancements in large language models (LLMs) have shown that while these models still do not match the full breadth of human cognition, they are making significant strides in specific areas such as machine translation (Fan 2020 and Zhu 2023). Unlike traditional machine translation (MT) tools such as Google Translate and DeepL, which rely solely on the source language text, LLMs possess the ability to follow instructions and can even learn from bilingual training examples provided within the prompt. This capability allows LLMs to provide context-sensitive translations with improved performance, making them as a more adaptable and user-friendly option to traditional methods (Qian and Kong, 2024A and 2024B).

Concepts serve as a common entity used across multiple cognitive functions, both in human minds and large language models. John McCarthy's insight, "To understand natural language is to understand the concepts in the language, not just the words," (McCarthy 2007) captures the fundamental cognitive principle that language is a tool for conveying and manipulating concepts, and true understanding requires engaging with those underlying mental constructs rather than simply processing linguistic symbols.

One important approach to guide the LLM-based translation outcome is to provide concept definitions and examples through prompt engineering. In section 2, by experimenting with many examples associated with common sense concepts (e.g., systems of measurement such as the metric system and the imperial system) and cross-cultural conceptualization, we show that, given proper prompts in the form of pre-editing analysis, concept-guided translation prompt instruction, and post-editing revision, LLMs can successfully reflect the relevant concepts in the translation outcome.

On the other hand, for more sophisticated translation concepts, we want to test whether performance of the concept-guided approach degrades as the concepts become more complex. In section 3, we test prompt-engineering-based implementations of multiple translation-related concepts using Chinese-to-English translation examples. By using hypothetical spaces to represent the complexity levels of various concepts, we verify that LLMs' translation performance degrades as concept complexity level increases. Also in section 3, we identify and analyze several challenges faced by the concept-guided approach: LLMs lack continuity to explain and practice sophisticated concepts; and compared with humans, LLMs lack graceful speed-accuracy tradeoff. These challenges stem from the fundamental differences between LLMs and human cognition. As artificial systems designed solely for generating text, LLMs' internal processes do not mimic or engage the complex cognitive functions and dynamic interactions typical of the human brain. Consequently, these limitations in simulating human-like cognition contribute to the observed deficiencies.

Nevertheless, LLMs have been shown to

exhibit human-like cognitive functions such as reasoning patterns (Lampinen 2024; Hagendorff 2023). Section 4 demonstrates that LLMs can use Chain-of-Thought (CoT) reasoning (Wei 2022) to overcome some challenges and enhance the performance. This is particularly effective for sub-problems that benefit from well-formed, consistent reasoning steps, ultimately leading to the correct conclusion. However, while humans use multiple cognitive functions to solve complex problems, LLMs can only project this complexity, often in fragmented form, as linguistic descriptions. Again, these broader perspectives and interconnections are critical strengths of the complex cognitive functions and dynamic interactions typical of the human brain.

Section 5 concludes our findings.

2 Advantages of Concept-guided Pre-editing Analysis, Translation Instruction, and Post-editing Revision

The role of concepts across multiple cognitive functions is multifaceted and central to human cognition: facilitating reasoning and generalization (William 2010), guiding decision-making by providing a framework for understanding and categorizing new information (Snowden 2007), integrating multiple features and dimensions of information (Villani 2019) to support flexible cognition (Ionescu 2006, Johnsson 2021), bridging episodic and semantic memory (Morton 2013), and creating functionally unitary representations that can be accessed and used across different cognitive tasks (Mahon 2009).

In this section, we showcase several examples where concept-guided prompting has significant advantage by using LLM-based pre-analyses, translation instructions and post-editing revisions.

2.1 LLMs Leverage Pre-editing Style Analysis as Translation Instructions

LLMs can generate style analysis reports for narratives. Table 1 presents an example using GPT4 to analyze an English narrative's style. In (Qian 2023), we compared these results with analyses by two human professional translators on the same text.

GPT4's analysis was not only more comprehensive and accurate but also faster than human analysts. Subsequently, a GPT4 prompt could use this pre-editing style analysis as a guide. For example, the prompt might be: "Translate the following text from English to Chinese,

incorporating applicable styles such as humor, self-deprecation, conversational tone, vivid descriptions, casual language and slang, honesty, anecdotal use, exaggeration for effect, direct address, and rhetorical questions." Using pre-editing auto-analysis as a translation style guide can significantly improve upon current methods that do not support style instruction.

User (Prompt)

Please read the story at <https://www.thecut.com/2020/03/book-excerpt-samantha-irbys-wow-no-thank-you.html> and **derive a list of the author's writing style.**

ChatGPT (Used Link Reader)

Based on the excerpt from Samantha Irby's essay collection "Wow, No Thank You.", here are some characteristics of her writing style:

1. **Humorous and Self-deprecating:** Irby often uses humor and self-deprecation to engage her readers. She makes fun of herself and her situations, which adds a light-hearted tone.
2. **Conversational Tone:** Her writing style is conversational, feeling personal and relatable.
3. **Vivid Descriptions:** Irby uses vivid and detailed descriptions to paint a clear picture of her experiences. This helps to immerse the reader in her stories.
4. **Casual Language:** She uses casual language, slang, and colloquialisms, which contributes to the conversational/informal tone.
5. **Honesty and Authenticity:** Irby's writing is characterized by its honesty and authenticity. She doesn't shy away from discussing personal and sometimes uncomfortable topics.
6. **Anecdotes:** She uses anecdotes to illustrate points and make writing more relatable.
7. **Exaggeration for Effect:** Irby often uses exaggeration to emphasize her points.
8. **Direct Address:** She frequently addresses the reader directly, which helps to establish a connection between her and the reader.
9. **Use of Rhetorical Questions:** Irby often uses rhetorical questions to engage the reader.
10. ...

Table 1: GPT4 analyzed the discourse-level writing style, with human-identified entries highlighted in bold (more details on this example can be found in Qian 2023).

2.2 LLMs Utilize Concept-guided Prompts to Guide Translation and Post-editing

Concepts are crucial in human cognition, allowing us to mentally group similar objects, events, and

ideas. Unlike traditional machine translation and editing methods, which focus on sentence-by-sentence and word-by-word changes, an LLM-based approach can leverage various concepts to guide translations and revisions more effectively.

For example, Table 2 illustrates how a GPT-4 prompt describes converting units of measure from the imperial system to the international system of units (SI) to enhance machine translation revisions. This allows for converting multiple types (such as inches, pounds, and gallons to meters, kilograms, and liters) across multiple sentences in a single instruction. This method represents a significant improvement, demonstrating that concept-guided instructions are major advancements over traditional approaches.

