Empowering Self-Learning of LLMs: Inner Knowledge Explicitation as a Catalyst

Shijue Huang[1,4*], **Wanjun Zhong**[2*], **Deng Cai**[2], **Fanqi Wan**[3], **Chengyi Wang**[2], **Mingxuan Wang**[2], **Mu Qiao**[2], **Ruifeng Xu**[1,4,5†]

[1]Harbin Institute of Technology, Shenzhen, China
[2]Bytedance Seed, China
[3] Sun Yat-sen University, China
[4]Guangdong Provincial Key Laboratory of Novel Security Intelligence Technologies
[5]Peng Cheng Laboratory, Shenzhen, China
joehsj310@gmail.com, wanjun@bytedance.com, xuruifeng@hit.edu.cn

Abstract

Self-learning of Large Language Models (LLMs) facilitates their advancement towards super-intelligence by training with self-synthesized experiences. However, a critical challenge is the amplification of hallucinations in generated data during iterative self-learning, underscoring the need for reliable data selection. To address this, we investigate the mechanism of *Inner Knowledge Explicitation*, which involves explicitly extracting the inner knowledge from memory of LLMs, to concurrently improves reasoning, and enables reliable self-learning data selection. This paper introduces a **S**elf **K**nowledge **E**xplicitation **Learn**ing (SKE-Learn) framework, which equips the LLMs with meta-skills to explicitly extract, verify and utilize inner knowledge for reasoning. By leveraging these meta-skills, SKE-Learn establishes a self-learning approach that ensures reliable selection of self-synthetic data. This approach enhances performance through iterative self-learning while mitigating the problem of hallucinations. Empirical results from six benchmarks demonstrate that *Inner Knowledge Explicitation* improves reasoning by serving as a more effective prompting method. Additionally, SKE-Learn, based on the verifiability of explicit knowledge, shows consistent performance improvements over multiple self-training iterations, with an average performance increase from 52.79% to 56.54% across all benchmarks. Furthermore, *Inner Knowledge Explicitation* provides explanation and intervention space during LLM's generation process.

Code — https://github.com/JoeYing1019/SKE-Learn

1 Introduction

Large language models (LLMs) (OpenAI et al. 2024; Dubey et al. 2024) have significantly advanced the field of natural language processing (NLP). However, as LLMs evolve rapidly, the traditional approach of collecting high-quality human annotations for model training struggles to meet the increasing scalability demands. Therefore, self-learning methods for LLMs (Wang et al. 2023b; Tu et al. 2024; Tao et al. 2024) have been proposed, allowing LLMs to learn autonomously from self-synthesized, large-scale training data.

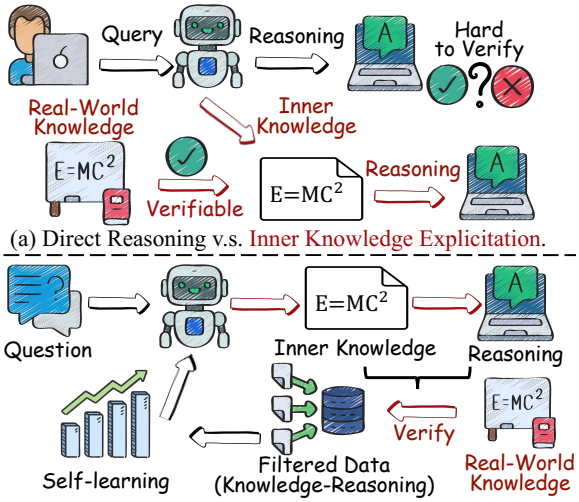


Figure 1: Illustration of (a) Direct Reasoning v.s. Inner Knowledge Explicitation; and (b) Self-learning with Verifiable Inner Knowledge.

This paradigm paves the way for automatic advancement of LLMs towards super-intelligence. Nonetheless, synthesized data often has the issue of hallucinations (Ji et al. 2023), which will be amplified during iterative self-training and leads to model collapse (Shumailov et al. 2024). Hence, reliable selection of self-learning data is of substantial significance to ensure learning robustness. However, previous methods (Huang et al. 2023; Lu et al. 2024) directly synthesize reasoning process, which is question-specific and challenging to verify without human annotations or more advanced models' judgments. In contrast, as knowledge has authorized resource for verification and is general for different questions, its correctness is more easily to be verified as shown in Figure 1 (a).

Inspired by human problem-solving, where relevant knowledge is recalled from memory before reasoning, equipping models with meta-skills to acquire, validate, and utilize their inner knowledge could enhance reasoning and ensure more reliable automatic data selection.

*These authors contributed equally.
†Corresponding author.

Thus, this paper aims to systematically investigate the mechanism of "*Inner Knowledge Explicitation*", focusing on two key questions: (1) Does the explicit extraction of knowledge from the inner memory of LLMs enhance their reasoning abilities? (2) Given that explicit knowledge is inherently verifiable, can it be leveraged to automatically select self-synthetic data for LLMs, thereby establishing a reliable self-learning system? Specifically to *Inner Knowledge Explicitation*, unlike traditional approaches that directly generate answers, we require LLMs to first extract explicit inner knowledge in natural language from their memory and then utilize it for reasoning, as illustrated in Figure 1 (b). The explicit inner knowledge not only improves reasoning performance as a guidance but also provides explanation for the utilized knowledge, which facilitates verification of data quality. Furthermore, we propose a **S**elf **K**nowledge **E**xplicitation **Learn**ing (SKE-Learn) framework, which empowers LLMs with meta-skills to explicitly extract, verify, and utilize inner knowledge for reasoning, and develop a self-learning approach using these skills for automatic reliable data selection. Specifically, SKE-Learn begins by extracting explicit inner knowledge from LLM's memory. This extracted knowledge then enhances followed reasoning process by providing high-level guidance on concepts or principles relevant to the given query.

Additionally, by leveraging the verifiability of explicit inner knowledge, SKE-Learn improves the reliability of self-synthetic data selection, alleviating the issue of hallucination in iterative training and concurrently enhancing model's meta-skills to utilize inner knowledge by establishing a self-learning approach. The self-learning approach in SKE-Learn comprises two stages: (1) Meta-skill Training: This stage focuses on developing model's meta-skills, including the self-extraction of explicit inner knowledge, knowledge-enhanced reasoning, and self-assessment that refers to the ability of self-evaluating explicit inner knowledge and reasoning process. (2) Iterative Self-training: In this stage, the model self-generates "question-knowledge-reasoning" traces for training. These traces are filtered based on model's self-assessment ability, which assesses the truthfulness of reasoning and explicit inner knowledge by comparing it with real-world knowledge as a verifiable reference. This iterative process strengthens the model's knowledge utilization and improves its reasoning capabilities over time.

