Challenges and Related Issues for Building Chinese Legal Ontology

Zhang Ni
Library of Sichuan University
Chengdu, P.R. China
e-mail: zhangni77@yeah.net

Pu Yifei
College of computer Science
Sichuan University
Chengdu, P.R. China
Puyifei_007@163.com

Wang Ping
Library of Sichuan University
Chengdu, P.R. China
e-mail: wangping@scu.edu.cn

Abstract—Legal ontologies have made significant progress during the past 20 years, however, few of these ontologies have gone beyond the stage of advanced and refined prototypes. When building legal ontology, the challenges are not only from the deficiency of the ontology techniques, but also from the inner legal systems. The paper analyzes the challenges for building Chinese legal ontology, such as the ambiguity of legal language, the deficiency of three-stage inference for court decisions and the limited role of cases in China. Based on the challenges, the proposed legal ontology should be circular ontology between normative documents and judicial cases. In addition, we discuss the issues for constructing legal ontology: the role of legal ontology, legal modules and basic elements for legal normative document.

Keywords—Legal ontology; fuzziness of law; three-stage deductive reasoning; case guidance system; legal module

I. INTRODUCTION

Multi-disciplinary integration can interpret the nature of problems from multiple dimensions. Since the rise of ontology studies in the 1990’s, legal ontology is attracting attention and the related searches are continuously increasing. According to McGuiness, with the maturity of the field of ontological engineering we can say that currently, legal ontologies have come of age [1]. From HYPO, subsequently developing into CATO, IBP, CABARET, and BankXX, there are more than 80 models. The scope of legal ontology has developed from the reasoning of a simple case to information retrieval and legal reasoning prefers a cognitive framework, and pays little attention to legal domain knowledge; thus, it is unsuitable for describing law and its related knowledge.

The Leibnitz Law Center then developed the LRI Core legal ontology retrieval system based on the amount of commonsense knowledge at the abstract and concrete levels, using Holland criminal law as an example [3]. This system not only can provide a framework to obtain a coherent view of a particular legal domain ontology but also allows the inheritance of well-defined terms [4]. The model has been applied in the E-court and E-power projects, however, it has some difficulty with detail description and reasoning because of the generalized terms.

2) Case-based reasoning

With the rapid development of addressing the uncertain and incomplete non-monotonic reasoning of artificial intelligence in the 1990s, legal reasoning and expert systems were also developed, and a range of case-based reasoning models emerged. Ashley designed the HYPO system used for American commercial secrets law in 1990. Rissland and Skalak developed the CABARET system for tax reduction in 1991. Hollatz built a discretionary model for spiritual compensation and traffic injury compensation by using radial-based Gaussian function networks in 1999 [5]. After observing 200 Germany non-pecuniary loss cases, Hollatz described the factual elements as injury type, duration, consequence (e.g. serious), gender, career change, other variable and medical damage.

From a judicial practice perspective, case-based legal ontology is often designed for specific legal implementation, which has the advantages of reducing discretionary diversity, enhancing the predictability and improving discretion unification. However, the expert systems and the discretion models are not widely
accepted in the legal field. For example, one small system ASSYST [6] is the rule-based expert advice system of the U.S. sentencing commission, which is implemented but stopped to update the system in 1996. Several possible reasons follow. First, the expert system is supposed to simulate the judicial decision making process by machine, but the judicial process is difficult to analog because of the complexity of natural language. Second, the discretionary factors should be hierarchical and explicit when building an expert system, but court decisions involve other considerations in addition to the statutes, such as economics, social, mental, and politics, etc. Third, the correlative of normative documents and the similarity of cases depend on legal reasoning, but legal reasoning techniques are far from completed. Fourth, case-based ontology emphasizes the similarity of judicial precedents and will remove the specification of cases. For these reasons, case-based reasoning ontology is still not accepted by traditional legal scholars.