<p>User (Prompt) Translate the following text segment from English to Chinese: Original English: ... ChatGPT The Chinese translation is: ... 他的身高是 5 英尺 10 英寸 (<i>His height is 5 feet and 10 inches</i>)...</p> <p>User (Prompt) Revise the translation to reflect the following concepts: <i>“convert the units of measure to the international system of units (SI)”</i> ChatGPT The Revised Chinese translation is: ... 他的身高是 1.78 米 (<i>His height is 1.78 meters</i>)...</p>

Table 2: GPT-4 prompt uses concept descriptions to automatically revise machine translation results, such as converting units from inches to meters.

Table 3 showcases an instance where GPT4 introduces a cultural difference issue in the prompt and guides a revision effort to address the issue. In such situations, human translators might only be able to fix one issue at a time. However, we can leverage concept-guided translation memory that can identify similar text segments based on concept similarity so that (conceptually) similar sentences can be processed in similar fashion (Qian and Kong, 2024B).

<p>User (Prompt) Translate the following text segment from English to Chinese:</p>
--

Original English: The last time I wore a Halloween mascot was in the second-grade costume parade at Lincoln Elementary School in 1986 (go dolphins).

The Chinese translation is:

ChatGPT

上一次我穿万圣节服装是在1986年林肯小学的二年级服装游行中（加油，海豚）。

User (Prompt)

Keep in mind that “dolphins” in this context refer to the school mascot. **Since most Chinese schools do not have mascot, please modify the translation to make the meaning more clear to Chinese audience.**

ChatGPT

上一次我穿万圣节服装是在1986年林肯小学的二年级服装游行中（加油，林肯小学的吉祥物海豚）。

Table 3: GPT4 Instructed translation revision: to reflect a cultural difference (most Chinese schools do not have mascot), GPT4 revision provides further elaboration (“林肯小学的吉祥物”—Lincoln elementary school’s mascot) in front of the word “海豚”.

2.3 Summary

For simple and straightforward concepts, concept-guided approaches, supported by prompt-engineering using LLMs, show significant benefits.

In the next section, we evaluate translation concepts with various complexities to evaluate the challenges faced by the concept-guided approaches.

3 Challenges Faced by the Concept-guided Approach

In this section, we test Chinese-to-English translation concepts with varying levels of difficulty. We found that while LLMs perform well with simple and straightforward concepts, they encounter significant challenges when handling more sophisticated concepts (Qian and Kong 2024A, Qian and Kong 2024B).

3.1 LLMs’ Performance Declines with Increasing Concept Complexity

In cognitive linguistics and mental spaces theory, a hypothetical space is a dynamically constructed partial and temporary conceptual domain used during thinking and speaking (Nolan, 2001). These mental spaces facilitate the representation of

complex scenarios, counterfactuals, and hypothetical situations.

Two translation concepts are compared for illustration:

“Subject omission” describes a linguistic phenomenon where a language permits the omission of subject if they can be inferred pragmatically or grammatically.

“Changing subject selection” involves choosing a different subject than the original one in the source language to enhance readability and fluency in the target language.

Applying concept 2 generally requires a larger hypothetical space than concept 1, as it often involves selecting from multiple possible subjects—nouns, noun phrases, or clauses—making it more complex.

We had used GPT4 to analyze the hypothetical space of ten translation concepts: (1) The term “pro-drop” describes a linguistic feature where a language permits the omission of subjects (and occasionally other pronouns) if they can be inferred pragmatically or grammatically. (2) The linguistic term “dislocation” describes the occurrence where the placement of words or phrases within a sentence varies considerably between two languages. (3) The concept of implicit conjunctions and the use of linking or transition words in Chinese contrasts with English. Chinese frequently omits conjunctions that are essential in English. (4) The practice of amalgamating clauses and brief sentences in Chinese to form an extended English sentence is noteworthy. (5) Deverbalization, a key concept in translation studies related to abstract meaning from the source language to effectively convey the intended message in the target language, becomes especially pertinent when translating between languages with markedly different structures, such as from Chinese to English. (6) Eliminating redundancy is a common practice in translating Chinese, enhances the readability and clarity of the target text. (7) Changing subject selection involves selecting a different subject than the one in the source text to enhance the readability and fluency of the English translation. (8) Changing perspective in the context of translating from Chinese to English involves altering the viewpoint or approach when converting text from one language to the other. (9) Changing parts of speech is essential to faithfully convey the meaning of the original while ensuring

adherence to English expressions, regardless of the corresponding parts of speech in Chinese. (10) Cultural translation aims to retain the intended meaning of a culturally sensitive message while considering cultural differences, in order to avoid confusion that may result from a direct translation.

GPT4 was used to generate a normalized score (ranged between 0 and 1) towards each concept in terms of the hypothetical spaces. A detailed concept description plus three examples prepared by human experts are provided in the prompt to set up the context for evaluation. Table 4 shows an example on deverbialization and the related complexity evaluation results generated by GPT4; The detailed GPT prompt and GPT-generated hypothetical space evaluation on all ten translation concepts are listed in Appendix 1.

User (Prompt)

Deverbalization is a concept in translation studies, particularly relevant in the context of translating between languages with significantly different structures, such as Chinese to English. The core idea of deverbalization is to move away from a direct, word-for-word translation, which often leads to awkward or inaccurate results due to the differences in grammar, syntax, and cultural contexts. ...

...

Three examples are listed below to illustrate the Chinese-to-English translation reflecting the concept above:

Chinese: 过去带动高速增长的主要是高投资，而高投资主要是三大需求来源，出口、基础设施建设、房地产。

English: The key driver of rapid growth has been investment, which is needed in three areas/goes to three areas, export, infrastructure, and real estate.

Chinese: 本土化人才了解中国市场上的需求和消费者的心理。

English: Local talents understand the market needs in China and know what customers want.

Chinese: 中国现在很多地区复工率很高。

English: Many areas in China have reopened their economies.

Evaluating the size of the "hypothetical space" for each of these translation concepts involves considering the range of interpretations, adaptations, and choices a translator might need to navigate when applying these concepts from Chinese to English. This hypothetical space represents the complexity and variability inherent in each concept, indicating the level of creativity, cultural understanding, and linguistic expertise required. Here's a rating for each concept, on a scale from 0 (minimal complexity) to 100 (maximum complexity), reflecting the size of their hypothetical spaces:

ChatGPT

Deverbalization (80/100): This concept requires deep understanding of both source and target languages, along with the ability to abstract and reconceptualize ideas in culturally appropriate ways. It involves a significant hypothetical space due to the need for creative interpretation and expression.

Changing Subject Selection (70/100): Adapting the subject for better readability and fluency can involve moderate to high complexity, especially when ensuring that the shift does not alter the intended meaning of the original text.

...

Table 4: GPT4 prompt and hypothetical space evaluation scores for the ten concepts (detailed prompt and GPT4-generated evaluation results are listed in Appendix 1).

