Extracting Case Law Sentences for Argumentation about the Meaning of Statutory Terms

Jaromir Savelka
Kevin D. Ashley

Intelligent Systems Program
University of Pittsburgh

jas438@pitt.edu

3rd Workshop on Argumentation Mining, ACL 2016, Berlin
August 12, 2016
Motivation

Task

Statutory Interpretation Data Set

Features for Predicting Interpretive Usefulness of Sentences

Predicting Usefulness

Conclusion
“Enterprise” means the related activities performed (either through unified operation or common control) by any person or persons for a common business purpose, and includes all such activities whether performed in one or more establishments or by one or more corporate or other organizational units including departments of an establishment operated through leasing arrangements, but shall not include the related activities performed for such enterprise by an independent contractor. [...]
“Enterprise” means the related activities performed (either through unified operation or common control) by any person or persons for a common business purpose, and includes all such activities whether performed in one or more establishments or by one or more corporate or other organizational units including departments of an establishment operated through leasing arrangements, but shall not include the related activities performed for such enterprise by an independent contractor. [...] 

Suppose there is a Thai restaurant in one part of the city and an Indian restaurant in another part both having a single owner.
29 U.S. Code 203 - Definitions

“Enterprise” means the related activities performed (either through unified operation or common control) by any person or persons for a common business purpose, and includes all such activities whether performed in one or more establishments or by one or more corporate or other organizational units including departments of an establishment operated through leasing arrangements, but shall not include the related activities performed for such enterprise by an independent contractor. [...]
“Enterprise” means the related activities performed (either through unified operation or common control) by any person or persons for a common business purpose, and includes all such activities whether performed in one or more establishments or by one or more corporate or other organizational units including departments of an establishment operated through leasing arrangements, but shall not include the related activities performed for such enterprise by an independent contractor. [...]
Motivation

Example Rule

No *vehicles* in the park.

Abstract rules in statutory provisions must account for diverse situations (even those not yet encountered).

⇒

Legislators use vague, open textured terms, abstract standards, principles, and values.
Abstract rules in statutory provisions must account for diverse situations (even those not yet encountered).

⇒
Legislators use vague, open textured terms, abstract standards, principles, and values.
Abstract rules in statutory provisions must account for diverse situations (even those not yet encountered).

⇒
Legislators use vague, open textured terms, abstract standards, principles, and values.
Abstract rules in statutory provisions must account for diverse situations (even those not yet encountered).

⇒
Legislators use vague, open textured terms, abstract standards, principles, and values.
Abstract rules in statutory provisions must account for diverse situations (even those not yet encountered).

⇒

Legislators use vague, open textured terms, abstract standards, principles, and values.
Abstract rules in statutory provisions must account for diverse situations (even those not yet encountered).

⇒
Legislators use vague, open textured terms, abstract standards, principles, and values.
Motivation

Example Rule

No *vehicles* in the park.

Abstract rules in statutory provisions must account for diverse situations (even those not yet encountered).

⇒

Legislators use vague, open textured terms, abstract standards, principles, and values.
**Motivation**

**Example Rule**

*No vehicles in the park.*

Abstract rules in statutory provisions must account for diverse situations (even those not yet encountered).

⇒

Legislators use vague, open textured terms, abstract standards, principles, and values.

When there are doubts about the meaning of the provision they may be removed by interpretation.
Motivation

Interpretation involves an investigation of how the term has been referred to, explained, interpreted or applied in the past.

Example Uses of the Term

i. Any mechanical device used for transportation of people or goods is a vehicle.

ii. A golf cart is to be considered a vehicle.

iii. To secure a tranquil environment in the park no vehicles are allowed.

iv. The park where no vehicles are allowed was closed during the last month.

v. The rule states: “No vehicles in the park.”

Going through the sentences is labor intensive because many sentences are useless and there is a large redundancy.
Presentation Overview

Motivation

Task

Statutory Interpretation Data Set

Features for Predicting Interpretive Usefulness of Sentences

Predicting Usefulness

Conclusion
Task

Ultimately we would like to generate the set of the useful sentences automatically.

Task Definition

Given the term of interest $t$, the statutory provision it comes from ($sp$), and a corpus of the available documents $DB$ generate a set of sentences $S (s_i \in DB)$ of the size $n$ that provides the most informative insight of how $t$ is used.
Task

Ultimately we would like to generate the set of the useful sentences automatically.

Task Definition

Given the term of interest $t$, the statutory provision it comes from $(sp)$, and a corpus of the available documents $DB$ generate a set of sentences $S (s_i \in DB)$ of the size $n$ that provides the most informative insight of how $t$ is used.

No vehicles in the park.
Task

Ultimately we would like to generate the set of the useful sentences automatically.

