A Possible Approach to Generalize NLP Models: Semantic Parsing with the “Term and Attribute” Structure

Anonymous ACL submission
(This PDF is converted from Word, the line number cannot be displayed on the right side)

Abstract

With the attempt to convert all legal rules into structured knowledge, the author has manually analyzed more than 20,000 legal rules since 2008. Through the extensive and recurrent analysis and deducing, the author discovered that the essence of human languages is to describe the relationship among Persons, Acts and Objects. Based on this finding, the author developed the “Term and Attribute” structure, which can convert all languages into structured knowledge.

1 Background

The author is qualified to practice laws in China and the State of New York, has some knowledge about natural language processing and knowledge graph but has no computer science background.

The author believed that legal rules are very important for human society but in most cases are too complicated for ordinary people to understand. All legal rules needs to be converted into structured knowledge for the purpose of, (1) to make rules of laws accessible and understandable by ordinary people, (2) to save human resources for the whole legal industry (note: it appears ridiculous when many lawyers race to interpret a new law or regulations, sometimes in different ways), and (3) to avoid ambiguity and duplicity in legal rules.

With this belief in mind, the author began to extensively analyze legal rules since 2008 with the aim to develop a uniform model to convert all legal rules into structured knowledge. As of January 31, 2025, the author has analyzed more than 20,000 legal rules, including legal rules of China, the United States and England.

2 The essence of human languages

Through the extensive and recurrent analysis and deducing, the author discovered that the essence of human languages is to describe the relationship among Persons, Acts and Objects.

There are six relationship combinations among Persons, Acts and Objects, namely,

#	Term 1	Term 2
1	Person	Person
2	Person	Act
3	Person	Object
4	Act	Act
5	Act	Object
6	Object	Object

Table 1: Relationship combinations among Persons, Acts and Objects

3 “Term and Attribute” Structure

3.1 Overview

From the semantic parsing perspective, the words in any language can be divided in two categories, namely, Substantial Terms and Modifier Terms. Substantial Terms means the words used to represent Persons, Acts and Objects. Modifier Terms means the words used to modify Substantial Terms.

Each legal rule describes a specific Attribute of a specific Term, and the value of the Attribute will be other Terms. Distinguished from the Triple in the Knowledge Graph, the Terms in a legal may have reciprocal relations. Namely, Term 1 may be the value of Attribute X of Term 2, and Term 2 may be the value of Attribute Y of Term 1.

3.2 Substantial Terms

3.2.1 Persons

Persons generally refers to any term that is used to represent a subject. For example, “seller”, “buyer”, “director”, “officer”, “company”, “congress”, “president”, “senator” and other persons in the legal rules.

In the real world, Persons can refer to a real natural person or a real entity.

3.2.2 Acts

Acts generally refers to any term that is used to represent an act. For example, “murder”, “sell”, “purchase”, “represent”, “resolve”, “decide”,

"perform", "indemnify" and other acts in the legal rules.

Most Acts in the legal rules and the real world refer to the same act, though an Act may have a specific meaning in a given context.

3.2.3 Objects

Objects generally refers to any term that is used to represent an object. For example, "property", "money", "residence", "procedure", "law", "rule", "judgment", "case", "causation" and other objects in the legal rules.

Most Objects in the legal rules and the real world refer to the same act, though an Object may have a specific meaning in a given context.

3.3 Modifier Terms

Modifier Terms can be roughly understood as Attributive, Adverbial and Complement word used to modify Substantial Terms. For example, "reasonable", "foreseeable", "necessary" and other modifier terms in the legal rules.

Most Modifier Terms in the legal rules and the real world refer to the same act, though a Modifier Term may have a specific meaning in a given context.

3.4 Meta Terms and Composite Terms

Meta Terms can roughly be understood as an origin of an specific expression, and in most case a single word in English language , for example, "company", "chairperson", "contract", "partnership" and "knowledge".

Composite Terms refers to the Terms that come from the Meta Terms with one or more rounds of combinations. (Note: the internal structure of this composite term is not explained in this Paper)

#	Example	Type
1	Company's chairman	Person
2	Defaulting party	Person
3	Formation of company	Act
4	Company provides security	Act
5	Refuse to grant	Act
6	Abuse power	Act
7	Legal service	Object
8	Foreseeable result	Object

Table 2: Example of Composite Terms

3.5 Attribute

3.5.1 Example 1

Legal Rule: a president candidate must be at least 45 years old.

This rule describes the age requirement of a president candidate, the value is ">= 45 years"

Term 1 = Person = president candidate

Term 2 = Object = ">= 45 years"

Term 2 is the "Age Requirement" of Term 1.

3.5.2 Example 1

Legal Rule: if Seller delays in delivery, it shall compensate Buyer for losses

This rule describes the consequence of "Seller delays in delivery" as well as the condition for "Seller shall compensate Buyer for losses"

Term 1 = Act = "Seller delays in delivery"

Term 2 = Act = "Seller shall compensate Buyer for losses"

Term 2 is the "Consequence" of Term 1, and Term 1 is the "Condition" of Term 2

4 Application

4.1 Interpretation of a given text

As this approach has been applied to more than 20,000 legal rules, it can be reasonably contemplated that it can be applied to convert any given text into structure knowledge.

4.2 Advanced Knowledge Graph

It can be used to interpret all legal rules in any given jurisdiction to generate the advanced knowledge graph, which can be further used to answer legal questions.

5 Limitation

As mentioned this approach was just be manually implemented by the author, but the data are stored in a computer readable format, significant work needs to be done before it can be implemented by language models.

References

No reference is used as this approach is totally new.