

Using LLMs and Formal Methods for Legal Reasoning

Ruzica Piskac¹, Scott Shapiro^{1,2}

¹Yale University

²Yale University & Yale Law School

Abstract

Lay users and, in complex cases, legal experts cannot easily understand legal texts and connections between different legal rules. Our work is focused on developing reliable and trustworthy legal chatbots (“lawbots”) to assist with legal inquiries. The paramount challenge with current Large Language Models (LLMs) is their propensity for “hallucinations,” or generating plausible but false or misleading information. To counteract this, we will use automated reasoning techniques. By applying these rigorous methods, we aim to ensure that the responses provided by our lawbots are accurate and reliable, thereby establishing a new standard of dependability in AI assistance for legal experts and democratizing legal knowledge for ordinary citizens.

Keywords novel applications, legal reasoning, compliance, using of automated reasoning in combination with generative AI techniques

Introduction

Legal codes are notoriously challenging to navigate. Consider the City of Portland’s zoning laws [2]. The file containing the zoning code is over 60 megabytes in size and contains tens of thousands of individual regulations and guidelines, each meticulously detailed and often interwoven with numerous others. For legal experts, searching through such an extensive and dense repository of regulations demands a considerable amount of time, expertise, and patience. For laypeople, attempting to decipher and understand these codes can be even more daunting, bordering on the impossible. This is not only due to the sheer volume of the material but also the specialized legal terminology that such documents typically exhibit. The challenge is further compounded by the need to cross-reference or understand the implications of one part of the code in the context of another.

We propose to devise and implement legal chatbots (“lawbots”) to automate the process of legal research and reasoning. Figure 1 illustrates one proposed use of such lawbots. These advanced AI-powered tools will be engineered to sift through vast and intricate legal codes, such as zoning laws, tax provisions, and government benefit regulations, at speeds exponentially faster than even the most skilled hu-

man lawyers. This remarkable capability stems from their ability to rapidly parse, analyze, and cross-reference dense legal texts that are traditionally time-consuming and laborious to navigate.

The current generation of chatbots are notoriously unreliable, with a tendency to “hallucinate,” i.e., assert false statements as true. Our proposal is to ensure the reliability and accuracy of these chatbots through application of automated reasoning. Rooted in rigorous logical and mathematical foundations, automated solvers will enable lawbots to process legal information with a high degree of precision. By systematically verifying and validating each step of the reasoning process, these lawbots minimize errors that can arise from misinterpretation or oversight.

We envision that the proposed lawbots will deliver results that are more accurate and comprehensive and will be delivered faster, than the users trying to find an explanation on their own. This precision is crucial in legal contexts, where the stakes are high and the nuances of law demand meticulous attention.

System Design for Lawbots

An overview of our idea on how to derive efficient lawbots is depicted in Figure 2. The user interacts with a lawbot by asking legal-related questions. In a naive version of a lawbot, an LLM would internally parse a user’s question and search for an answer in the existing body of legal texts. However, the users typically ask more evolved questions where a simple search is not enough. Large language models are notoriously unreliable when it comes to answering questions involving complex reasoning [7]. To overcome this major obstacle, we plan to leverage the power of existing automated reasoning tools, such as satisfiability modulo theory (SMT) solvers [3]. To utilize SMT solvers in answering the users’ questions, we first translate their questions into the first-order logic formalism, such as SMT-LIB [1]. This is the input format used for SMT solvers. To be able to answer the users’ questions by using SMT solvers, we also need to translate legal texts into the same formalism. As the number of legal texts can be really large, this translation is expected to be a time- and energy-consuming process. However, legal regulations change very slowly, so the translation of legal regulations will need to be completed only once.

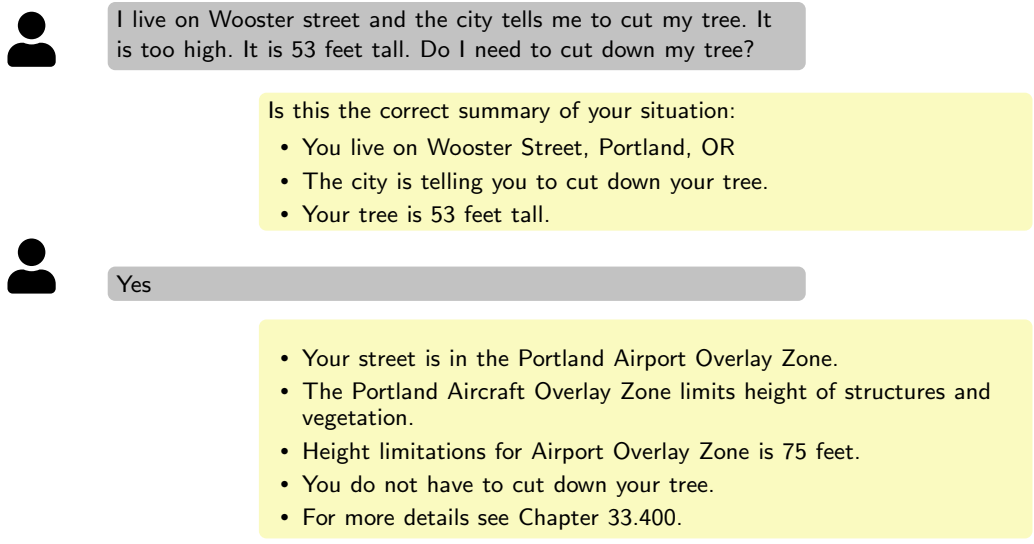


Figure 1: An example use of our lawbots: a user seeking clarification on local tree height regulations from a chatbot tailored to Portland’s zoning law

