Legal Knowledge and Information Systems
G. Sileno et al. (Eds.)

doi:10.3233/FAIA230991

LawGiBa – Combining GPT, Knowledge Bases, and Logic Programming in a Legal Assistance System

Ha Thanh Nguyen [a] Randy Goebel [b] Francesca Toni [c] Kostas Stathis [d] Ken Satoh [a]
[a] *National Institute of Informatics (NII), Japan*
[b] *University of Alberta, Alberta Machine Intelligence Institute, Canada*
[c] *Imperial College London, UK*
[d] *Royal Holloway, University of London, UK*

Abstract. We present LawGiBa, a proof-of-concept demonstration system for legal assistance that combines GPT, legal knowledge bases, and Prolog's logic programming structure to provide explanations for legal queries. This novel combination effectively and feasibly addresses the hallucination issue of large language models (LLMs) in critical domains, such as law. Through this system, we demonstrate how incorporating a legal knowledge base and logical reasoning can enhance the accuracy and reliability of legal advice provided by AI models like GPT. Though our work is primarily a demonstration, it provides a framework to explore how knowledge bases and logic programming structures can be further integrated with generative AI systems, to achieve improved results across various natural languages and legal systems.

Keywords. interactive demonstration, ChatGPT, legal advice, knowledge base, logic programming, Prolog

1. Introduction

Large language models (LLMs) like GPT (Generative Pre-trained Transformer) [1,2] have garnered immense recent interest for their impressive ability to generate relevant and potentially accurate answers to natural language queries. But, although these models have been applied successfully in many domains, they face significant challenges when dealing with critical discipline areas such as law. LLMs often lack the necessary structural knowledge and logical reasoning required to provide trustworthy and detailed information, which is essential for providing accurate legal advice [3,4].

To address this limitation, we present LawGiBa, a legal assistance system that combines GPT, legal knowledge bases, and Prolog's logic programming structure to offer accurate legal advice tailored to specific queries. LawGiBa enhances the traditional LLM approach with a knowledge base and a logical reasoning function, which increases the reliability and accuracy of the information and guidance provided. We develop a prototype of LawGiBa using the COLIEE Competition on Legal Information Extraction/Entailment Task 3 and Task 4 data [5]. This interactive demonstration showcases

the advantages of incorporating a knowledge base and logical reasoning in AI systems, particularly for critical domains like law.

2. System Description

Our demonstration system is a web-based legal intelligence demo that uses GPT-4 [6,2] to answer questions about the Japanese Civil Code. The user interface consists of a dropdown to select predefined examples, checkboxes to enable the knowledge base and the logic programming function, a read-only text area for the question, and a submit button. Users can choose to enable the knowledge base or the logic programming function, which influences GPT-4's strategy to generate answers. The system also includes tooltips for displaying Civil Code articles when they are referenced by GPT-4's explanations in instances where logic programming is not enabled.

2.1. Input Interface

The input interface (illustrated in Figure 1) consists of a dropdown menu containing a list of examples. Users can select an example, and the corresponding question text appears in a read-only text area. Users can also choose to enable the knowledge base and/or logic programming options by checking the respective checkboxes, before submitting the question.

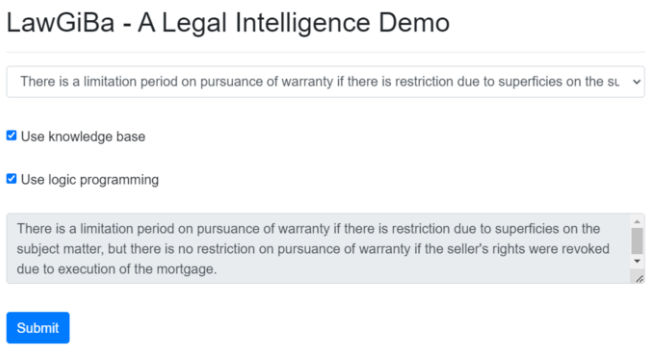

Figure 1. Input interface with example selection, knowledge base and logic programming checkboxes.

2.2. GPT-4 Answer Generation

Once the user submits a question, GPT-4 generates an answer following the strict format of "Answer—Explanation", where the "Answer" can be either "Y" or "N" and the "Explanation" provides a text rationale for the given answer. In our system, GPT-4's strategy for generating answers depends on whether the user has enabled the knowledge base or logic programming options.

If the knowledge base is enabled, GPT-4 will consider the provided Civil Code articles when generating its answer. If logic programming is enabled, GPT-4 will convert the knowledge base (legal articles) and the question into Prolog code, which allows its execution as a logic program.