Machine translation evaluation employed human experts to assign 5-point Likert scores to GPT-4's English-to-Chinese translation outcomes of ten concepts. The scale scores used were; Strongly Disagree = -2; Disagree = -1; Neither Agree nor Disagree=0; Agree=1; Strongly Agree = 2. The experiment used a dataset of 40 English-Chinese sentence pairs, with four pairs per concept (detailed in Appendix 2). Using 1-in-4 leave-one-out cross-validation, the GPT-4 model was given three Chinese-English sentence pairs (as few-shot training examples) and the concept definition in the prompt, and then generated the English translation for the fourth pair's Chinese sentence. This process was repeated to produce four distinct translations for each Chinese sentence, all of which were subject to human evaluation.

Figure 1 displays human-generated Likert scores for concept-guided machine translations (top plot) and GPT-4-generated scores for hypothetical space complexity (bottom plot). Generally, GPT-4 translation performance declines as the complexity of the concepts increases, indicating that LLMs struggle with more sophisticated concepts involving larger hypothetical spaces. An exception is the "cultural translation" concept, which, unlike others focused on syntax and semantics, falls into a unique category. This suggests that GPT-4's complexity scores for hypothetical space may not accurately reflect challenges in cultural translation.

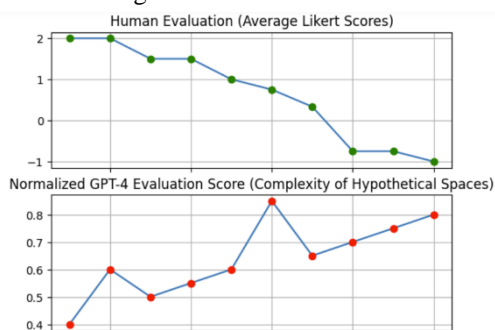


Figure 1: As GPT4's hypothetical space evaluation scores increase (normalized from 0 to 1), its performance (Likert Scores generated by human evaluators) degrades. The X-axis lists ten concepts, from left to right: eliminating redundancy, prodrop, implicit conjunction and linking, combining clauses and short sentences, dislocation, cultural translation, changing part of speech, subject selection, perspective, and deverbalization.

3.2 LLMs Lack Continuity

The "continuity" of AI models, particularly LLMs, refers to their ability to maintain coherent memory, reasoning, and contextual understanding over time. The lack of continuity in LLM models is a significant challenge that raises concerns about their ability to achieve human-like intelligence and self-awareness (Lieberman, 2024).

Table 5 presents an example where GPT4 contradicts itself. The concept under focus is "Pro-drop" that refers to the grammatical feature where

subjects/pronouns are omitted if their meaning can be inferred from context. In the example, GPT4 shows that it has epistemological knowledge about the concept. In addition, with the help of the definition and few-shot examples, GPT4 identifies that the concept should be applied and states that the subject is missing in the original Chinese text, yet it identifies a subject in its direct translation and retains this subject in the final translation, where it also claims to add a subject absent from the Chinese version. This example highlights the inconsistencies and logical flaws of large language models (LLMs) when dealing with complex concepts.

<p>User (Prompt) What is the concept of "Pro-drop" in the context of Chinese-to-English translation?</p> <p>ChatGPT The concept of "Pro-drop" (short for pronoun dropping) refers to the grammatical feature in certain languages where pronouns that are syntactically expected can be omitted if their meaning is understood from context. This characteristic is particularly relevant in the context of translating from Chinese to English because Chinese is a Pro-drop language, whereas English generally is not. In Chinese, speakers often omit subjects and sometimes objects when these are clear from the context or previously mentioned in the discourse...</p>
<p>User (Prompt) The concept of "pro-drop" refers to the ability of a language to omit subjects (and sometimes other pro-nouns) when they are pragmatically or grammatically inferable. Chinese is a pro-drop language, meaning that subjects can be and often are omitted when they are un-derstood from context. In contrast, English is a "non pro-drop" language, meaning that subjects are almost always explicitly stated, as English need the presence of subjects for clarity and grammatical correctness. Three examples are listed below to illustrate the Chinese-to-English translation reflecting the concept above: Chinese: 对传统那些没有比较优势和竞争优势产业, 继续给予必要保证。 English: Traditional industries without comparative advantage and competitiveness should be given the support.</p>

Chinese: 要重视发挥家庭、学校和社区作用, 帮助儿童抵御恐怖极端思想侵蚀。

English: Attention should be paid to the roles of families, schools, and communities in preventing children from being influenced by extreme terrorism.

Chinese: 需要货币政策传导机制来保证资金流动到位。

English: We need transmission mechanism of monetary policy to ensure the availability of capital flows.

If you are an expert-level translator, do you think we should apply the same concept on the following Chinese text segment?

Chinese: 各地积极性高涨。

Please give a score between 0 and 100 (0 means the concept definitely should not be applied on translating the sentence while 100 means that the concept definitely should be applied on translating the sentence.

If the score is above 75, what should be the proper translation reflecting the concept. Please generate a proper English translation reflecting the concept.

ChatGPT

The Chinese sentence "各地积极性高涨" can be translated while considering the pro-drop concept. This sentence, when directly translated, means "Enthusiasm is high everywhere." In a non pro-drop language like English, it's necessary to specify the subject for clarity. However, in this Chinese sentence, the subject is dropped because it's understood from the context.

In this case, I would give a score of around 85, indicating that the pro-drop concept should be applied in translation. The English translation, reflecting this concept and the inherent meaning of the sentence, could be:

"Enthusiasm is high in all regions."

Here, the subject "enthusiasm" is explicitly stated to maintain clarity in English, which is a non pro-drop language. The translation captures the essence of the original Chinese sentence while adhering to the syntactic requirements of English.

Table 5: GPT-4 "understands" the "pro-drop" concept; with the help of concept definition and few-shot examples, it identifies that the concept should be applied (a subject needs to be added for English) but it contradicts itself by keeping the same subject.

Cognitive dissonance, a psychological theory proposed by Leon Festinger (Festinger 1962), suggests that individuals experience discomfort when holding conflicting beliefs, attitudes, or behaviors. This discomfort, known as dissonance, drives individuals to seek consistency between their beliefs and actions to reduce the associated stress. Cognitive dissonance involves several cognitive functions: detecting discrepancies (attention and memory), experiencing physiological arousal and discomfort (emotion regulation), perception and confirmation (perception), and rationalization and justification (decision making).

LLMs are powerful tools for generating text and understanding language, but they lack the ability to leverage human cognitive functions, such as cognitive dissonance, to maintain consistency.

3.3 LLMs Lack Graceful Speed-Accuracy Tradeoff

The speed-accuracy tradeoff, as described in (Heltz, 2014), refers to the intricate relationship between an individual's preference for responding slowly with fewer errors versus responding quickly with more errors.