3) Mission-based legal domain practical ontology

Some practical ontologies are designed for the implementation of certain legal or administrative affairs. For example, a tax declaration expert system was developed in Holland, and a prior-based electronic voting system was introduced in France [7]. Elaws is an expert suggestion system that was designed for workers and small-scale employment acts in the U.S.A. Parmendies is a public opinion system of legal bills in England [8]. CLIME is an ontology for maritime law proposed by Boer, Hoekstra and Winkel. A legal ontology of U.S. Internal Revenue Code (IRC) was constructed, which includes the three following modules: automatic audit, online help by the robot and FAQ (Frequently Asked Questions) concerning revenue problems. The Ontology of Professional Judicial Knowledge (OPJK) provides an online access to a FAQ system, which benefits judicial training and retrieval for the common person.

The function of this ontology is simple, but the related domain is centralized. In practice, such small and mission orientation ontologies have fine feedback [9].

This paper is structured as follows. Sect. 1 introduces the overview of legal ontologies. Sect. 2 discusses the challenges for building a legal ontology that includes the ambiguity of legal language, the deficiency of three-stage inference for court decisions and the limited role of cases in China. In Sect. 3, we discuss the issues in constructing legal ontology, such as role of legal ontology, legal modules and basic elements for a normative document.

II. CHALLENGES FOR BUILDING LEGAL ONTOLOGY

The three types of legal ontologies have made significant progress during the past 20 years; however, few of these ontologies have gone beyond the stage of advanced and refined prototypes, scalable and ready for reuse [10]. This situation may be caused by the bottleneck and deficiency of the ontology techniques and the challenges from inner legal systems.

First, semantic techniques are faced with considerable challenges when translating natural languages to an intelligent machine [11]. Natural language is complex and varied, which includes different language styles, such as old and modern language, online and formal language, and foreign and local dialect. For the same event, semantic techniques will lead to different expressions. Thus, the great challenges confronting web3.0 is handling the essence of things, clarifying their relations, and applying reasoning techniques and legal system to facilitate user feedback.

Second, ontology is the commonly shared concepts set. Legal concepts appear explicit, but the content is abundant and not easy to control. The relation between concepts also cannot be expressed by simple logic such as is-a, kind-of, composed-of, etc. Legal concepts appear as a well-structured system with explicit concepts that exists independently of individual desires and beliefs, and legislation and judicial decisions are the strength balance of all parties. Legal concepts are closely related to politics, economics, morals, etc. In Latin, law is noted by jus or lex. Jus is relatively abstract, including legal, rights, justice and fairness, etc. In China, law is the integration of an obscure words: Tian, which includes the abstract meaning of law and presents the highest moral standards. Ancient regulations lay particular stress on criminal law and form a specific Chinese ancient law advocating "Zun and kiss", where children respect their elders or are threatened with repression. If law is confined to the statutes and regulations and something eternal in law is ignored, it will cause dilemmas such as many regulations being unused, legislation and judicial decisions unravelling, and legal education not conforming to legal practice.

Third, legal ontology retrieval is an analog for human thinking, and logical reasoning is the bridge for statutes and judicial precedents. Legal reasoning logic is not complete, which restricts the practice of legal ontology.

Finally, the construction of a legal ontology requires the merging of multiple subjects, and the separation of subjects has hindered exchange and cooperation. The problem is obvious in the Chinese legal field because most law scholars adhere to self-contained research methods and consciously or unconsciously reject the knowledge integration of other disciplines. The major legal ontologies now are designed by computer researchers who do not fully understand legal concepts and their relations or the operation of the law. Therefore, the lack of a communication platform also restricts the practicability of a legal ontology.

A. Challenge 1. The fuzziness of the law

Fensel suggests that an ontology should include four characteristics. 1) Conceptualization is the abstract model of an objective world. 2) Explicit indicates that the concepts and their relations are precisely defined. 3) Formal means the exact mathematical description. 4) Shared suggests that the knowledge is commonly accepted by users. Thus, the precondition for building an ontology is the terminology and its explicit relations. However, the wording of a legislative definition is usually insufficient to enable a full characterization of the defined concepts. The law has an ambiguous existence including much vagueness and uncertainty, which is the essential feature.
of law. Law is always fuzzy, and it signifies ambiguous knowledge [12].

