Legal Semantic Web - A Recommendation System

Gaurav Kant Shankhdhar, Pursuing M.Tech, Dept. of Software Engineering, Babu Banarasi Das University, Lucknow, UP, India,

V.K. Singh, Asst. Prof. & Head - I.T., Babu Banarasi Das National Institute of Technology & Mgmt, Lucknow, India,

M. Darbari, PhD, Assoc. Prof., Department of CSE,BBDU, Lucknow, India,

ABSTRACT

The project Legal Semantic Web - A Recommendation System makes use of the Semantic Web and it is used for the proactive legal decision making. With the help of web semantics, a lawyer handling a new case can filter out similar cases from the court case repository implemented using RDF (Resource Description Framework), and from here he can extract the judgments done on those similar cases. In this way he can better prepare himself with similar judgments in his hands which will guide him to an improved argumentation. The role of web semantics here is that it introduces intelligent matching of the court case details. The search is not only thorough but also accurate and precise to the maximum level of attainment with the use of ontology designed exclusively for this purpose.

Keywords


1. INTRODUCTION

There have been considerable changes in the field of law in the past decades which have made the area of law more complex. As a result, practice of law, sentencing and legal drafting have grown in pace. It is pertinent to mention here that 5 year growth percentage of lawyers is 19% [1] (see figure 1). Tot. advocates and Adv. Enrolled are in multiple of ten thousands. And also, the production of statutes, codes, legal rules and rulings... - has also increased.

Step 1: To search the vast corpus of law and identify the rules in it, that is relevant to the given legal query. Rules include sections and proposed punishment.

Step 2: Interpret their legal meaning by giving due consideration as to how different rules interact and how they have been interpreted in the past.

Step 3: Consider how those rules apply to the specific query. Lawyers use search engines and commercial databases to deal with information in an automatic way, and as such helps to reduce the information overload, while retaining the relevant information.

The lawyer’s most common task is to give legal advice and perform argumentation in courts. The steps which the lawyer performs in conducting this are [7]:

a): Management of the complexity and types of legal knowledge (related to the law and law objects).

b): Practical and reasonable ways to store, retrieve and structure a monolithic legal information pertaining technical people.

This has led to two of the main problems which are

Fig 1: 5 Yr Growth % of Lawyers in India

The legal domain has been in the past and even up till now has been dealt with the application of Artificial Intelligence amounting to huge contribution. According to Pamela N. Gray in 1997 [2] mentioned that the theory and the tools of Artificial Legal Intelligence had developed in corresponding leaps, with the following progression of themes: (i) legal language, (ii) deontic logic, (iii) rule processing, (iv) case processing, (v) Stratification of reasoning, (vi) procedural reasoning, (vii) co-ordination of multiple tasks.

Knowledge Engineering[3] was also the source for the growing interest in the development and use of legal ontologies [4].

This research is mainly related to the use of The Semantic Web[5][6] and its application to enable a proactive legal recommendation for the lawyer. Lawyers, handle information in order to filter out the relevant rules and regulations, acts and sections and anticipate what the future judgment might be. This is done in order to get a better understandability of the case at hand and prepare them with better argumentation in court. Because they are dependent on such a huge and varied legal corpus they are vulnerable to the Information Overload phenomenon. The Semantic Web aims to enable machines to deal with information in an automatic way, and as such helps to reduce the information overload, while retaining the relevant information.

Legal Web, Legal Intelligence had developed in corresponding leaps, with the application of Artificial Intelligence amounting to huge contribution. According to Pamela N. Gray in 1997 [2] mentioned that the theory and the tools of Artificial Legal Intelligence had developed in corresponding leaps, with the following progression of themes: (i) legal language, (ii) deontic logic, (iii) rule processing, (iv) case processing, (v) Stratification of reasoning, (vi) procedural reasoning, (vii) co-ordination of multiple tasks.

Knowledge Engineering[3] was also the source for the growing interest in the development and use of legal ontologies [4].

This research is mainly related to the use of The Semantic Web[5][6] and its application to enable a proactive legal recommendation for the lawyer. Lawyers, handle information in order to filter out the relevant rules and regulations, acts and sections and anticipate what the future judgment might be. This is done in order to get a better understandability of the case at hand and prepare them with better argumentation in court. Because they are dependent on such a huge and varied legal corpus they are vulnerable to the Information Overload phenomenon. The Semantic Web aims to enable machines to deal with information in an automatic way, and as such helps to reduce the information overload, while retaining the relevant information.