Extensive experiments demonstrate that while LLMs possess extensive knowledge, explicit extraction of inner knowledge still significantly enhances their reasoning performance. The self-learning approach of SKE-Learn alleviates amplified hallucinations and improves model's reasoning capabilities iteratively, resulting in an average performance increase from 52.79% to 56.54% across six benchmarks. Additionally, the *Inner Knowledge Explicitation* also offers explanations and intervention space in LLM generation. Contributions of this work are as follows:

- We systematically investigate *Inner Knowledge Explicitation* and show that extracting explicit inner knowledge from LLMs' memory enhances reasoning performance.
- We propose a SKE-Learn framework to improve self-learning of LLMs, utilizing verifiable knowledge for reliable data selection, ultimately alleviating amplified hallucinations and improving effectiveness in self-training.
- Comprehensive experiments across six benchmarks reveal that both the *Inner Knowledge Explicitation* mechanism and the SKE-Learn self-learning approach elicit reasoning abilities, and provide better interpretability in model knowledge utilization.

2 Preliminary

In this section, we define the concept of *Inner Knowledge Explicitation* in the scope of this paper. Specifically, *Inner Knowledge Explicitation* refers to the mechanism that LLMs explicitly output their inner knowledge relevant to given queries in natural language. Here, we limit the knowledge scope to the relevant concepts, theories, principles, laws, and factual information regarding domains of given queries.

3 Methodology

In this section, we introduce the meta-skills emphasized in SKE-Learn and the prompting method with *Inner Knowledge Explicitation* (§3.1). Follow-up with a self-learning approach equips LLMs with meta-skills on explicitly extracting, verifying, and utilizing inner knowledge, thereby enabling reliable self-synthetic training data selection and improving model's performance (§3.2). Figure 2 shows the overall workflow.

3.1 Meta-skills and Prompting Method with Inner Knowledge Explicitation

Given a question q and its corresponding reference knowledge $\tilde{k}$, SKE-Learn incorporates three key meta-skills:
(1) **Self-extraction of Inner Knowledge**, which involves extracting relevant explicit inner knowledge k from the model's memory in response to q:

$$k = \mathcal{M}(q, p_{\text{extract}}), \tag{1}$$

(2) **Knowledge-enhanced Reasoning**, referring to the ability generate reasoning r based on the extracted knowledge:

$$r = \mathcal{M}(q, k, p_{\text{reason}}), \tag{2}$$

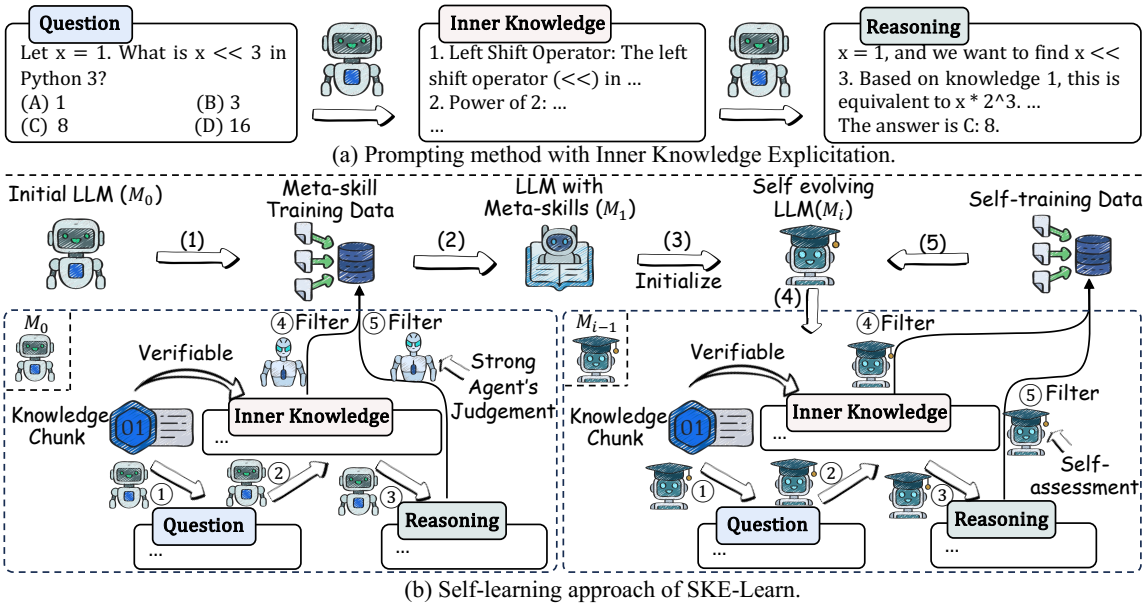
(3) **Self-assessment**, which is the ability to self-evaluate the quality of the reasoning process, and inner knowledge taking authoritative real-world knowledge as a reference:

$$s^k = \mathcal{M}(q, k, \tilde{k}, p^k_{\text{score}}), \tag{3}$$

$$s^r = \mathcal{M}(q, k, r, p^r_{\text{score}}), \tag{4}$$

where $\mathcal{M}$ represents model; s^k and s^r are scores for knowledge and reasoning; p_{extract}, p_{reason}, p^k_{score} and p^r_{score} are related prompts for these meta-skills.

As illustrated in Figure 2 (a), the prompting method with *Inner Knowledge Explicitation* leverages the aforementioned first two meta-skills by extracting inner knowledge and then applying it for reasoning. Similar to evidence-based reasoning (Howick, Glasziou, and Aronson 2010; Gupta et al. 2022), SKE-Learn enables the model to quote explicit inner knowledge as evidence, thereby reducing hallucinations and enhancing overall quality of generated outputs.



(a) Prompting method with Inner Knowledge Explicitation.

(b) Self-learning approach of SKE-Learn.

Figure 2: Illustration of (a) prompting method with *Inner Knowledge Explicitation*, which first extracts inner knowledge then applies it for reasoning; and (b) self-learning approach of SKE-Learn, comprising meta-skill training and iterative self-training.

3.2 Self-learning Approach of SKE-Learn

To further boosting model's reasoning ability and reinforcing meta-skills, we design a self-learning approach. This approach consists of two stages: meta-skill training and iterative self-training as depicted in Figure 2 (b).

Meta-skill Training This stage focuses on referencing meta-skills including self-extraction of explicit inner knowledge, knowledge-enhanced reasoning, and self-assessment. Assuming the availability of a substantial unsupervised knowledge corpus $\mathcal{C}$, we aim to construct several data collections $\mathcal{D}^K_{\text{meta}}$, $\mathcal{D}^R_{\text{meta}}$, and $\mathcal{D}^S_{\text{meta}}$ for different meta-skills, the former two include instances of high-quality "question-knowledge-reasoning" traces (i.e. q_i, k_i, r_i) and the latter one additionally comprises corresponding knowledge chunks and scores for the knowledge and the reasoning (i.e. $c_i, q_i, k_i, r_i, s^k_i, s^r_i$). As illustrated in the left part of Figure 2 (b), our process begins by sampling knowledge chunk c_i from $\mathcal{C}$. The initial model $\mathcal{M}_0$ then generates a question q_i that requires reasoning about c_i, and further generates associated knowledge k_i and reasoning r_i.

