ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/321257442

Question analysis for Vietnamese legal question answering

Conference Paper · October 2017

DOI: 10.1109/KSE.2017.8119451

CITATIONS
8

READS
381

4 authors, including:

Ngo Xuan Bach
Posts and Telecommunications Institute of Technology
44 PUBLICATIONS **325** CITATIONS

SEE PROFILE

Tu Minh Phuong
Posts and Telecommunications Institute of Technology
74 PUBLICATIONS **817** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project Social Media Text Processing View project

Project Vietnamese POS Tagging for Social Media Text View project

All content following this page was uploaded by Tu Minh Phuong on 05 January 2019.

The user has requested enhancement of the downloaded file.

2017 9th International Conference on Knowledge and Systems Engineering(KSE)

Question Analysis for Vietnamese Legal Question Answering

Ngo Xuan Bach*†, Le Thi Ngoc Cham*, Tran Ha Ngoc Thien*, Tu Minh Phuong*†
*Department of Computer Science, Posts and Telecommunications Institute of Technology, Vietnam
†Machine Learning & Applications Lab, Posts and Telecommunications Institute of Technology, Vietnam
{bachnx, chamltn, thienthn, phuongtm}@ptit.edu.vn

***Abstract*—This paper presents a study on analyzing questions in legal domain for Vietnamese language, which is an important step in building an automated question answering system for the domain. We focus on questions about transportation law – the law with arguably the largest number of violations and thus is the most asked about. Given a legal question in natural language, our goal is to extract important information such as Type of Vehicle, Action of Vehicle, Location, and Question Type. We model the question analysis task as a sequence labeling problem and present a CRF-based method to deal with it. Experimental results on a corpus consisting of 1678 Vietnamese questions show that our method can extract 16 types of information with high precision and recall.**

I. Introduction

Building automatic question answering (QA) systems, which can give the direct and exact answers to natural language questions, has been a long standing goal in Artificial Intelligence. A lot of work on this area has been conducted over several decades [1], [5], [6], [9], [22]. The methods for building QA systems can be divided into two main classes, namely retrieved-based and knowledge-based. The retrieved-based approach [9], [30], or Information Retrieval (IR) based, relies on the huge amounts of information available as text on the Web. Given a user's question, the method analyzes the question to determine the likely answer type, such as a person, a location, or an organization, and formulates queries to send to a search engine. The search engine returns suitable documents, which are ranked according to the relevance to the queries. Finally, candidate answers are extracted from the documents and ranked. The knowledge-based approach [4], [5], [6] instead builds a semantic representation of the query. The representation is then used to query over a structured database of facts to find the suitable answer. Because each approach has its own advantages, practical QA systems can exploit a hybrid method [10].

Question analysis plays a very important role in building a good question answering system. The aim of this process is to understand the user's question. In a retrieved-based QA system, question analysis consists of two main steps: 1) question classification (to determine answer type); and 2) query formulation (extract key terms in the question to formulate queries). In a knowledge-based QA system, question analysis is the process of mapping the natural language question into a semantic representation like a logical form. A system that fails in the question analysis stage could not find a desire answer.

Question 1: đi ô tô không nhường đường cho xe thô sơ khi chuyển hướng phạt bao nhiêu tiền? (How much are motorists fined for failing to give away to rudimentary vehicles when changing lanes?)
Expected answer: 100000 đồng đến 200000 đồng (from 100,000 to 200,000 Dong)

Question 2: người đi xe máy có nồng độ cồn là 0,3 miligam/1 lít khí thở có bị phạt không? (Are motorcyclists with an alcohol level of 0.3 milligrams per liter of breathing fined?)
Expected answer: Có (Yes)

Fig. 1: Examples of questions about Vietnamese transportation law and expected answers.

