Conceptual Metaphor Theory as a Prompting Paradigm for Large Language Models

Oliver Kramer

Computational Intelligence Group
Department of Computer Science
University of Oldenburg, Germany
oliver.kramer@uni-oldenburg.de

Abstract. Large language models (LLMs) show strong reasoning abilities but often produce explanations that are inconsistent or difficult to interpret. In cognitive psychology and linguistics, Conceptual Metaphor Theory (CMT) describes how abstract concepts gain structure through systematic mappings from concrete experiential domains. Building on this framework, we introduce a CMT-based prompting paradigm that encodes explicit source–target mappings into prompt instructions, guiding models through structured, stepwise inference rather than ad hoc reasoning. A benchmark spanning metaphor identification, domain-specific reasoning, teaching, and reading comprehension demonstrates that CMT prompting improves coherence, interpretability, and conceptual depth across multiple LLM architectures. These results highlight the potential of cognitively grounded prompting to bridge statistical language modeling and human-like reasoning.

1 Introduction

LLMs exhibit impressive abilities in text generation, translation, and problem solving, yet their underlying reasoning processes remain difficult to interpret. Although they produce fluent outputs, models frequently struggle to maintain consistent logic or convey abstract concepts in a structured and transparent manner. As a result, prompt engineering has become central to improving reasoning quality. Methods such as Chain-of-Thought prompting [1] and Tree of Thoughts [2] encourage multi-step inference, but these techniques rely primarily on procedural heuristics rather than principles grounded in human cognition.

To provide a cognitively motivated alternative, we draw on CMT from cognitive linguistics [3, 4]. CMT explains how people understand abstract domains by mapping them onto familiar physical experiences, as in metaphors such as *time is money* or *knowledge is a journey.* These mappings influence reasoning, learning, and communication [5], offering structured patterns through which humans interpret complex ideas.

Embedding this framework into LLM prompting enables models to reason metaphorically—systematically transferring relational structure from concrete *source domains* to abstract *target domains.* Building on prior work in cognitive prompting [6], we propose a prompting paradigm that makes these mappings explicit, shaping model outputs through a psychologically grounded mechanism rather than ad hoc textual cues. This integration connects cognitive linguistics with modern prompt engineering, aiming to produce explanations that are clearer, more coherent, and conceptually aligned with human reasoning.

2 Conceptual Metaphor Theory

CMT, introduced by Lakoff and Johnson [3, 4], treats metaphors not as stylistic ornaments but as core organizing principles of human cognition. In this framework, a *source domain* refers to a concrete, familiar experience—such as money, journeys, or physical movement—while a *target domain* denotes an abstract concept like time, knowledge, or progress. Mappings between these domains transfer relational structure: in the metaphor *"time is money"*, notions of scarcity, expenditure, and value are projected onto temporal experience. Research in cognitive psychology shows that such metaphorical mappings shape reasoning and interpretive judgments [5], influencing how people understand and communicate complex ideas.

A central claim of CMT is that abstract reasoning is grounded in embodied experience. The human mind draws on perceptual, motor, and spatial schemas—such as containment, movement, and balance—to scaffold higher-level thought. This grounding explains why expressions like *rising tension*, *a warm relationship*, or *moving forward with an idea* feel intuitive across cultures: they mirror patterns derived from bodily interaction with the physical world. Metaphors thus function as cognitive compression mechanisms, providing structured, navigable analogies for otherwise abstract conceptual spaces.

In the context of artificial intelligence, metaphors have primarily been used to analyze or classify figurative language. Our work takes a different direction by employing CMT as a *prompting paradigm*: rather than detecting metaphors, the model is guided to reason through explicit source–target mappings. Building on earlier efforts in cognitive prompting [6], this perspective leverages a long-standing theory of conceptual organization to shape the structure and interpretability of LLM reasoning.

3 CMT-Based Prompting

In CMT-based prompting, the mapping principle between concrete and abstract domains is encoded directly into the model's instructions, transforming ordinary text generation into structured, metaphor-driven reasoning. A typical CMT prompt follows three key steps:

1. **Adopt a metaphorical reasoning stance:** The model is instructed to act as a cognitive agent that interprets abstract ideas through metaphors grounded in concrete experience.