Compared to humans, LLMs were found unable to manage the speed-accuracy tradeoff as effectively (Goodman, 2024). In the context of translation, we tested multiple concepts and had similar findings. Table 6 presents a case study where GPT-4 was taught the concept of "changing subject selection"—selecting a subject different from the one in the source text (e.g., Chinese) to improve the readability and fluency of the target language (e.g., English) translation—using an instruction followed by three examples. In theory, any noun, noun phrase, or noun clause can serve as the subject in English. However, the training examples emphasized the initial possessive noun phrases in Chinese, consistently selecting the possessor noun as the English subject. Most human translators, upon reviewing these examples, would likely choose the possessor noun (迁移-migration) as the subject in their English translation. In contrast, GPT-4 failed to make this choice. Unlike humans, who can make a well-judged speed-accuracy tradeoff based on a few examples, LLMs' performance appears less adept and sometimes even seems irrational or confused.

User (Prompt)

For Chinese-to-English translation, changing subject selection involves selecting a subject different from the one in the source text to improve the readability and fluency of the English translation.

Below are three examples:

1. 技术的进步极大地改变了我们的生活方式。(direct translation: The advancement of **technology** has greatly changed our way of life.)

English translation: **Technology** have greatly advanced to change our way of life.

2. 老年人本身的免疫功能就相对于年轻人要弱一些。(direction translation: The immune function of the **elderly** is weaker than that of the young.)

English translation: Unlike younger people, **the elderly** have weaker immune systems.

3. 未来20年气候变化的威胁将超过恐怖主义。(direct translation: The threat of **climate change** will be more than terrorism in the next 20 years.)

English: In the next two decades, **climate change** will pose a greater threat than terrorism.

... Please generate a proper Chinese-to-English translation, and a proper English translation reflecting the concept and the few-shot examples :

Chinese: 迁移的流向和形式也都发生了很大的变化。(direct translation: The patterns and forms of **migration** have also undergone great changes.)

ChatGPT

English: The patterns and forms of migration have undergone significant changes.

Table 6: GPT-4 shows less graceful speed-accuracy tradeoff by not following the patterns illustrated by the three teaching examples.

Human’s ability to have graceful speed-accuracy tradeoff is supported by several cognitive functions such as information processing dynamics (involving perception, attention, memory, and problem-solving), response strategy (decision making), and performance assessment (metacognition).

LLMs are powerful tools for generating text and understanding language, but they lack the ability to leverage human cognitive functions to achieve graceful speed-accuracy tradeoff.

3.4 Summary

This section highlights three challenges in the concept-guided approach: (1) LLMs’ performance declines as the complexity of the required

conceptual space increases; (2) LLMs struggle with continuity in explaining and practicing advanced concepts; and (3) LLMs lack the ability of graceful speed-accuracy tradeoff. The challenges stem from the absence of comprehensive human-like cognitive functions in LLMs.

4 Chain-of-Thought (CoT) Reasoning Provides a Structured Way for LLMs to Support Sophisticated Concepts

In this section, we explore methods to address the limitations identified previously. We focus on Chain-of-Thought (CoT) reasoning prompting as a solution and evaluate the pros and cons of this approach.

4.1 CoT Demonstrates Detailed Reasoning Steps

Reasoning is a fundamental cognitive function that involves drawing conclusions, solving problems, and making decisions and encompasses various cognitive functions. Chain-of-thought (CoT) prompting can improve the performance of large language models (LLMs) on complex reasoning tasks (Wei 2022). The key idea behind CoT is to provide the LLM with examples that demonstrate step-by-step reasoning (intermediate reasoning steps) towards the final results, rather than just giving the final answer directly. This guides the LLM to construct a logical pathway from the original input to the final solution, rather than relying solely on its internal knowledge base which may have gaps or biases.

Compared to simple prompting without step-by-step reasoning, CoT prompting proves itself as effective solutions to complex translation concepts. Below are some examples in which CoT prompting shows significant benefits.

4.2 CoT Addresses the Issue of Lacking Continuity

Although fully implementing human-like cognitive dissonance is challenging, the linguistic part of its rationalization and justification component can be implemented through CoT reasoning to address the issue of continuity.

Table 7 shows that CoT prompting solves the “lack of continuity” example we described in section 3.2 and Table 5 related to the “Pro-drop”. An additional reasoning step is added to reflect the

omitted subject in each teaching examples, and during the generation phase, GPT4 is able to infer the omitted subject for the test example and generate the correct translation outcome that reflects the “pro-drop” concept correctly.

User (Prompt)

The concept of "pro-drop" refers to the ability of a language to omit subjects (and sometimes other pronouns) when they are pragmatically or grammatically inferable. Chinese is a pro-drop language, meaning that subjects can be and often are omitted when they are understood from context. In contrast, English is a "non pro-drop" language, meaning that subjects are almost always explicitly stated, as English need the presence of subjects for clarity and grammatical correctness.

Three examples are listed below to illustrate the Chinese-to-English translation reflecting the concept above:

Original Chinese: 对传统那些没有比较优势和竞争优势产业，继续给予必要保证。

Omitted subject: 传统产业 (traditional Industries)

English Translation: Traditional industries without comparative advantage and competitiveness should be given the support.

Original Chinese: 要重视发挥家庭、学校和社区作用，帮助儿童抵御恐怖极端思想侵蚀。

Omitted subject: 注意力 (attention or extra attention)

English: Attention should be paid to the roles of families, schools, and communities in preventing children from being influenced by extreme terrorism.

Original Chinese: 需要货币政策传导机制来保证资金流动到位。

Omitted subject: 我们 (we) or 中央银行 (central bank)

English: We need transmission mechanism of monetary policy to ensure the availability of capital flows.

If you are an expert-level translator, do you think we should apply the same concept on the following Chinese text segment?

Chinese: 各地积极性高涨。

Please give a score between 0 and 100 (0 means the concept definitely should not be applied on translating the sentence while 100 means that the concept definitely should be applied on translating the sentence.

If the score is above 75, what should be the omitted subject and the proper translation reflecting the concept. Please generate a proper English translation reflecting the concept.

ChatGPT

I would score this 90. The concept of "pro-drop" should indeed be applied here as the subject is contextually inferable.

Omitted subject: 各地人民 (People in various regions)

Proper Translation: "People in various regions are highly enthusiastic."

Table 7: By incorporating an additional CoT reasoning step, GPT4 addresses the issue of lacking continuity.

4.3 CoT Addresses the Issue of Lacking Speed-Accuracy Tradeoff

Although fully implementing human-like speed-accuracy tradeoff is challenging, the linguistic part of its information processing and response strategy components can be implemented through CoT reasoning to address the issue of continuity.

Table 8 shows that CoT prompting solves the “lacking speed-accuracy tradeoff” example we described in section 3.3 and Table 6 related to the “subject changing”. In each teaching example, multiple reasoning steps are added to reflect the direct translation, the subject phrase, and the alternative subject selection using the possessor noun in the subject phrase. Then during the generation phase, GPT4 is able to infer the same reasoning steps for the test example and have human-like speed-accuracy tradeoff and generate the corresponding translation outcome.