In this work, we consider the problem of building a Vietnamese question answering system for transportation law. In Vietnam, people use their personal cars and motorbikes frequently, making the knowledge about transportation regulations important for most population. At the same time, a large number of transportation law violations has been recorded every day, mainly due to the ignorance or lack of law understanding. Because most drivers do not prefer to read and learn law in their original forms, the existence of QA system in natural language is a very useful solution in improving the awareness and understanding of transportation law for Vietnamese drivers. Figure 1 shows examples of questions and expected answers about Vietnamese transportation law. The English translations are given in the brackets.

We focus on the question analysis stage, which aims to extract key information of a given legal question. Such important information is then used to query over a knowledge base to find the answer. We present a method to extract key information using Conditional random fields (CRFs) [12]. We also introduce an annotated corpus consisting of 1678 questions about Vietnamese transportation law. Experimental results show the effectiveness of the proposed method.

The rest of this paper is organized as follows. Section II reviews related work. Section III introduces our framework for building a question answering system for Vietnamese transportation law. Our method for analyzing legal questions is presented in Section IV. Section V describes our dataset and experimental results. Finally, Section VI concludes the paper and discusses future work.

II. Related Work

In this section, we review related work on legal text processing and question answering in the legal domain.

Legal text processing. Over the past two decades, a lot of work on legal text processing has been done. Studies on legal text processing can be divided into several topics, including Legal Ontology Learning [19], [27], Legal Information Extraction [14], [28], Legal Semantic Annotation [7], [20], Automatic Identification of Legal Terms [17], Legal Knowledge Modeling [16], Legal Argumentation [29], Legal Automatic Summarization [11], Analyzing Logical Structures of Legal Texts [2], [3], [23], Reference Resolution in Legal Texts [25], and Question Answering in the Legal Domain [15], [18], [24], [26]. A comprehensive review of legal text processing can be found in Bach et al. [3].

Question answering in the legal domain. Little research has been conducted on question answering in the legal domain, despite its wide uses and applications. Paulo et al. [18] present a QA system for Portuguese juridical documents. Their approach bases on a computational linguistic framework and uses extensive sources: syntactic analysis followed by semantic analysis; and finally, a semantic/pragmatic interpretation using ontology and logical inferences. Monroy et al. [15] describe a question answering system for Spanish at the shallow level by using graphs. The system can output a set of articles related to the question based on the similarity, i.e. TF-IDF, between documents via terms in documents. Tomura [24] presents a study on building a question answering system for Japanese legal texts. The system can deal with five types of questions using the requisite-effectuation structures of law sentences. Tran et al. [26] present a study on exploiting reference information to build a question answering system for Japanese legal texts. Unlike the work of Tomura [24], which only can find the answer in a single document, their system exploits the reference information among multiple legal documents to find answers. This approach also uses requisite-effectuation structures of legal sentences and some effective similarity measures to support finding correct answers without training data.

III. A Question Answering System for Vietnamese Transportation Law

In this section, we introduce our question answering system for Vietnamese transportation law. The overall framework of the system is illustrated in Figure 2, which consists of three main modules: Automatic Speech Recognition, Question Analysis, and Answer Extraction.

- **Automatic Speech Recognition**: this module recognizes the user's spoken question and converts it to textual representation.
- **Question Analysis**: this module analyzes the question in natural language text to extract key information, which will be utilized to extract the answer.
- **Answer Extraction**: this module exploits important information of the question to find out a suitable answer using a knowledge base.

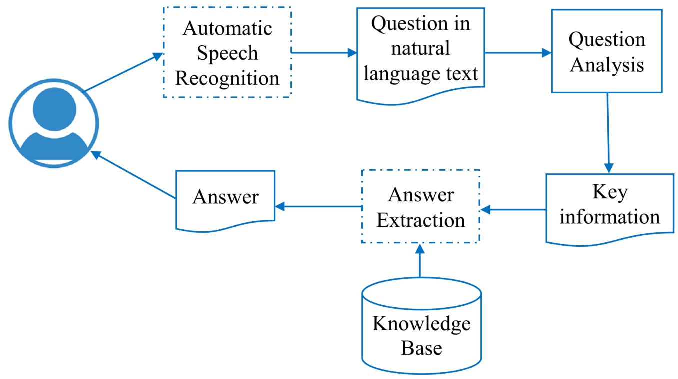

Fig. 2: A QA Framework for Vietnamese Transportation Law.