The lawyer’s most common task is to give legal advice and perform argumentation in courts. The steps which the lawyer performs in conducting this are [7]:

Step 1: To search the vast corpus of law and identify the rules in it, that is relevant to the given legal query. Rules include sections and proposed punishment.

Step 2: Interpret their legal meaning by giving due consideration as to how different rules interact and how they have been interpreted in the past.

Step 3: Consider how those rules apply to the specific query. Lawyers use search engines and commercial databases to deal with step 1. Step 2 and 3 are still mostly carried out by the lawyers.

Eg: if there is a query for “murder by a sharp weapon”. The lawyer in step 1 will first search for the rulings for this query taking help of a Legal Library or a search engine that will return the related acts and rules. This is accomplished by help of computers.

But in step 2, the task of lawyer begins when he compares and contrasts the various rulings and finds the relations between the
rulings. Like for “murder by a sharp weapon” previous judgments might be “life imprisonment” or “sentence to death”.

The lawyer goes deeper in researching to find out in which cases there is “life imprisonment” and in which cases it is “sentence to death”. He finds out what sections were imposed on these two separately. In step 3 he decides what can be the appropriate rulings that can be applied to the new case. All this is done by the lawyer himself searching the journals, books and other legal paraphernalia without the aid of any computerized or electronic media that can help the lawyer in his search for interrelated rulings, acts and regulations.

2. LITERATURE REVIEW

The aim, here, behind this literature study is in order to know of the previous research done in this direction. After examining several studies dedicated to the topic, it was concluded that most of the studies dealt with broader topics with references to most adequate information to build systems relating to Legal Advice or Legal Recommendation. This section mainly covers various reasoning approaches in legal domain and ideas to incorporate reasoning in the Legal Searches.

2.1 Types of Legal Analysis Reasoning Approaches [8]

The reasoning approaches usually made use of in legal reasoning process are:

a) Rule based analysis  
b) Precedent analysis or Analogy method  
c) Textual Reasoning & Legislative Intent  
d) Policy Based Reasoning  
e) Tradition Reasoning

2.1.1 Rule Based Analysis

In Rule base reasoning an answer is reached by establishing and applying a rule of law. It declares “X is the answer because that answer will encourage desirable results for our society and discourage undesirable results [8]. Rule based analysis can derive from a case or a statute. Break the rule into the separate elements to be established and then match the facts and circumstances of problem with each element of the rule to see if the element is proven.

2.1.2 Precedent Analysis or Analogy Method

Analogical reasoning reaches an answer by showing direct factual similarities between governing case law and the clients’ facts. It declares “R obtained as answer due to facts of this case are such as X vs. Y and R was the result there” [8]. Analogy Reasoning draws parallels between the factual situations and the cases that have already been decided. Cases consists of facts, so compare and contrast of facts and circumstances of case with decided cases lead us to similarities between the two. If key facts are similar, the analogy is drawn. If the key facts are varied then a different rule must be applied.

Reasoning by analogy is most often used in case analysis. Case analysis is a method used to predict the applicability of prior opinions to present controversy.

2.1.3 Textual Reasoning & Legislative Intent

In this reasoning read and reread the statute and focus on its exact language. As in the rule based reasoning break the statute into separate elements to establish and then match the facts and circumstances of the problem with each element of the statute to see if the element is proven. Once the statute has been broken into elements, the statutory words must be understood and interpreted [8].

2.1.4 Policy Based Reasoning

In Policy based reasoning the most appropriate answer is selected keeping in mind which answer would be the best for the society at large. It declares “X is the answer because that answer will encourage desirable results for our society and discourage undesirable results [8]. Policy arguments appeal to future consequences that follow from adopting a certain rule. The court first predicts the consequences that will flow from giving the law one interpretation or another then declares which set of consequences more consistent with underlying values of law [9]. Policy arguments are found both in case and statute but have few parameters.

