SEAS: A Retrieval Optimization Framework for LLMs by Using Supplemental Enhancement of Action Segments

Anonymous ACL submission

Abstract

Utilizing the Retrieval-Augmented Generation (RAG) framework with large language models for question answering often results in low retrieval precision and recall rates. A solution to address this issue involves retrieving external knowledge at various granularities. However, this strategy typically suffers from decreased precision in coarse-grained retrieval and omissions in fine-grained retrieval. To overcome these challenges, we introduce a framework, named Supplemental Enhancement of Action Segments (SEAS). SEAS utilizes few-shot prompting to extract action segments from texts, which are then used to enhance the retrieval of texts. Due to the complexity and specialization of knowledge, the RAG method is particularly important in domain-specific tasks. For example, in the legal field, retrieving legal articles is a critical task, and external legal knowledge can be introduced to mitigate hallucinations in the completion of legal tasks. Thus, we conducted Law Retrieval Task experiments on a Japanese legal dataset and Legal Question Answering Task experiments on a Chinese legal dataset. The results indicate that SEAS significantly outperforms all baseline methods. The code will be available on GitHub.

1 Introduction

When using large language models (LLMs) for question answering, the Retrieval-Augmented Generation (RAG) framework (Lewis et al., 2020) has become one of the most popular frameworks for reducing hallucinations (Zhang et al., 2023). Despite its advantages, the RAG framework often faces challenges in retrieval precision and recall (Gao et al., 2023) due to text complexity—long texts introduce irrelevant information, while short texts lack essential features. Recent studies have explored various strategies to enhance retrieval, including adjustments in retrieval granularity (Ram et al., 2023) and retrieval frequency (Izacard et al., 2022). Our research explores effective retrieval granularity within the RAG framework. Subsequently, we propose a method that integrates various retrieval granularities to improve retrieval precision, thereby improving the performance of generation tasks.

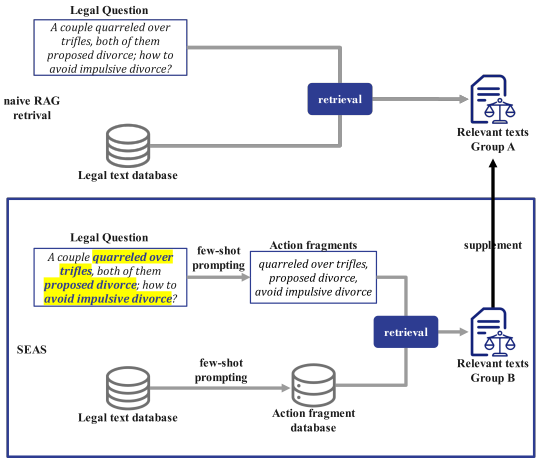


Figure 1: Overview of Supplemental Enhancement of Action Segments: (1) Extracting action segments from the questions and legal texts using few-shot prompting. (2) Supplementing and enhancing the retrieval results of complete legal texts with these action segments.

There are two main challenges in this work: (1) defining and extracting effective retrieval granularity, and (2) balancing the trade-offs between granularity levels.

To address these challenges, we introduce a new framework, Supplemental Enhancement of Action Segments (SEAS) (see Figure 1), which utilizes "action segments" as fine-grained retrieval units. The concept of "action segments" is derived from the jurisprudential concept of "legal actions (Rechtgeschaefte)", which refers to actions implemented by individuals that produce legal effects. For example, in the legal question "*A couple quarreled over trifles, both of them proposed divorce; how to avoid an impulsive divorce?*" the action segments

identified are "*quarreled for trifles,*" "*proposed divorce,*" and "*avoid impulsive divorce.*" These segments serve as fine-grained document chunks that supplement the retrieval of coarse-grained chunks, encompassing complete questions and legal texts.