Task Definition

Given the term of interest \( t \), the statutory provision it comes from \((sp)\), and a corpus of the available documents \( DB \) generate a set of sentences \( S (s; \in DB) \) of the size \( n \) that provides the most informative insight of how \( t \) is used.

No vehicles in the park.

Any mechanical device used for transportation of people or goods is a vehicle. A golf cart is to be considered a vehicle. To secure a tranquil environment in the park no vehicles are allowed.
Hypotheses

( H0: A sentence may be reliably evaluated in terms of its usefulness for an interpretation of the term from a specific statutory provision. )
Hypotheses

( H0: A sentence may be reliably evaluated in terms of its usefulness for an interpretation of the term from a specific statutory provision. )

H1: By using an appropriate list of linguistic features about/in the sentence it is possible to automatically evaluate how useful the sentence is for an interpretation of the term.
Hypotheses

(H0: A sentence may be reliably evaluated in terms of its usefulness for an interpretation of the term from a specific statutory provision. )

H1: By using an appropriate list of linguistic features about/in the sentence it is possible to automatically evaluate how useful the sentence is for an interpretation of the term.

(H2: By using the information about the interpretive usefulness of a sentence we can outperform existing systems, that do not use the information, in the task of retrieving the n best sentences for the interpretation of the term. )
Presentation Overview

Motivation

Task

**Statutory Interpretation Data Set**

Features for Predicting Interpretive Usefulness of Sentences

Predicting Usefulness

Conclusion
**Court decisions** are an ideal source of sentences interpreting statutory terms.

For our corpus we selected three terms from different provisions of the United States Code:

1. “independent economic value” (18 U.S. Code § 1839(3)(B))
2. “identifying particular” (5 U.S. Code § 552a(a)(4))
3. “common business purpose” (29 U.S. Code § 203(r)(1))

For each term we have collected a small set of sentences by extracting all the sentences mentioning the term from the top 20 court decisions retrieved from Court Listener.¹

In total we assembled a small corpus of 243 sentences.

¹. [https://www.courtlistener.com/]
Two expert annotators classified the sentences into four categories according to their usefulness for the interpretation:
Two expert annotators classified the sentences into four categories according to their usefulness for the interpretation:

1. **high value** – sentence intended to define or elaborate on the meaning of the term
Two expert annotators classified the sentences into four categories according to their usefulness for the interpretation:

1. **high value** – sentence intended to define or elaborate on the meaning of the term

2. **certain value** – sentence that provides grounds to elaborate on the term’s meaning
Two expert annotators classified the sentences into four categories according to their usefulness for the interpretation:

1. **high value** – sentence intended to define or elaborate on the meaning of the term
2. **certain value** – sentence that provides grounds to elaborate on the term’s meaning
3. **potential value** – sentence that provides additional information beyond what is known from the provision the term comes from
Two expert annotators classified the sentences into four categories according to their usefulness for the interpretation:

1. **high value** – sentence intended to define or elaborate on the meaning of the term
2. **certain value** – sentence that provides grounds to elaborate on the term’s meaning
3. **potential value** – sentence that provides additional information beyond what is known from the provision the term comes from
4. **no value** – no additional information over what is known from the provision
Two expert annotators classified the sentences into four categories according to their usefulness for the interpretation:

1. **high value** – sentence intended to define or elaborate on the meaning of the term
2. **certain value** – sentence that provides grounds to elaborate on the term’s meaning
3. **potential value** – sentence that provides additional information beyond what is known from the provision the term comes from
4. **no value** – no additional information over what is known from the provision

inter-annotator agreement: .746
weighted kappa: .66
Statutory Term Interpretation Data Set

Summary statistics about the annotated corpus:

<table>
<thead>
<tr>
<th>Term</th>
<th># HV</th>
<th># CV</th>
<th># PV</th>
<th># NV</th>
<th># Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ind. economic val.</td>
<td>2</td>
<td>5</td>
<td>40</td>
<td>5</td>
<td>52</td>
</tr>
<tr>
<td>Identifying part.</td>
<td>6</td>
<td>8</td>
<td>40</td>
<td>17</td>
<td>71</td>
</tr>
<tr>
<td>C. business purp.</td>
<td>20</td>
<td>26</td>
<td>51</td>
<td>23</td>
<td>120</td>
</tr>
<tr>
<td>Total</td>
<td>28</td>
<td>39</td>
<td>131</td>
<td>45</td>
<td>243</td>
</tr>
</tbody>
</table>

HV  high value
CV  certain value
PV  potential value
NV  no value

<table>
<thead>
<tr>
<th></th>
<th>high</th>
<th>certain</th>
<th>potential</th>
<th>no</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>19</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>certain</td>
<td>15</td>
<td>12</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>potential</td>
<td>2</td>
<td>27</td>
<td>105</td>
<td>11</td>
</tr>
<tr>
<td>no</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>36</td>
</tr>
</tbody>
</table>
Presentation Overview