In this work, we focus on the second module, i.e. Question Analysis. Nowadays, good performance automatic speech recognition systems are available on most smart phones. We leave the last module, i.e. Answer Extraction, for future work. Our system can be classified as knowledge-based QA in the sense that it first processes the question and then queries over a knowledge base to find the answer.

IV. Question Analysis as a Sequence Labeling Problem

This section presents our method for analyzing legal questions. We first show how to model the task as a sequence labeling problem. We then introduce linguistic features and describe a powerful sequence learning model, CRFs [12], which will be used to solve the task.

A. Method

The goal of question analysis is to determine question type and extract the key information from the input legal question. Let's consider examples in Figure 3. In the first question, we want to extract three types of information: Type of Vehicle, Action of vehicle, and Question Type, which are phrases indicated by TV, A, and QT tags, respectively. We will call such phrases *important phrases* in the rest of this paper. In the second question, we want to extract four types of information: Type of Vehicle, Alcohol Concentration, Value, and Question Type, which are indicated by TV, AC, V, and QT tags, respectively.

We model the task as a sequence labeling problem, in which the input question is considered as a sequence of elements. Unlike English words, the boundaries of a Vietnamese word are not easily defined based on white spaces. In Vietnamese, each word may consist of one or more syllables separated by white spaces. Therefore, each element in the input sequence can be a syllable or a word, depending on the modeling method. Figure 4 illustrates an example of how to tag an input question in a syllable-based model (upper) and a word-based model (lower) using IOB notation[1]. In this notation, the first element (syllable or word) of an important phrase

[1] I, O, and B means Inner, Outside, and Beginning, respectively.

Question 1:
đi <TV>ô tô</TV> <A>không nhường đường cho xe thô sơ khi chuyển hướng</A> <QT>phạt bao nhiêu tiền</QT>?

(How much are motorists fined for failing to give away to rudimentary vehicles when changing lanes?)

Question 2:
người đi <TV>xe máy</TV> có <AC>nồng độ cồn</AC> là <V>0,3 miligam/1 lít khí thở</V> <QT>có bị phạt không</QT>

(Are motorcyclists with an alcohol level of 0.3 milligrams per liter of breathing fined?)

Fig. 3: Examples of legal questions and key information.

expressing key information is tagged by B. Other elements of that important phrase are tagged by I, while an element not included in any important phrase is tagged by O. For example, a B-TV tag indicates the first element (B) of an important phrase which tell us about the type of vehicle (TV).

In this work, we consider 16 types of key information (phrases expressing key information are non-overlapping): Action of vehicle (A), Alcohol concentration (AC), Annotation (ANO), Driving license (DL), Location (L), Question type (QT), Speed (SD), Traffic instructor (TI), Traffic Light (TL), Traffic participant (TP), Type of vehicle (TV), Value (V), Additional information about vehicle (IF1), Additional information about traffic light (IF2), Additional information about traffic participant (IF3), and Additional information in general (IF4). The list of 16 types and their examples are shown in Figure 5. Therefore, totally we have 33 tags: 16 tags with prefix "B", 16 tags with prefix "I", and tag "O".

B. Features

To build the sequence labeling model for the task, we exploit lexical features and simple syntactic features, i.e. part-of-speech (POS) tags, as follows:

- **N-grams**: we use unigrams, bigrams, and trigrams of syllables (syllable-based model) and words (word-based model) extracted in a window size of $2 \times W + 1$, surrounding the current element (W elements on the left hand side and W elements on the right hand side).
- **POS tags**: similarly, we also use POS tags of syllables and words in a window size of $2\times W+1$, surrounding the current element. Because POS tags are assigned to words, in a syllable-based model, we assign the same POS tag for all syllables of a word.