2.1.5 Tradition Reasoning

Tradition reasoning achieves an answer by justifying a story that calls forth that result. It declares “X is the answer because that is the way things have been always been done”. The U.S Supreme court has identified “tradition” as a principal test for determining our fundamental rights. Justice Scalia has authored “at least 53 opinions that relied expressly on tradition to resolve constitutional issues [8][10].

2.2 Ontology for the Legal Domain

The legal ontology applications range from information systems to knowledge based systems. The level of detail of knowledge representation is directly proportional to the level of ‘intelligence’ required. Some examples of applications of legal ontologies are:

2.2.1 Information retrieval. By encoding knowledge with meaning, concepts and the relations, it greatly empowers the users of information retrieval systems. Information about level of specificity of concepts can help the user to find information relevant to his query. Examples of relevant publications and projects are Matthijssen (1999), who introduces an interface between the lay user and a legal database, LOIS (which stands for ‘Lexical Ontologies for legal Information Sharing’, cf. Dini et al. 2005) and BEST (which stands for ‘Batna Establishment using Semantic web Technology’, cf. Van Laarschot et al. 2005). [11]

2.2.2 Translation of legal documents. By making the meaning of legal terms explicit, it can help in translating legal documents from one language into another one. A framework for this specific purpose has been developed by Termorshuizen-Arts (2003). [11]

2.2.3 Automated classification and summarizing. Each document or entity is given a unique identification which makes it easy to be classified.

2.2.4 Decision support and decision making. Legal (procedural) regulations often contain decision structures that allow making certain decisions or qualifications. Although such structures can be modeled in relatively simple decision trees, such decision trees still require user intervention on making a choice in each step. An ontology can be used to encode not only the decision steps, but also the contents of the decision rules. Advantages of using an ontology in such a case are supposed to include consistency of the modeling activity result and the re-usability of the underlying ontology for other modeling activities. Although the models underlying case-based systems are seldom called ‘ontologies’, they can be regarded as such. The model underlying a sentencing system described in Oskamp (1998)[11] is just one example of this. It contains a model with two main constituents: facts and factors. Factors are
subjective qualifications of (sets of) objective facts. By the
type of these entities, they constitute an ‘ontology’ of the
arguments underlying sentencing decisions.

2.2.5 Agent technology. Agents are assumed to allow for
intelligent autonomous communication between different
computer systems. For such communication, the modeling of
rules governing that communication is necessary. As in the case
of decision support and decision making, such modeling can be
supported by an underlying ontology. Potential practical
applications of agents in the legal domain are automated dispute
resolution by negotiation and the controlled exchange of
sensitive data, for instance in electronic legal record.

The present work is also inspired by The Double Role of
Ontologies in Information Science Research, 2007,[12], The
European Legal Semantic Web [15], Building a Semantically
Rich Legal Case Repository in OWL [4].

3. PROPOSED SOLUTION

3.1 Where Semantic Web Comes In

Semantic Web will enable applications to dig deeper into steps
2 and 3 and will partially help the lawyer in advance to know
what will be the near judgment.

As in the legal ontology, Rules will be related through their
interactions. There will be sections and their meanings along
with the relationship with other sections eg., section 304-B, 306
and 309 of The Indian Penal Code [12] are related by being
sections related to “suicide”, so a search for a section will also
include other related sections. So in this way the lawyer will be
able to get the interactions between the rulings and also how the
rulings have been interpreted in the past. All this is done by a
well knit ontology made through Protege 4.2. Finally the graph
consisting of the relevant Rules. Sections imposed along with
their interrelatedness is extracted through SPARQL [13] query
and is used to get the required data needed by the lawyer.

3.2 A Bit More of Semantic Web in Relation
to Law

Leveraging the higher degree of organization of legal data
resting on top of the legal ontology and the possibility of
drawing inferences from the data, a semantic legal query system
should be able to do more than merely retrieve information
based on keywords selected by a human agent. In a world of
perfect formalization, an application could carry out the
interpretation of the search items and the logical extraction of
records autonomously. With a search tool such as existing
common search engines or indexed searches made available by
the current Indian legal search engines, the legal researcher can
input key words into a search and get in return a (usually long)
list of documents which contain, or are indexed by, those key
words. The present legal knowledge data is in natural language,
which machines do not understand, a machine does not
understands properties and relationships so for the legal
researcher must still manually examine the documents to find
the substantive information. The need is the searchers like, for
example, a legal researcher searches for cases in which a
company has the property of being in the role of applicant or
appellant or where a lawyer is in the relationship of
representing a client.