User (Prompt)

For Chinese-to-English translation, changing subject selection involves selecting a subject different from the one in the source text to improve the readability and fluency of the English translation.

Below are three examples:

1. Original Chinese: 技术的进步极大地改变了我们的生活方式。

Direct translation: The advancement of technology has greatly changed our way of life.

The subject phrase: The progress of technology

English translation with alternative subject selection using the possessor noun in the subject phrase: Technology have greatly advanced to change our way of life.

2. Original Chinese: 老年人本身的免疫功能就相对于年轻人要弱一些。

Direction translation: The immune function of the elderly is weaker than that of the young.

The subject phrase: The immune function of the elderly

English translation with **alternative subject selection using the possessor noun in the subject phrase**:

Unlike younger people, the elderly have weaker immune systems.

3. Original Chinese: 未来 20 年气候变化的威胁将超过恐怖主义。

Direct translation: The threat of climate change will be more than terrorism in the next 20 years.

The subject phrase: The threat of climate change

English translation with **alternative subject selection using the possessor noun in the subject phrase**: In the next two decades, climate change will pose a greater threat than terrorism.

Please follow the same pattern to translate the following Chinese sentence:

Original Chinese: 迁移的流向和形式也都发生了很大的变化。

ChatGPT

The sentence "迁移的流向和形式也都发生了很大的变化" directly translates to "The directions and forms of migration have also undergone significant changes." Here, the subject phrase is "The directions and forms of migration."

Applying the pattern of using the possessor noun as the subject, the translation would be: "Migration has changed significantly in its directions and forms."

Table 8: GPT4 achieves a more effective speed-accuracy tradeoff through the application of Chain of Thought (CoT) reasoning, as demonstrated by three illustrative teaching examples. In each teaching example, multiple reasoning steps are added to reflect the direct translation, the subject phrase, and the alternative subject selection using the possessor noun in the subject phrase. Then during the generation phase, GPT4 can infer the same reasoning steps and generate the correct result.

4.4 Summary

We show that LLMs can use Chain-of-Thought (CoT) reasoning to overcome some challenges and enhance the performance of the concept-guided approach. This is particularly effective for problems that benefit from well-formed, consistent reasoning steps, ultimately leading to the correct conclusion.

On the other hand, while humans use multiple cognitive functions to solve complex problems, LLMs can only model this complexity, often in fragmented form, as linguistic description. This approach may work for some challenges but falls

short for others. It demands unrealistic expectations to predefine linguistic solution steps and risks overlooking the broader context and interconnections between cognitive aspects—areas where human cognition excels.

5 Conclusion

Humans outperform large language models (LLMs) because human cognition involves a range of cognitive functions and their dynamic interaction. This study explores how integrating human cognition through concept-guided instruction and few-shot teaching can guide LLMs to improve translation outcomes. We first demonstrate that for simple and widely used concepts, concept-guided approaches offer significant benefits. We then tested prompt engineering with Chinese-to-English translation examples, using GPT-4 to estimate the complexity of various concepts and human experts to evaluate translation outcomes' quality. Our preliminary findings show that LLM translation performance declines as concept complexity increases.

We also identify several challenges in the concept-guided approach: LLMs struggle with continuity in explaining and practicing sophisticated concepts due to the lack of human-like cognitive functions, such as cognitive dissonance. Additionally, LLMs lack a graceful speed-accuracy tradeoff because they do not possess the cognitive functions such as dynamic information processing, response strategies, and performance assessment that humans do.

However, LLMs can mitigate some of these challenges by using Chain-of-Thought (CoT) reasoning, which is especially effective for problems requiring consistent, well-structured reasoning steps. Despite this, LLMs can only model cognitive complexity as fragmented linguistic descriptions, unlike humans, who excel at understanding broader contexts and interconnections between cognitive aspects. This limitation makes it unrealistic to expect LLMs to effectively define solution steps automatically, a task that humans can accomplish by naturally considering critical context.

The findings from this preliminary study on the English-to-Chinese language pair indicate that further research is necessary to validate the key results.

References

- Bao, C.Y. (2022). Class notes on Chinese-to-English translation, “Advanced Translation and Interpretation” course, Translation and Interpretation Program, Middlebury Institute of International Studies
- Chaiken, S., & Trope, Y. (Eds.). (1999). *Dual-process theories in social psychology*. Guilford Press.
- Cook, N. D. (2018). The triadic roots of human cognition: “Mind” is the ability to go beyond dyadic associations. *Frontiers in Psychology*, 9, 1060.
- Daniel, K. (2017). *Thinking, fast and slow*.
- Fan, A. (2020). Introducing the first AI model that translates 100 languages without relying on English. Meta. Accessed 28th April 2024. <https://about.fb.com/news/2020/10/first-multilingual-machine-translation-model>.
- Fenton, K. (2024). *Psy 210-002: Introduction to Psychology*. Openstax Pressbook.
- Festinger, L. (1962). Cognitive dissonance. *Scientific American*, 207(4), 93-106.
- Goodman, Noah (2024). Reasoning in Human and Machine Intelligence. Keynote Presentation, AAAI Spring Symposium on Empowering Machine Learning and Large Language Models with Domain and Commonsense Knowledge, Palo Alto, CA, USA.
- Hagendorff, T., Fabi, S., & Kosinski, M. (2023). Human-like intuitive behavior and reasoning biases emerged in large language models but disappeared in ChatGPT. *Nature Computational Science*, 3(10), 833-838.
- Harris, C. L. (2006). Language and cognition. *Encyclopedia of cognitive science*, 10(0470018860), s00559.
- Heitz, R. P. (2014). The speed-accuracy tradeoff: history, physiology, methodology, and behavior. *Frontiers in neuroscience*, 8, 86875.
- Ionescu, M., & Ralescu, A. (2006, July). Multidimensional Conceptual Spaces for Summarization. In *2006 IEEE International Conference on Fuzzy Systems* (pp. 560-565). IEEE.
- Johnsson, L. (2021). Multidimensional property supplementation: A method for discovering and describing emergent qualities of concepts in grounded theory research. *Qualitative Health Research*, 31(1), 184-200.
- Lampinen, A. K., Dasgupta, I., Chan, S. C., Sheahan, H. R., Creswell, A., Kumaran, D., ... & Hill, F. (2024). Language models, like humans, show content effects on reasoning tasks. *PNAS nexus*, 3(7).
- Lazzara, J. (2020). *Chapter 7 Psy 2E: Thinking, Language, and Problem Solving*. Openstax Pressbook.
- Lieberman, Henry (2024). The Next Grand Challenge for AI: Making Better Mistakes. Keynote Presentation, AAAI Spring Symposium on Empowering Machine Learning and Large Language Models with Domain and Commonsense Knowledge, Palo Alto, CA, USA.
- Mahon, B. Z., & Caramazza, A. (2009). Concepts and categories: a cognitive neuropsychological perspective. *Annual review of psychology*, 60(1), 27-51.
- Morton, N. W. (2013). Interactions between episodic and semantic memory. Technical report, Vanderbilt Computational Memory Lab.
- McCarthy, J. (2007). *What is artificial intelligence*.
- Neisser, U. (2014). *Cognitive Psychology: Classic Edition*. Psychology Press.
- Perlovsky, L. (2011). Language and cognition interaction neural mechanisms. *Computational Intelligence and Neuroscience*: 1-13.
- Qian, M., Wu, H.Q., Yang, L., and Wan, A. (2023). Augmented Machine Translation Enabled by GPT4: Performance Evaluation on Human-Machine Teaming Approaches. In *Proceedings of the First Workshop on NLP Tools and Resources for Translation and Interpreting Applications*. pages 20–31, Varna, Bulgaria.
- Qian, M., Kong, C.Q. (2024A). Exploring the Gap: The Challenge of Achieving Human-like Generalization for Concept-based Translation Instruction Using Large Language Models. *Symposium on Human-Like Learning. AAAI 2024 Spring Symposium Series*, Stanford University. Palo Alto, USA.
- Qian, M., Kong, C.Q. (2024B). Enabling Human-centered Machine Translation Using Concept-based Large Language Model Prompting and Translation Memory, In *Proceedings of 4th International Conference on Artificial Intelligence in HCI (AI-HCI), Part of HCI International Conference*, Washington DC, USA.
- Shapiro, L. A., & Spaulding, S. (Eds.). (2014). *The Routledge handbook of embodied cognition*.
- Snowden, D. J., & Boone, M. E. (2007). A leader's framework for decision making. *Harvard business review*, 85(11), 68.
- Villani, C., Lugli, L., Liuzza, M. T., & Borghi, A. M. (2019). Varieties of abstract concepts and their multiple dimensions. *Language and Cognition*, 11(3), 403-430.
- Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., ... & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems*, 35, 24824-24837.