C. Conditional Random Fields

This section gives a brief introduction to Conditional Random Fields (CRFs), which serves as the learning method to build our sequence labeling model. Although any sequence learning method can be used in our model, we choose CRFs due to their popularity and effectiveness among natural language processing and data mining communities.

Conditional random fields (CRFs) [12], [21] are undirected graphical models, which define the probability of a label sequence y given an observation sequence x as follows:

$$P(y|x,\lambda,\mu) = \frac{1}{Z(x)} exp(SUM)$$

where $SUM = \sum_j \lambda_j f_j(y_{i-1}, y_i, x, i) + \sum_k \mu_k g_k(y_i, x, i)$; $f_j(y_{i-1}, y_i, x, i)$ is a transition feature function, which is defined on the entire observation sequence x and the labels at positions i and $i-1$ in the label sequence y; $g_k(y_i, x, i)$ is a state feature function (or node feature), which is defined on the entire observation sequence x and the label at position i in the label sequence y; λ_j and μ_k are parameters of the model, which are estimated in the training process; and $Z(x)$ is a normalization factor. In this work, x is a sequence of words (in the word-based model) or syllables (in the syllable-based model), and y is a sequence of tags such as B-TV and I-TV.

V. Experiments

A. Dataset

The set of questions about Vietnamese transportation law has been collected from the following sources. First, we collect questions from Web pages for driver license test: car riders[2] and motorcycle riders[3]. In addition, questions are generated from rules stated in documents available in online collections of legal documents[4], and official government website[5].

Table I shows statistics of our dataset consisting of 1678 questions in two main classes: Fine and Instruction.

1) **Fine**: Questions about fines for breaking the law, which are further divided into three sub-categories: Car, Motorbike, and Others. Below are examples.
   - How much are car riders fined for failing to give way to rudimentary vehicles? (Car)
   - How much are motorcycle riders fined for using the umbrella? (Motorbike)
   - How much are bicycle riders fined for failing to follow the signposts? (Others)
2) **Instruction**: Questions about legal instructions, which are further divided into four sub-categories: Concept, Situation, Driving License, and Regulation. Below are examples.
   - What sort of roads are there in the populated areas? (Concept)
   - How do the drivers use the vehicle lights when facing oncoming traffic at night? (Situation)
   - Are rudimentary drivers in need of driving license? (Driving License)
   - Is the act of roads usurpation strictly forbidden? (Regulation)

Two Vietnamese people were asked to annotate important information in each question. Questions with disagreement

[2]http://vnexpress.net/interactive/2016/thi-sat-hach-lai-xe
[3]http://www.thuexemientrung.net/150-cau-hoi-thi-bang-lai-xe-may.html
[4]http://thuvienphapluat.vn
[5]http://vanban.chinhphu.vn

Syllable-based model																
Question	đi	ô	tô	không	nhường	đường	cho	xe	thô	sơ	...	phạt	bao	nhiêu	tiền	?
Tags	O	B-TV	I-TV	B-A	I-A	I-A	I-A	I-A	I-A	I-A	...	B-QT	I-QT	I-QT	I-QT	O

Word-based model												
Question	đi	ô_tô	không	nhường	đường	cho	xe_thô_sơ	...	phạt	bao_nhiêu	tiền	?
Tags	O	B-TV	B-A	I-A	I-A	I-A	I-A	...	B-QT	I-QT	I-QT	O

Fig. 4: Examples of IOB tags in syllable-based and word-based models.

