Textual Processing of Legal Cases

Adam Wyner
Department of Computing Science
University of Aberdeen

Law School
University of Aberdeen
May 31, 2013
Overview

• My background.
• State-of-play.
• Set the context.
• Spectrum of topics.
• Crowdsourcing legal case annotation.
My Background – AI and Law since 2001

- PhD in Linguistics (Cornell, 1994) and PhD in Computer Science (King's College London 2008).
- Contract concepts formalised in a logic and a computer program.
- Textual information extraction in legal cases and regulations.
- Legal case ontologies (objects and their relationships).
- Legal argumentation in case law and judicial hierarchy.
- Policy-making support tools.
- LegalRuleML – marking up the law to make it machine readable.
State of Play – Lots of Change

• Internet.
• Government information needs (transparency, efficiency....).
• Open Data and Big Data.
• Making the law machine-readable in UK National Archives and USA.
• More powerful search and reasoning tools.
• Economic change in firms, billing, hiring.... (Susskind).
• Blogs:
  
  http://legalinformatics.wordpress.com/
  http://computationallegalstudies.com/

May 31, 2013
Set the Context

- Support tools for information extraction and decision support.
- Formal analysis of the law and legal system.
- Decompositional analysis; incremental, partial solutions. (Think about how large, complex, multi-party legal problems are analysed and implemented).
- Individual v Collaborative, Collective, distributed legal work.
- Not robot lawyers, or fully automated justice, or 3 Laws.

<table>
<thead>
<tr>
<th>Law Applied to Computers</th>
<th>Computers Applied to Law</th>
</tr>
</thead>
<tbody>
<tr>
<td>privacy, software patents, data management, copyright....</td>
<td>knowledge extraction, representation, management; reasoning; networks.</td>
</tr>
</tbody>
</table>
Example

• Present an example available now that can be used for teaching, research, and practice.
Parts and Wholes

• Lots of open, unstructured legal data – cases, statues, regulation, guidance.

• How to find the information you want?

• Want to make cases and legislation searchable, but by concepts and relationships, not just by strings. Queries:

  – IBM is the plaintiff, Seventh Circuit Court is the jurisdiction, program code is the trade secret, and the decision upholds a lower court decision in favour of IBM.

  – The holdings by judge X and with respect to cause of action A against defendant Z
To analyse a legal case, legal professionals annotate the case into its constituent parts. The analysis is summarised in a case brief. However, the current approach is very limited:

- Analysis is time-consuming and knowledge-intensive.
- Case briefs may miss relevant information.
- Case analyses and briefs are privately held.
- Case analyses are in paper form, so not searchable over the Internet.
- Current search tools are for text strings or meta-tags, not concepts and relationships in the text.
Solution – Crowdssource Annotation

• Create a richly annotated, open source corpus for:
  – Online search.
  – Semantic web applications and information extraction.
  – Machine learning.

• Used by:
  – Law school students learning case analysis.
  – Legal professionals in identifying relevant cases.
  – Researchers in legal informatics.
Who Does This? Familiarity?

- Annotating done by law school students or legal professionals.
- Uses the same annotations that are learned in law school and applied in practice. No need to teach the annotations.
- Interannotator agreement and adjudication tools.
- Creation of a "Gold Standard".
- Make this a 'learning' tool – compare and contrast annotations, explaining why.
Project Website

• [http://wyner.info/LanguageLogicLawSoftware/?p=1315](http://wyner.info/LanguageLogicLawSoftware/?p=1315)

• The website has further information and a video on how to use the tool.
The annotations that are searchable in the corpus.
Arguments for a particular party, using the most general notion:

• for Appellee, for Appellant, for Defendant, for Plaintiff.
Facts

Legal and procedural facts:

• **Cause of Action** – the specific legal theory upon which the plaintiff brings the suit.
• **Defenses raised by Defendant** – the defendant's defenses against the cause of action.
• **Legal Facts** – the legally relevant facts of the case that are used in arguing the issues.
• **Remedy requested by Plaintiff** – what the plaintiff asks the court to grant.
Indexes

Various indicative information:

• Case Citation – the citation of the particular case being annotated.
• Court Address – the address of the court.
• Hearing Date – the date of the hearing.
• Judge Name – the names of the judge, annotated one at a time.
• Jurisdiction – the legal jurisdiction of the case.
Issues

Issues before the court:

• Procedural Issue – what the appellee claims that the lower court did wrong.
• Substantive Issue – the point of law that is in dispute (legal facts have their own annotation).
Legal Roles

The role of the parties in the case:

• Appellee, Appellee’s Lawyer, Appellant, Appellant’s Lawyer, Defendant, Defendant’s Lawyer, Plaintiff, Plaintiff’s Lawyer.