Williams, J. J., & Lombrozo, T. (2010). The role of explanation in discovery and generalization: Evidence from category learning. *Cognitive science*, 34(5), 776-806.

Zhu, W., Liu, H., Dong, Q., Xu, J., Huang, S., Kong, L., ... & Li, L. (2023). Multilingual machine translation with large language models: Empirical results and analysis. arXiv preprint arXiv:2304.04675.

Appendix 1: Complexity Evaluation on the Ten Translation Concepts Using Hypothetical Spaces

Below are the GPT-4 analyses on the hypothetical spaces (complexity), the black/regular font words are the prompt and red/italicized font words are the generated outcomes by the GPT4.

Below are definitions and few-shot examples of ten concepts related to Chinese-to-English Translation:

1. Deverbalization is a concept in translation studies, particularly relevant in the context of translating between languages with significantly different structures, such as Chinese to English. The core idea of deverbalization is to move away from a direct, word-for-word translation, which often leads to awkward or inaccurate results due to the differences in grammar, syntax, and cultural contexts. Instead, the translator interprets the meaning of the source text at a deeper, more abstract level, and then re-expresses this meaning in the target language.

Three examples are listed below to illustrate the Chinese-to-English translation reflecting the concept above:

Chinese: 过去带动高速增长的主要是高投资，而高投资主要是三大需求来源，出口、基础设施建设、房地产。

English: The key driver of rapid growth has been investment, which is needed in three areas/goes to three areas, export, infrastructure, and real estate.

Chinese: 本土化人才了解中国市场上的需求和消费者的心理。

English: Local talents understand the market needs in China and know what customers want.

Chinese: 中国现在很多地区复工率很高。

English: Many areas in China have reopened their economies.

2. For Chinese-to-English translation, changing subject selection involves selecting a subject different from the one in the source text to improve the readability and fluency of the English translation.

Three examples are listed below to illustrate the Chinese-to-English translation reflecting the concept above:

Chinese: 资源环境约束边界临近，最典型的例子就是雾霾。

English: Resources and environment almost reach the limit with smog being the biggest concern.

Chinese: 老年人本身的免疫功能就相对于年轻人要弱一些。

English: Unlike younger people, the elderly have weaker immune systems.

Chinese: 迁移的流向和形式也都发生了很大的变化。

English: Migration was happening in different directions and ways.

3. The concept of "pro-drop" refers to the ability of a language to omit subjects (and sometimes other pro-nouns) when they are pragmatically or grammatically inferable. Chinese is a pro-drop language, meaning that subjects can be and often are omitted when they are understood from context. In contrast, English is a "non pro-drop" language, meaning that subjects are almost always explicitly stated, as English need the presence of subjects for clarity and grammatical correctness.

Three examples are listed below to illustrate the Chinese-to-English translation reflecting the concept above:

Chinese: 各地积极性高涨。

English: There is a lot of discussion and enthusiasm on this issue.

Chinese: 要重视发挥家庭、学校和社区作用，帮助儿童抵御恐怖极端思想侵蚀。

English: Attention should be paid to the roles of families, schools, and communities in preventing children from being influenced by extreme terrorism.

Chinese: 对传统那些没有比较优势和竞争优势产业，继续给予必要保证。

English: Traditional industries without comparative advantage and competitiveness should be given the support.

4. Changing the perspective in the context of Chinese-to-English translation refers to adapting the viewpoint or approach when translating text from Chinese into English. This concept is crucial due to significant differences between the two languages in terms of grammar, syntax, idiomatic expressions, and cultural contexts.

Three examples are listed below to illustrate the Chinese-to-English translation reflecting the concept above:

Chinese: 中国经济已经触底，中速增长平台初步确定，将来会逐步进入中速增长期，也就是由高速增长阶段转向高质量发展阶段。

English: China's economy has touched the bottom. It is set for medium-speed growth, which will last for a period of time. There will be more focus on quality instead of speed.

Chinese: 随着经济发展，可能是不同时期有不同热门的行业。

English: Along with economic development, sectors could become popular at different times.

Chinese: 本地化产生了多语言的需求。

English: Location involves translation into multiple languages.

5. For Chinese-to-English translation, Changing parts of speech is to faithfully convey the meaning of the original at the same time accords with English means of expression irrespective of the part of speech of a word in Chinese.

Three examples are listed below to illustrate the Chinese-to-English translation reflecting the

concept above:

Chinese: 存货、出口和生产性投资逐步进入回升期，将对基础设施投资减速形成对冲。

English: Investments in inventory, export, and manufacturing are bouncing back, offsetting the decline in infrastructure investment.

Chinese: 我是研究农村经济开始的。

English: I started my career as a researcher in rural economy.

Chinese: 委员们提出了许多很有操作性、针对性的意见建议。

English: Members of the committee have made a number of suggestions that are specific and easy to implement.