Since $\mathcal{M}^0$ may struggle with self-assessment in the initial phase, we draw inspiration from Zheng et al. (2023) and leverage a more advanced model (i.e. GPT-4), $\mathcal{J}$, to evaluate the knowledge k_i and reasoning r_i at this stage. Specifically, $\mathcal{J}$ utilizes the knowledge chunk c_i as a direct and verifiable reference to assess the correctness and reliability of k_i. It determines whether k_i accurately reflects the question's requirements and covers all necessary aspects, ultimately assigning a score s^k_i to k_i. For the reasoning r_i, $\mathcal{J}$ evaluates its factual accuracy, logical coherence, and application of the provided explicit knowledge, producing a score s^r_i for r_i. To mitigate the risk of knowledge distillation from used advanced model at this stage, we have $\mathcal{M}^0$ synthesize all the knowledge and reasoning traces, using the more advanced model $\mathcal{J}$ solely for correctness evaluation and filtering. Thus, the knowledge and reasoning are originally acquired by $\mathcal{M}^0$, the only injected capability is critical scoring.

Subsequently, we select the high-quality sets of k_i and r_i based on pre-defined threshold to construct the meta-skill training data $\mathcal{D}^K_{\text{meta}}$ for self-extraction of inner knowledge and $\mathcal{D}^R_{\text{meta}}$ for knowledge-enhanced reasoning, respectively. For the meta-skills of self-assessment, a small portion of the data $\mathcal{D}^S_{\text{meta}}$ is uniformly sampled from both s^k_i and s^r_i for training. To further enhance the diversity of the questions, we also incorporate a small number of existing questions during this process to construct the knowledge and reasoning process in a similar manner. Utilizing this meta-skill training data, we then develop a new model $\mathcal{M}^1_{\text{meta}}$ with followed training objective:

$$\begin{aligned}\mathcal{L}_{\text{meta}} = &-\mathbb{E}_{(q_i,k_i)\sim\mathcal{D}^K_{\text{meta}}}\log(k_i|q_i)\\ &-\mathbb{E}_{(q_i,k_i,r_i)\sim\mathcal{D}^R_{\text{meta}}}\log(r_i|q_i,k_i)\\ &-\mathbb{E}_{(c_i,q_i,k_i,r_i,s^k_i,s^r_i)\sim\mathcal{D}^S_{\text{meta}}}[\log(s^k_i|q_i,k_i,c_i)+\\ &\log(s^r_i|q_i,k_i,r_i)] \end{aligned} \tag{5}$$

Iterative Self-training. Since $\mathcal{M}^1_{\text{meta}}$ possesses better meta-skills in self-extraction of inner knowledge, knowledge-enhanced reasoning, and self-assessment, it can engage in iterative self-training through self-synthesizing data and selecting high-quality ones by self-assessment capability. This process further strengthens these meta-skills and enhances model's overall reasoning performance.

Given an unsupervised knowledge chunk c_j, we use $\mathcal{M}^1_{\text{meta}}$ to self-synthesize "question-knowledge-reasoning" traces, similar to the meta-skill training phase, thereby obtaining question q_j, knowledge k_j and reasoning process r_j.

Unlike the meta-skill phase that uses a stronger model for scoring, here we leverage the self-assessment capability of $\mathcal{M}^1_{\text{meta}}$ to provide scores for k_j and r_j, denoted as s^k_j and s^r_j, respectively. Then high-quality k_j and r_j are selected to construct a new set of self-evolving training data $\mathcal{D}^K_{\text{evol}}$ and $\mathcal{D}^R_{\text{evol}}$, which are used to train a more proficient model, $\mathcal{M}^2$. The training objective of this stage is:

$$\begin{aligned}\mathcal{L}_{\text{evol}} = &-\mathbb{E}_{(q_j,k_j)\sim D^K_{\text{evol}}} \log(k_j|q_j) \\ &- \mathbb{E}_{(q_j,k_j,r_j)\sim D^R_{\text{evol}}} \log(r_j|q_j,k_j) \end{aligned} \tag{6}$$

During this stage, we iteratively obtain models $\mathcal{M}^i$ trained on data synthesized by $\mathcal{M}^{i-1}$. The self-extracted knowledge of $\mathcal{M}^{i-1}$ is validated by its self-assessment ability with ready-made authoritative references from real-world knowledge chunks. This validation alleviates hallucinated data being used for self-learning by establishing a more reliable data selection process. Moreover, since all data originates from the model itself, this process ensures both reliability and automation. As a result, at each round, the model can progressively refine its knowledge boundaries and meta-skills from experiences of the previous round's model, leading to iterative improvements in performance.

4 Experiments

In this section, we verify whether *Inner Knowledge Explicitation* can elicit reasoning and facilitate self-learning (§4.2), whether knowledge injection occurs in self-learning (§4.3), and provide detailed analyses about the meta-skills evolution, interpretability, intervention space and hallucination alleviation brought by *Inner Knowledge Explicitation* (§4.4).

4.1 Experimental Setup

Implementation Details. We fine-tune Llama3-8B (Dubey et al. 2024) on 100,000 instances of Magpie data[1] (Xu et al. 2024b), and derived an instruct model as the backbone model of whole experiments, namely Llama3-8B-Magpie. All models are trained with full parameters for 2 epochs, using batch size of 32, learning rate of 2e-5, with 100 warmup steps. In inference phase, we set all temperatures as 0 to ensure better reproducibility. All experiments are performed on eight NVIDIA A100-SXM4-80GB GPUs.

Comprising meta-skill training, we conduct totally four rounds of iterative training. Following Yuan et al. (2024), we mix a proportion of general data[2] and the scoring data from the meta-skill training stage at each round to maintain both general responsiveness and self-assessment capability. Training data details for each round are provided in Table 1.

The unsupervised knowledge corpora are sourced from a collection based on Wikipedia[3] , and the small proportion of existing questions leveraged in this stage are drawn from a dataset focused on STEM[4]. In meta-skill training, we use GPT-4 with version `gpt-4-0125-preview` as judge model. The scores for knowledge and reasoning process both range from 0 to 10, with higher scores indicating better quality. The score threshold for selecting data is 8.

Model	Knowledge	Reasoning	Scoring	General
$\mathcal{M}^0 \rightarrow \mathcal{M}^1_{\text{meta}}$	5,000	5,000	4,237	25,000
$\mathcal{M}^1_{\text{meta}} \rightarrow \mathcal{M}^2$	100,35	10,135	42,37	25,000
$\mathcal{M}^2 \rightarrow \mathcal{M}^3$	23,503	22,450	4,237	25,000
$\mathcal{M}^3 \rightarrow \mathcal{M}^4$	38,069	42,672	4,237	25,000

Table 1: Training data details in our self-learning approach.