Motivation

Task

Statutory Interpretation Data Set

Features for Predicting Interpretive Usefulness of Sentences

Predicting Usefulness

Conclusion
In this category a sentence can be assigned one of the following labels:

- same provision
- same section
- different section
- different jurisdiction
- unknown

**Example**

The full text of §1839(3)(B) is: “[...]”. [...] Every firm other than the original equipment manufacturer and RAPCO had to pay dearly to devise, test, and win approval of similar parts; the details unknown to the rivals, and not discoverable with tape measures, had considerable “independent economic value ... from not being generally known”.
Features: Similarity

In this category a sentence can be assigned one of the following labels:

- same
- similar
- related
- different

18 U.S. Code §1839

[...] the information derives independent economic value, actual or potential, from not being generally known to, and not being readily ascertainable through proper means by, the public;

17 U.S. Code §116

[...] posted in the establishment in a prominent position where it can be readily examined by the public;
Features: Syntactic Importance

In this category a sentence can be assigned one of the following labels:

- dominant
- important
- not important
Features: Assignment or Contrast

In this category a sentence can be assigned one of the following labels:

- another term is a specific case of the term of interest
- the term of interest is a specific case of another term
- the term of interest is the same as another term
- the term of interest is not another term
- no assignment

Example: Another term a specific case

The Fifth Circuit has held that the profit motive is a common business purpose if shared.
Features: Feature Assignment

In this category a sentence can be assigned one of the following labels:

- the term of interest is a feature of another term
- another term is a feature of the term of interest
- no feature assignment

Example: Another term is a feature

However, Reiser concedes in its brief that the process has independent economic value.
In this category a sentence can be assigned one of the following labels:

▶ standard sentence
▶ citation
▶ quoted expression
▶ heading
▶ footnote

Example: Heading

A. Related Activities and Common Business Purpose.

Example: Footnote

[5] [...] However, in view of the ‘common business purpose’ requirement of the Act, we think [...]
In this category a sentence can be assigned one of the following labels:

- application of law to factual context
- applicability assessment
- statement of fact
- statement of law
- interpretation of law
- general explanation or elaboration
- reasoning statement
- holding
- other
Features: Attribution

In this category a sentence can be assigned one of the following labels:

- judge
- legislator
- party to the dispute
- witness
- expert
- other

Example: Party to the dispute

In support of his contention that Gold Star Chili and Caruso’s Ristorante constitute an enterprise, plaintiff alleges that Caruso’s Ristorante and Gold Star Chili were engaged in the related business activity [...].
Presentation Overview

Motivation

Task

Statutory Interpretation Data Set

Features for Predicting Interpretive Usefulness of Sentences

Predicting Usefulness

Conclusion
We conducted an experiment to confirm H1, i.e., we investigated if the interpretive value of a sentence can be predicted automatically.

The goal is to assign each sentence with one of the four labels:

1. high value
2. certain value
3. potential value
4. no value

As features we used the eight linguistic categories:

1. source
2. similarity
3. syntactic importance
4. assignment or contrast
5. feature assignment
6. structural placement
7. rhetorical role
8. attribution
Experiment

We randomly divided the sentences into the training set (2/3) and the test set (1/3).

As classification models we used:

1. Most frequent class (baseline)
2. Naïve Bayes
3. SVM
4. Random Forest

Because the dataset is small we repeated the experiment 100 times.

In each run we evaluated the performance on the test set as well as performed a 10-fold cross validation on the training set.
Results

Mean results from 100 runs of a classification experiment:

<table>
<thead>
<tr>
<th>Classifier</th>
<th>CV</th>
<th>STD</th>
<th>TEST</th>
<th>STD</th>
<th>SIG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most frequent</td>
<td>.545</td>
<td>.025</td>
<td>.531</td>
<td>.049</td>
<td>–</td>
</tr>
<tr>
<td>Naïve Bayes</td>
<td>.544</td>
<td>.037</td>
<td>.611</td>
<td>.066</td>
<td>no</td>
</tr>
<tr>
<td>SVM</td>
<td>.633</td>
<td>.044</td>
<td>.657</td>
<td>.066</td>
<td>no</td>
</tr>
<tr>
<td>Random Forest</td>
<td>.677</td>
<td>.033</td>
<td>.696</td>
<td>.042</td>
<td>yes</td>
</tr>
</tbody>
</table>

CV 10-fold cross validation on the training set
STD standard deviation
TEST validation on the test set
SIG statistical significance

<table>
<thead>
<tr>
<th></th>
<th>high</th>
<th>certain</th>
<th>potential</th>
<th>no</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>certain</td>
<td>2</td>
<td>7</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>potential</td>
<td>1</td>
<td>5</td>
<td>33</td>
<td>2</td>
</tr>
<tr>
<td>no</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>13</td>
</tr>
</tbody>
</table>
## Results: Features