First, legislative language is fuzzy. Law has experienced a process from self-governing rules to legislative statutes, and legal language plays an important role. Legal language is symbolic and translates the proper court decision of which legislators may be aware to statutes by inductive reasoning. Concerning countries such as China, which has vast territory, a variety of nationalities and multiple languages, language blending and changing aggravates the complexity and multiplicity. Thus, there are many obscure areas in legal language [13].

Second, the written language is bound to the context. When applying statutes to specific cases, written language is also an interpretation process. Different judges may have different interpretations for the same legal concepts, and in practice, different interpretation methods may be adopted, such as literary, contextual, systematic, legal spirit, the law’s purpose, comparative, social, etc. For example, Wang Hai is a common person sued when fighting counterfeits. Wang Hai brought a large amount of flawed products and double compensation must be awarded to consumers according to the Chinese Consumer Protection Act. However, judges have different understandings of the concept of consumer. Some judges support double compensation to Wang Hai by considering him as consumer, whereas some judges oppose compensation because Wang Hai knew the flaws and intended to defraud.

Third, the legal system is a definite system, not a closed system. The legal system has numerous ties with morality, religion, economics, politics, etc. Some moral or religious rules permeated the legal field and became legal principles. However, there are still some conflicts. In environment law, for example, some enterprises prefer being fined and will continue to pollute the environment, and this action is not illegal but immoral.

Fourth, classifying laws in logically distinct categories has always been a major task of legal philosophy. However, the legal branches are not entirely distinct from one another. For example, to which legal division is labor law affiliated: civil, economic, administrative or social law?

Finally, punishment is open-textured. Based on the roles of the statutes as proposing, adjusting or warning, judges may often use vague judicial language in their discretion, such as pleaded guilty, better attitude, bad social influence, enormous damage, etc. to express consciously and unconsciously his or her understanding of the content of the statutes.

The application of legal rules is often contingent on the satisfaction of a particular concept. However, this feature of law is doomed because the pursuit of the explicit in law is not practical. Legal concepts contribute more on a qualitative definition whereas judicial precedents contribute more to a quantitative definition. The method of categorization is a real-life approach. This approach has a hierarchy, but the hierarchy is flexible. Thus, the combination of legal concepts and cases can well define the nature of a legal concept.

B. Challenge 2. The deficiency of three-stage deductive reasoning for court decision

Three-stage deductive reasoning is the main adopted reasoning method in statute law countries. Three-stage deductive reasoning takes law and regulations as major premises, legal facts as minor premises, and then obtains the judgment of the action. Legislation translates life’s realities to statutes through inductive reasoning using legal language and principles, and court decisions interpret the statutes and apply them to certain events. These two processes have significant information entropy. The legislative purpose is the requirement of legislators. However, the judicial process is not similar to Ariadne's thread and enables the judge to decide according to the expectation of legislators. Dworkin considers the judicial process thinking strategies that combine legal knowledge, time information and thinking method. Legislation is not scientific and legal application is the communication process between law and facts [14]. Thus, the legal concepts are formed in the process of making statutes and judicial practice.

First, the lack of communication between legislation and judicial decisions will lead to their separation, and the law’s enforcement deviates from the legislation purpose. In China, for example, legislation took policy as center in 1980s, and the country wanted to build the legal system in 1990s. When the legislative effect was surveyed in 2000s, we find that the statutes were weakly implemented and that some regulations overlap or are competing, conflicting or have never been used since they were enacted. Some legislators attempted to graft other countries’ regulations to China’s, but the well-designed rules cannot be implemented in China. Some deputies of the National People's Congress are rubber stamped in their legislative voting. In law practice, there are still some problems. For example, the judges are veterans and college graduates without legal experiences, and the judiciary is not independent, especially in the areas of finance and judicial selection. Moreover, the law-executor is influenced by traditional legal ideology to emphasize criminal law, including extorting confessions by torture.

