

1 Conceptual Organization of Case Law Knowledge Bases, Carole Hafner (Adam Wyner)

The legal case base contains thousands of cases, all represented as text. In the last forty years or so, the corpora of cases have been digitally represented, enabling digital searches. With new technologies come new opportunities. For the case base, this was the opportunity to search quickly and efficiently through the cases for particular information, facilitating the identification of relevant cases for academic study or legal argument. One prominent approach since the 1960s was simply Boolean searches, using “and”, “or”, and “not”, for matching strings; companies such as LexisNexis and Westlaw provided legal firms with databases and search facilities, showing it to be a commercially viable approach. To the current day, string search continues to be the primary search approach used by web-based search engines, e.g. Google, albeit vastly improved with the use of pre-indexation of webpages on the Internet.

However, string searches are limited in a range of respects. In [2], several of these limitations are highlighted, and an alternative approach is proposed - concept indexation and concept search. Among the limitations, it requires some expertise to know what search terms to use, it is not feasible to know all the relevant combinations of terms, and there is an over-estimation of the quality of the results. A better approach would be to annotate the cases with conceptual cover terms, which are semantically meaningful generalisations over the textual particulars and variant expressions - textual differences would be homogenised, and relationships between textual elements could more easily be identified. Then searches would be with respect to fewer terms or related terms, providing a better, more accurate rate of returned cases. However, to annotate the cases requires that the textual content be annotated. Prior to the annotation task, we must provide some “map” of the relevant legal concepts and issues and their relationships to the legal case corpora.

The paper proposes an organization of a case law knowledge base in terms of three interacting components. First, there is domain knowledge model that defines the basic concepts of a case law domain; these may be the actors, events and relationships from a particular domain, e.g. intellectual property or personal injury liability. Second, a legal case frame that represents the properties, roles and relationships of elements of the case and which the case instantiates. And finally, Issue/Case Discrimination Trees that represent the significance of each case relative to a model of the normative relationships of the legal domain, e.g. shephardised relationships and case-based reasoning. In current terminology, the first two would be ontologies, an ontology of the domain and of legal cases, which the last reflects reasoning.

Case notes from Annotated Law Reports are used to derive the various components, among which we find:

Legal basis: Plaintiff, Defendant, Cause of Action

Background: description of underlying events, facts, and undisputed legal concepts.

Issues: disputed legal questions that must each be decided.

Procedural context: e.g. case on appeal, the lower court’s decision, and the grounds of appeal.

Holdings: decisions on all the issues and which side prevails.

Legal Theory: decision rules used in the case.

Cases then can be annotated with the particulars. When we search, we can search by concept, not string. For instance, we search for cases the defendant was a pet owner, the cause of action was personal liability injury, the background facts included injury caused by the pet, and the decision was in favour of the plaintiff.

Beyond simply annotating the particulars of the case, we must see how the particulars are tied to the decision since the cases that one wants to extract from the case base share particulars, but also the reasoning from the particulars to the decision; that is, the system needs case-based reasoning. The rules that the system uses must deal with conflicting rules, exceptions, negative queries, and “customary factors that may influence an outcome” rather than necessary and sufficient factors. In [2], Issue/Case Discrimination Trees are introduced for this purpose, where there are nodes for legal issues and nodes for factors that normatively influence the decision in the case. There are several ways nodes are connected to make the system more nuanced and articulated. Finally, cases are linked to the tree, where the cases are associated with issues and factors, and the direction of the decision. The idea is that by using the Trees, one can enter in basic information about a case, then retrieve closely related and relevant cases (or vary the information test alternative results); in other words, it uses the conceptual representation of the cases along with a system of expert legal reasoning about the cases.

[2] represents an early effort to use conceptual representations of cases together with expert case-based reasoning to support information extraction and case searching. Related work at the time and subsequently, e.g. [1, 4, 3], developed various ways of legal case-based reasoning relative to the facts as well as the relation to the textual source. However, the conceptual representation of the domain concepts and of the case frame seem to be a distinct aspect of [2] that was not taken much further in subsequent research.

One of the reasons for this, I believe, is that conceptual annotation is a knowledge and labour intensive task - the knowledge bottleneck, which is the problem of getting the textual data into a marked up form for conceptual retrieval. There is currently some progress along these lines, using well-developed, powerful text analytic tools [5]. Indeed, reviewing [2] has helped to identify some of the legal conceptual constants that can be targetting for future text analytic studies.

References

- [1] K. Ashley. *Modelling Legal Argument: Reasoning with Cases and Hypotheticals*. Bradford Books/MIT Press, Cambridge, MA, USA, 1990.
- [2] C. Hafner. Conceptual organization of case law knowledge bases. In *ICAIL '87: Proceedings of the 1st International Conference on Artificial Intelligence and Law*, pages 35–42, New York, NY, USA, 1987. ACM.

- [3] E. L. Rissland, K. D. Ashley, and L. K. Branting. Case-based reasoning and law. *The Knowledge Engineering Review*, 20:293–298, 2006.
- [4] R. O. Weber, K. D. Ashley, and S. Brüninghaus. Textual case-based reasoning. *Knowledge Engineering Review*, 20(3):255–260, 2005.
- [5] A. Wyner. Towards annotating and extracting textual legal case elements. *Informatica e Diritto: Special Issue on Legal Ontologies and Artificial Intelligent Techniques*, 19(1-2):9–18, 2010.