

Structured Legal Argumentation with LLMs: A Study in Landlord-Tenant Law

Gregoire Fournier¹ Daniel W. Linna Jr.²

¹Department of Mathematics, Statistics, and Computer Science,
University of Illinois Chicago, USA

²Pritzker School of Law & McCormick School of Engineering,
Northwestern University, USA

March 3, 2025



NORTHWESTERN
UNIVERSITY

Legal
Argumentation
with LLMs:
Landlord-Tenant
Law

Gregoire Fournier
and Daniel W.
Linna Jr.

LLMs and Law

Our approach

Examples

Analysis and
Limitations

LLMs in Legal Contexts

- ▶ Extract structured representations from legal texts to support expert systems [Janatian et al., 2023].
- ▶ Annotate legal texts [Savelka and Ashley, 2023].
- ▶ Explain legal concepts to professionals [Savelka et al., 2023b].
- ▶ Analyze court opinions for the interpretation of legal concepts [Savelka et al., 2023a].
- ▶ Apply tax law [Nay et al., 2023].

Legal
Argumentation
with LLMs:
Landlord-Tenant
Law

Gregoire Fournier
and Daniel W.
Linna Jr.

LLMs and Law

Our approach

Examples

Analysis and
Limitations

Why do we need interpretability in the legal context?

- ▶ **Trust and Transparency:** For legal practitioners and other users to be able to understand the model's reasoning process, verify and trust it.
- ▶ **Accountability:** Judicial decisions must be clear and justifiable.
- ▶ **Expanding access to justice:** To make it easier for laypeople to engage with and assess AI-generated legal content.

- ▶ A framework for legal reasoning (IRAC):
 - Issue,
 - Rule,
 - Analysis,
 - Conclusion.
- ▶ Semi-structured legal reasoning [Westermann, 2024].
- ▶ Structured legal reasoning [Fournier and Linna, 2024].

Pitfalls of LLMs generated legal reasoning

Legal
Argumentation
with LLMs:
Landlord-Tenant
Law

Gregoire Fournier
and Daniel W.
Linna Jr.

LLMs and Law

Our approach

Examples

Analysis and
Limitations

- ▶ **False or misleading legal information:** an LLM may invent laws, precedents, events ...
- ▶ **Lack of interpretability:** What is the LLM's answer based on?
- ▶ **Hard to verify:** Why is the output correct?

Research Question

Legal
Argumentation
with LLMs:
Landlord-Tenant
Law

Gregoire Fournier
and Daniel W.
Linna Jr.

LLMs and Law

Our approach

Examples

Analysis and
Limitations

Main Inquiry

Can LLMs with *Context Augmentation* and *Chain of Thought* prompting generate **accurate, factual, relevant,** and **comprehensive** legal reasoning for **Landlord-Tenant** problems?

Chain of thought

Chain of thought (CoT) [Wei et al., 2023] is a prompting technique that consists in generating intermediate steps to a problem, gradually reaching a final answer. Improvements include:

- ▶ Creating prompts to yield more logical arguments [Press et al., 2023],
- ▶ Adding contextual information at different stages [Trivedi et al., 2022],
- ▶ Leveraging several reasoning paths [Wang et al., 2023].

Chain of thought for legal reasoning

Having the LLM generate the reasoning step-by-step:

- ▶ Allows a user to review and understand the output.
- ▶ Increases the possibility that a user can assess and correct the argument.
- ▶ Provides valuable assistance to laypersons in drafting documents such as demand letters, which can help expand access to justice.

Context augmentation

The screenshot displays the American Legal Publishing eCodeALP website. The top navigation bar includes the American Legal Publishing logo, a search bar, and links for Login, Select Language, and Resources. The breadcrumb trail indicates the location: IL > Chicago > Municipal Code of ... > 5-12-010 Title, purpose and scope. The left sidebar contains a table of contents for the 5-12-010 Title, with the following sections: CHAPTER 5-12 CHICAGO FAIR HOUSING REGULATIONS, CHAPTER 5-10 TENANT OPPORTUNITY TO PURCHASE WOODLAWN PILOT PROGRAM, CHAPTER 5-11 TENANT OPPORTUNITY TO PURCHASE BLOCK (606) DISTRICT PILOT PROGRAM, CHAPTER 5-12 RESIDENTIAL LANDLORD AND TENANTS, and 5-12-010 Title, purpose and scope. The main content area shows the 5-12-010 Title, purpose and scope, with sections on Exclusions, Definitions, Tenant responsibilities, Landlord's right of access, Remedies for improper denial of access, and Landlord's responsibility to. The page also features a search bar, login, and language selection options.

Figure: Chicago's Residential Tenant-Landlord Ordinance (RLTO) is 12-page regulation for landlord-tenant relationships.