Second, the outcome of the law is multi-dimensional, and cannot be easily expressed. Human knowledge often limits and lags the realities. A proper mechanism is necessary to translate the related legal issues of new cases into the statutes.

Third, people may not have a unified understanding of a legal provision because of the limit of legal techniques and legal language, even for judges who firmly understand the law. Considering Chinese spiritual compensation rules, for example, article 82 indicates that proper spiritual compensation should be given. What is proper? For the same case, the compensation awarded in the first trial was RMB 10,000, whereas the award became RMB 210,000 in the appeals court.

Fourth, the sentence is not only a judicial interpretation process by judges but also a process referring to legal principles and the evaluation criteria of convicted criminals. Judicial interpretation is a legal document with Chinese characteristics and often is expressed by interpretation or regulation or official reply, etc. Generally, judicial interpretation has two roles: the
interpretation of vague and ambiguous statutes and unification of discretion

Finally, if the three-stage reasoning ignores the corresponding background, the verdict may deviate from the fundamental legal values. Legislation is an integrated regulation network combined with numerous statutes and rules under specific guides [14]. Law application must refer to the legislative background and other rules to extract the implicit related meaning.

Therefore, three-staged reasoning is not the only or the main method for application of the law (Kaufmann 2004). A legal concept cannot be acquired only through logical reasoning, and the proper concepts can be discovered by the integration of inductive and deductive reasoning with the help of judicial cases.

C. Challenge 3. Case Guidance system in China

Judicial precedents are the micro of the judicial system, which is the effective measure for explaining the statutes and forming the legal blindness in the common law countries. However, the role of case law has been underestimated for some time in civil law countries where legal values, principles, logic and concepts are emphasized. Judicial precedents have played a positive role in Chinese history, for example, royal decisions in the Qin Dynasty, referring cases in the Han Dynasty, and legal precedents in the Yuan Dynasty. The condition changed at the end of the Qing Dynasty. With the integration of the two legal systems, how can the judicial cases be translated to directly refer precedents in civil law countries? Case law has been reevaluated in civil law countries, which plays an important role in Germany and Japan.

In China, the role of case law is gradual strengthening. First, the openness of judicial cases is much better than before. Guided by the traditional legal ethics that the law has unpredictable power, provided it is not published, the statutes and judicial precedents could not be obtained easily. However, this condition has gradually changed because the statutes and judicial cases can currently be acquired in many ways, such as the opening of information by the courts, legal research websites, news reports, the public searching platform, baidu or google, and commercial cases databases westlaw or lexis, etc. Second, people have paid more attention to legal cases: even legislation and its modification is promoted by some cases. For example, the cases of Ma Dandan, Se Xianglin, and Sun Zhigang impelled the modification of national compensation law. The case of Guangzhou Xuting and the case of Chengdu Xuting often appear in newspapers headlines. People are concerned about judicial cases, and the trial consequences for similar cases are questioned.

For the judicial cases to be directly adopted as the discretionary rules in civil law countries, this should be handled with caution. First, using the judicial precedents as possible reused rules, it may be some certain expression similar to regulations. What is the form of the expression? Second, what judicial cases can be used as precedents? The selection of cases is important in civil law countries because the case law cannot be admitted. Third, do the judicial precedents have the same role as other regulations? When case law conflicts with the regulations, which one should be adopted?

China established a unique judicial system – a case guidance system based on these questions. The system was officially confirmed in 2010, which addresses similar cases to provide judges a reference. The system connects cases that are similar in legal relation, main points of the case, case facts and arguments focus. Regulations on case guidance in China were enacted by the Supreme People’s Procuratorate and the Supreme People’s Court in 2010 respectively. The two regulations have confirmed the role of judicial precedents. Case guidance in China is not the same as stare decisis in the USA. First, the guidance cases are chosen only by Supreme People’s Court. Second, the cases are chosen according to certain procedures, such as recommending, review, discussion and announcement procedures. Third, the choosing of a guide case must satisfy the rules of extensive attention, related regulation that is not concrete, typical, complex or new emerging cases, and other cases with great influence.