2.3. Output Interface

The output interface displays GPT-4's answer and explanation. GPT-4's answer is displayed as "Correct" or "Incorrect" based on the provided label for the example. When logic programming is not enabled, the explanation text will include tooltips for any Civil Code articles that GPT-4 has referenced. Users can hover over the article numbers to see the tooltip containing the specific article content. In cases where logic programming is enabled, the output will display Prolog code (shown in Figure 2) generated by GPT-4, which can be used by users or other systems to verify the answer with a Prolog engine.

Figure 2. Output interface displaying GPT-4's generated answer and explanation with Prolog code.

Overall, we believe that the system provides a user-friendly, interactive way to explore legal questions by leveraging the powerful GPT-4 model, a knowledge base of Civil Code articles, and optional logic programming to enhance its explanatory performance.

3. Observations and Discussions

Enhancing Expert Decision-Making: The LawGiBa system allows legal experts and AI practitioners to evaluate GPT-4's responses not only at the binary answer level but also with the provided explanations. The system contributes to improving the decision-making process of these professionals, enhancing the overall effectiveness of incorporating AI into the legal domain.

Tackling Hallucination Issues: Our work with LawGiBa can be considered a small step towards addressing hallucination issues in generative large language models (LLMs). Although this demo paper does not extensively explore this direction, we hope that it provides an initial framework to encourage further interest and research among both researchers and practitioners, emphasizing the importance of continuing the pursuit of more reliable and accurate AI models.

Utilizing Knowledge Base and Logic Programming: The interactive demo demonstrates how using a domain-specific knowledge base can help GPT avoid generating baseless answers and thus reduce hallucination occurrences. Employing a logic programming

format not only helps formulate better explanations but also aids in debugging the hidden reasoning process within LLMs, which are often regarded as black boxes. Note that how to automate or even semi-automate the LLM debugging with explanations remains a significant challenge.

Challenges in Legal Complexity: Despite the system being given a knowledge base and the ability to use logic programming to produce explanations, there are cases where the model still provides incorrect answers. This highlights the complexity of the legal domain and emphasizes that much work remains to be done to achieve optimal results. An important step in this direction is a detailed analysis of errors, such as done for the COLIEE competition domain [7].

Limitations in Prolog Code Generation: During some trials of the system, we observed that the generated Prolog code may not always be semantically or syntactically correct. This underlines another area for future research and improvement, emphasizing the need to refine code generation process to provide more accurate outputs for users and systems.

4. Conclusions

We have introduced LawGiBa, a proof-of-concept legal assistance system that combines GPT, a knowledge base, and Prolog's logic programming structure to improve the accuracy and reliability of legal advice provided by AI models. The interactive demonstration highlights the importance of incorporating a knowledge base and logical reasoning in critical domains like law. Although our work is primarily a demonstration, it sets the stage for future research to explore how knowledge and reasoning can be integrated with generative AI to achieve improved results across various languages and legal systems.

Acknowledgements

This work was supported by the AIP challenge funding related with JST, AIP Trilateral AI Research, Grant Number JPMJCR20G4.

References

[1] Radford A, Narasimhan K, Salimans T, Sutskever I, et al.. Improving language understanding by generative pre-training. OpenAI; 2018.

[2] OpenAI R. GPT-4 technical report. arXiv. 2023:2303-08774.

[3] Nguyen HT, Toni F, Stathis K, Satoh K. Beyond Logic Programming for Legal Reasoning. arXiv preprint arXiv:230616632. 2023.

[4] Nguyen HT, Goebel R, Toni F, Stathis K, Satoh K. How well do SOTA legal reasoning models support abductive reasoning? arXiv preprint arXiv:230406912. 2023.

[5] Kim MY, Rabelo J, Goebel R, Yoshioka M, Kano Y, Satoh K. COLIEE 2022 Summary: Methods for Legal Document Retrieval and Entailment. In: JSAI International Symposium on Artificial Intelligence. Springer; 2022. p. 51-67.

[6] Ouyang L, Wu J, Jiang X, Almeida D, Wainwright C, Mishkin P, et al. Training language models to follow instructions with human feedback. Advances in Neural Information Processing Systems. 2022;35:27730-44.

[7] Yoshioka M, Kano Y, Kiyota N, Satoh K. Overview of Japanese Statute Law Retrieval and Entailment Task at COLIEE-2018; 2018. Available from: https://api.semanticscholar.org/CorpusID:208223588.