GRAPH-GRPO-LEX

Contract Graph Modeling and Reinforcement Learning with Group Relative Policy Optimization

Moriya Dechtiar, Daniel Martin Katz, Mari Sundaresan, Sylvain Jaume, Hongming Wang

IEEE ICDM 2025 DMAIL Workshop | Washington, DC

Harvard, Illinois Tech, Georgetown, MIT

Objectives & Abstract

Represent contracts as semantic graphs using a detailed ontology

Train an LLM to extract nodes/edges via SFT and RL / gated GRPO

Leverage graph metrics (density, depth, centrality) and create the "Contract linter"

sample

Background & Motivation

Contracts are key document that underlie most commercial activity

Contracts come in a variety of flavors including —

- Non Disclosure Agreements
- Property Leases
- Employment Agreements
- Loan Agreements
- Licensing Agreements
- Etc.

Contracts feature clauses, subclauses, etc.

Background & Motivation

Modern Contracts are both lengthy and complex

Both the creation and review of contracts is both time consuming and prone to errors

There has been a variety of efforts to use technology including AI to assist in both the creation and review of contracts

Knowledge Graphs (KGs) have been one vehicle to represent the underlying complexity of legal documents

Background & Motivation

Prior KG work inspires graph-based representations for legal documents

Graph + RL allows automated extraction and analysis, overcoming manual bottlenecks

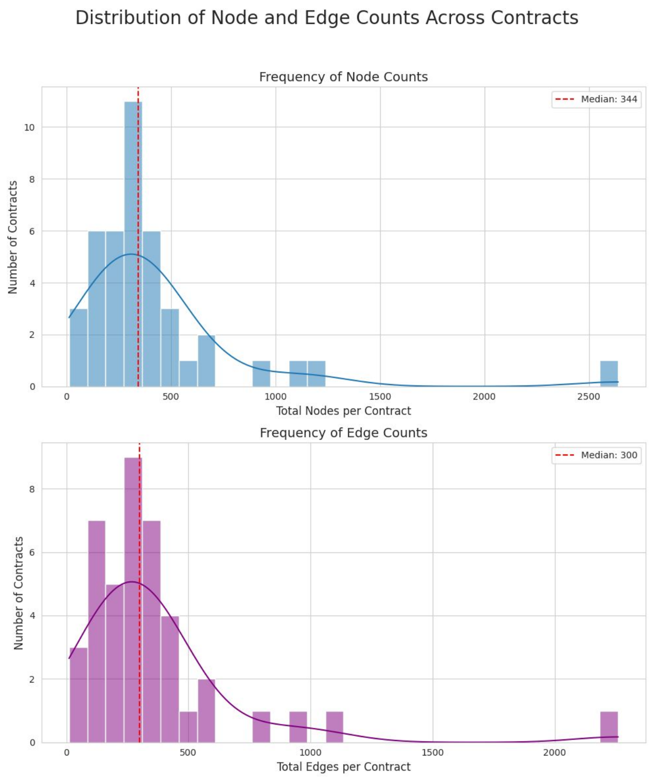
Added benefit of applying graph metrics for quality and risk assessment

Dataset & Corpus

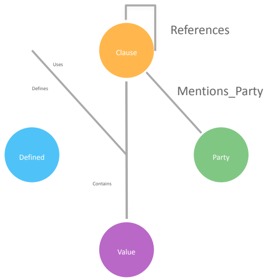
43 CUAD contracts dataset ≈ 1,600 clauses

Alt-test validation:
LLM labels over human annotators
(ω≈0.99, AAP≈0.907)