We evaluate SEAS through two experiments: (1) Japanese Law Retrieval from COLIEE's Task 3[1], and (2) Chinese Legal Question Answering[2]. Our experiments show that SEAS outperforms in the both tasks. In the Japanese Law Retrieval task, SEAS boosts the retrieval performance of three embedding models: BAAI/bge-large-en-v1.5[3], OpenAI text-embedding-3-small[4] and text-embedding-3-large[4]. In the Chinese Legal Question Answering task, SEAS enhances the performance of GPT-3.5-Turbo+RAG and GPT-4+RAG, with Accuracy (ACC) (Yue et al., 2023) improvements of 2.9% and 2%, respectively.

In this paper, we propose a new framework SEAS to enhance the RAG performance, the main contributions of the framework are as follow:

- To the best of our knowledge, this is the first method that uses **action segments** as retrieval granularities to enhance retrieval performance in legal domain.
- SEAS combines retrievals of document chunks at different granularities, exploring optimization paths in the RAG framework.

2 Related work

The RAG framework in LLM generation tasks suffers from low precision and recall rates in retrieval (Lewis et al., 2020). Research on the granularity of RAG retrieval (Khandelwal et al., 2019; Nishikawa et al., 2022; Kang et al., 2023) and chunking strategies (Langchain, 2023; Yang, 2023) seeks to improve precision and efficiency by using text chunks of varying sizes. Our framework improves overall precision by combining retrieval results of different granularities.

Determining appropriate granularity and obtaining fine-grained document chunks are two key challenges in our framework. Several studies (Min et al., 2023; Kamoi et al., 2023; Chen et al., 2023a) have investigated semantic representations of text at the propositional level. Building on this foundation, Chen et al. (2023b) have effectively utilized propositions as retrieval units. Inspired by these advancements, our approach integrates propositions from legal texts—action segments—as fine-grained retrieval units to address the first challenge.

Recent research has leveraged content generated by LLMs for retrieval and enhancement tasks (Gao et al., 2023), as demonstrated in studies by Wang et al. (2023), Yu et al. (2022), and Cheng et al. (2023). These studies highlight the innovative use of data sources within the RAG framework. Inspired by these developments, we have employed few-shot prompting with LLMs to extract fine-grained document chunks, thus addressing the second challenge.

3 Method

In this section, we introduce the details of our proposed framework–Supplemental Enhancement of Action Segments (SEAS), as illustrated in Figure 1. First, we devise the Action Segment Extraction (Section 3.1), which extracts action segments from legal texts via few-shot prompting. Then, using these action segments, we implement the Supplemental Enhancement (Section 3.2). This process supplements the retrieval results of complete text chunks with the results of action segments to produce the final relevant legal texts for the legal question.

3.1 Action Segment Extraction

Action Segment Extraction involves extracting text which describes actions from a legal text database and legal questions. We use few-shot (3-shot) prompting to extract text describing actions from each legal article in the database, thereby creating an action segment database. Similarly, we use a similar few-shot (3-shot) prompting method to extract text describing actions from the legal questions. The configurations of the 3-shot prompting can be found in Appendix A.

3.2 Supplemental Enhancement

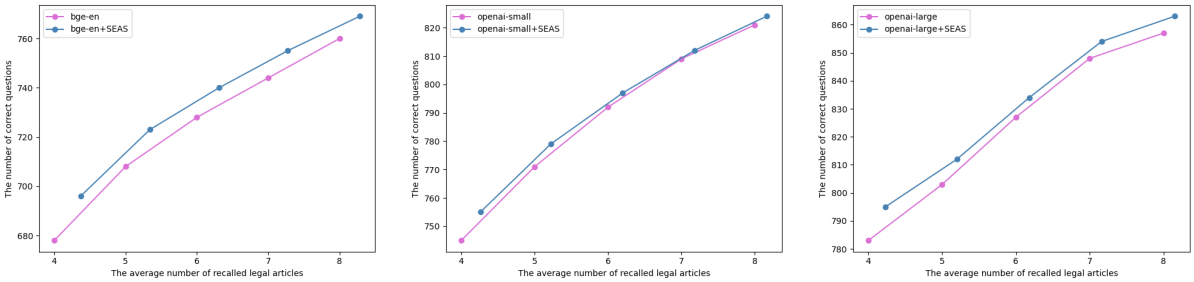
First, we use an embedding model to encode the legal texts and legal questions, retrieving texts (relevant text group A in figure 1) relevant to the legal question and selecting the top X legal texts. Next, we use the same embedding model to encode the