Tag	Meaning	Example	Translation
A	Action of vehicle	Vượt, dừng, đỗ	Overtake, reverse, park
AC	Alcohol concentration	Nồng độ cồn	Alcohol concentration
ANO	Annotation	Khái niệm đường chính	Motor way, main road
DL	Driving license	Giấy phép hạng A1, A2	License class A1, A2
IF1	Additional information about vehicle	Quái tải trọng, 30 chỗ ngồi	Overload, 30 seats
IF2	Additional information about traffic light	Màu đỏ, tín hiệu vàng	Red, yellow
IF3	Additional Information about traffic participant	Đội mũ bảo hiểm, lái xe	Wear helmet, drive
IF4	Additional information	Giờ, đường dốc, giơ tay	Bypass, ring road
L	Location	Ngã tư, chân cầu vượt	Crossroads, fork
QT	Question type	Phạt bao nhiêu	How much
SP	Speed	Tốc độ	Speed
TI	Traffic instructor	Cảnh sát	Traffic police
TL	Traffic light	Đèn đỏ, biển báo	Signpost
TP	Traffic participant	Người đi bộ, hành khách	Pedestrian, motorists
TV	Type of vehicle	Ô tô, xe máy, xe đạp	Car, motorbike, bicycle
V	Value	10km/h	10km/h

Fig. 5: List of tags and their meaning.

TABLE I: Statistics of the dataset

Question class	Subclass	Quantity
Fine	Car	375
	Motorbike	345
	Others	152
Instruction	Concept	73
	Situation	270
	Driving License	134
	Regulation	329
Total		**1678**

between two annotators were examined and corrected by the third annotator. To measure the inter-annotator agreement on assigning the same tags, we used the Cohen's kappa coefficient [8]. The Cohen's kappa coefficient of our dataset was 0.84, which can be interpreted as almost perfect agreement. The distribution of annotated tags is shown in Figure 6. The most frequent tags includes QT (Question Type), TV (Type of Vehicle), and A (Action) .

B. Experimental Settings

We divided the dataset into ten subsets of (nearly) equal size and conducted 10-fold cross-validation. To measure the performance of labeling models we used precision, recall, and the F_1 score, which were computed for each kind of tag and for overall as well. Let's consider an example of tag TV (Type of Vehicle). Suppose that A and B are the set of segments that the system predicted as TV and the set of segments labeled as TV by human, respectively. The precision, recall, and the F_1 score for tag TV can be computed as follows[6] (similarly for other tags):

$$Precision = \frac{|A \cap B|}{|A|}, Recall = \frac{|A \cap B|}{|B|}, \text{and}$$

$$F_1 = \frac{2 * Precision * Recall}{Precision + Recall}.$$

We conducted experiments to compare the following models:

- **Word-based models**: we conducted experiments on three models with window size of 3, 5, and 7, respectively. For each model, we investigated two variations, i.e. with and without part-of-speech (POS) tags.
- **Syllable-based models**: similarly, we also conducted experiments on three models with two variations.

We used vnTagger[7] [13] to do word segmentation and POS tagging.

[6] To judge whether or not a segment in A corresponds to one in B we used exact matching.

[7] Software available at http://mim.hus.vnu.edu.vn/phuonglh/softwares/vnTagger.

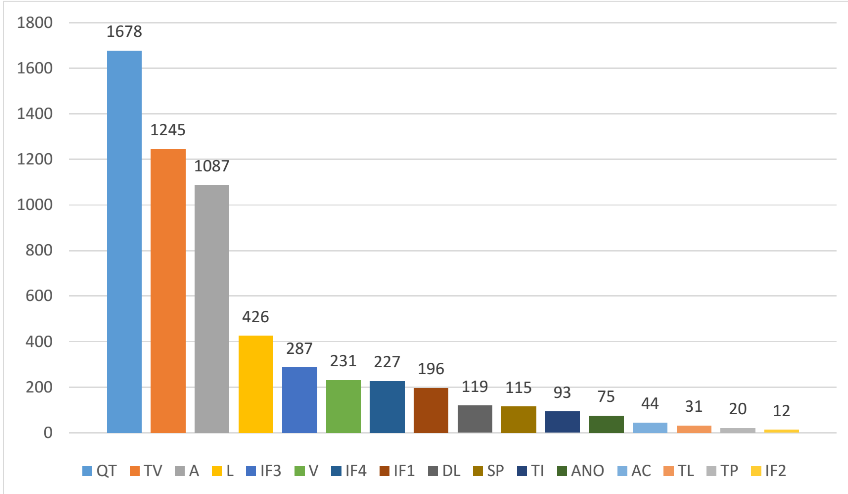

Fig. 6: The distribution of tags in our corpus.