Statistic	Value
Contracts	43
Clauses	≈1,600
Node types	4
Edge types	6



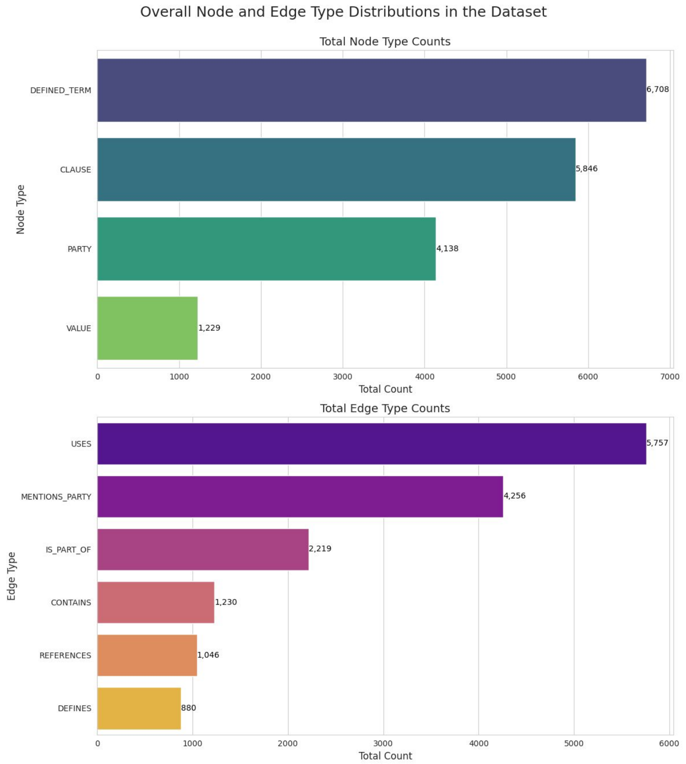
Graph Ontology & Metrics



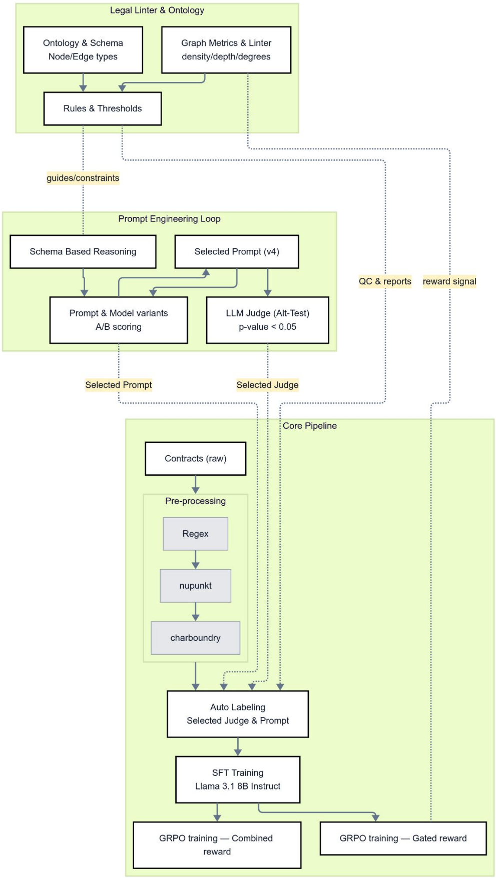
Nodes: Clause, Defined Term, Party, Value

Edges:
Structural (Is Part Of, Contains),
Semantic (Defines, Uses, References)

Metrics: Density, Depth, Centrality, k-Core, Orphans/Leaves, Articulation Points



Pipeline & Methodology



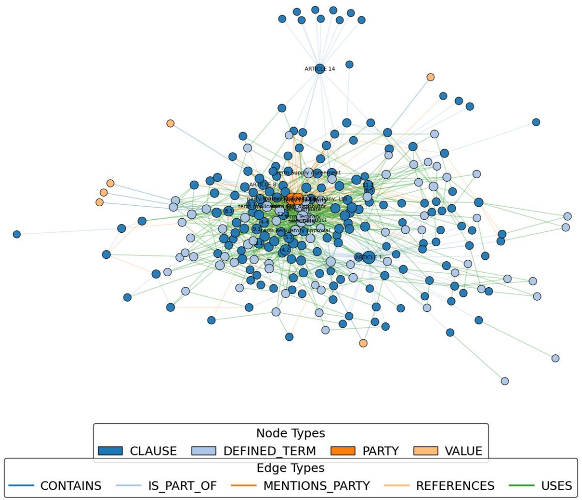
NUPUNKT/CHARBOUNDARY
Prompt Engineering to choose llm
Alt-Test LLM vs Human Labels
SFT Llama-3.1-8B with QLoRA
GRPO Custom Reward Function

[parsing]