[1]Competition on Legal Information Extraction/Entailment: https://sites.ualberta.ca/~rabelo/COLIEE2024/

[2]The data comes from DISC-LawLLM-eval + open resource web.

[3]https://huggingface.co/BAAI/bge-large-en-v1.5

[4]Openai embedding model: https://platform.openai.com/docs/guides/embeddings



(a) bge-en (b) openai-small (c) openai-large

Figure 2: Evaluation results for the Japanese Law Retrieval task from COLIEE. The baselines are three embedding models: bge-en, openai-small, and openai-large. This figure shows the retrieval performance of these models after enhancement with SEAS.

action segment database along with the action segments extracted from legal questions, aiming to retrieve and select the top Y action segments from legal texts (relevant text group B in figure 1). We then identify the original legal texts corresponding to these action segments, integrate these Y legal texts with the initially selected X legal texts, and perform deduplication to compile the final F set of $[X, X+Y]$ relevant legal texts, refer to Equation 1.

$$F = X \cup Y, X \subsetneqq A, Y \subsetneqq B \tag{1}$$

4 Experiment

We conducted two experiments to evaluate the retrieval effectiveness of SEAS. The first experiment directly tested the retrieval effectiveness of SEAS, while the second experiment evaluated its impact on downstream task performance. The first experiment used the law retrieval dataset from COLIEE Task 3. This task involves extracting a subset of Japanese Civil Code Articles to answer a yes/no Japanese legal bar exam question, assessing the retrieval effectiveness of SEAS. In the second experiment, Chinese Legal Question Answering, we used SEAS to retrieve relevant articles and added them to the prompt for GPT-3.5-turbo (OpenAI, 2022) and GPT-4 (OpenAI, 2023) to generate answers to Chinese legal questions, evaluating the quality of the answers to assess the impact of SEAS on downstream generation tasks. The configurations of the few-shot prompting used in the experiments can be found in Appendix A.

4.1 Japanese Law Retrieval

We evaluate the enhancement effects of SEAS on embedding models. In this experiment, a legal question can be related to multiple articles. In this dataset, a legal question is related to no more than 6 articles, with an average of 1.28 articles per question. The model identifies several articles related to the question.

Datasets The data comes from Task 3 (English version) of the COLIEE 2023 datasets (Goebel et al., 2024), including the Japanese Civil Code texts and 1,097 yes/no Japanese legal bar exam questions from the training set, along with the articles relevant to each question.

Evaluation Metrics Our testing shows that LLMs have the capability to select the correct articles, making it particularly crucial that the retrieved articles comprehensively cover all relevant laws. Consequently, we evaluate retrieval effectiveness by counting the number of questions n for which the retrieved articles cover all relevant articles:

$$n = \sum_{i=1}^{1097} P(i) \tag{2}$$

$$P(x) = \begin{cases} 1 & Q_x \subseteq R_x \\ 0 & Q_x \nsubseteq R_x \end{cases} \tag{3}$$

where i represents the number of the question, $P(x)$ is a function to count, Q_x represents the set of relevant articles for the x-th question, R_x represents the set of retrieved articles for the x-th question in Equation (2, 3).

Baselines We choose three embedding models as baselines: BAAI/bge-large-en-v1.5 (output dimension 1024), OpenAI text-embedding-3-small (output dimension 1536), and text-embedding-3-small (output dimension 3072). We compare the number of questions for which the relevant articles are correctly retrieved before and after enhancement with SEAS.