Benchmarks and Evaluation Metrics. We conducted experiments across a wide range of popular benchmarks, including general examination benchmarks such as MMLU (Hendrycks et al. 2021), AGIEval (Zhong et al. 2024), and ARC (encompassing ARC-E and ARC-C) (Clark et al. 2018). Additionally, we evaluated on the comprehensive reasoning benchmark BBH (Suzgun et al. 2023) and the knowledge question-answering benchmark Natural Questions (NQ)[5] (Kwiatkowski et al. 2019). All evaluation metrics are accuracy, with corresponding evaluation scripts derived from OpenCompass (Contributors 2023).

Baselines. To verify the performance of prompting method with *Inner Knowledge Explicitation*, some zero-shot **Prompting-based Approaches** utilizing LLM's inner knowledge implicitly are compared, including: (1) Zero-shot (Brown et al. 2020) generates answers directly. (2) Chain-of-Thought (Wei et al. 2022) generates step-by-step thoughts. (3) Plan-and-Solve (Wang et al. 2023a) conducts planning before reasoning. (4) Rephrase-and-Respond (Deng et al. 2023) instructs LLMs to rephrase the question. (5) System-2-Attention (Weston and Sukhbaatar 2023) instructs LLMs to remove irrelevant information from prompts. (6) Thread-of-Thought (Zhou et al. 2023) features an enhanced thought inducer. (7) Analogical (Zhou et al. 2023) automatically generates exemplars. (8) Re-Reading (Xu et al. 2024a) repeats the question twice.

For the **Self-learning Approach**, we introduce a self-learning baseline denoted as $\tilde{\mathcal{M}}$, which is trained by typical self-learning manner that self-synthesizes Chain-of-Thought format "question-reasoning" trace and selects high-quality data by solely self-rewarding on reasoning process (Yuan et al. 2024), with an equivalent amount of data created from all knowledge corpora used in SKE-Learn.

4.2 Main Result

The main results are demonstrated in Table 2, we can observe that:

Inner Knowledge Explicitation Elicits Reasoning Ability. Compared to existing prompting methods that utilize knowledge implicitly, SKE-Learn demonstrates superior performance across all benchmarks, outperforming the most competitive baseline (i.e., Thread-of-Thought) by an average of

[1]The first 100,000 entries from https://huggingface.co/datasets/Magpie-Align/Magpie-Pro-300K-Filtered.

[2]Sampled 25,000 instances from aforementioned Magpie data.

[3]https://huggingface.co/datasets/fmars/wiki_stem.

[4]https://huggingface.co/datasets/cfahlgren1/swti-stem-20k.

[5]We use the standard open-domain splits as per previous studies (Lewis, Stenetorp, and Riedel 2021; Wang et al. 2024a).

Methods	MMLU	BBH	ARC-E	ARC-C	AGIEval	NQ	Average
Prompting-based Approaches							
Zero-shot (Brown et al. 2020)	54.81	16.82	83.46	69.28	45.93	24.04	49.06
Chain-of-Thought (Wei et al. 2022)	48.06	24.60	68.77	60.24	44.99	17.76	44.07
Plan-and-Solve (Wang et al. 2023a)	50.01	24.67	67.05	55.21	45.67	12.00	42.44
Rephrase-and-Respond (Deng et al. 2023)	46.14	22.46	71.38	61.69	47.97	21.94	45.26
System-2-Attention (Weston and Sukhbaatar 2023)	52.34	14.35	72.90	63.40	46.60	17.04	44.44
Thread-of-Thought (Zhou et al. 2023)	56.32	25.54	87.5	71.50	44.04	23.24	51.36
Analogical (Yasunaga et al. 2024)	54.32	23.54	73.74	61.95	42.29	27.26	47.18
Re-reading (Xu et al. 2024a)	52.19	**29.90**	81.82	70.73	**48.35**	21.16	50.69
SKE-Learn ($\mathcal{M}^0$)	**57.21**	27.56	**85.73**	**71.93**	44.39	**29.89**	**52.79**
Self-learning Approaches							
Self-learning Baseline ($\tilde{\mathcal{M}}$)	54.78	25.19	82.74	71.42	45.86	19.28	49.88
SKE-Learn ($\mathcal{M}^1_{\text{meta}}$)	57.94	32.49	85.94	73.46	44.85	31.44	54.35
SKE-Learn ($\mathcal{M}^2$)	58.90	34.17	87.21	75.85	46.42	31.58	55.69
SKE-Learn ($\mathcal{M}^3$)	**59.45**	35.16	87.67	76.19	46.18	31.52	56.03
SKE-Learn ($\mathcal{M}^4$)	58.84	**35.25**	**89.60**	**76.37**	**47.12**	**32.05**	**56.54**
Ablation Study							
SKE-Learn ($\mathcal{M}_{\text{S}}$)	56.70	25.57	80.68	70.82	45.90	28.73	51.40
SKE-Learn ($\mathcal{M}_{\text{G}}$)	57.42	28.31	86.95	74.23	45.62	30.69	53.87
SKE-Learn ($\mathcal{M}_{\text{S+G}}$)	57.06	26.80	86.15	73.38	45.65	29.00	53.01
SKE-Learn ($\mathcal{M}_{\text{CPT}}$)	56.97	28.03	86.62	71.16	44.64	30.75	53.03
SKE-Learn ($\mathcal{M}_{\text{CPT+S+G}}$)	56.70	26.33	80.72	65.02	43.34	24.93	49.51

Table 2: Main results across various benchmarks. The Average score represents the mean performance across all benchmarks. Scores that are **bolded** indicate the highest performance among same setting.

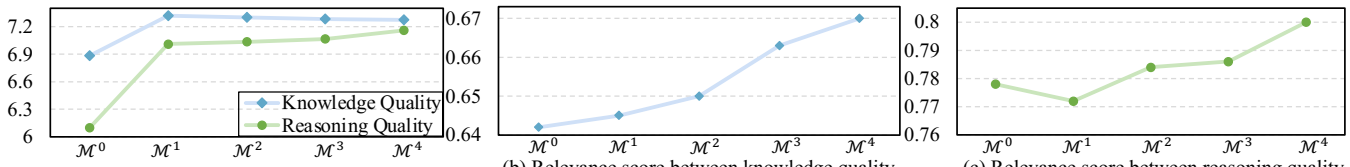
1.43%. We attribute this improvement to the role of *Inner Knowledge Explicitation*, which enables LLMs to accurately recall the necessary underlying concepts and principles. Explicitly generating these in natural language further strengthens the understanding of this information, thereby enhancing the subsequent reasoning process. Additionally, we observe that the Zero-shot approach outperforms several prompting techniques, such as Chain-of-Thought, in certain tasks. We hypothesize that this is due to the training data of Llama3-8B-Magpie, which primarily consists of long answers with detailed analyses. As a result, the Zero-shot outputs, which also include detailed explanations, resemble a variant of Chain-of-Thought prompting. This similarity may make the Zero-shot approach even more effective, given its closer alignment with the training data format.