2. **Identify source and target domains:** The prompt directs the model to recognize both domains, where the *source* represents a familiar, tangible

experience and the *target* denotes an abstract concept to be understood through this mapping.

3. **Infer meaning through mapping:** Relational properties and dynamics from the source domain are projected onto the target, yielding structured and interpretable reasoning.

The prompt itself functions as a cognitive scaffold, reminding the model that abstract reasoning can be grounded in concrete analogies. An illustrative version of such a CMT-based instruction is shown below:

```
As a cognitive agent utilizing Conceptual Metaphor Theory,
you can interpret abstract concepts (target domains)
through more concrete experiences (source domains).

Source and Target Domains:
- Source Domain: The concrete or physical experience from which we draw metaphorical expressions.
- Target Domain: The abstract concept we aim to understand through the source domain.

Examples:
1. He has a heart of stone:
   - Source Domain: Stone (hard and unfeeling)
   - Target Domain: A person's heart (emotions)
   - Inference: The person is unfeeling, mirroring the hardness of stone.

2. Knowledge is a journey:
   - Source Domain: Journey (paths, obstacles, destinations)
   - Target Domain: Learning or understanding
   - Inference: Learning involves progress toward goals
     and overcoming obstacles to reach insight.
```

This prompt structure embodies the three steps above, domain identification, mapping, and inference, enhancing interpretability and conceptual coherence. Through such explicit metaphorical framing, LLMs can reason more transparently and human-like across explanation, analogy construction, and metaphor interpretation tasks.

Here is a clean rewrite of the entire section, with all forbidden references removed and replaced by wording that stands on its own. Only your allowed citations are used (Lakoff1, Lakoff2, thibodeau2011metaphors, cp, wei2022chain, tree, figlang).

4 CMT Benchmark

To assess the impact of CMT-based prompting, we constructed a benchmark of 100 tasks organized into four categories, each targeting a different facet of metaphor-guided reasoning.

Metaphor Identification and Mapping (MIM). These tasks evaluate the model's ability to uncover and articulate conceptual correspondences in figurative expressions. For example, given

"Her words were a sharp dagger, cutting through his confidence."

a CMT-prompted model identifies the source domain (*dagger*, an instrument that causes harm) and the target domain (*words*, as vehicles of influence). It then infers that speech can inflict emotional damage much like a physical weapon. By making the source–target distinction explicit, the prompt encourages structured and interpretable reasoning.

Domain-Specific Reasoning (DSR). These tasks examine how the model uses metaphorical structure to explain abstract or technical concepts. For instance, when asked,

"How does the immune system protect the body from infections?"

the model may map *immune system* $\rightarrow$ *defense army*: soldiers corresponding to white blood cells, enemies to pathogens, and weapons to antibodies. Such analogies demonstrate how concrete schemas organize complex mechanisms into intuitive explanatory structures.

Explanation and Teaching Tasks (ETT). These tasks measure how well the model communicates specialized knowledge in accessible form, linking to educational studies on metaphor and conceptual transfer in learning [5]. For example,

"How does a neural network learn?"

may yield the analogy *a neural network is like a team of musicians tuning to harmony*, where musicians correspond to neurons, tuning to weight adjustment, and harmony to low error. Training thus becomes the process of achieving collective resonance, a vivid metaphor aiding non-expert understanding.

Reading Comprehension of Metaphors (RCM). These tasks assess interpretive reasoning in extended, literary contexts, comparable to figurative-language understanding benchmarks such as FigLang [7]. For instance,

"Her voice was a lighthouse guiding him through the storm of uncertainty."

the model identifies the source domain (*lighthouse*, a steady beacon) and target domain (*voice*, a source of reassurance), concluding that her voice provides guidance and stability amid confusion. Such explicit mappings turn figurative language into transparent reasoning steps.

5 Experimental Analysis

We conduct an experimental analysis to evaluate how CMT-based prompting influences reasoning quality across different task types and model architectures. Each category comprised 25 tasks, resulting in a balanced benchmark of 100 items. For every task, responses from baseline and CMT-prompted models were evaluated by a large reference model (Llama3-70B) acting as an expert assessor. Evaluations followed three criteria: structural interpretation, coherence of explanation, and mapping accuracy. Each criterion was rated on a five-point scale from 1 (poor) to 5 (excellent), allowing a consistent comparison of reasoning quality across models. This setup enabled a controlled assessment of how explicit metaphorical scaffolding affects reasoning depth, interpretability, and conceptual alignment across diverse task types.