Mean results of classification experiment where each line reports the performance when the respective feature was removed:

<table>
<thead>
<tr>
<th>Features</th>
<th>CV</th>
<th>STD</th>
<th>TEST</th>
<th>STD</th>
</tr>
</thead>
<tbody>
<tr>
<td>all</td>
<td>.677</td>
<td>.033</td>
<td>.696</td>
<td>.042</td>
</tr>
<tr>
<td>-source</td>
<td>.519</td>
<td>.05</td>
<td>.586</td>
<td>.046</td>
</tr>
<tr>
<td>-semantic relationship</td>
<td>.675</td>
<td>.031</td>
<td>.694</td>
<td>.049</td>
</tr>
<tr>
<td>-syntactic importance</td>
<td>.532</td>
<td>.028</td>
<td>.521</td>
<td>.047</td>
</tr>
<tr>
<td>-structural placement</td>
<td>.695</td>
<td>.033</td>
<td>.708</td>
<td>.047</td>
</tr>
<tr>
<td>-rhetorical role</td>
<td>.687</td>
<td>.033</td>
<td>.695</td>
<td>.049</td>
</tr>
<tr>
<td>-attribution</td>
<td>.657</td>
<td>.034</td>
<td>.671</td>
<td>.048</td>
</tr>
<tr>
<td>-assignment/contrast</td>
<td>.668</td>
<td>.032</td>
<td>.669</td>
<td>.045</td>
</tr>
<tr>
<td>-feature assignment</td>
<td>.662</td>
<td>.032</td>
<td>.684</td>
<td>.047</td>
</tr>
</tbody>
</table>
Presentation Overview

Motivation

Task

Statutory Interpretation Data Set

Features for Predicting Interpretive Usefulness of Sentences

Predicting Usefulness

Conclusion
The ultimate aim is to develop and test a fully functional and well described framework supporting interpretation of statutory terms.

We would like to further develop the component for predicting interpretive value of a sentence.

We would also like to focus on the other constituents of the processing pipeline.

As the next step we would like to add new data to the corpus to have 1,000–2,000 sentences.

Then, we would like to test H2 (outperform the existing systems in the retrieval of the $n$ best sentences).
Conclusion

We have shown that:

▶ a sentence may be reliably evaluated in terms of its usefulness for an interpretation of a selected statutory term. (0.746 inter-annotator agreement, 0.66 weighted kappa)

▶ by using the selected linguistic features it is possible to automatically evaluate how useful a sentence is for an interpretation of a selected statutory term. (0.696 agreement with gold s.)

... confirming H0 and H1.

Therefore, we have suggested a feasibility of the framework for the computational support for interpretation of statutory terms sketched in this talk.
“Enterprise” means the related activities performed (either through unified operation or common control) by any person or persons for a common business purpose, and includes all such activities whether performed in one or more establishments or by one or more corporate or other organizational units including departments of an establishment operated through leasing arrangements, but shall not include the related activities performed for such enterprise by an independent contractor. […]
“Enterprise” means the related activities performed (either through unified operation or common control) by any person or persons for a common business purpose, and includes all such activities whether performed in one or more establishments or by one or more corporate or other organizational units including departments of an establishment operated through leasing arrangements, but shall not include the related activities performed for such enterprise by an independent contractor. [...]
“Enterprise” means the related activities performed (either through unified operation or common control) by any person or persons for a **common business purpose**, and includes all such activities whether performed in one or more establishments or by one or more corporate or other organizational units including departments of an establishment operated through leasing arrangements, but shall not include the related activities performed for such enterprise by an independent contractor. […]

**List of Interpretive Sentences**

The **“common business purpose”** requirement is not defined in the Act.

The utilization of a common service does not by itself establish a **common business purpose** shared by the owners of separate businesses.

Activities are performed for a **common business purpose** if they are “directed toward the same business objective or to similar objectives in which the group has an interest.”

In a situation such as this, in which the Court has concluded that there are no related activities, the fact of common ownership of the two businesses clearly is not sufficient to establish a **common business purpose**.

The Fifth Circuit has held that the profit motive is a **common business purpose** if shared.
Thank you!

Questions, comments and suggestions are welcome now or any time at jas438@pitt.edu.
Kevin D Ashley and Vern R Walker, *From information retrieval (ir) to argument retrieval (ar) for legal cases: Report on a baseline study.*


Jordan Daci, *Legal principles, legal values and legal norms: are they the same or different?*, Academicus International Scientific Journal 02 (2010), 109–115.


Raquel Mochales-Palau and M Moens, *Study on sentence relations in the automatic detection of argumentation in legal cases*, FRONTIERS IN ARTIFICIAL INTELLIGENCE AND APPLICATIONS 165 (2007), 89.