Model	ACC	CPL	CLR
General LLMs			
GPT-3.5-turbo	1.97	1.83	2.71
GPT-4	2.10	2.10	3.09
Chinese Legal LLMs			
DISC-LawLLM	2.43	2.22	3.08
Tongyi Farui	3.12	2.93	3.52
General LLMs with RAG			
GPT-3.5-turbo + RAG	3.14	2.76	3.70
GPT-4 + RAG	3.50	3.30	4.11
(SEAS; Ours)			
GPT-3.5-turbo + RAG + SEAS	3.23	2.82	3.72
GPT-4 + RAG + SEAS	**3.57**	**3.33**	**4.14**

Table 1: Evaluation results for the Chinese Legal Question Answering task. Baselines are General LLMs, Chinese Legal LLMs and General LLMs with RAG Framework. This figure shows the effectiveness of SEAS-enhanced LLMs in generating task outcomes.

Results The experiment evaluates the effectiveness of integrating the top 1 article retrieved by the SEAS with the top 4, 5, 6, 7, and 8 articles retrieved by the BAAI/bge-large-en-v1.5, text-embedding-3-small, and text-embedding-3-small models (see Figure 2). Upon incorporating the supplemental articles retrieved by SEAS, all three models improved performance, correctly retrieving relevant articles for more questions. Notably, the enhancement effect of SEAS was greatest for bge-en, followed by openai-large, and least for openai-small.

4.2 Chinese Legal Question Answering

In this section, we evaluate the impact of the SEAS-enhanced legal retrieval model on downstream question answer tasks. Specifically, we use BAAI/bge-large-zh-v1.5[5] as Chinese embedding model. The quality of these generated answers is evaluated by using the DISC-LawLLM-eval (Yue et al., 2023) method. The DISC-LawLLM-eval method involves inputting the question, the generated answer, and a reference answer into LLMs. Considering the proficiency of powerful LLMs like GPT-4 in aligning with human judgments—demonstrating more than 80% consistency (Zheng et al., 2023)—we employ GPT-4 to evaluate the quality of the generated answers.

Datasets The dataset comprises 222 Chinese civil law text questions along with their reference answers (Chinese version), including 62 questions from DISC-LawLLM-eval and 160 questions from Chinese legal consultations, justice-related publications, and other sources.The detailed sources of the data can be found in Appendix B.

Evaluation Metrics We use the evaluation metrics from DISC-LawLLM-eval(Yue et al., 2023), including accuracy, completeness and clarity. (1) **Accuracy (ACC):** The consistency of the content and semantics of the answer with the reference answer. (2) **Completeness (CPL):** The answer do not omit any details compared to the reference answers. (3) **Clarity (CLR):** The juridical logic analysis of the answer is rigorous and clear, and the sentences are well-organized.

Baselines We select GPT-3.5-turbo, GPT-4, GPT-3.5-turbo + RAG, GPT-4 + RAG, and the Chinese legal large language models DISC-LawLLM (Yue et al., 2023) and Tongyi Farui[6] (commercial model) as baselines.

Results The experiment evaluates the effectiveness of integrating the top 3 articles retrieved by the SEAS with the top 3 articles retrieved by BAAI/bge-large-zh-v1.5 for generating answers. (see Table 1). The answers generated after supplementing with the top 3 articles retrieved by SEAS surpassed those generated by the unenhanced LLMs and the LLMs with RAG. Notably, GPT-4+RAG+SEAS achieved the highest performance, surpassing the generation effects of the Tongyi Farui and DISC-LawLLM models. Specifically, the ACC of SEAS-enhanced GPT-3.5-turbo+RAG increased by 2.9%, and the ACC of SEAS-enhanced GPT-4+RAG increased by 2%.

5 Conclusion

In this work, we propose Supplemental Enhancement of Action Segments (SEAS). It generates fine-grained retrieval units—action segments—as retrieval granularity for legal domain questions through few-shot prompting. And then, these segments are used to supplement and enhance the retrieval results of coarse-grained retrieval units—complete legal texts. Our framework combines the advantages of different granularity document chunks, optimizing the retrieval process. Experimental results show that SEAS improves the retrieval performance of various embedding models and guides downstream LLMs to generate better answers. We hope this work provides insights into optimizing the retrieval process within the RAG framework and can be applied to real-world scenarios.