There is also a need for multi-dimensional search in which one
thing can be searched that may be governed by other factors
also. Eg., There should be a search engine delivering not just
those judgments which are relevant for IPC 302, (here it will
give the searches of all the document URIs which have IPC
302), but just those judgments that refer to this legislation as
valid on a specific point in time like “murder by strangulation
between 1990 and 2000” [14]. Here on one axis lie the murder
cases and on the other lie the time period.

The underlying problem is that legal textual information is
expressed in natural language. What literate people read as
meaningful words and sentences appear to a computer as just
strings of ones and zeros. While a search can be made for the
string Applicant or Appellant, there are no (widely available)
searches for a string that represents an individual or company
who bears the role of Applicant or Appellant. To make
language on the Web more meaningful and structured,
additional content must be added to the source material, which
is where the Semantic Web and Natural Language
Processing come into play. A semantic web RDF[15] graph
grows inside out. It expands as it comes in conjunction with
other graphs, so in the world of semantic web there may be an
expanding net of relationships and properties which make for
the understandability of the machines.

Fig 2: shows a scrap of how the structure of entities is laid
down using RDF. Here it is important to mention that now
the machines know the relationships and properties of the
objects and can perform intelligent searches.

The problem can be solved by the use of triples[16]. A triple
has a “subject”, a “property” and an “object” The Applicant,
here a subject has all the companies connected by the property
hasCompanyRegNo” which has “CompanyRegNo” as an
object. Besides this there are also triples consisting of:
Company----hasRegNo----RegNo, and Company----
hasName---
Name

So, from the applicant, the company registration number is
fetched and by backward tracking from company registration
number, the company subject is obtained from which the
company name is retrieved.

3.3 IPC Sections

Legal semantic web- A Recommendation System will first do
the keywords searching which is the task as specified in step 1.
The legal meaning is interpreted by the RDF triples containing
the case details. One simple example is that when a lawyer
searches for “section 302” will also get the results for the
intrinsic search for “murder” and vice versa. Similarly by
searching “section 307” will get the results for “attempt
to murder”. This is done because these searches mean the same
thing. In this way the machine understands the inherent
meaning of the query. This can also be stated like if there is no
provision for searching “murder”, the query for “murder” will
result in nothing. If the need is to make the search flexible
enough to search for legal sections like “302” as well as
sections written in English Language, like “Murder” or
“Attempt to murder”, the need is to establish some relationship
between these entities. This is done as follows: if there is a
search for “murder”, it returns the SectionID through which
CriminalCaseID is obtained and the required case is returned.
Similarly if “302” is searched, the SectionId and through it, the
CriminalCaseID is returned. This accomplishes the step 2 and 3 as specified above. This is shown in Figure 3.

4. MAKING A SEMANTICALLY RICH LEGAL ONTOLOGY IN OWL[17]

The retrieval of conceptual information from legal documents depends on the construction of a knowledge representation of the document. A number of interesting research works on legal case knowledge representation have been proposed including frame–like structures, semantic nets and dimensions. However some limitations exist in these works. For instance, some render little inference capabilities, some ignore contextual information essential to conceptual retrieval and some give no consideration to semantic interoperability. This work addresses these limitations by using an open standard ontology language and refined ontology architecture. The ontology is easy to maintain, reuse, extend and renders rich inference and reasoning capabilities. In addition, a framework is proposed in our work in order to integrate heterogeneous knowledge representations and to reduce the manual effort required for the annotation process by enabling semi-automation of the annotation process.

4.1. Legal Case Ontology Design

The model created by our approach encodes both domain dependent and domain independent semantics into the case text. Contextual features are well captured therefore richer information is disclosed to applications to enable conceptual searching. In addition, OWL is used in coding our model in considering that it is an open and standardized web ontology description language and it has a rich set of available inference mechanisms. This section details the four steps used in our approach in building a legal case ontology.