TABLE II: Results of word-based models

Window size	POS tags	Precision(%)	Recall(%)	F_1(%)
3	NO	93.44	90.84	92.12
	YES	93.41	91.32	92.35
5	NO	**93.92**	91.12	92.50
	YES	93.74	**91.61**	**92.66**
7	NO	93.80	90.95	92.35
	YES	93.76	91.48	92.60

TABLE III: Results of syllable-based models

Window size	POS tags	Precision(%)	Recall(%)	F_1(%)
3	NO	93.21	91.06	92.12
	YES	93.44	91.71	92.57
5	NO	93.38	91.22	92.28
	YES	**93.84**	**92.05**	**92.94**
7	NO	93.66	91.66	92.65
	YES	93.66	92.00	92.82

TABLE IV: Experimental results on each individual tag

Tag	Precision(%)	Recall(%)	F_1(%)
A	88.20	89.14	88.66
AC	95.78	95.78	95.78
ANO	85.58	60.57	68.82
DL	97.97	99.20	98.54
IF1	94.64	87.35	90.67
IF2	100.00	73.33	82.67
IF3	88.08	75.06	80.91
IF4	85.77	74.34	79.51
L	92.23	92.71	92.44
QT	96.03	94.81	95.42
SP	95.61	91.44	93.23
TI	99.05	95.53	97.24
TL	86.95	60.35	68.59
TP	80.00	60.67	67.21
TV	97.32	98.78	98.04
V	98.18	99.26	98.71
Overall	**93.84**	**92.05**	**92.94**

C. Results

The overall results of word-based models and syllable-based models are shown in Tables II and III, respectively. We have some remarks as follows.

- For both models, we got the best results with the window size of 5. The word-based model and the syllable-based model achieved 92.66% and 92.94% in the F_1 score, respectively.
- Using POS tags improved the performance of both models with all window sizes, i.e. 3, 5, and 7. However, the difference was not significant.
- In comparison with word-based models, syllable-based models achieved slightly better performance in all settings.

From the experimental results, we can conclude that the information of word segmentation and POS tags was not so effective. A reason may be the segmentation and tagging tools [13], which were trained on general text, performed poorly on legal documents.

The experimental results on each individual tag of our best model, i.e. the syllable-based model with a window size of 5 and the POS information, are shown in Table IV. Tags with a high F_1 score, i.e. more than 90%, include AC (95.78%), DL (98.54%), IF1 (90.67%), L (92.44%), QT (95.42%), SP (93.23%), TI (97.24%), TV (98.04%), and V (98.71%). Tags with a low F_1 score, i.e. less than 70%, include ANO (68.82%), TL (68.59%), and TP (67.21%).

Among three tags with a low F_1 score, TL (31 times) and TP (20 times) seldom appear in the corpus. The case of tag ANO is quite special. This tag is used for concepts and annotations in definition questions. There are various types of concepts and annotations, which vary from questions to questions. This make the system difficult to deal with them.