[5] https://huggingface.co/BAAI/bge-large-zh-v1.5

[6] https://tongyi.aliyun.com/farui

Limitation

Despite SEAS being a model-agnostic framework that can be combined with other components, our study is limited in demonstrating generalizability across different types or scales of embedding models. Additionally, although the framework focuses on improving RAG retrieval and is domain-agnostic, our experiments were conducted only on two legal datasets, lacking tests in other domains. While SEAS effectively retrieves text chunks through few-shot prompting with LLMs, the generation cost becomes significant when the datasets are large.

Ethics Statement

The acquisition of data for this study was conducted with the explicit permission of the publisher, ensuring full compliance with all applicable legal and ethical standards. This project was conducted as a collaborative effort, and we fully acknowledge the contributions of each collaborator, ensuring a transparent and ethical process throughout the entire collaboration.

References

Sihao Chen, Hongming Zhang, Tong Chen, Ben Zhou, Wenhao Yu, Dian Yu, Baolin Peng, Hongwei Wang, Dan Roth, and Dong Yu. 2023a. Sub-sentence encoder: Contrastive learning of propositional semantic representations. *ArXiv*, abs/2311.04335.

Tong Chen, Hongwei Wang, Sihao Chen, Wenhao Yu, Kaixin Ma, Xinran Zhao, Dong Yu, and Hongming Zhang. 2023b. Dense x retrieval: What retrieval granularity should we use? *ArXiv*, abs/2312.06648.

Xin Cheng, Di Luo, Xiuying Chen, Lemao Liu, Dongyan Zhao, and Rui Yan. 2023. Lift yourself up: Retrieval-augmented text generation with self memory. *ArXiv*, abs/2305.02437.

Yunfan Gao, Yun Xiong, Xinyu Gao, Kangxiang Jia, Jinliu Pan, Yuxi Bi, Yi Dai, Jiawei Sun, Qianyu Guo, Meng Wang, and Haofen Wang. 2023. Retrieval-augmented generation for large language models: A survey. *ArXiv*, abs/2312.10997.

Randy Goebel, Yoshinobu Kano, Mi-Young Kim, Juliano Rabelo, Ken Satoh, and Masaharu Yoshioka. 2024. Overview and discussion of the competition on legal information, extraction/entailment (coliee) 2023. *The Review of Socionetwork Strategies*, 18(1):27–47.

Gautier Izacard, Patrick Lewis, Maria Lomeli, Lucas Hosseini, Fabio Petroni, Timo Schick, Jane A. Yu, Armand Joulin, Sebastian Riedel, and Edouard Grave. 2022. Few-shot learning with retrieval augmented language models. *ArXiv*, abs/2208.03299.

Ryo Kamoi, Tanya Goyal, Juan Diego Rodriguez, and Greg Durrett. 2023. Wice: Real-world entailment for claims in wikipedia. In *Conference on Empirical Methods in Natural Language Processing*.

Minki Kang, Jin Myung Kwak, Jinheon Baek, and Sung Ju Hwang. 2023. Knowledge graph-augmented language models for knowledge-grounded dialogue generation. *ArXiv*, abs/2305.18846.

Urvashi Khandelwal, Omer Levy, Dan Jurafsky, Luke Zettlemoyer, and Mike Lewis. 2019. Generalization through memorization: Nearest neighbor language models. *ArXiv*, abs/1911.00172.

Langchain. 2023. Recursively split by character.

Patrick Lewis, Ethan Perez, Aleksandara Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Kuttler, Mike Lewis, Wen tau Yih, Tim Rocktäschel, Sebastian Riedel, and Douwe Kiela. 2020. Retrieval-augmented generation for knowledge-intensive nlp tasks. *ArXiv*, abs/2005.11401.

