Case Frames as Contextual Mappings to Case Law in BestPortal

Rinke Hoekstra \textsuperscript{a,b} Arno Lodder \textsuperscript{c} Frank van Harmelen \textsuperscript{a}

\textsuperscript{a} AI Department, VU University Amsterdam, The Netherlands
\textsuperscript{b} Faculty of Law, University of Amsterdam, The Netherlands
\textsuperscript{c} Computer Law Institute, Faculty of Law, VU University Amsterdam, The Netherlands

Abstract

This paper introduces case frames as a way to provide a more meaningful structure to vocabulary mappings used to bridge the gap between laymen and legal descriptions of court proceedings. Case frames both reduce the ambiguity of queries, and improve the ability of users to formulate good quality queries. We extend the BestMap ontology with a formalisation of case frame based mappings in OWL 2, present a new version of BestPortal, and show how case frames impact retrieval results compared to simple contextual mappings and a direct full text search.

1 Introduction

Publishing metadata as Linked Data does not automatically accommodate the fact that different communities may, and most often will, use different vocabularies to describe the same data. This problem frequently arises in Law, where government, citizens and businesses form different speech communities but have a shared interest [4]. This is a problem for improving access to legal information for people from a different speech community: laymen will be unable to phrase their queries in such a way that relevant documents are retrieved: a search on dog, child and bite may retrieve the one case in which a dog bit a child, but not any other cases where owners of animals are liable for damages caused to a third party. In previous work we therefore argued for a translation approach to legal information disclosure: queries to a corpus of legal documents are translated from ordinary language to their legal equivalent. This approach was implemented in the first version of BestPortal; a search engine for court proceedings.\footnote{BestPortal is accessible to the public through http://semweb.cs.vu.nl/best-portal.}

In this paper, we introduce the use of case frames to provide more meaningful structure to the translations used in BestPortal.

Approach

In the translation approach to legal information retrieval [3] the queries posed by laymen are translated to the queries a legal professional would use to find relevant information. It sets this work apart from initiatives that aim to provide a formal definition of types of liability for the purpose of (semi) automatic liability attribution. Translating layman descriptions to a legal characterisation of a case is not trivial [5]. We specify BestMap mappings [2] between everyday descriptions of cases and the terminology a legal professional would use to find relevant case law pertaining to that case [3]. These mappings are contextual: they focus on the co-occurrence of concepts in the description rather than one to one correspondence between everyday and legal concepts.

Users of BestPortal express their query by selecting everyday ‘laymen’ concepts. A standard OWL reasoner is then used to determine whether the individual matches a mapping defined in the ontology. If this is the case, any restrictions on that class using the legal vocabulary will be inferred to hold for the individual as well. We applied this methodology in the development of mappings for the legal area of unlawful acts. Once we obtain a description expressed using legal concepts, the lexical manifestations of these concepts – a combination of weighted phrases, a vector or fingerprint – are used as a query on an indexed corpus of case law. For instance, a query for “strict liability for animals” will include the phrase “the animal’s own energy”, which indicates that the volatility of the animal is discussed in a document. This distinguishes our
Table 1: Retrieval results for “Dog bites child” vs “Child bites dog”, top 20 hits.

<table>
<thead>
<tr>
<th></th>
<th>Rechtspraak.nl</th>
<th>w/o Case Frames</th>
<th>Case Frames</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>“Dog bites child”</td>
<td>1</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>“Child bites dog”</td>
<td>0</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

approach from more traditional concept-based search where search is performed on the basis of the word senses of synonyms of the concept label. However, [3] does not take into account what roles concepts play in a case description: there is no distinction between a biting dog and a biting child. In this paper we make these role explicit in a ‘case frame’ [1].

The unlawfulness of an action is determined by identifying the individual participants in the action, the situation in which the case takes place, and its consequences. Legal descriptions identify unlawfulness, duty of care, justification, exclusion of guilt, ground for attribution, strict liability, liberating circumstance, liable person. These roles were implemented in a new version of BestPortal, featuring client-side JavaScript, RDFa annotations and integration with GeoNames and DBpedia.

**Evaluation and Discussion** A first small evaluation shows that using case frames has a significant effect on retrieval quality over the simple contextual mappings. Table 1 compares results for our example queries, “Dog bites child, resulting in damage” and “Child bites dog, resulting in damage”, achieved via a naive full text query on Rechtspraak.nl, and a query on the same corpus via simple mappings and via case frames.\(^2\) Rechtspraak.nl could only find a limited number of applicable cases (a total of four), and could not distinguish between the two queries. The one correct hit concerned a case where a woman was bitten by a dog. Using simple mappings improves matters somewhat – the total number of results rose to over 140 thousand – but still produces the same results for both queries, and generates high number of false positives. Case frames have a significant positive effect on the first query: false positives dropped to 30%, with no cases of animal molest. The top 6 results all concerned strict liability for animals. However, case frames do not perform as well for the second query, where most results concerned child custody cases\(^4\) and none of the results were about animals.

Although the quality of search results can be much improved, even this very limited evaluation shows the potential for case frame based conceptual search as compared to the simple mappings. The difference with a traditional naive fulltext search on the same corpus is even more striking and shows that the current Rechtspraak.nl requires users to make an effort in coming up with the right formulation.

**References**


\(^2\)Our full text search string was “dog bites child damage”. \(^3\) Results were capped at a maximum of 20. \(^4\)These form the bulk of court proceedings that concern children.