Fine vs. Instruction. We also measured the performance of the proposed method on two types of questions, i.e. Fine and Instruction. In all models and variations, we got the better results on Fine type. The best model achieved 95.33% and 90.32% in the F_1 score on two types, respectively (see Table

TABLE V: Experimental results on Fine and Instruction questions

	Fine			Instruction		
Tag	**Pre(%)**	**Re(%)**	**F_1(%)**	**Pre(%)**	**Re(%)**	**F_1(%)**
A	90.89	91.18	**91.03**	82.20	84.58	83.37
AC	95.24	95.24	95.24	100.00	100.00	**100.00**
ANO	-	-	-	87.04	60.26	71.21
DL	83.33	83.33	83.33	98.57	100.00	**99.28**
IF1	76.92	66.67	71.43	95.91	88.65	**92.13**
IF2	-	-	-	100.00	75.00	85.71
IF3	83.72	69.68	76.06	92.52	79.53	**85.53**
IF4	86.96	68.97	76.92	85.44	75.84	**80.36**
L	94.35	93.94	**94.14**	90.00	91.59	90.79
QT	99.80	99.80	**99.80**	92.53	90.26	91.38
SP	-	-	-	96.52	95.69	96.10
TI	-	-	-	98.89	95.70	97.27
TL	66.67	61.54	64.00	100.00	57.89	**73.33**
TP	100.00	76.92	**86.96**	66.67	44.44	53.33
TV	99.50	99.50	**99.50**	92.01	96.94	94.41
V	100.00	100.00	**100.00**	97.01	98.78	97.89
Overall	**95.96**	**94.70**	**95.33**	**91.51**	**89.16**	**90.32**

V for more detail). Instruction type contains questions about concepts, situations, and regulations, which are complicated and difficult.

VI. Conclusion and Future Work

We have presented a study on analysis of Vietnamese questions on transportation law. We described a method using Conditional random fields to extract 16 types of important information from law questions. We also introduced an annotated corpus for the task consisting of 1678 Vietnamese questions and conducted experiments to show the effectiveness of the proposed method.

For the future work, we plan to build a complete question answering system for Vietnamese transportation law. A possible solution is exploiting key information, which has been extracted in this work, to query over a knowledge base of Vietnamese transportation law. Investigating other types of law rather than transportation law is another interesting direction for further research.

Acknowledgment

This work was supported in part by the Ministry of Information and Communications of Vietnam under grant No. DT.024/17.

References

[1] J. Andreas, M. Rohrbach, T. Darrell, and D. Klein. Learning to Compose Neural Networks for Question Answering. In *Proceedings of NAACL*, pp. 1545–1554, 2016.

[2] N.X. Bach, N.L. Minh, and A. Shimazu. RRE Task: The Task of Recognition of Requisite Part and Effectuation Part in Law Sentences. *International Journal of Computer Processing Of Languages (IJCPOL)*, 23(2), pp. 109–130, 2011.

[3] N.X. Bach, N.L. Minh, T.T. Oanh, and A. Shimazu. A Two-Phase Framework for Learning Logical Structures of Paragraphs in Legal Articles. *ACM Transactions on Asian Language Information Processing (ACM TALIP)*, 12(1), Article 3, 2013.

[4] J. Berant, A. Chou, R. Frostig, and P. Liang. Semantic parsing on freebase from question-answer pairs. In *Proceedings of EMNLP*, pp. 1533–1544, 2013.

[5] J. Berant and P. Liang. Semantic parsing via paraphrasing. In *Proceedings of ACL*, pp. 1415–1425, 2014.

[6] A. Bordes, S. Chopra, and J. Weston. Question Answering with Subgraph Embeddings. In *Proceedings of Conference on Empirical Methods on Natural Language Processing (EMNLP)*, pp. 615–620, 2014.

[7] R. Brighi, L. Lesmo, A. Mazzei, M. Palmirani, and D. Radicioni. Towards semantic interpretation of legal modifications through deep syntactic analysis. In *Proceedings of JURIX*, pp. 202–206, 2008.

[8] J. Cohen. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 1960.

[9] L. Dong, F. Wei, M. Zhou, and K. Xu. Question Answering over Freebase with Multi-Column Convolutional Neural Networks. In *Proceedings of ACL-IJNLP*, pp. 260–269, 2015.

[10] D.A. Ferrucci. IBM's Watson/DeepQA. *SIGARCH Computer Architecture News*, 39(3), 2011.