Sewon Min, Kalpesh Krishna, Xinxi Lyu, Mike Lewis, Wen tau Yih, Pang Wei Koh, Mohit Iyyer, Luke Zettlemoyer, and Hannaneh Hajishirzi. 2023. Factscore: Fine-grained atomic evaluation of factual precision in long form text generation. *ArXiv*, abs/2305.14251.

Yuichi Nishikawa, A. Holobar, Kohei Watanabe, Tetsuya Takahashi, Hiroki Ueno, Noriaki Maeda, Hirofumi Maruyama, Shinobu Tanaka, and Allison S. Hyngstrom. 2022. Detecting motor unit abnormalities in amyotrophic lateral sclerosis using high-density surface emg. *Clinical Neurophysiology*, 142:262–272.

OpenAI. 2022. Introducing chatgpt.

OpenAI. 2023. Gpt-4 technical report.

Ori Ram, Yoav Levine, Itay Dalmedigos, Dor Muhlgay, Amnon Shashua, Kevin Leyton-Brown, and Yoav Shoham. 2023. In-context retrieval-augmented language models. *Transactions of the Association for Computational Linguistics*, 11:1316–1331.

Yile Wang, Peng Li, Maosong Sun, and Yang Liu. 2023. Self-knowledge guided retrieval augmentation for large language models. In *Conference on Empirical Methods in Natural Language Processing*.

Sophia Yang. 2023. Advanced rag 01: Small-to-big retrieval.

W. Yu, Dan Iter, Shuohang Wang, Yichong Xu, Mingxuan Ju, Soumya Sanyal, Chenguang Zhu, Michael Zeng, and Meng Jiang. 2022. Generate rather than retrieve: Large language models are strong context generators. *ArXiv*, abs/2209.10063.

Shengbin Yue, Wei Chen, Siyuan Wang, Bingxuan Li, Chenchen Shen, Shujun Liu, Yuxuan Zhou, Yao Xiao, Song Yun, Wei Lin, Xuanjing Huang, and Zhongyu Wei. 2023. Disc-lawllm: Fine-tuning large language models for intelligent legal services. *ArXiv*, abs/2309.11325.

Yue Zhang, Yafu Li, Leyang Cui, Deng Cai, Lemao Liu, Tingchen Fu, Xinting Huang, Enbo Zhao, Yu Zhang, Yulong Chen, Longyue Wang, Anh Tuan Luu, Wei Bi, Freda Shi, and Shuming Shi. 2023. Siren's song in the ai ocean: A survey on hallucination in large language models. *ArXiv*, abs/2309.01219.

Lianmin Zheng, Wei-Lin Chiang, Ying Sheng, Siyuan Zhuang, Zhanghao Wu, Yonghao Zhuang, Zi Lin, Zhuohan Li, Dacheng Li, Eric P. Xing, Haotong Zhang, Joseph Gonzalez, and Ion Stoica. 2023. Judging llm-as-a-judge with mt-bench and chatbot arena. *ArXiv*, abs/2306.05685.

A Configurations of Few-shot Prompting

A.1 Prompting Used in The SEAS Framework

Tables (2, 4, 3, 5) show the prompting settings in the SEAS framework. Tables 2 and 4 present the prompting settings for the Japanese Law Retrieval task, with Table 2 showing the settings for *"Extracting legal action segments for questions"* and Table 4 for *"Extracting legal action segments for articles."* Table 3 and 5 are for the Chinese Legal Question Answering task, respectively showing the settings for *"Extracting legal action segments for questions"* and *"Extracting legal action segments for articles."*

A.2 Prompting Used in LLM generates answers

Table 6 presents the prompting used by the LLM to generate answers to questions in the Chinese Legal Question Answering task, utilizing the extracted action segments.

A.3 Prompting Used in Evaluation

Table 7 shows the prompting used to evaluate the effectiveness of the LLM-generated answers in the Chinese Legal Question Answering task.