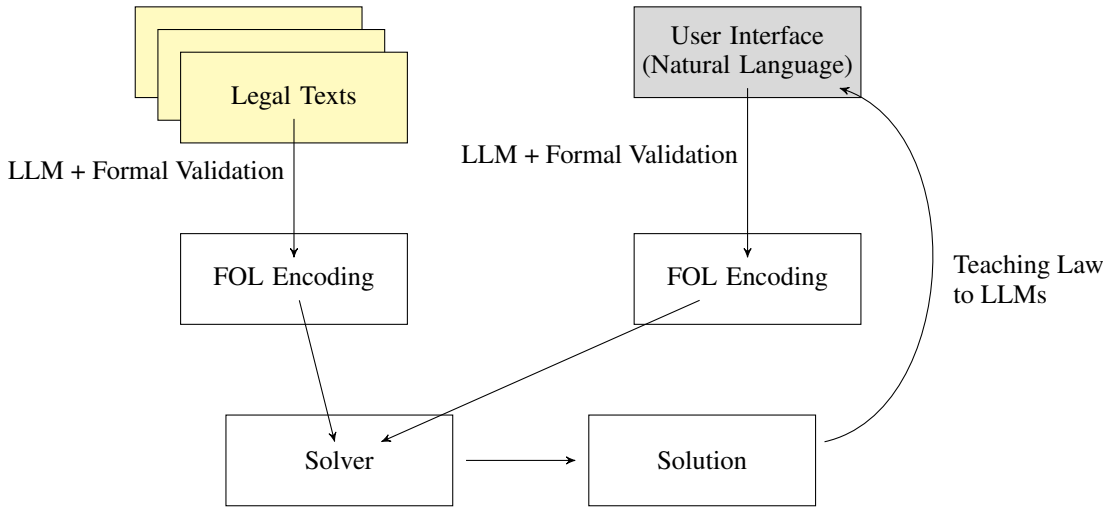


Figure 2: An overview of a system for lawbots

Related Work and Conclusions

Legal scholars have long dreamed of applying Artificial Intelligence to legal reasoning [6]. With the growth of the applicability and importance of generative AI, this dream is coming closer to reality. Even articles in the popular press report on the impact of AI on legal practice [4]. Recently OpenAI introduced a new version of ChatGPT that enables regular users to write their own specialized chatbots, in the form of wrappers that are built on top of ChatGPT [5]. While users may be able to build their own legal chatbots, they still need legal expertise to fine-tune the wrappers themselves.

There are already highly specialized AI-based startups focused on legal domain. Spellbook, tailored for legal work, particularly contract drafting, offers a workflow improvement but is not easily adaptable for legal reasoning about

regulatory materials. Casetext, through its “CoCounsel” feature, facilitates document review, legal research, memos, deposition preparation, and contract analysis, relying on large language models for inference. We note that no startup is using any sorts of formal validation of their results. By using LLMs, which are unreliable inference engines, these existing systems give users no reason to trust their outputs.

We believe that lawbots holds the potential to democratize access to legal knowledge. By breaking down the barriers of legal jargon and the intricacies of legal systems, they make legal understanding more accessible to practicing lawyers, paralegals, and even lay users. This democratization can empower individuals and businesses to better navigate legal systems, understand their rights and obligations, and make more informed decisions without the prohibitive costs and time traditionally associated with legal consultation.

References

- [1] Barrett, C.; Fontaine, P.; and Tinelli, C. 2016. The Satisfiability Modulo Theories Library (SMT-LIB). www.SMT-LIB.org.
- [2] City of Portland, Oregon. 2023. Title 33 Planning and Zoning. <https://www.portland.gov/code/33>.
- [3] de Moura, L. M.; and Bjørner, N. S. 2011. Satisfiability modulo theories: introduction and applications. *Commun. ACM*, 54(9): 69–77.
- [4] Economist, T. 2023. Generative AI could radically alter the practice of law. <https://www.economist.com/business/2023/06/06/generative-ai-could-radically-alter-the-practice-of-law>.
- [5] Heath, A. 2023. OpenAI is letting anyone create their own version of ChatGPT. <https://www.theverge.com/2023/11/6/23948957/openai-chatgpt-gpt-custom-developer-platform>.
- [6] Sunstein, C. R. 2001. Of Artificial Intelligence and Legal Reasoning. *University of Chicago Law School, Public Law and Legal Theory Working Paper Series*, November.
- [7] Sutcliffe, G.; McKeown, J.; and Steen, A. 2023. A Chat with Bard. In Schulz, S.; Korovin, K.; and Rawson, M., eds., *Proceedings of the 14th International Workshop on the Implementation of Logics*, volume 10329 of *Easy-Chair Preprint*, 299–314. Easychair.