6. The linguistic concept dislocation refers to the phenomenon where the positioning of words or phrases in a sentence differs significantly between two languages. This can pose challenges in translation, particularly between languages with distinct syntactic structures, like Chinese and English.

Chinese: 常态化防控和应急处置放松了要求。

English: The requirements for daily prevention and control, as well as emergency response, have been relaxed.

Chinese: GDP比重，非典时期才影响26%，这(次)是占到74%。

English: 26% of the GDP was affected during the SARS outbreak. But this time, the proportion is 74%.

Chinese: 我知道这个目标目前在国际上是很多国家所没有制定的。

English: I'm aware that many countries around the world have not yet set this goal.

7. The concept of implicit conjunctions and linking/transition words for Chinese-to-English translation: Chinese often omits conjunctions that would be necessary in English. This is because Chinese relies heavily on context and the inherent logic of the sentence structure. For instance, two

clauses might be placed side by side without a conjunction, yet a native Chinese speaker would understand their relationship based on the context. That usually is not the case for English.

Three examples are listed below to illustrate the Chinese-to-English translation reflecting the concept above:

Chinese: 学校决定精简机构，减少系科。

English: The school decided to streamline the administration by reducing the number of departments.

Chinese: 比如，欧洲老龄化程度高、死亡率高。

English: For instance, the death rate in Europe is very high due to a large aging population.

Chinese: 提炼出中国智慧，总结出中国方案。

English: Draw some lessons so that we can come up with Chinese wisdoms and solutions.

8. The concept of combining clauses and short sentences in Chinese to create a longer English sentence. English allows for complex sentences with multiple clauses, often linked by conjunctions, which can express detailed nuances and conditions. Chinese, on the other hand, favors brevity and conciseness.

Three examples are listed below to illustrate the Chinese-to-English translation reflecting the concept above:

Chinese: 跨国公司的生产经营是在本土，但是它的生产经营战略是全球化的。

English: Multinational corporations are based in their home countries but with global operations and strategies.

Chinese: 本轮疫情呈现出多地暴发、波及范围广、确诊人数多等特点。

English: Recently, the pandemic took place in many locations with extensive scale and cases.

Chinese: 俄罗斯的改革比中国晚起步10年左右，但他们的金融抑制指数已经到了0.4，时间跨度更短、下降速度更快。

English: Russia started its reform 10 years later

than China, but it has already reached a financial repression index of 0.4 in a shorter period of time and at a faster pace.

9. Eliminating redundancy is a common practice for Chinese-to-English because redundancy is commonly seen in the Chinese language. Although some redundancies may serve as emphasis, many could be eliminated without compromising the intended meaning. Removing unnecessary repetition during translation could improve the readability and clarity of the target text.

Three examples are listed below to illustrate the Chinese-to-English translation reflecting the concept above:

Chinese: 全世界气候变化的原因来看，大概主要就是有两个原因。

English: Climate change has two factors.

Chinese: 水电是包含在可再生能源当中的，世界各国都把水电作为可再生能源的一部分。

English: Hydropower is considered a source of renewable energy by all countries.

Chinese: 从需求侧来讲，中国经济过去长时间的高增长，主要是高投资在支撑，而高投资主要有三大需求在支撑，也就是基础设施建设、房地产和出口。

English: On the demand side, the driver of China's economic growth has long been high investment required in three key areas: infrastructure, real estate, and exports.

10. Cultural translation aims to show the cultural differences while respecting the source culture as well.

Chinese: 有“歇歇脚、缓一缓”的心态。

English: Some people tend to let down their guard.

Chinese: 美联储把利率降到零，弹药全部用完。

English: There isn't much space as the Fed has lowered interest rate to zero.

Chinese: 我要是不来，那也是非常遗憾的。我这次来，即使被打得鼻青脸肿也在所不惜。

English: I would regret it if I didn't come, so I'm here today no matter what happens.

These concepts (constrained by both definitions and examples) have different levels of complexity when they are applied. Can you provide a rating evaluation (minimum 0 and maximum 100) in terms of the size of typical “hypothetical spaces” that need to be considered when the concepts are applied?

Evaluating the size of the "hypothetical space" for each of these translation concepts involves considering the range of interpretations, adaptations, and choices a translator might need to navigate when applying these concepts from Chinese to English. This hypothetical space represents the complexity and variability inherent in each concept, indicating the level of creativity, cultural understanding, and linguistic expertise required. Here's a rating for each concept, on a scale from 0 (minimal complexity) to 100 (maximum complexity), reflecting the size of their hypothetical spaces:

Deverbalization (80/100): This concept requires deep understanding of both source and target languages, along with the ability to abstract and reconceptualize ideas in culturally appropriate ways. It involves a significant hypothetical space due to the need for creative interpretation and expression.

Changing Subject Selection (70/100): Adapting the subject for better readability and fluency can involve moderate to high complexity, especially when ensuring that the shift does not alter the intended meaning of the original text.

Pro-drop (60/100): Translating from a pro-drop language like Chinese to a non-pro-drop language like English requires careful addition of explicit subjects without making the text feel redundant or

altering its meaning, presenting a moderate hypothetical space.

Changing the Perspective (75/100): Adapting the viewpoint requires a strong grasp of cultural and contextual nuances as well as linguistic fluency. The hypothetical space is large because it involves rethinking and sometimes reimagining the original message from a different angle.

Changing Parts of Speech (65/100): This task involves creatively maintaining the original's meaning while changing the grammatical structure, which can be complex due to the flexibility required in reconceptualizing the message.

Dislocation (60/100): Addressing syntactic differences involves understanding the functional and positional flexibility of elements in both languages. The hypothetical space is moderate, as it often has clear-cut rules but also requires attention to readability and flow.

Implicit Conjunctions and Linking (50/100): While the concept itself may seem straightforward, applying it effectively requires a nuanced understanding of how ideas are connected in English, presenting a moderate hypothetical space.

Combining Clauses and Short Sentences (55/100): This concept requires skill in creating complex sentences that are still clear and coherent, involving a moderate level of complexity in the hypothetical space due to the need to balance detail with readability.

Eliminating Redundancy (40/100): While often technical, this task requires less creative interpretation than others, presenting a smaller hypothetical space. The challenge lies in discerning which redundancies are stylistic versus those adding unnecessary repetition.

Cultural Translation (85/100): This concept encompasses a vast hypothetical space due to the need to convey cultural nuances and contexts accurately and respectfully, requiring deep cultural knowledge and sensitivity.

These ratings reflect the relative complexity and breadth of choices translators must navigate within each concept, acknowledging that individual translation tasks might present unique challenges that could affect these evaluations.

Table 9: Hypothetical spaces (complexity)
Analyses for the ten translation concepts using
GPT-4.

Appendix 2: English-Chinese sentence pairs (four pairs per concept) used for 1-in-4 leave-one-out cross-validation evaluation

Below is a dataset of 40 English-Chinese sentence pairs (four pairs per concept) that was used for 1-in-4 leave-one-out cross-validation evaluation of GPT-4 generated translation.