Self-learning Framework Iteratively Improves Performance. Compared to the self-learning baseline $\tilde{\mathcal{M}}$, which was trained on a large set of QA pairs, the meta-skill training model ($\mathcal{M}^1$) demonstrates a 4.47% improvement in average performance while utilizing 90% less training data. Subsequent iterative training yields even greater performance gains. The SKE-Learn self-learning framework shows continuous improvement across all benchmarks when compared to the non-trained model ($\mathcal{M}^0$), with consistent performance enhancements throughout all training rounds. Specifically, the meta-skill training stage ($\mathcal{M}^1_{\text{meta}}$) leads to an average score increase of 1.56%, while the fourth training iteration ($\mathcal{M}^4$) achieves an average improvement of 3.75% compared with $\mathcal{M}^0$. Notably, results on certain tasks, such as BBH, exhibit a substantial enhancement of 7.69% after completing the entire self-learning process. These findings suggest that self-learning, based on *Inner Knowledge Explicitation*, effectively enhances the model's reasoning performance. Furthermore, the incorporation of self-assessment and the verifiability of explicit knowledge ensures a more reliable data selection process, thereby further boosting the performance of self-learning.

4.3 Ablation Study

We conduct a comprehensive ablation study to verify to what extent the potential knowledge injection affects final performance of self-learning. Our analysis focuses on the effects of (1) newly sampled general data, (2) externally sourced scoring data from GPT-4, and (3) referenced unsupervised knowledge corpus. Specifically, we use our backbone model, Llama3-8B-Magpie, to establish baselines: (1) $\mathcal{M}_{\text{S}}$, fine-tuned on scoring data; (2) $\mathcal{M}_{\text{G}}$, fine-tuned on sampled general data; (3) $\mathcal{M}_{\text{S+G}}$, fine-tuned on both scoring and sampled general data; (4) $\mathcal{M}_{\text{CPT}}$, continually pre-trained on Llama3-8B using 103,096 chunks of knowledge chunks accumulated throughout self-learning process; and (5) $\mathcal{M}_{\text{CPT+S+G}}$, which evaluates combined effect of all data by fine-tuning $\mathcal{M}_{\text{CPT}}$ with both scoring and sampled general data.

The ablation results, presented in the *Ablation Study* part of Table 2, reveal that incorporating these three types of external data leads to only marginal improvements. Nonetheless, all baseline models that trained on these data still perform worse than our first round training model ($\mathcal{M}^1_{\text{meta}}$). This suggests that the observed performance gains are not due to knowledge injection, but rather stem from the development of meta-skills in acquiring, verifying, and utilizing explicit inner knowledge, which are progressively reinforced through a high-quality self-learning process.



(a) GPT-4 quality score of knowledge and reasoning.

(b) Relevance score between knowledge quality and final performance.

(c) Relevance score between reasoning quality and final performance.

Figure 3: (a) GPT4 quality scores of knowledge and reasoning during iteration; (b) and (c) Relevance scores between knowledge/reasoning quality and final performance.

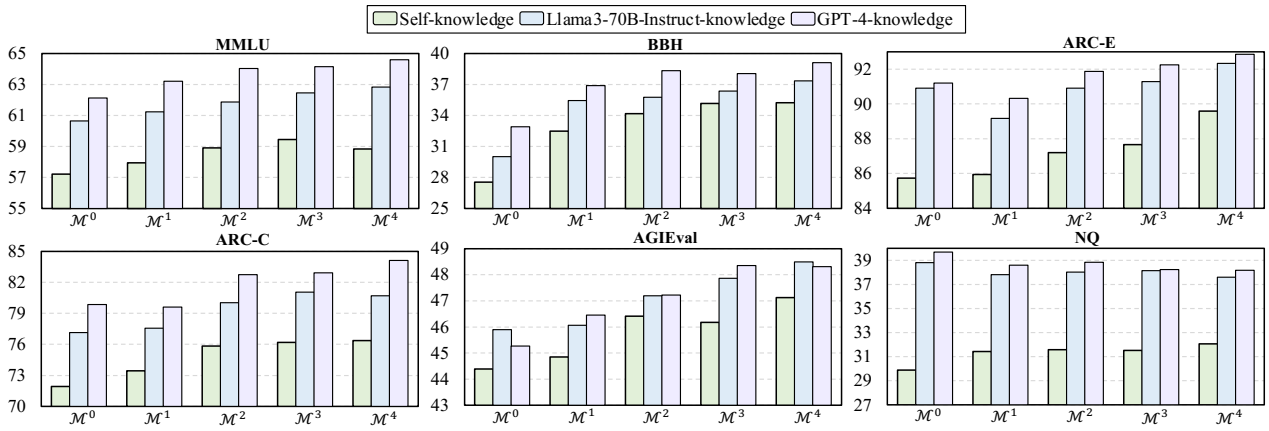


Figure 4: Results of knowledge from our model self, Llama3-70B-Instruct and GPT-4.

4.4 Analysis

To deepen understanding of *Inner Knowledge Explicitation* and SKE-Learn, we answer the following questions: (1) How do meta-skills evolve throughout iterative self-training? (2) Can explicit knowledge serve as an explanation of final performance? (3) Can explicit knowledge offer opportunities for intervening in generation process? (4) How does our self-learning approach alleviate hallucination?

Answer 1: Meta-Skills Improve Through Self-Learning. To investigate the development of meta-skills during iterative self-learning, we randomly select 500 data points from six benchmarks, and collect the generated knowledge and reasoning process at each iteration, resulting in a total of 15,000 data points (3,000 per iteration). Then these data are evaluated for quality using GPT-4 via LLM-as-a-Judge prompting (Zheng et al. 2023), in alignment with the meta-skill training stage. As shown in Figure 3 (a), self-learning significantly enhances both knowledge and reasoning quality scores. Notably, the knowledge quality score stabilizes after the meta-skill training, indicating that the model has effectively learned to extract accurate knowledge during this phase. As a result, subsequent training using self-synthesized data, which does not incorporate external knowledge, will not further enhance knowledge quality significantly. However, the reasoning quality score continues to improve, suggesting that the model's capacity to utilize inner knowledge is progressively refined through self-learning.

Answer 2: Inner Knowledge Explicitation Brings Interpretability. Intuitively, if a model can accurately extract and apply its inner knowledge to perform knowledge-enhanced reasoning, its final performance should be closely related to the quality of inner knowledge and reasoning. To assess whether *Inner Knowledge Explicitation* contributes to interpretability, we evaluate the relevance between the quality of the knowledge/reasoning and the model's final performance using the aforementioned GPT-4 score data. Specifically, we calculate the following relevance scores:

$$\mathcal{R}_k = \frac{\sum_{i=1}^{n} \mathbb{1}[I_{(r_i)} = 1, s_i^k >= 8] + [I_{(r_i)} = 0, s_i^k < 8]}{\sum_{i=1}^{n} 1}, \quad (7)$$

$$\mathcal{R}_r = \frac{\sum_{i=1}^{n} \mathbb{1}[I_{(r_i)} = 1, s_i^r >= 8] + [I_{(r_i)} = 0, s_i^r < 8]}{\sum_{i=1}^{n} 1}, \quad (8)$$

where $\mathcal{R}_k$ and $\mathcal{R}_r$ are the relevance scores between the quality of knowledge/reasoning and final performance; $I_{(r_i)}$ represents the final result of correct (1) or wrong (0); $\mathbb{1}[x]$ is 1 if its required condition x is satisfied, otherwise 0. As shown in Figure 3 (b), the relevance between knowledge and final performance suggests that explicit knowledge provides a partial explanation for the correctness or error in model response. Moreover, this relevance score improves with iterative training, indicating that the interpretability of explicit knowledge enhances over time. Additionally, Figure 3 (c) demonstrates that the relevance between knowledge-enhanced reasoning and final performance also strengthens during self-learning. We attribute this to the reason that the process of quoting explicit knowledge in knowledge-enhanced reasoning will make model's reasoning more trustworthy.