We ask the LLM to refer solely to text of the RLTO to avoid giving false legal information.

Legal
Argumentation
with LLMs:
Landlord-Tenant
Law

Gregoire Fournier
and Daniel W.
Linna Jr.

LLMs and Law

Our approach

Examples

Analysis and
Limitations

Prompt to generate legal reasoning

Legal
Argumentation
with LLMs:
Landlord-Tenant
Law

Gregoire Fournier
and Daniel W.
Linna Jr.

LLMs and Law

Our approach

Examples

Analysis and
Limitations

Input Exposition: *[Provide a detailed description of the legal scenario, including relevant facts, context, and specific issues at stake.]*

Figure: Step 1: Input Exposition

Prompt to generate legal reasoning

Legal
Argumentation
with LLMs:
Landlord-Tenant
Law

Gregoire Fournier
and Daniel W.
Linna Jr.

LLMs and Law

Our approach

Examples

Analysis and
Limitations

Input Exposition: *[Provide a detailed description of the legal scenario, including relevant facts, context, and specific issues at stake.]*

Task: Generate structured legal arguments based on the exposition provided.

Figure: Prompt to generate legal reasoning

Tenant and Landlord Laws

Legal
Argumentation
with LLMs:
Landlord-Tenant
Law

Gregoire Fournier
and Daniel W.
Linna Jr.

LLMs and Law

Our approach

Examples

Analysis and
Limitations

Input Exposition: *[Provide a detailed description of the legal scenario, including relevant facts, context, and specific issues at stake.]*

Task: Generate structured legal arguments based on the exposition provided.

Tenant and Landlord Laws considered are the Residential Landlord and Tenant Ordinance of the City of Chicago provided below:

[Text of the RLTO.]

Figure: Context augmentation

Prompt to generate legal reasoning

Input Exposition: *[Provide a detailed description of the legal scenario, including relevant facts, context, and specific issues at stake.]*

Task: Generate structured legal arguments based on the exposition provided.

Tenant and Landlord Laws considered are the Residential Landlord and Tenant Ordinance of the City of Chicago provided below:

[Text of the RLTO.]

Output Format:

Argument 1:

Exposition: Summarize relevant facts from the input that relate to this argument.

Specific Law: Identify a specific law or statute that applies to the scenario.

Why This Law Applies: Explain how and why this law is relevant to the facts presented.

Conclusion: State the conclusion derived from the application of this law to the facts provided.

[Continue generating arguments as necessary, each focusing on a different applicable law.]

Legal
Argumentation
with LLMs:
Landlord-Tenant
Law

Gregoire Fournier
and Daniel W.
Linna Jr.

LLMs and Law

Our approach

Examples

Analysis and
Limitations

Methodology

The output was evaluated by lawyer at the **LCBH**¹, who is an expert in landlord-tenant law, using the metrics:

- ▶ **Accuracy** for a given legal scenario measures how closely the set of generated arguments aligns with the true or expected answer.
- ▶ **Comprehensiveness** measures how well one given argument coherently and concisely addresses the relevant aspects of the input legal scenario regarding the legal requirement cited.
- ▶ **Factuality** assesses whether an argument originates from the text of the RLTO.
- ▶ **Relevance** evaluates whether the argument logically relates to the legal scenario.

Law Center for Better Housing (LCBH)

The LCBH is a *non-profit law firm* that aims at **protecting renters' rights** in the Chicago area.

Rentervention, one of its programs, provides free resources to help tenants enforce their rights, including a *chatbot that helps renters find the correct information* for their issue.

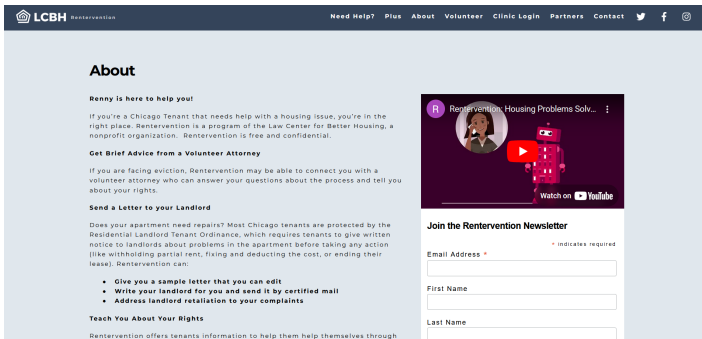


Figure: Rentervention's website.

Legal
Argumentation
with LLMs:
Landlord-Tenant
Law

Gregoire Fournier
and Daniel W.
Linna Jr.

LLMs and Law

Our approach

Examples

Analysis and
Limitations

Scenarios

Table: Scenarios and their sources.