B Data Source of Chinese Legal Question Answering

There is currently no high-quality Chinese civil law subjective question answering dataset. To accomplish the Chinese Legal Question Answering task, we have constructed a dataset comprising 222 question-answer pairs of Chinese civil law subjective questions. We invited law students to screen 62 civil law questions from the 300 subjective questions in the DISC-LawLLM-eval dataset. Since the number was insufficient, we asked the law students to collect high-quality question banks, which included questions from the National Unified Legal Professional Qualification Examination, the "Civil Code of the People's Republic of China" knowledge competition, university civil law exam question banks, and "100 Questions on the Civil Code—100 Questions Closely Related to People's Lives." After removing duplicates from the collected civil law questions, we obtained an additional 160 questions. These questions, together with the 62 from the DISC-LawLLM-eval dataset, form the Chinese civil law subjective question-answer dataset used in this study.

Please identify the roles, actions in the inquiry, as illustrated below:
Example1:A contract of sales concluded by a minor may not be rescinded if it relates to daily life, even in cases the consent of the parental authority is not obtained.
Answer: {"Roles":" a minor", "Action":" A contract of sales concluded by a minor"}
Example2: Demand damages arising from latent defect in the subject matter of a pledge is not secured by a pledge, when the contract clause is not concluded.
Answer: {"Roles":" none", "Action":" Demand damages arising from latent defect in the subject matter of a pledge"}
Example3: In cases where two buildings which are owned by one person that the distinction of principal and accessory cannot be made between become one building due to the construction, if mortgages are created with respect to the two buildings respectively, the mortgages shall own each mortgage on the building in proportion to the respective price continuedly.
Answer: {"Roles":"one person", "Action":"the distinction of principal and accessory cannot be made between become one building due to the construction" }
Question:

Table 2: Extracting legal action segments for questions in Japanese Law Retrieval

请你找出提问中的角色、行为、提问，举例如下：
举例1：吃了上错桌的菜需要付钱吗？回答：{“角色”:”无”,”行为”:”吃了上错桌的菜”,”提问”:”需要付钱吗？”}
举例2：被家暴了应该如何申请离婚?回答：{“角色”:”无”,”行为”:”被家暴”,”提问”:” 如何申请离婚？”}
举例3：商品房实际销售价格高于备案价格的，房屋买卖合同是否有效?回答：{“角色”:”商品房”,”行为”:” 实际销售价格高于备案价格”,”提问”:” 房屋买卖合同是否有效?”}
下面，请你找出下列提问中的角色、行为、提问，若"角色"有多个，用列表分开表示,只有一个也用列表表示，"行为"有多个，用逗号连接成一句话，仅用json回答，不需要其他解释：
提问：

Table 3: Extracting legal action segments for questions in Chinese Legal Question Answering

Please extract the "actions" from the legal articles.
Definition of actions: The specific deeds or steps taken by someone, generally composed of the predicate and the object in a sentence. A legal article may contain one or more actions, and often the appearance of a subject indicates the presence of an action. Here are a few examples:
Example 1: Article 5: (1) A minor must obtain the consent of the minor's legal representative to perform a juridical act; provided, however, that this does not apply to a juridical act for merely acquiring a right or being released from an obligation.(2) A juridical act in contravention of the articles of the preceding paragraph is voidable.(3) Notwithstanding the articles of paragraph (1), a minor may freely dispose of property that the legal representative has permitted the minor to dispose of for a specified purpose, to an extent that falls within the scope of that purpose. The same applies if the minor disposes of property that the legal representative has permitted the minor to dispose of without specifying a purpose.
Result: actions:[" perform a juridical act", " contravention of the articles of the preceding paragraph"," dispose of property that the legal representative has permitted the minor to dispose of for a specified purpose", " disposes of property that the legal representative has permitted the minor to dispose of without specifying a purpose"]
Example 2: Article 14: (1) If the grounds prescribed in the main clause of Article 11 cease to exist, the family court must rescind the decision for the commencement of curatorship at the request of the person in question, that person's spouse, that person's relative within the fourth degree of kinship, the guardian of a minor, the supervisor of a minor's guardian, the curator, the curator's supervisor, or a public prosecutor.(2) At the request of a person provided for in the preceding paragraph, the family court may rescind all or part of the decision referred to in paragraph (2) of the preceding Article.
Result: actions:["rescind the decision for the commencement of curatorship", " rescind all or part of the decision referred to in paragraph (2) of the preceding Article"]
Example 3: Article:(1) If the owner of a first thing attaches a second thing that the owner owns to the first thing to serve the ordinary use of the first thing, the thing that the owner attaches is an appurtenance.(2) An appurtenance is disposed of together with the principal thing if the principal thing is disposed of.
Result: actions: ["attaches a second thing that the owner owns to the first thing to serve the ordinary use of the first thing"," is disposed of"]
Below, please analyze the following legal article. Please note, please strictly follow the example format to generate the answer, do not omit any words in your response, and you do not need to start your response with "the result is" or break it into multiple lines.
Article:

Table 4: Extracting legal action segments for articles in Japanese Law Retrieval

请从法律条款中提取“行为”。
行为的定义：某人采取的具体行为或步骤，通常由句子中的谓语和宾语组成。一个法律条款可能包含一个或多个行为，且通常主语的出现表明存在一个行为。
以下是几个例子：
举例1：法条：民事主体从事民事活动，应当遵循自愿原则，按照自己的意思设立、变更、终止民事法律关系。
结果：行为：["从事民事活动","遵循自愿原则","设立、变更、终止民事法律关系"]
举例2：自然人从出生时起到死亡时止，具有民事权利能力，依法享有民事权利，承担民事义务。
结果：行为：["具有民事权利能力","享有民事权利","承担民事义务"]
举例3：营利法人的权力机构、执行机构作出决议的会议召集程序、表决方式违反法律、行政法规、法人章程，或者决议内容违反法人章程的，营利法人的出资人可以请求人民法院撤销该决议。但是，营利法人依据该决议与善意相对人形成的民事法律关系不受影响。
结果：行为: ["违反法律、行政法规、法人章程","违反法人章程","请求人民法院撤销该决议"]
下面，请分析以下法律条款。请注意，请严格按照示例格式生成答案，不要遗漏任何单词，并且不要将答案分成多行。
法条：

Table 5: Extracting legal action segments for articles in Chinese Legal Question Answering

请参考以下的法条回答问题。回答应遵守以下要求：回答中不要引用下面未出现的法条；回答中要引用法条编号和其中相关的内容；问题中如果提问者表明了身份信息（如所在行业、职业、职位等），回答还应考虑这些身份信息，为其提供更准确、更具有可参考性的回答；
问题中如果存在多种情形的，回答应对可能存在的各种情形进行分类讨论；回答不要超过300字。
法条：
问题：

Table 6: LLM generates answers (Use to leverage action segments in generating answers with LLM)

You are a professional, impartial, and strict scorer. You will be given a question, a pending scored answer, and a reference answer. Please rate the pending scored answer based on the reference answer in the following aspects:
’accuracy’: The content of pending scored answer conforms to reference answer in semantic, especially noting the content of the law, the facts in question and the conclusion. ’completeness’: Compared to reference answer, pending scored answer does not miss any details in reference answer. Do not let the length of the answer influence your judgment. ’clarity’: the logic of pending scored answer is rigorous and clear, and the sentences are well-organized. If Accuracy and Completeness is bad, Clarity should also be bad.
score should be from 1 (lowest)-5 (highest). Your rating should be strict enough, and do not easily give full scores. Pay attention to the difference between laws mentioned in the reference answer and the pending scored answer. In your response, you should only include a JSON object, with keys being the aspects and values being the scores. Do not include any additional information or explanation.
Question:
Reference answer:
Pending scored answer:

Table 7: Evaluation (Use to evaluate the answers generated by LLM)