Four models were tested: Llama3.2, Phi3, Gemma2, and Mistral, both with and without CMT prompts. Table 1 reports the mean relative improvement of CMT-prompted performance over the baseline, expressed as percentage change with respect to the base score.

Category	Llama3.2	Phi3	Gemma2	Mistral
MIM	+10.4	+12.2	+7.3	+16.4
DSR	+13.8	+14.0	+20.0	+13.7
ETT	+23.8	+20.3	+19.6	+16.5
RCM	+11.3	+137.1	+4.5	+19.3

Table 1: Relative improvement (%) of CMT-prompted performance compared to baseline across models and categories.

Across all categories, CMT prompting led to consistent gains, typically between 10 and 25 percent. The largest relative improvements appeared in ETT, where explicit analogies helped organize complex ideas into coherent structures. DSR also benefited, confirming that metaphor-based mappings support the explanation of abstract mechanisms through concrete analogies. MIM improved across all models, indicating that CMT prompts strengthen the recognition of underlying conceptual correspondences. RCM showed higher relative variance, with particularly strong proportional gains for Phi3, suggesting that smaller models profit most from explicit cognitive scaffolding.

Llama3.2 and Gemma2 achieved the highest absolute performance levels, while Phi3 showed the strongest relative increase, and Mistral exhibited moderate but consistent gains. These results indicate that the benefit of CMT prompting depends on model size and internal reasoning architecture.

Qualitatively, CMT-prompted responses were clearer and more interpretable. When explaining economic inflation, a baseline model produced a generic description, whereas the CMT-prompted version compared inflation to air filling a balloon, clarifying its gradual and expansive nature. In literary comprehension tasks, the metaphorical framework encouraged models to explain how figurative expressions shape meaning rather than simply restating text, resulting in more transparent and human-like reasoning.

6 Conclusions

This work introduced a CMT-based prompting strategy for LLMs, framing abstract reasoning tasks through concrete source domains and guiding inference via structured metaphorical mappings. Across benchmarks, this approach improved coherence, interpretability, and explanatory quality, indicating that metaphoric scaffolding enhances the cognitive depth and transparency of LLM reasoning. CMT-based prompting therefore bridges the statistical fluency of language models with the structured, analogy-driven reasoning characteristic of human cognition, offering a lightweight but principled method for shaping model behavior.

Future work can extend the benchmark to multimodal and cross-linguistic settings and explore the development of *machine-metaphor thinking*, systems capable of constructing new metaphors autonomously by drawing systematic parallels between domains. Such capabilities may support more flexible reasoning, richer explanation strategies, and ultimately a deeper integration of cognitive science insights into the design of next-generation prompting frameworks.

The open-source implementation of CMT prompting is available here.

References

[1] Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed Chi, Quoc V. Le, and Denny Zhou. Chain-of-thought prompting elicits reasoning in large language models. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 35, pages 24877–24891, 2022.

[2] Shunyu Yao, Dian Yu, Jeffrey Zhao, Ian Shafran, Tom Griffiths, Yuan Cao, and Karthik Narasimhan. Tree of thoughts: Deliberate problem solving with large language models. In *Advances in Neural Information Processing Systems*, volume 36, pages 11809–11822, 2023.

[3] George Lakoff and Mark Johnson. *Metaphors We Live By.* University of Chicago Press, 1980.

[4] George Lakoff. The contemporary theory of metaphor. In *Metaphor and Thought*, pages 202–251. Cambridge University Press, 1993.

[5] Paul H. Thibodeau and Lera Boroditsky. Metaphors we think with: The role of metaphor in reasoning. *PLoS ONE*, 6(2):e16782, 2011.

[6] Oliver Kramer and Jill Baumann. Unlocking structured thinking in language models with cognitive prompting. In *European Symposium on Artificial Neural Networks (ESANN)*, pages 657–662, 2025.

[7] Tuhin Chakrabarty, Smaranda Muresan, and Nanyun Peng. Figlang 2020 shared task: Understanding figurative language. In *Workshop on Figurative Language Processing*, pages 1–11. Association for Computational Linguistics, 2021.