• General – buyer/seller, employer/employee, landlord/tenant, etc.
Relevant information not covered by the other annotations:

- Appellee, Appellee’s Lawyer, Appellant, Appellant’s Lawyer, Defendant, Defendant’s Lawyer, Plaintiff, Plaintiff’s Lawyer.
- General – buyer/seller, employer/employee, landlord/tenant, etc.
Various parts of the legal decision:

• Concurring Opinion.
• Dicta – commentary about the judgement and holding, but not part of the rationale.
• Dissenting Opinion.
• Holding – the rule of law or legal principle that was applied in making the judgement.
Reasoning Outcomes II

Various parts of the legal decision:

• Judgement – the court’s final decision about the rights of the parties, the court’s response to a party’s request for relief, and bearing on prior decisions (e.g. affirmed, reversed, remanded, etc.).

• Rationale – the court’s analysis of the issues and the reasons for the holding.
728 F.2d 818

220 U.S.P.Q. 167

AMERICAN CAN COMPANY, Plaintiff-Appellee,
v.

Ishwar MANSUKHANI, et al., Defendants-Appellants.

No. 82-2004.

United States Court of Appeals, Seventh Circuit.
have discovered such formulations, this is not a patent action and "obviousness" is not a
benchmark. See Ferroline Corp. v. General Aniline & Film Corp., 207 F.2d 912 (7th Cir.)

4 The compositions of series 400 inks were not known to others in the industry. Steps
taken to guard their secrecy--notably confidentiality agreements employees were reque-
signer with M & T and with Whittaker. There was no dispute as to the value of the forma-
tion or with respect to the time and effort spent in developing them. Duplication was not so
as to deprive the series 400 inks of trade secret status. A few sophisticated competi-

May 31, 2013
Wyner, Text Processing
Aberdeen Law School
Results of Annotation

• The annotators do their tasks and complete the project.
• Carry out inter-annotator agreement analysis.
• Curate the disagreements to create a Gold Standard corpus. Use this for machine learning.
• Search the annotations using an online tool, e.g. ANNIC.
Searching Annotations

{Legal Roles.Feature=="Plaintiff"}

Corpus: Entire datastore  Annotation set: All sets

Results:  Context size:  Next page of 50 results

Metals Company 279 F.2d 337  I. Jonas SPECINER, Trustee in Bankruptcy of ABC

Legal Roles.Feature

Configure
Searching Annotations

{Indexes.Feature=="Case Citation"}
FAQ

- How easy is it to learn to use the tool?
- What if I don't agree with some of your annotations or features?
- Can I get the results of my annotations?
- How many cases must I do?
- How much time will it take?
- What if I have a problem with using the tool or if the tool is buggy?
- How thoroughly should I annotate the cases?
- Do we track good annotators and bad annotators?
Other Tasks Using Similar Tools

• Regulatory rules - the rules that define a regulation such as 'if...then' or 'must'.
• Legal case factors for case based reasoning - the mitigating or aggravating circumstances that contribute to the decision.
• Argumentation patterns.
Other Todos

- Legal historical analysis.
- Applying a rule - heuristics and circumstances.
- Interpretive rules.
- Analysis of statutes, regulations, and guidance.
- Jury support tools.
- eDiscovery.
- Evidence.
- Crowdsourcing legislation.
- Predicting case decisions with respect to justices.
- Linking legal information and visualisations.
• Individual activities.
• Discussion group.
• Interdisciplinary funding opportunities.
Thanks for your attention!

- Contacts:
  - Adam Wyner, azwyner@abdn.ac.uk
- Crowdsourcing Legal Cases Website:
  [http://wyner.info/LanguageLogicLawSoftware/?p=1315](http://wyner.info/LanguageLogicLawSoftware/?p=1315)