A Case Study on Legal Case Annotation

Adam Wyner, University of Aberdeen
Wim Peters, University of Sheffield
Daniel Katz, Michigan State University

JURIX 2013
Bologna Italy

University of Aberdeen

The University of Aberdeen was founded in 1495 when William Elphinstone, Bishop of Aberdeen, petitioned Pope Alexander VI on behalf of James IV, King of Scots to create King's College. This makes it Scotland's third-oldest university (after the University of St. Andrews and the University of Glasgow) and fifth-oldest in the English-speaking world.

Wyner, Peters, and Katz
JURIX-2013

Overview

- A study on legal case annotation exercise with law school students.

- Variety of legal annotation types.

- Online tools for annotation and analysis.

- Results – IAA and Gold Standard creation.

- Queries

- Discussion

Wyner, Peters, and Katz
JURIX-2013

Introduction

- A preliminary study of annotating a corpus of legal cases.

- Background – previous papers on annotation.

- Objectives.

- Uses online tools.

- Results – data analysis and observations.

- Future work.

Wyner, Peters, and Katz
JURIX-2013

Objectives - Research

- Legal case content analysis – what is in legal cases and how is the content expressed?

- Fine-grained textual information extraction from across a corpus of cases.

- Create a gold standard for machine learning.

Wyner, Peters, and Katz
JURIX-2013

Objectives - Law Peoples

- Engage with law faculty and students in an activity they already carry out and in a manner familiar to them.

- Stream together individual annotation efforts.

- Create a gold standard and materials for human learning.

- Open source, transparent, curatable, reproducible vs. legal information service providers.

Wyner, Peters, and Katz
JURIX-2013

Corpus

- 20 cases from the CATO case base (of 140 cases) were annotated.

- 10 curated to a Gold Standard.

- All cases on appeal and bear in intellectual property.

- Various jurisdictions and dates.

- 231,555 tokens, various size (11KB-74KB; average 33KB).

Method - Annotations - What

- *32 annotations commonly used in law schools to annotate legal cases and using their common definitions.*

- Our case analysis task is the online, tool based version of what law school students do with cases.

Annotations - Facts and Indexes

Facts - legal and procedural facts:

- Cause of Action - the specific legal theory upon which the plaintiff brings the suit.
- Legal Facts - the legally relevant facts of the case that are used in arguing the issues.

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.

Annotations - Issues, Roles, Procedural History

Issues - the issues before the court:

- Procedural Issues - what the appellee claims that the lower court did procedurally wrong.
- Substantive Issues - the point of law that is in dispute.

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.

Procedural History - the disposition of the case with respect to the lower court(s):

- Appeal Information - who appealed and why they appealed.

Annotations - Reasoning Outcomes

Reasoning Outcomes - various parts of the legal decision:

- Holding - the rule of law or legal principle that was applied in making the judgement; the new legal ground that the court is covering in this case.
- Judgement - Given the holding and the corresponding rationale for the holding, the judgement is 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.
- Rationale - the court's analysis of the issues and the reasons for the holding.

Wyner, Peters, and Katz
JURIX-2013

Method - Tool, Who, How

- *Teamware* tool of General Architecture for Text Engineering (GATE) – a web-based application for annotator to annotate text given a list of annotations.

- 3 paid law school students over the summer (time was tracked by tool).

- Overseen by law school faculty (Katz).

- Coached how to use the tool.

Wyner, Peters, and Katz
JURIX-2013

Method - Blog

- http://wyner.info/LanguageLogicLawSoftware/index.php/2012/05/01/crowdsourced-legal-case-annotation/

- Blog contains background, justification of approach, annotation types, instructional videos on using the online tool, a FAQ, and a questionnaire about user experience.

Analysis

- Look at text span overlap – partial, total, non-overlap. Find high textual overlap for some annotations and low overlap for others.

Wyner, Peters, and Katz
JURIX-2013

Analysis - Overlaps

- High full agreement correlates with conceptual simplicity. Minor variations – space, punctuation....

- High partial agreement indicates complexity.

- Totals (non-overlaps):
  - Rationale: 816
  - LegalFact: 690
  - CauseOfAction: 84
  - Holding: 156

Table 1. Partial versus full agreement between pairs of annotators

Annotation Type	partial	full
Rationale	525	11
LegalFacts	342	21
Judgement	65	12
CauseOfAction	20	0
Holding	8	0
Plaintiff	6	25
CaseCitation	10	59
Appellant	16	26
AppellantsLawyer	2	48
DecisionDate	3	14
HearingDate	0	19
JudgeName	2	89
DefendantsLawyer	0	22
AppelleesLawyer	0	53
Defendant	16	32
Appellee	11	26
Jurisdiction	5	20