Scenario Number	Description	Source
1	Heating and mold problems	Claude
2	Rent increase	Claude
3	Privacy concerns	Claude
4	Property damage	Claude
5	Cockroach infestation	Rentervention
6	Eviction notice	Rentervention
7	Landlord asking about crime	Rentervention
8	Roof leaking	Rentervention
9	Landlord taking photos inside	Rentervention
10	Wear and tear	Authors

Example 1: Input

2 Scenario 2: Rent increase

2.1 Input

I began renting my one-bedroom apartment on May 1st, 2022, with a reasonable rent for the area. On April 15th, 2023, about two weeks before my lease was set to renew, I received an email from my landlord, Mr. Johnson, informing me that he would be raising the rent by 8% for the upcoming year due to increasing property taxes and maintenance costs.

Caught off guard by this increase, I spent the next few days looking at other rental listings in the area and found that the new rate was still competitive for the neighborhood. On April 20th, I emailed Mr. Johnson to ask if there was any room for negotiation. He responded on April 22nd, offering to reduce the increase to 7% if I signed a two-year lease instead of one.

Example 1: Output

- **Argument 1:**

- **Exposition:** The tenant received an email from the landlord on April 15th, 2023, informing them of an 8% rent increase, citing increased property taxes and maintenance costs. The tenant found the new rate competitive but sought to negotiate. On April 22nd, the landlord offered to reduce the increase to 7% if the tenant signed a two-year lease.
- **Specific Law:** Chicago Residential Landlord and Tenant Ordinance (RLTO) Section 5-12-130(j)(2)
- **Why This Law Applies:** Section 5-12-130(j)(2) stipulates that for any residential tenancy of six months to three years, the landlord must notify the tenant in writing at least 60 days before the termination date of the rental agreement if they intend to increase the rent.
- **Conclusion:** The landlord's email on April 15th, notifying the tenant of a rent increase effective May 1st, does not comply with the 60-day notice requirement. Therefore, the tenant may argue that the rent increase is invalid due to insufficient notice.

This argument is **accurate, relevant, factual** and **comprehensive**. This is what a lawyer would advise their client.

Example 1: Output

- **Argument 2:**

- **Exposition:** The landlord provided a revised offer to reduce the rent increase to 7% if the tenant signed a two-year lease.
- **Specific Law:** Chicago Residential Landlord and Tenant Ordinance (RLTO) Section 5-12-140(a)
- **Why This Law Applies:** Section 5-12-140(a) states that no rental agreement may provide that the tenant agrees to waive or forgo rights, remedies, or obligations provided under this chapter.
- **Conclusion:** The tenant could argue that the landlord's offer to reduce the rent increase contingent on signing a two-year lease may be seen as an attempt to make the tenant waive their right to proper notice of rent increase, thus violating Section 5-12-140(a).

- ▶ This argument is **relevant** and **factual**.
- ▶ However, the interpretation of the law is incorrect; therefore, it is **not accurate nor comprehensive**.

Example 1: Output

- **Argument 3:**

- **Exposition:** The tenant has been renting since May 1st, 2022, and he landlord cited increased property taxes and maintenance costs as reasons for the rent increase.
- **Specific Law:** Chicago Residential Landlord and Tenant Ordinance (RLTO) Section 5-12-150
- **Why This Law Applies:** Section 5-12-150 prohibits landlords from taking retaliatory actions against tenants for exercising their rights or seeking repairs.
- **Conclusion:** If the tenant had previously requested repairs or exercised other rights, the tenant could argue that the rent increase is retaliatory, especially if the landlord cannot substantiate the claimed increases in property taxes and maintenance costs.

- ▶ The argument is **factual**.
- ▶ It is **comprehensive**, since the reasoning is sound.
- ▶ However, it is **neither relevant nor accurate**, as it ignores the main problem to invent another one.

- ▶ 54 out of 55 generated arguments were **factual**. The one mistake is due to the model's poor reformulation of an RLTO section.

- ▶ In eight of the ten scenarios, the generated reasoning was **accurate**. The LLM was able to identify the legal requirement to apply, justify why it was applicable, and reach the correct conclusion.
- ▶ The outputs for Scenarios #7 and #9 were **not accurate**. In both cases, the law to apply is not in the RLTO, which highlights a limitation of our method.
 - Scenario #7: landlord asking for a criminal record.
 - Scenario #9: tenant's right to privacy.

Relevance of arguments

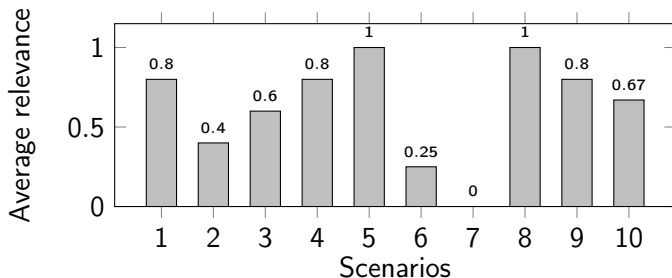


Figure: The average relevance of the arguments generated.