Pro-drop

Chinese: 各地积极性高涨。

English: There is a lot of discussion and enthusiasm on this issue.

Chinese: 要重视发挥家庭、学校和社区作用, 帮助儿童抵御恐怖极端思想侵蚀。 English:

Attention should be paid to the roles of families, schools, and communities in preventing children from being influenced by extreme terrorism.

Chinese: 对传统那些没有比较优势和竞争优势产业, 继续给予必要保证。

English: Traditional industries without comparative advantage and competitiveness should be given the support.

Chinese: 需要货币政策传导机制来保证资金流动到位。

English: We need transmission mechanism of monetary policy to ensure the availability of capital flows.

Dislocation

Chinese: 常态化防控和应急处置放松了要求。

English: The requirements for daily prevention and control, as well as emergency response, have

been relaxed.

Chinese: GDP比重, 非典时期才影响26%, 这(次)是占到74%。

English: 26% of the GDP was affected during the SARS outbreak. But this time, the proportion is 74%.

Chinese: 我知道这个目标目前在国际上是很多国家所没有制定的。

English: I'm aware that many countries around the world have not yet set this goal.

Chinese: 当地语言以前中国公司不太注重。

English: Chinese companies didn't pay enough attention to local languages.

Implicit conjunctions and linking/transition words

Chinese: 学校决定精简机构, 减少系科。

English: The school decided to streamline the administration by reducing the number of departments.

Chinese: 比如, 欧洲老龄化程度高、死亡率高。

English: For instance, the death rate in Europe is very high due to a large aging population.

Chinese: 提炼出中国智慧, 总结出中国方案。

English: Draw some lessons so that we can come up with Chinese wisdoms and solutions.

Chinese: 本地化是跨国公司进驻中国, 打通语言通道的关键。

English: Localization is an important tool to break the language barrier as they enter the Chinese market.

Combining clauses and short sentences in Chinese to create a longer English sentence

Chinese: 跨国公司的生产经营是在本土, 但是它的生产经营战略是全球化的。

English: Multinational corporations are based in their home countries but with global operations and strategies.

Chinese: 本轮疫情呈现出多地暴发、波及范围广、确诊人数多等特点。

English: Recently, the pandemic took place in many locations with extensive scale and cases.

Chinese: 俄罗斯的改革比中国晚起步10年左右, 但他们的金融抑制指数已经到了0.4, 时间跨度更短、下降速度更快。

English: Russia started its reform 10 years later than China, but it has already reached a financial repression index of 0.4 in a shorter period of time and at a faster pace.

Chinese: 说到可再生能源，首先是可再生，就说明它是资源量非常大的，可以永续利用，同时它的污染很小、分布很广。

English: Speaking of renewable energy, the word “renewable” suggests that these energy sources are abundant, sustainable, and widely distributed with very little pollution.

Deverbalization

Chinese: 过去带动高增长的主要是高投资，而高投资主要是三大需求来源，出口、基础设施建设、房地产。

English: The key driver of rapid growth has been investment, which is needed in three areas/goes to three areas, export, infrastructure, and real estate.

Chinese: 本土化人才了解中国市场上的需求和消费者的心理。

English: Local talents understand the market needs in China and know what customers want.

Chinese: 中国现在很多地区复工率很高。

English: Many areas in China have reopened their economies.

Chinese: 中国对外开放的步伐也在不断加快，我们要向外商提供更开放、更透明的、更有力的投资环境。

English: As China further opens up to the outside world, we should provide a more open, transparent, and helpful investment environment.

Eliminating redundancy

Chinese: 全世界气候变化的原因来看，大概主要就是有两个原因。

English: Climate change has two factors.

Chinese: 水电是包含在可再生能源当中的，世界各国都把水电作为可再生能源的一部分。

English: Hydropower is considered a source of renewable energy by all countries.

Chinese: 从需求侧来讲，中国经济过去长时间的高增长，主要是高投资在支撑，而高投资主要有三大需求在支撑，也就是基础设施建设、房地产和出口。

English: On the demand side, the driver of China’s economic growth has long been high investment required in three key areas: infrastructure, real estate, and exports.

Chinese: 我们需要的是强劲的反弹，所以需要10倍的努力来推动反弹。

English: We need to achieve a robust rebound through ten times of effort.

Changing subject selection

Chinese: 资源环境约束边界临近，最典型的例

子就是雾霾。

English: Resources and environment almost reach the limit with smog being the biggest concern.

Chinese: 老年人本身的免疫功能就相对于年轻人要弱一些。

English: Unlike younger people, the elderly have weaker immune systems.

Chinese: 迁移的流向和形式也都发生了很大的变化。

English: Migration was happening in different directions and ways.

Chinese: 未来20年气候变化的威胁将超过恐怖主义。

English: In the next two decades, climate change will pose a greater threat than terrorism.

Changing the perspective

Chinese: 中国经济已经触底，中速增长平台初步确定，将来会逐步进入中速增长期，也就是由高速增长阶段转向高质量发展阶段。

English: China's economy has touched the bottom. It is set for medium-speed growth, which will last for a period of time. There will be more focus on quality instead of speed.

Chinese: 随着经济发展，可能是不同时期有不同热门的行业。

English: Along with economic development, sectors could become popular at different times.

Chinese: 本地化产生了多语言的需求。

English: Location involves translation into multiple languages.

Chinese: 大多数国家对创业投资都是采取一种比较鼓励的政策。

English: Many countries have policies that encourage venture capital investment.

Changing Parts of Speech

Chinese: 存货、出口和生产性投资逐步进入回升期，将对基础设施投资减速形成对冲。

English: Investments in inventory, export, and manufacturing are bouncing back, offsetting the decline in infrastructure investment.

Chinese: 我是研究农村经济开始的。

English: I started my career as a researcher in rural economy.

Chinese: 委员们提出了许多很有操作性、针对性的意见建议。

English: Members of the committee have made a number of suggestions that are specific and easy to implement.

Chinese: 城镇化就是阶段性的，到了一定阶段一定会减速。

English: Urbanization takes place in different stages. It will slow down at a certain stage.

Cultural translation

Chinese: 有“歇歇脚、缓一缓”的心态。

English: Some people tend to let down their guard.

Chinese: 美联储把利率降到零，弹药全部用完。

English: There isn't much space as the Fed has lowered interest rate to zero.

Chinese: 我要是不来，那也是非常遗憾的。我这次来，即使被打得鼻青脸肿也在所不惜。

English: I would regret it if I didn't come, so I'm here today no matter what happens.

Chinese: “大干快上”本身没有错，关键是干什么、上什么。

Table 10: A dataset of 40 English-Chinese sentence pairs (four pairs per concept) was used for 1-in-4 leave-one-out cross-validation evaluation.