Wyner, Peters, and Katz
JURIX-2013
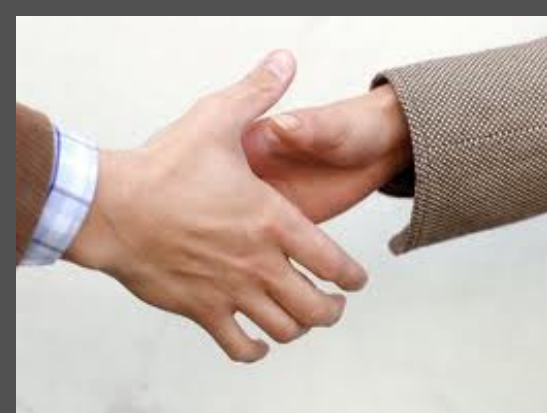
Analysis - IAA

- *Inter Annotator Agreement* because lots of partial overlaps, no Gold Standard, negative examples.

- Aggregated pair-wise Precision, Recall, and F-measures between annotators over all documents in the corpus. Assumes one in pair is "correct".
  - Precision: correct out of retrieved (accuracy).
  - Recall: correctly retrieved out of correct (coverage).

- Strict (favours strict overlap); lenient (favours partial overlap)

Table 2. Observed agreement between annotators

Annotation Type	P Strict	R Strict	F1 Strict	P Lenient	R Lenient	F1 Lenient
Indexes	0.84	0.82	0.83	0.93	0.90	0.91
CaseCitation	0.86	0.79	0.82	1.0	0.92	0.96
Jurisdiction	0.84	0.79	0.81	1.0	0.94	0.97
HearingDate	0.70	0.90	0.79	0.70	0.90	0.79
DecisionDate	0.74	0.45	0.56	0.89	0.55	0.68
LegalRoles	0.79	0.67	0.72	0.92	0.79	0.85
JudgeName	0.82	0.87	0.85	0.84	0.89	0.87
Plaintiff	0.76	0.64	0.69	0.94	0.79	0.86
Defendant	0.68	0.58	0.63	0.96	0.82	0.88
DefendantsLawyer	0.81	0.43	0.56	0.81	0.43	0.56
Appellant	0.59	0.52	0.55	0.86	0.76	0.81
AppellantsLawyer	0.90	0.81	0.86	0.94	0.85	0.89
Appellee	0.65	0.59	0.62	0.88	0.80	0.83
AppelleesLawyer	0.78	0.88	0.83	0.78	0.88	0.83
Facts	0.53	0.67	0.59	0.37	0.47	0.41
CauseOfAction	0	0	0	0.33	0.4	0.36
LegalFacts	0.1	0.06	0.06	0.36	0.45	0.4
ReasoningOutcomes	0.05	0.04	0.04	0.34	0.3	0.32
Holding	0	0	0	0.12	0.89	1
Judgement	.01	0.1	0.1	0.47	0.50	0.49
Rationale	0.22	0.18	0.2	0.22	0.18	0.2

Wyner, Peters, and Katz
JURIX-2013

Analysis – Curation

- Strategies
  - Majority "vote". Problem for annotations with too much variety.
  - Arbitration by annotators. Problem for students (though good for studying).
  - Arbiter. Needs an expert.
- For unproblematic annotations used majority; for problematic used arbiter.
- Produces a Gold Standard.

Wyner, Peters, and Katz
JURIX-2013

Analysis – Curation

Table 3. Observed agreement between annotators

Annotation Type	P Strict	R Strict	F1 Strict	P Lenient	R Lenient	F1 Lenient
Holding	0.03	0.33	0.06	0.18	0.61	0.25
Judgement	0.33	0.08	0.13	0.67	0.1	0.35
CauseOfAction	0	0	0	0.33	0.4	0.36
LegalFacts	0.1	0.06	0.06	0.36	0.45	0.4
Rationale	0.22	0.18	0.2	0.22	0.18	0.2

Shows little agreement between students and arbiter.

Room for didactic intervention.

Wyner, Peters, and Katz
JURIX-2013

Other Results

- Examine annotations *in situ* in the text.
- Examine annotations *distributed across* the texts.

Wyner, Peters, and Katz
JURIX-2013

Annotations within a Text

759 F2d 1053 Defiance Button Machine Company v. C & C Metal Products Corp

Wyner, Peters, and Katz
JURIX-2013

Annotations across Texts

- All the texts are indexed.

- Search across the texts using the semantic annotations.

- Simple or complex query patterns.

