Using Legal Ontology for Query Enhancement in Generating a Document Summary

M. Saravanan., B. Ravindran and S. Raman

Department of Computer Science & Engineering, IIT Madras, India

Abstract. In this paper, we have proposed a novel comprehensive structural framework for the construction of ontology from a given legal corpus for the purpose of query enhancement to pick relevant documents for generating a summary. We evaluated our system with queries given by legal experts and non-experts and found that ontology-based query enhancement system results are significantly better than standard Microsoft Windows search queries.

Keywords. Legal Ontology, Information Retrieval, Query enhancement, Text Summarization

Introduction

In this paper, we explore an interesting issue related to the Information Retrieval (**IR**) task of processing of user's query to pick the relevant judgments. The retrieved documents are summarized with term distribution model [1] and the results are presented in the form of a user understandable summary. In this work, we describe the construction of legal ontology that is useful in designing a legal knowledge base to answer queries related to legal cases [2]. The purpose of the knowledge base is to help in understanding the terms in a user query by the way of establishing a connection to legal concepts and exploring all possible related terms and relationships. Our proposed ontology-based document summarizer can assist the legal community to read the gist of the related cases instead of full judgments that have bearing to their present case. We have considered three sub-domains: rent control, income tax and sales tax for the legal ontology construction. The basic components of our newly created ontology structural framework are given in Table 1.

Table 1. Basic components of legal ontology framework

Components	*Description of components*
Groups	Sub-domains considered in this study
Person	Two contending parties appear before the court to resolve their conflict; one of the parties will be the subject. The object will be either a thing or a person.
Things	An object can be either a thing or animate beings including humans. The thing can be either corporeal or incorporeal. Subject action on the object or with respect to object has an impact or consequence on the other contending party.
Event	The ultimate /last facts in a process (series) of facts that have given rise to the conflict/that have triggered conflict between the contending parties. The conflict is about the duty/obligation (or right) of contending parties.
Facts	This constitutes the process. Subject's actions in relation tot an object consequently constitute the process that has or is likely to have or is apprehended to have an impact on the contending parties. Facts and circumstances of the case that are serious / relevant make up the process.
Acts	Courts always deal with application of law to the give facts. The law applied is extracted from the relevant statutes / provisions / rules / regulations / articles / judicial interpretations that are called acts.

1. IR Oriented Legal Ontology

It is now widely recognized that constructing a domain model, or ontology, is an important step in the development of knowledge based systems. The newly created ontology follows an innovative framework which covers all basic details available in a legal document. The components of our framework were identified and initial ontology has been created based on the dictionary of terms and case ontology created in the early stage. Case ontology defines a hierarchical structure which is composed of the concept described in legal cases. To develop this, we had several discussions with legal communities and also on experiments with many documents related to different sub-domains for generalization performances. An initial ontology has six important top-level components of knowledge in the legal domain – *group*, *person*, *things*, *event*, *facts* and

acts – which are defined in Table 1. We present a top-down approach in the construction of sub-components by using domain knowledge. A top-down approach starts with the definition of the most general concepts in the domain along with subsequent specialization of those concepts. There are different kinds of relations like *is-a, related-to, composed-of* etc., which are used in the formation of entire ontological structure to describe the relationship between terms with other terms and its semantics. An ontology together with a set of individual instances of classes constitute a *Knowledge Base*. During the time of construction of ontology, we have included the features such as simple word features, handling of multiple words, different words with same meaning, word with multiple meaning and also considered the high level abstraction instead of low level abstraction of terms defined in the ontological structure. We implemented this ontology in protégé (http://protege.stanford.edu), a graphical ontology editor tool. We have used many of the plug-ins of this tool for different presentations and that also stores the knowledge base in XML representation. We have also implemented a software environment to help a legal user to query the knowledge base with his experience to extract the relevant judgments that have bearing to their present case.

Ontology evaluation is an important issue that must be addressed if ontologies are to be widely adopted in the information retrieval applications [3]. In general, ontology has been employed to achieve better precision and recall in the text retrieval systems [4]. So we have employed Precision (**P**), Recall (**R**) and F-measure (**F**) to evaluate the results of our method with human generated ideal search results and also compared with query results of Microsoft Windows search technique which has been considered as a baseline in this paper. The findings which are given in Table 2 show that our ontology-based methods (with and without query enhancements) yields a relative improvement of 25% compared to the baseline on legal user queries and 15% on non-legal users. This difference in success rate is due to the usage of complex legal terms by experts on their query. The integration of many features for a term in the newly developed knowledge base shows the excellent improvement in query results. The improvement of results produced by ontology-based method over the baseline appears to be statistically significant in all measures considered in this study. It has been observed that there is an increase in precision score in ontology-based with and without query enhancement system. This indicates the expert usage of our software environment by legal experts. The legal experts are given an option to choose the framework terms which are relevant to their query. This may be a reason for getting a better precision score in our methods. At the end, we have employed a term distribution model [1] for extraction of important sentences from the retrieved documents to generate coherent and concise outputs that incorporate important features of the legal document, namely the *ratio decidendi* and *final decision*.

Table 2. Precision, Recall and F-measures for comparison of methods

Method/type of user	**Legal User**			**Non-legal user**		
	P	**R**	**F**	**P**	**R**	**F**
Baseline	0.561	0.724	0.632	0.683	0.772	0.724
Without enhancement	0.718	0.917	0.805	0.834	0.842	0.837
With enhancement	0.829	0.967	0.893	0.879	0.919	0.899

2. Conclusion

Our ontology-based methods consistently outperform start of the art search method, even when it is forced to work for all three sub-domains considered in this study. An intuitive explanation for the performance of our ontology-based system is that it provides a knowledge base which had a huge collection of terms and its relationships and other related features which is used for better enhancements of query terms. Our structural framework can also be expanded with the addition of terms by adding new documents and from different sub-domains in future course of time.

References

[1] M. Saravanan, B. Ravindran, and S. Raman, Improving legal document Summarization using graphical models, *Legal Knowledge and Information Systems, JURIX 2006: The Nineteenth Annual Conference*, Paris, 2006, IOS Press, pp.51-60.

[2] A. Valente and J. Breuker, Making Ends Meet: Conceptual Models and Ontologies in Legal Problem Solving, *Proceedings of the XI Brazilian AI Symposium (SBIA'94)*, 1994, pp.1-15.

[3] Asunción Gómez-Pérez. Ontology Evaluation, *Handbook on Ontologies*, (2004), 251-274.

[4] Nicola Guarino, Formal Ontology and Information Systems, *Proceedings of FOIS'98*, Trento, Italy, 6-8, June 1998, IOS Press, Amsterdam, 1998, 3 -15.