[11] C. Grover, B. Hachey, I. Hughson, and C. Korycinski. Automatic summarization of legal documents. In *Proceedings of the International Conference on Artificial Intelligence and Law (ICAIL)*, pp. 243–251, 2003.

[12] J. Lafferty, A. McCallum, and F. Pereira. Conditional random fields: probabilistic models for segmenting and labeling sequence data. In *Proceedings of the International Conference on Machine Learning (ICML)*, pp. 282–289, 2001.

[13] H.P. Le, A. Roussanaly, T.M.H. Nguyen, and M. Rossignol. An Empirical Study of Maximum Entropy Approach for Part-of-Speech Tagging of Vietnamese Texts. In *Proceedings of TALN*, 2010.

[14] L. McCarty. Deep semantic interpretations of legal texts. In *Proceedings of the International Conference on Artificial Intelligence and Law (ICAIL)*, pp. 217–224, 2007.

[15] A. Monroy, H. Calvo, and A. Gelbukh. NLP for Shallow Question Answering of Legal Documents Using Graphs. In *Proceedings of CICLing*, pp. 498–508, 2009.

[16] M. Nakamura, S. Nobuoka, and A. Shimazu. Towards translation of legal sentences into logical forms. In *Proceedings of the International Workshop on Juris-informatics(JURISIN)*, 2007.

[17] K. Pala, P. Rychly, and P. Smerk. Automatic identification of legal terms in Czech legal texts. *Semantic Processing of Legal Texts*, pp. 83–94, 2010.

[18] Q. Paulo, I.P. Rodrigues. A question-answering system for Portuguese juridical documents. In *Proceedings of the International Conference on Artificial Intelligence and Law (ICAIL)*, pp. 256–257, 2005.

[19] J. Saias and P. Quaresma. A methodology to create legal ontologies in a logic programming based web information retrieval system. *Law and the Semantic Web*, pp. 185–200, 2005.

[20] P. Spinosa, G. Giardiello, M. Cherubini, S. Marchi, G. Venturi, and S. Montemagni. NLP-based metadata extraction for legal text consolidation. In *Proceedings of the International Conference on Artificial Intelligence and Law (ICAIL)*, pp. 40–49, 2009.

[21] C. Sutton and A. McCallum. An Introduction to Conditional Random Fields for Relational Learning. *MIT Press*, 2006.

[22] M. Tan, C. Santos, B. Xiang, and B. Zhou. Improved Representation Learning for Question Answer Matching. In *Proceedings of ACL*, pp. 464–473, 2016.

[23] K. Tanaka. About Semantic Function of the Legal-Effect's Restrictive Part. *Natural Language*, 21, pp. 1–8, 1998 (in Japanese).

[24] K. Tomura. Study on question answering system for laws. *Technical report*, School of Information Science, Japan Advanced Institute of Science and Technology, 2013.

[25] O.T. Tran, B.X. Ngo, M.L. Nguyen, and A. Shimazu. Automated reference resolution in legal texts. *Artificial Intelligence and Law*, 22(1), pp. 29–60, 2014.

[26] O.T. Tran, B.X. Ngo, M.L. Nguyen, and A. Shimazu. Answering Legal Questions by Mining Reference Information. *New Frontiers in Artificial Intelligence*, LNAI 8417, pp. 214–229, 2014.

[27] S. Walter and M. Pinkal. Automatic extraction of definitions from German court decisions. In *Proceedings of the Workshop on Information Extraction Beyond The Document, COLING*, pp. 20–28, 2006.

[28] S. Walter. Linguistic description and automatic extraction of definitions from German court decisions. In *Proceedings of LREC*, 2008.

[29] A. Wyner, R. Palau, M. Moens, and D. Milward. Approaches to text mining arguments from legal cases. *Semantic Processing of Legal Texts*, pp. 60–79, 2010.

[30] X. Yao and B.V. Durme. Information extraction over structured data: Question answering with freebase. In *Proceedings of ACL*, pp. 956–966, 2014.

View publication stats