Wyner, Peters, and Katz
JURIX-2013

Cause of Action Examples

{CauseOfAction}

Corpus: Entire datastore   Annotation set: All sets

Results: ⊙────────   Context size: ⊙───

🔍 Search   ✖ Clear   ➡ Next page of 50 results

Context	breached a confidential relationship by misappropriating the trade secrets of their former employer, Temple
Token.string	breached a confidential relationship by misappropriating the trade secrets of their former employer, Temple ➖
CauseOfAction	➖
➕ Configure	

Page 1 (41 results)   💾 Export

Left context	Match	Right context	Features
Get 2 Tags Free". 9	Zoecon Industries so...mpetition agreement.	On August 19, 1980, the	Token.string=Zc
Get 2 Tags Free". 9	Zoecon Industries so...ompetition agreement .	On August 19, 1980, the	Token.string=Zc
name, this use constituted a	misappropriation of ... a misrepresentation	to the trade. Despite this	Token.string=mi
of its claim of alleged	conversion	of its customer lists and	Token.string=co
Ink Company (Carter's) and alleging	unfair competition a...ure of trade secrets	by Speedry to Carter in	Token.string=ur
on the basis of an	unjust enrichment su...a and the conversion	thereof to Mr. Durant, not	Token.string=ur
an accounting from, Carter's, asserting	unfair competition a...ure of trade secrets	by appellants and subsequent use	Token.string=ur
breached a confidential relationship by	misappropriating the trade secrets	of their former employer, Temple	Token.string=mi
seeks injunctive relief against defendants'	use of its trademark	DEFIANCE and trade name Defiance	Token.string=us
(10th Cir.1978). An action for	unauthorized use or ...ure of trade secrets	sounds in tort; privity of	Token.string=ur
suspicious that Smalling may have	infringed its patent	or at least stepped over	Token.string=inf

Annotation Type	Count
link	43
CauseOfAction	41
li	40
script	40
Appellee	38
DefendantsLawyer	36
Plaintiff	33
h2	30
Jurisdiction	29
DecisionDate	22
HearingDate	21

Global   One ite

Cause of Action & Holding

Wyner, Peters, and Katz
JURIX-2013

Holdings XML Export

```
<TR><TD>854, 94 L.Ed. 1097 (1950). </TD>
<TD>While the competing heaters solved the same industrial problems and
reflected certain design similarities, there were also substantial design
differences. Based on the record before us, we cannot say the finding of the
court below that plaintiff failed to meet its burden of proving infringement is
clearly erroneous. See Becker v. Webcor, Inc., 289 F.2d 357 , 360-61 (7th Cir.
1961), Cert. denied, 368 U.S. 970, 82 S.Ct. 445, 7 L.Ed.2d 398 (1962)</TD>
<TD>. 24 BS&amp;B's other claims of</TD>
<TD>Token.string=While, the, competing, heaters, solved, the, same, industrial,
problems, and, reflected, certain, design, similarities,, there, were, also,
substantial, design, differences., Based, on, the, record, before, us,, we,
cannot, say, the, finding, of, the, court, below, that, plaintiff, failed, to,
meet, its, burden, of, proving, infringement, is, clearly, erroneous., See,
Becker, v., Webcor,, Inc.,, 289, F.2d, 357, ,, 360-61, (7th, Cir., 1961),,
Cert., denied,, 368, U.S., 970,, 82, S.Ct., 445,, 7, L.Ed.2d, 398</TD>
<TD>{Holding}</TD>
<TD>black-sivalls-bryson-v-keystone-steel-fabrCURATED.html.xml2.
xml_00031___1378741220497___3984</TD>
<TD>annotator1</TD>
</TR>
```

```
<TR><TH>Left context</TH>
<TH>Match</TH>
<TH>Right context</TH>
<TH>Features</TH>
<TH>Query</TH>
<TH>Document</TH>
<TH>Annotation set</TH>
```

Wyner, Peters, and Katz
JURIX-2013

Discussion - Research

- Cases "in the wild" have a range of drafting styles:
  - Causes of action presented as issues or narratively;
  - Holdings presented as intermediate clarifications on a point or as a refinement of the law;
  - Scope of legal facts;
  - Rationales similar to holdings;
  - Legal facts part of rationale.
- How are commercial case briefs made? What evidence is there of their accuracy?

Wyner, Peters, and Katz
JURIX-2013

Discussion – Teaching and Learning

- Students need training and quality materials.
- Didactic intervention
  - Prepare gold standard for training.
  - Break the tasks down.
  - Student exercises in approximating the gold standard.
  - Foster discussion, exchange of information, and close textual analysis.

Wyner, Peters, and Katz
JURIX-2013

Discussion - Next

- Next steps
  - Decompose the task
    - Facts (Cause of Action, Legal Facts)
    - Issues (Procedural, Substantive)
    - Reasoning Outcomes (Holdings, Judgements, Rationales)
  - Extract and analyse the elements for linguistic indicators

Wyner, Peters, and Katz
JURIX-2013

Thanks

- Questions
- Comments

Wyner, Peters, and Katz
JURIX-2013