We distinguish two types of scenarios:

- Scenarios #2,6,7 where most arguments were not relevant to the correct legal reasoning;
- The other scenarios where most arguments were.

Question: Can we identify arguments that are much more relevant than the other applicable ones?

Comprehensiveness of the reasoning

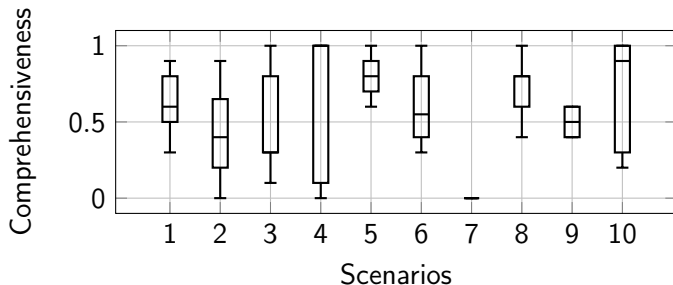


Figure: Boxplot representing the comprehensiveness of the arguments for the ten scenarios.

In every scenario but #7 and #9, the model successfully provided the correct legal reasoning in one of its arguments. The model, however, did not perform well at generating only high-quality arguments.

Conclusion

Key Contributions:

- ▶ We employed *Context Augmentation* and *Chain of Thought (CoT) instructions* with GPT-4o to generate legal arguments for **Landlord-Tenant** issues.
- ▶ We demonstrated the potential of LLMs to produce **accurate, relevant, factual, and comprehensive** legal reasoning.

Identified Limitations:

- ▶ Inability to identify legal issues beyond the provided context.
- ▶ Challenges in assessing the **relevance** of generated arguments.

Directions for Future Research

► Integrating Non-Expert Users:

- Investigate the role non-experts can play in the legal-reasoning process.
- Study models of human-in-the-loop systems to better incorporate diverse inputs.

► Designing Collaborative AI Systems:

- Explore methodologies for crafting AI systems that enhance user interaction and feedback.
- Focus on improving reasoning capabilities through user input in iterative development cycles.

References I



Fournier, G. and Linna, D. W. (2024).

Structured legal argumentation with llms: A study in landlord-tenant law.

In International Conference on Legal Knowledge and Information Systems, Jurix.



Janatian, S., Westermann, H., Tan, J., Savelka, J., and Benyekhle, K. (2023).

From text to structure: Using large language models to support the development of legal expert systems.

In Legal Knowledge and Information Systems. IOS Press.

References II



Nay, J. J., Karamardian, D., Lawsky, S. B., Tao, W., Bhat, M., Jain, R., Lee, A. T., Choi, J. H., and Kasai, J. (2023).

Large language models as tax attorneys: A case study in legal capabilities emergence.

CoRR.



Press, O., Zhang, M., Min, S., Schmidt, L., Smith, N. A., and Lewis, M. (2023).

Measuring and narrowing the compositionality gap in language models.

In Findings of the Association for Computational Linguistics: EMNLP.

References III



Savelka, J. and Ashley, K. (2023).

The unreasonable effectiveness of large language models in zero-shot semantic annotation of legal texts.

Frontiers in Artificial Intelligence.



Savelka, J., Ashley, K. D., Gray, M. A., Westermann, H., and Xu, H. (2023a).

Can GPT-4 support analysis of textual data in tasks requiring highly specialized domain expertise?

In *ICAIL*.

References IV



Savelka, J., Ashley, K. D., Gray, M. A., Westermann, H., and Xu, H. (2023b).

Explaining legal concepts with augmented large language models (gpt-4).


CoRR.



Trivedi, H., Balasubramanian, N., Khot, T., and Sabharwal, A. (2022).


Interleaving retrieval with chain-of-thought reasoning for knowledge-intensive multi-step questions.

References V

-  Wang, X., Wei, J., Schuurmans, D., Le, Q. V., Chi, E. H., Narang, S., Chowdhery, A., and Zhou, D. (2023).

Self-consistency improves chain of thought reasoning in language models.

In The Eleventh International Conference on Learning Representations, ICLR.

-  Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q., and Zhou, D. (2023).

Chain-of-thought prompting elicits reasoning in large language models.

References VI



Westermann, H. (2024).

Dallma: Semi-structured legal reasoning and drafting with large language models.

In *GenLaw*.

Legal
Argumentation
with LLMs:
Landlord-Tenant
Law

Gregoire Fournier
and Daniel W.
Linna Jr.

LLMs and Law

Our approach

Examples

Analysis and
Limitations