Answer 3: Inner Knowledge Explicitation Facilitates Interventions in the Generation Process. Enhancement of meta-skills on knowledge utilization enables the model to

Question: How many three-digit positive integers are there?	(A) 899	(B) 900	(C) 1000	(D) 999

Knowledge and reasoning from $\mathcal{M}^0$:
Knowledge: (1) A three-digit positive integer is an integer between 100 and 999, inclusive. (2) The number of three-digit positive integers is equal to the number of ways to choose 3 digits from 9 digits (0-9, excluding 0 as the first digit) and arrange them in a specific order.
Reasoning: Based on knowledge 1, a three-digit positive integer is an integer between 100 and 999, inclusive. Based on knowledge 2, we need to choose 3 digits from 9 digits and arrange them in a specific order. The number of ways to choose 3 digits from 9 digits is equal to $9 \times 9 \times 9 = 729$. The number of ways to arrange 3 digits in a specific order is equal to $3! = 6$. Therefore, the total number of three-digit positive integers is therefore $729 \times 6 = 4356$. The answer is therefore D: 999.

Knowledge and reasoning from $\mathcal{M}^4$:
Knowledge: (1) A three-digit positive integer is an integer between 100 and 999, inclusive. (2) The number of integers in a given range can be calculated by subtracting the smallest value from the largest value, plus 1.
Reasoning: Based on knowledge 1, a three-digit positive integer is an integer between 100 and 999, inclusive. Based on knowledge 2, the number of three-digit positive integers can be calculated by subtracting the smallest three-digit integer (100) from the largest three-digit integer (999), plus 1. The answer is B. 900.

Table 3: A case that completes by models before and after self-learning.

better adjust their outputs by giving modified correct knowledge. Thus, this improves the chance of knowledge intervention. In this study, we simulate this intervention by employing more stronger LLMs to generate the corresponding knowledge. Specifically, we use Llama3-70B-Instruct and GPT-4 for this purpose, integrating their knowledge into our model. As illustrated in Figure 4, results indicate that knowledge derived from more powerful models improves performance. This demonstrates that explicit knowledge can act as an interface for modifying the model's reasoning process. By providing diverse knowledge inputs, the model's reasoning can be adjusted accordingly. Additionally, iterative training shows that results based on knowledge from different models all consistently improve, suggesting that the meta-skills of knowledge-enhanced reasoning are progressively enhanced through iterative self-learning.

Answer 4: Inner Knowledge Explicitation alleviates amplified hallucinations. To investigate whether our self-learning approach mitigates amplified hallucinations, we present a case study illustrated in Table 3. Given the same question, the model before self-learning ($\mathcal{M}^0$) incorrectly hallucinates a flawed method for calculating the number of three-digit positive integers, leading to erroneous reasoning. However, after iterative self-learning, the refined model $\mathcal{M}^4$ generates a correct calculation method and effectively applies this knowledge to arrive at the accurate answer. This demonstrates that our self-learning approach successfully leverages the verifiable nature of knowledge for reliable data selection, thereby reducing amplified hallucinations.

5 Related Work

Prompting-based methods for LLMs. Zero-shot prompting has been extensively investigated to enhance reasoning abilities in LLMs. Some strategies focus on improving LLMs' comprehension of the given query (Deng et al. 2023; Xu et al. 2024a). Other approaches explicitly generate the reasoning thoughts within LLMs (Wei et al. 2022; Zhou et al. 2023). Additionally, certain methods assign specific roles to the LLMs within the prompt (Zheng, Pei, and Jurgens 2023; Wang et al. 2024b). In contrast to these techniques, which leverage the inner knowledge of LLMs implicitly, our approach focuses on the explicit extraction of this inner knowledge from LLMs' memory, and find that this approach can enhance model's reasoning performance.

Self-learning methods for LLMs. Self-learning methods that enable LLMs to autonomously learn from self-generated experiences are rapidly advancing, with a crucial aspect of the selection of high-quality data. Some studies rely on external metrics selection (Singh et al. 2024; Qiao et al. 2024; Ulmer et al. 2024), while others use internal consistency measures or model-inherent criteria to achieve similar goals (Huang et al. 2023; Yuan et al. 2024; Lu et al. 2024). Different from these approaches, our work takes a distinct approach by utilizing the verifiability of explicit knowledge to improve data quality verification, alleviating amplified hallucinations.

Knowledge-enhanced methods for LLMs. Knowledge-enhanced approaches have been extensively explored to improve the capabilities of LLMs. Some studies focus on retrieving knowledge from external sources using Retrieval-Augmented Generation (RAG) techniques to enhance LLM performance (Li, Yuan, and Zhang 2024; Fatehkia, Lucas, and Chawla 2024; Wiratunga et al. 2024). Other research has demonstrated that incorporating external knowledge from more advanced LLMs can improve commonsense reasoning (Liu et al. 2022; Fu et al. 2023). In contrast to these methods that rely on external knowledge, SKE-Learn leverages self-extracted inner knowledge from the LLM's own memory to enhance reasoning performance.

6 Conclusion

In this paper, we investigate the *Inner Knowledge Explicitation* mechanism, which explicitly extracts inner knowledge from the memory of LLMs. To this end, we propose a **S**elf **K**nowledge **E**xplicitation **Learn**ing (SKE-Learn) framework that enhances LLMs' meta-skills to explicitly extract, verify and utilize inner knowledge for reasoning. Leveraging the verifiability of explicit knowledge, SKE-Learn establishes a self-learning approach that ensures the reliable selection of self-synthetic data, thereby mitigating amplified hallucinations and enhancing the effectiveness of self-training. Experimental results across six benchmarks demonstrate that *Inner Knowledge Explicitation* elicits reasoning capabilities of LLMs. SKE-Learn allows LLMs to iteratively self-improve, achieving an average performance increase from 52.79% to 56.54% across all benchmarks. Moreover, the explicit knowledge provides explanation and intervention space during LLM's generation process.