The case guidance system has significant meaning in improving judicial quality when the good and bad judges are intermingled and their territory is vast with great differences in social conditions. However, the performance of the system is not satisfactory. First, the number and the scope of guidance cases are limited. The system concentrates on cases in the intermediate or Supreme Court. Second, the legal force of the guidance cases is not as strong as stare decisis, and guidance cases are only referred to and not directly quoted in court. Third, the utility of the guidance case is not high. By observing the pilot courts of Sichuan province, we found that the only 0.58% of cases adopted the guidance cases. Fourth, how to refine the main points of the cases have not been established yet because it is difficult to describe a set of facts and legal consequences in only two or three words.

The establishment of the case guidance system in China indicates that judicial precedents have been gradually accepted in China. With an increasing number of judicial cases, building a powerful searching platform and approach is urgent. This paper focuses on building a powerful legal ontology retrieval model.

III. RELATED ISSUES ON LEGAL ONTOLOGY

Considering the challenges, the legal ontology should be a circular ontology between normative documents and judicial cases. A three-layer ontology model is proposed by Dumontier [15]. The top layer describes the category, attributes and forms through common concepts that are independent of certain domains or problems, such as space, time, physical object, events, action, etc. The second layer is domain ontology, which details the top-layer concepts, and it forms from the static knowledge in the domain. The third layer is application ontology that describes the related concepts and knowledge in a certain domain. The general three approaches for constructing the legal ontology model are as follows. 1) The Top-down method is created by asking domain experts to agree on a unique point of view in their specializations; 2) The Bottom-up method is made when the domain concepts and terminology are extracted from appropriate documents to induce, cluster and generalize to compose
an ontology. 3) The Middle-out method begins with a group of core concepts and expands the concept models. The three methods have their own advantages in building an ontology, but the Top-down and Bottom-up methods are used more than the Middle-out method.

For a normative ontology, the bottom-up approach is used, and all of the required elements are extracted from appropriate documents to compose the ontology. We try to locate the related normative, such as legal principle, acts, regulations, directive, judicial interpretations and industry regulations, etc. For case ontology, the top-down approach is used for and the assumption on the how the case is determined to contribute to the ontology. We examine the most similar judicial cases by comparison with the new case.

A. The role of this legal ontology

In general, the legal ontology retrieval has two roles: 1) to facilitate exchange and re-use of information in knowledge bases that may be distributed over the Internet. 2) to manage relations and distinctions among concept types [16].

As a domain ontology, it should: 1) organize and build the domain information; 2) reason and solve domain problems; 3) make semantic retrieval and query; 4) integrate and exchange domain information; and 5) help the user understand the domain knowledge [17]. Therefore, the roles of the proposed legal ontology are as follows. 1) The legal ontology should semantically identify and locate the main point of the new case that is a standardized brief to the case, according to the regulations of civil, administrative and penalty cases issued by Chinese Supreme Court; 2) The ontology should also list the normative documents and assume the correlative according to their contents and then range the related provisions in terms of the explicit legal consequence. The principle for finding the normative documents is trying to obtain the relatively clearer rules by more related norms. 3) Because the law cannot usually provide a ready norm, the case provides a good opportunity to determine how to apply the law and regulations to actual cases. The ontology should suppose the referee basis for certain cases, obtain the discretionary factors of the cases, and arrange cases as an analogy of these cases, then propose a suggestion for deciding the case.