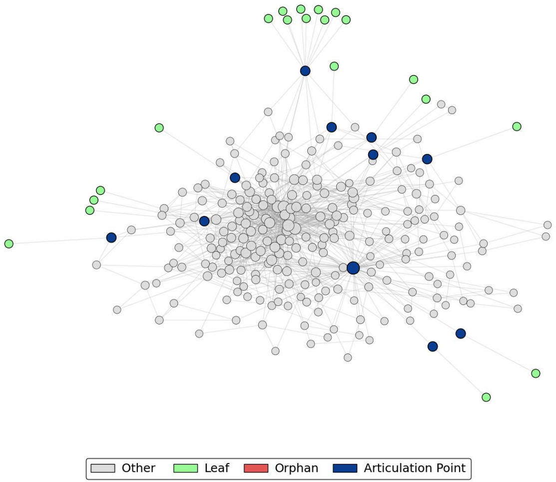
Case Study: Zogenix Agreement

Nodes	Density	Edges	Articulation pts.	Depth	Leaves
257	0.014	916	11	6	97

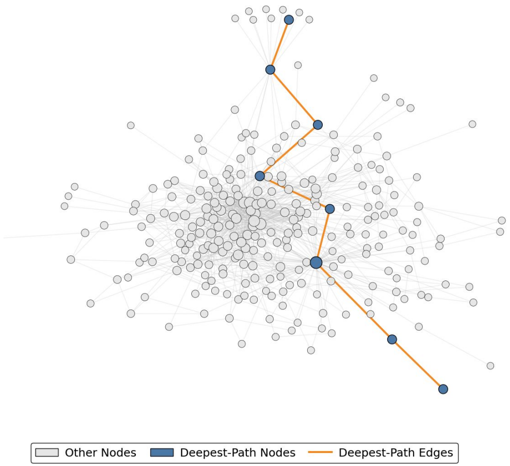
ZogenixInc Distributor Agreement Graph



ZogenixInc Distributor Agreement - Leaves, Orphans, and Articulation Points



ZogenixInc Distributor Agreement - Deepest Path length=7



[sample]

Supervised Fine-Tuning (SFT) Results

Baseline model: Meta-Llama-3.1-8B-Instruct + QLoRA (rank 8, α=16)

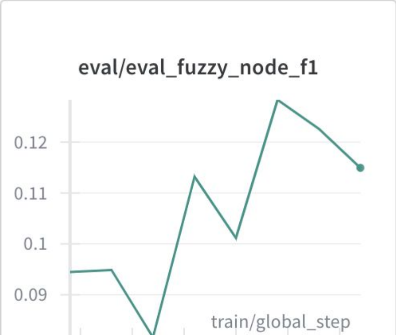
Strict micro-F1: 0.661; Fuzzy micro-F1: 0.693

Invalid JSON rate: 0.0 (deterministic decoding)

Metric	Train	Eval
strict_micro_precision	0.625	0.5394
strict_micro_recall	0.6875	0.8541
strict_micro_f1	0.6547	0.6612
fuzzy_micro_precision	0.6292	0.5569
fuzzy_micro_recall	0.7	0.9166
fuzzy_micro_f1	0.6627	0.6929
invalid_json_rate	0.0	0.0

Challenges & Lessons

High-temperature sampling yields invalid JSON ⟶ lower T to 0.3 and reward partial validity
Greedy decoding over-generates ⟶ add JSON stopper to truncate correctly
Base tokenizer produces gibberish ⟶ ASCII clamp filters non-ASCII tokens
Compute constraints: single A100 GPU; group size limited to 4

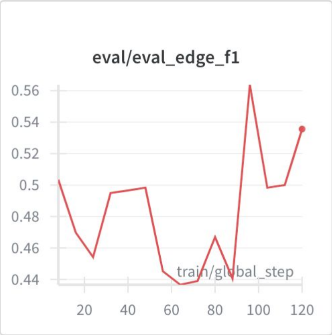


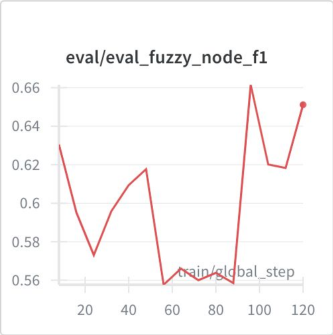
Gated GRPO Training

Composite reward:
structure → strict/fuzzy F1 → embedding similarity → edit distance

Sample size 4; LR=2e−6; T=0.4; Top-p=0.9;
Rep. penalty=1.2; JSON stopper & ASCII clamp

Metric	GRPO-Eval	SFT-Eval
strict_micro_precision	0.806	0.5394
strict_micro_recall	0.790	0.8541
strict_micro_f1	0.798	0.6612
fuzzy_micro_precision	0.817	0.5569
fuzzy_micro_recall	0.802	0.9166
fuzzy_micro_f1	0.809	0.6929
invalid_json_rate	0.02	0.0





Conclusions & Contributions

1. Developed a comprehensive legal ontology
2. Built a clause-centric pipeline for node/edge extraction and assembly
3. Validated Prompt Engineered LLM labels via Alt-test
4. Created rich metrics - The Contrat Linter
5. Designed a composite reward function for contract graphs
6. Introduced gated GRPO training, achieving significant gains over SFT

Future Work & Outlook

Integrate graph-based KGs with RAG frameworks for grounded contract Q&A

Generate Contract Linter Reports for risk assessments

Apply gated RL for drafting assistance and negotiation analysis

Translate Linter metrics to Graph Judges and inspect uncertainty based error signal for selective refinement - Coming Soon!