Acknowledgments

This work was partially supported by the National Natural Science Foundation of China 62176076, Natural Science Foundation of GuangDong 2023A1515012922, the Shenzhen Foundational Research Funding JCYJ20220818102415032, the Major Key Project of PCL2021A06, Guangdong Provincial Key Laboratory of Novel Security Intelligence Technologies 2022B1212010005.

References

Brown, T. B.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell, A.; Agarwal, S.; Herbert-Voss, A.; Krueger, G.; Henighan, T.; Child, R.; Ramesh, A.; Ziegler, D. M.; Wu, J.; Winter, C.; Hesse, C.; Chen, M.; Sigler, E.; Litwin, M.; Gray, S.; Chess, B.; Clark, J.; Berner, C.; McCandlish, S.; Radford, A.; Sutskever, I.; and Amodei, D. 2020. Language models are few-shot learners. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*, NIPS '20. Red Hook, NY, USA: Curran Associates Inc. ISBN 9781713829546.

Clark, P.; Cowhey, I.; Etzioni, O.; Khot, T.; Sabharwal, A.; Schoenick, C.; and Tafjord, O. 2018. Think you have Solved Question Answering? Try ARC, the AI2 Reasoning Challenge. arXiv:1803.05457.

Contributors, O. 2023. OpenCompass: A Universal Evaluation Platform for Foundation Models. https://github.com/open-compass/opencompass.

Deng, Y.; Zhang, W.; Chen, Z.; and Gu, Q. 2023. Rephrase and Respond: Let Large Language Models Ask Better Questions for Themselves. arXiv:2311.04205.

Dubey, A.; Jauhri, A.; Pandey, A.; and *et al.* 2024. The Llama 3 Herd of Models. arXiv:2407.21783.

Fatehkia, M.; Lucas, J. K.; and Chawla, S. 2024. T-RAG: Lessons from the LLM Trenches. arXiv:2402.07483.

Fu, Y.; Peng, H.; Ou, L.; Sabharwal, A.; and Khot, T. 2023. Specializing Smaller Language Models towards Multi-Step Reasoning. arXiv:2301.12726.

Gupta, V.; Bhat, R. A.; Ghosal, A.; Shrivastava, M.; Singh, M.; and Srikumar, V. 2022. Is My Model Using the Right Evidence? Systematic Probes for Examining Evidence-Based Tabular Reasoning. *Transactions of the Association for Computational Linguistics*, 10: 659–679.

Hendrycks, D.; Burns, C.; Basart, S.; Zou, A.; Mazeika, M.; Song, D.; and Steinhardt, J. 2021. Measuring Massive Multitask Language Understanding. *Proceedings of the International Conference on Learning Representations (ICLR)*.

Howick, J.; Glasziou, P.; and Aronson, J. 2010. Evidence-based mechanistic reasoning. *Journal of the Royal Society of Medicine*, 103: 433–41.

Huang, J.; Gu, S.; Hou, L.; Wu, Y.; Wang, X.; Yu, H.; and Han, J. 2023. Large Language Models Can Self-Improve. In Bouamor, H.; Pino, J.; and Bali, K., eds., *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, 1051–1068. Singapore: Association for Computational Linguistics.

Ji, Z.; Lee, N.; Frieske, R.; Yu, T.; Su, D.; Xu, Y.; Ishii, E.; Bang, Y. J.; Madotto, A.; and Fung, P. 2023. Survey of Hallucination in Natural Language Generation. *ACM Comput. Surv.*, 55(12).

Kwiatkowski, T.; Palomaki, J.; Redfield, O.; Collins, M.; Parikh, A.; Alberti, C.; Epstein, D.; Polosukhin, I.; Devlin, J.; Lee, K.; Toutanova, K.; Jones, L.; Kelcey, M.; Chang, M.-W.; Dai, A. M.; Uszkoreit, J.; Le, Q.; and Petrov, S. 2019. Natural Questions: A Benchmark for Question Answering Research. *Transactions of the Association for Computational Linguistics*, 7: 452–466.

Lewis, P.; Stenetorp, P.; and Riedel, S. 2021. Question and Answer Test-Train Overlap in Open-Domain Question Answering Datasets. In Merlo, P.; Tiedemann, J.; and Tsarfaty, R., eds., *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, 1000–1008. Online: Association for Computational Linguistics.

Li, J.; Yuan, Y.; and Zhang, Z. 2024. Enhancing LLM Factual Accuracy with RAG to Counter Hallucinations: A Case Study on Domain-Specific Queries in Private Knowledge-Bases. arXiv:2403.10446.

Liu, J.; Liu, A.; Lu, X.; Welleck, S.; West, P.; Le Bras, R.; Choi, Y.; and Hajishirzi, H. 2022. Generated Knowledge Prompting for Commonsense Reasoning. In Muresan, S.; Nakov, P.; and Villavicencio, A., eds., *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 3154–3169. Dublin, Ireland: Association for Computational Linguistics.

Lu, J.; Zhong, W.; Huang, W.; Wang, Y.; Zhu, Q.; Mi, F.; Wang, B.; Wang, W.; Zeng, X.; Shang, L.; Jiang, X.; and Liu, Q. 2024. SELF: Self-Evolution with Language Feedback. arXiv:2310.00533.

OpenAI; Achiam, J.; Adler, S.; and *et al.* 2024. GPT-4 Technical Report. arXiv:2303.08774.

Qiao, S.; Zhang, N.; Fang, R.; Luo, Y.; Zhou, W.; Jiang, Y.; Lv, C.; and Chen, H. 2024. AutoAct: Automatic Agent Learning from Scratch for QA via Self-Planning. In Ku, L.-W.; Martins, A.; and Srikumar, V., eds., *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 3003–3021. Bangkok, Thailand: Association for Computational Linguistics.

Shumailov, I.; Shumaylov, Z.; Zhao, Y.; Papernot, N.; Anderson, R.; and Gal, Y. 2024. AI models collapse when trained on recursively generated data. *Nature*, 631: 755 – 759.

Singh, A.; Co-Reyes, J. D.; Agarwal, R.; Anand, A.; Patil, P.; Garcia, X.; Liu, P. J.; Harrison, J.; Lee, J.; Xu, K.; Parisi, A.; Kumar, A.; Alemi, A.; Rizkowsky, A.; Nova, A.; Adlam, B.; Bohnet, B.; Elsayed, G.; Sedghi, H.; Mordatch, I.; Simpson, I.; Gur, I.; Snoek, J.; Pennington, J.; Hron, J.; Kenealy, K.; Swersky, K.; Mahajan, K.; Culp, L.; Xiao, L.; Bileschi, M. L.; Constant, N.; Novak, R.; Liu, R.; Warkentin, T.; Qian, Y.; Bansal, Y.; Dyer, E.; Neyshabur, B.; Sohl-Dickstein, J.; and Fiedel, N. 2024. Beyond Human Data: Scaling Self-Training for Problem-Solving with Language Models. arXiv:2312.06585.