4.2 Grouping similar concepts

On the topmost lies the murder case details like applicant/appellant, opposite party, date of judgment, intention of crime, IPC sections imposed, judgment, weapon if used and other case attributes. Each section is uniquely identified by a SectionID and SectionID is related to SectionNUM and SectionDesc which hold the IPC section number like 324 IPC and its description like “Voluntarily causing hurt by dangerous weapons or means”. The Sections are mainly grouped into 3 categories:

4.3 Identifying Classes and Properties

A class provides an abstraction mechanism for grouping resources with similar characteristics [18], whilst a property is often used to identify the non hierarchical relationships between domain and range (denoted as R(Domain,Range)). OWL defines two types of properties: data property and object property. Data property is an alias of attribute while object property is a binary relationship between two classes.

4.4 Classes of Legal Ontology : This Legal Ontology shows the relation between super class and sub class where Thing represents Legal management systems.
4.5 **Properties**: Properties are instances of the class rdf:Property. In the RDF graph, the property represents the predicate and describes a relation between subject resources and object resources.

The most important class of the Legal Ontology Hierarchy (Figure 4) is the ID class as it connects all the other classes. ID uniquely identifies a murder case, e.g., 1001

**ID** connecting other classes through properties is shown in the following figure 7.

4.6 **Relatedness of Sections**: Sections and their descriptions are held by section ID, e.g., sec-id1, sec-id2 and so on. So the interrelatedness of the sections is also the interrelatedness between section IDs. Here in the figure sec-id1 holds section 302 IPC and also its description that is “murder”. The related section IDs are linked to each other by a property called “hasRelatedSEC_ID”.

---

**Fig 4**: Classes In The Legal Ontology

**Fig 5**: Properties or Predicates in the Legal Ontology

**Fig 6**: Property Assertions in the Legal Ontology
4.7 Weapons Hierarchy

The above figure 8 shows the weapon Ontology with various classifications.

4.8 Sparql Query: SPARQL query for related sections to IPC 302

(The output is shown in figure. 9)
SELECT *
where {
?element
<http://www.semanticweb.org/rs/ontologies/2013/9/CriminalCase-25#hasSectionNum> "SEC_NUM".
?element
<http://www.semanticweb.org/rs/ontologies/2013/9/CriminalCase-25#hasSectionID> ?Related_Sec_ID.
}
SELECT ?Related_Sec_ID
where {
?element
<http://www.semanticweb.org/rs/ontologies/2013/9/CriminalCase-25#hasRelatedSEC_ID> ?Related_Sec_ID.
?element
<http://wwwsemanticweb.org/rs/ontologies/2013/9/CriminalCase-25#hasSectionNum>
}

4.9. Sparql Querying for finding case details including judgment. Here, section IPC 302 is taken. (The output is shown in figure. 10)
where {
?element
<http://www.semanticweb.org/rs/ontologies/2013/9/CriminalCase-25#hasID> ?caseID.
?element
?element
?element
?element
}
SELECT ?Related_Sec_ID
where {
?element
<http://www.semanticweb.org/rs/ontologies/2013/9/CriminalCase-25#hasRelatedSEC_ID> ?Related_Sec_ID.
}
5. CONCLUSION

The paper highlights generation of Legal Semantic Web Recommendation System which uses the concept of Semantic Web for the proactive legal decision making.

With the help of web semantics, a lawyer handling a new case can filter out similar cases from the court case repository implemented using RDF (Resource Description Framework) and can extract the judgments done on similar cases. The lawyer is prepared with similar judgments lists which can guide him to an improved argumentation.

The role of web semantics here is to intelligently match the court case details. The search is not only thorough but also accurate and precise with the use of ontology designed using Protégé.

6. FUTURE SCOPE

The future scope of our work is to apply the potential of Knowledge Representation[19,20,21] along with reasoning in the Web context. The important thing here is to detect the possibilistic outcome. In future we will try to integrate the records of Indian courts (starting from small session courts to district to High Courts finally leading to the Supreme Courts) gets collaborated and mapped through legal ontologies. This knowledge floating over the internet will be priceless and enable the legislative, judicial and even common man to draw conclusive legal output that it needs.[22]

7. REFERENCES

[11] L. Momers (B) Centre for Law in the Information Society, Leiden University, Leiden, The Netherlands e-mail: l.momers@law.leidenuniv.nl
[18] WWW.w3.org/TR/owl-features as on 5-1-2014.