B. The ontology modules

Jurisprudence is a knowledge hierarchy using the legal system as a study object including thousands of legal concepts and their relations. In the U.S. Internal Revenue Code, for example, more than 50,000 concepts and regulations are involved. If we try to build a unified legal ontology, the magnanimity contents and the specialized legal ontology language will bring some realistic obstacles for extension, re-use and entertainment of the legal ontology. Applying modular design to ontology construction is the mainstream strategy for building an ontology [18]. Ontology module not only can uniform the ontology description style but also ready for reuse. A legal ontology module decomposes a complex domain to some relatively independent but correlative sub-domains from the top down according to the proper principle and then builds a module for the sub-domain called an ontology module. However, we will meet several problems when creating a legal ontology, such as the ambiguity of legal demand, the specifics of legal knowledge and the mass of related legal information. How should the legal ontology module be divided?

A legal system can be classified in many ways. 1) According to the adjusted object and adjusted methods of law, a legal system can be classified as Constitutional, Civil Law, Criminal Law, Administrative or Procedural law or it can be classified as Criminal, Civil, Commercial, Intellectual property, Maritime, administrative, state compensation law and law enforcement. A more detailed legal system is classified as Constitutional, Administrative, Civil and Family, Commercial, Economic, Labor, Social security, Military, Environmental, Criminal, or Procedural law. In addition, the legal departments have crossed. 2) According to the law promulgate departments, a legal system can be classified as law enacted by the National People's Congress or Standing Committee, administrative law enacted by the State Council, military regulations enacted by the Central Military Committee, interpretations enacted by the Supreme People's Court or the Supreme People's Procuratorate, local regulations enacted by local government, or administrative directives enacted by commissions under the State Council. 3) According to the nature of illegal case, a legal system is classified as American, British, Canadian, Chinese law, etc. 4) According to the nature of illegal case, a legal system is classified as criminal, civil, or administrative cases, etc. 5) According to the deciding court, cases can be classified as a decision by the United States Supreme Court, Federal court, or published or unpublished decisions by local courts, etc.

This classification is undoubtedly important for systematically and browsing legal resource, but what strategy should be used to integrate the legal resources and how should the legal module be divided? In general, the principles for an ontology module are simplicity, intelligibility, relative independence and easy re-use. An ontology module should be introduced directly or indirectly to the task database. The independence of the module is preserved to the greatest extent, and the loop quotation is avoided to reduce the coupling factor and the communication among the modules [19].

![Figure 1. Legal domain ontology module of medical dispute](image)

Medical disputes, for example, involve constitutional, civil law and criminal law as well as administrative and local regulation, civil procedural law and directives, etc.
Considering the legal principle and compensation method, it is better to divide the module into the following three parts: administrative penalty, criminal punishment and civil compensation. In addition, the module can be divided into physical compensation and spiritual compensation for civil compensation. The module should not be divided further, such as punishments of warning, suspended licenses, or cancelled licenses as administrative penalties.

C. Basic elements for a normative document

Normative documents are a family of a vast amount of rules and include constitutions, laws, regulations, directives, judicial interpretations and other normative documents. What are the basic elements for a normative document? The normative document should have the normal format.

The basic components of normative documents in the FOI-law are normative, world, responsibility, creative and meta-legal knowledge, whereas in the LRI-core ontology, they are physical, mental and social knowledge. However, the two ontologies divide the knowledge more from philosophy and less from the legal domain angle. Saravanadas divides the legal ontology into persons, things, events, facts, and acts [20]. In this paper, we obtain the basic components of legal normative documents from the commonly used catalog of legal concepts in the sub-domain of spiritual compensation in medical disputes, a partial view of the legal ontology is given in Fig. 2 [21].

Figure 2. An extract of the legal ontology framework on selected sub-domain.

IV. CONCLUSION

The main contribution of this work is to analyze the challenges of building a legal ontology system, such as the ambiguity of legal language, the deficiency of three-stage inference for court decisions and the limited role of cases in continental countries such as China. The basic components for legal concepts are concluded as subject, object, relationship, facts and the other concepts. In addition, we discussed the issues for building legal ontology; the role of legal ontology, legal modules and basic elements for legal normative document.

ACKNOWLEDGMENT

The work supported by the Soft Science Project of Sichuan Province of China under Grant 2013ZSR0010.

REFERENCES