Suzgun, M.; Scales, N.; Schärli, N.; Gehrmann, S.; Tay, Y.; Chung, H. W.; Chowdhery, A.; Le, Q.; Chi, E.; Zhou, D.; and Wei, J. 2023. Challenging BIG-Bench Tasks and Whether Chain-of-Thought Can Solve Them. In Rogers, A.; Boyd-Graber, J.; and Okazaki, N., eds., *Findings of the Association for Computational Linguistics: ACL 2023*, 13003–13051. Toronto, Canada: Association for Computational Linguistics.

Tao, Z.; Lin, T.-E.; Chen, X.; Li, H.; Wu, Y.; Li, Y.; Jin, Z.; Huang, F.; Tao, D.; and Zhou, J. 2024. A Survey on Self-Evolution of Large Language Models. arXiv:2404.14387.

Tu, T.; Palepu, A.; Schaekermann, M.; Saab, K.; Freyberg, J.; Tanno, R.; Wang, A.; Li, B.; Amin, M.; Tomasev, N.; Azizi, S.; Singhal, K.; Cheng, Y.; Hou, L.; Webson, A.; Kulkarni, K.; Mahdavi, S. S.; Semturs, C.; Gottweis, J.; Barral, J.; Chou, K.; Corrado, G. S.; Matias, Y.; Karthikesalingam, A.; and Natarajan, V. 2024. Towards Conversational Diagnostic AI. arXiv:2401.05654.

Ulmer, D.; Mansimov, E.; Lin, K.; Sun, L.; Gao, X.; and Zhang, Y. 2024. Bootstrapping LLM-based Task-Oriented Dialogue Agents via Self-Talk. In Ku, L.-W.; Martins, A.; and Srikumar, V., eds., *Findings of the Association for Computational Linguistics ACL 2024*, 9500–9522. Bangkok, Thailand and virtual meeting: Association for Computational Linguistics.

Wang, C.; Cheng, S.; Guo, Q.; Yue, Y.; Ding, B.; Xu, Z.; Wang, Y.; Hu, X.; Zhang, Z.; and Zhang, Y. 2024a. Evaluating open-QA evaluation. In *Proceedings of the 37th International Conference on Neural Information Processing Systems*, NIPS '23. Red Hook, NY, USA: Curran Associates Inc.

Wang, L.; Xu, W.; Lan, Y.; Hu, Z.; Lan, Y.; Lee, R. K.-W.; and Lim, E.-P. 2023a. Plan-and-Solve Prompting: Improving Zero-Shot Chain-of-Thought Reasoning by Large Language Models. In Rogers, A.; Boyd-Graber, J.; and Okazaki, N., eds., *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2609–2634. Toronto, Canada: Association for Computational Linguistics.

Wang, Y.; Kordi, Y.; Mishra, S.; Liu, A.; Smith, N. A.; Khashabi, D.; and Hajishirzi, H. 2023b. Self-Instruct: Aligning Language Models with Self-Generated Instructions. In Rogers, A.; Boyd-Graber, J.; and Okazaki, N., eds., *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 13484–13508. Toronto, Canada: Association for Computational Linguistics.

Wang, Z.; Mao, S.; Wu, W.; Ge, T.; Wei, F.; and Ji, H. 2024b. Unleashing the Emergent Cognitive Synergy in Large Language Models: A Task-Solving Agent through Multi-Persona Self-Collaboration. In Duh, K.; Gomez, H.; and Bethard, S., eds., *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, 257–279. Mexico City, Mexico: Association for Computational Linguistics.

Wei, J.; Wang, X.; Schuurmans, D.; Bosma, M.; Ichter, B.; Xia, F.; Chi, E. H.; Le, Q. V.; and Zhou, D. 2022. Chain-of-thought prompting elicits reasoning in large language models. In *Proceedings of the 36th International Conference on Neural Information Processing Systems*, NIPS '22. Red Hook, NY, USA: Curran Associates Inc. ISBN 9781713871088.

Weston, J.; and Sukhbaatar, S. 2023. System 2 Attention (is something you might need too). arXiv:2311.11829.

Wiratunga, N.; Abeyratne, R.; Jayawardena, L.; Martin, K.; Massie, S.; Nkisi-Orji, I.; Weerasinghe, R.; Liret, A.; and Fleisch, B. 2024. CBR-RAG: Case-Based Reasoning for Retrieval Augmented Generation in LLMs for Legal Question Answering. In Recio-Garcia, J. A.; Orozco-del Castillo, M. G.; and Bridge, D., eds., *Case-Based Reasoning Research and Development*, 445–460. Cham: Springer Nature Switzerland. ISBN 978-3-031-63646-2.

Xu, X.; Tao, C.; Shen, T.; Xu, C.; Xu, H.; Long, G.; and guang Lou, J. 2024a. Re-Reading Improves Reasoning in Large Language Models. arXiv:2309.06275.

Xu, Z.; Jiang, F.; Niu, L.; Deng, Y.; Poovendran, R.; Choi, Y.; and Lin, B. Y. 2024b. Magpie: Alignment Data Synthesis from Scratch by Prompting Aligned LLMs with Nothing. arXiv:2406.08464.

Yasunaga, M.; Chen, X.; Li, Y.; Pasupat, P.; Leskovec, J.; Liang, P.; Chi, E. H.; and Zhou, D. 2024. Large Language Models as Analogical Reasoners. In *The Twelfth International Conference on Learning Representations*.

Yuan, W.; Pang, R. Y.; Cho, K.; Li, X.; Sukhbaatar, S.; Xu, J.; and Weston, J. 2024. Self-Rewarding Language Models. arXiv:2401.10020.

Zheng, L.; Chiang, W.-L.; Sheng, Y.; Zhuang, S.; Wu, Z.; Zhuang, Y.; Lin, Z.; Li, Z.; Li, D.; Xing, E.; Zhang, H.; Gonzalez, J. E.; and Stoica, I. 2023. Judging LLM-as-a-Judge with MT-Bench and Chatbot Arena. In Oh, A.; Naumann, T.; Globerson, A.; Saenko, K.; Hardt, M.; and Levine, S., eds., *Advances in Neural Information Processing Systems*, volume 36, 46595–46623. Curran Associates, Inc.

Zheng, M.; Pei, J.; and Jurgens, D. 2023. Is "A Helpful Assistant" the Best Role for Large Language Models? A Systematic Evaluation of Social Roles in System Prompts. arXiv:2311.10054.

Zhong, W.; Cui, R.; Guo, Y.; Liang, Y.; Lu, S.; Wang, Y.; Saied, A.; Chen, W.; and Duan, N. 2024. AGIEval: A Human-Centric Benchmark for Evaluating Foundation Models. In Duh, K.; Gomez, H.; and Bethard, S., eds., *Findings of the Association for Computational Linguistics: NAACL 2024*, 2299–2314. Mexico City, Mexico: Association for Computational Linguistics.

Zhou, Y.; Geng, X.; Shen, T.; Tao, C.; Long, G.; Lou, J.-G.; and Shen, J. 2023. Thread of Thought Unraveling Chaotic Contexts. arXiv:2311.08734.