Towards Annotating and Extracting Textual Legal Case Elements

ADAM WYNER


1. INTRODUCTION

In common law contexts, judges and juries decide a legal case to follow previously decided cases (precedents) rather than legislation as in civil law contexts. The set of such cases is the legal case base. Legal professionals must find, analyse, and reason with and about cases drawn from the case base in the course of arguing for a decision in a current undecided case. A range of elements of cases may be relevant to query and extract such as the citation index, participants, locale, jurisdiction, representatives, judge, prototypical fact patterns (factors), applicable law, and others. Commercial providers of legal information allow legal professionals to search the case base by keywords and meta data. However, the case base and search tools are proprietary, of limited, non-extensible functionality, and are restricted access. Moreover, no provider works with Semantic Web functionalities such as ontologies or rich XML annotations, nor are natural language processing techniques applied to the cases to support analysis to acquire information.

Text annotation of unstructured linguistic information is a significant, difficult aspect of the “knowledge bottleneck” in legal information processing. In this paper, we apply natural language processing tools to textual elements in cases, which are unstructured text, to produce annotated text, from which information can be extracted, thus contributing to overcoming the bottleneck. The extracted information can then be submitted to further processes. Where the annotations are associated with an ontology along

* The Author is Research Associate at University of Liverpool, Department of Computer Science.
1 Correspondence to Adam Wyner adam@wyner.info.
with an associated case based reasoner\(^3\), then we make progress towards a
textual case based reasoning system which enables processing from natural
language case decisions in the case base to generated decisions in novel cases\(^4\).
However, this paper focuses on the initial development in annotating cases
with respect to case elements.

The paper is a feasibility study for future research on information extrac-
tion of case elements\(^5\). In this paper, we focus on case elements rather than
case factors\(^6\).

In Section 2., we discuss background and materials. In Section 3., we
present the methodology, which uses the General Architecture for Text En-
gineering (GATE) system, sample components of system, sample results,
and a work flow for further refinement\(^7\). Finally, in Section 4., we review
the paper and outline future work to evaluate and improve our results.

2. BACKGROUND AND MATERIALS

Legal case based reasoning with factors has been a topic of central con-
cern in artificial intelligence and law. For our purposes, there are two main
branches of research. One branch, knowledge representation and reasoning
systems, requires a knowledge base that is constructed by manual analysis\(^8\).

---


\(^5\) Contact the Author for materials.


\(^7\) For GATE, see http://gate.ac.uk/.