

Analyzing Legal Texts

Research Agenda:

- From Statistical Parsing
- To Semantic Interpretation
- To Information Extraction

L. T. McCarty, "Deep Semantic Interpretations of Legal Texts," *Proceedings of the Eleventh International Conference on Artificial Intelligence and Law*, pages 217-224 (ACM Press, 2007).

		cases	words	sentences	parse trees	null parses
SupCt	May 1999	9	35,796	1,213	1,141	5.94%
2dCirc	May 1999	31	66,333	2,341	2,170	7.30%
	June 1999	26	108,778	3,983	3,800	4.59%
3dCirc	May 1999	15	71,038	2,443	2,286	6.43%
	June 1999	30	148,320	5,382	5,132	4.65%
		111	430,265	15,362	14,529	5.42%

Preprocessing cases:

- remove embedded citations.
- decompose extremely long sentences (> 100 words).
- compute sentence boundaries and part-of-speech tags.

Partially automated, but manual effort is still required.

Acknowledgment: Stephen Max


```

cat 'SBAR'(whpp) :: W^(E^(P^Term & S)/Tense) *->
  head 'WHPP' :: W^E^P^Term,
  comp 'S'(Tense) :: E^S.

cat 'WHPP' :: W^Subj^P^Term *->
  head 'IN'(_Word) :: Obj^Subj^P^Term,
  args 'WHNP' :: W^Obj.

cat 'WHNP' :: W^Term *->
  head 'WDT'(_Word) :: W^Term.

```

Semantics of 'WDT' and 'WHNP': $W^{nterm}(which, W, [])$

Semantics of 'IN': $Obj^{Subj^P^pterm}(in, P, [Subj, Obj])$

Unify: $Obj = nterm(which, W, [])$

$Term = pterm(in, P, [Subj, Obj])$

Semantics of 'WHPP':

$W^{Subj^P^pterm}(in, P, [Subj, nterm(which, W, [])])$

```

cat 'SBAR'(whpp) :: W^(E^(P^Term & S)/Tense) *->
  head 'WHPP' :: W^E^P^Term,
  comp 'S'(Tense) :: E^S.

cat 'WHPP' :: W^Subj^P^Term *->
  head 'IN'(_Word) :: Obj^Subj^P^Term,
  args 'WHNP' :: W^Obj.

cat 'WHNP' :: W^Term *->
  head 'WDT'(_Word) :: W^Term.

```

Semantics of 'S': $E^{\text{sterm}(\text{claims}, E, [_, _])}$

Unify: $\text{Term} = \text{pterm}(\text{in}, P, [E, \text{nterm}(\text{which}, W, [])])$

$\text{Tense} = [\text{present}]$

Semantics of 'SBAR':

$W^{\left(E^{\left(P^{\text{pterm}(\text{in}, P, [E, \text{nterm}(\text{which}, W, [])]) \right)} \right.}$ &
 $\left. \left. \text{sterm}(\text{claims}, E, [_, _]) \right) / [\text{present}] \right)$

Structured Casenote:

...

The traditional case brief focuses on the procedural context first: Who is suing whom, and for what? What is the plaintiff's legal theory? What facts does the plaintiff allege to support this theory? How does the defendant respond? How does the trial court dispose of the case? What is the basis of the appeal? What issues of law are presented to the appellate court? How does the appellate court resolve these issues, and with what justification?

Information Extraction:

- Explicit Rules
- Supervised Learning
- Semi-Supervised Learning

Parsing the Text of the Uniform Commercial Code

Timothy J. Armstrong
L. Thorne McCarty

Workshop on
Natural Language Engineering of
Legal Argumentation

ICAAIL'09
Barcelona

Why Parse Statutes?

- To extract their logical structure.
- To refine the semantics of the domain.
- To develop a domain ontology.

Problems:

- Syntax is convoluted and unnatural.
- There is no tagged corpus (currently).

Preprocessing Statutes

Hierarchical Outline Structure:

- Remove labels from the text.
- Store outline with pointers into the text.

Statutory Cross References:

- Replace with synthetic proper nouns.
- After the parse, restore original references.

Coordinated Conjunctions

83 parses, 3.9 minutes, none correct:

"Check" means (i) a draft, other than a documentary draft, payable on demand and drawn on a bank or (ii) a cashier's check or teller's check.

100 parses, 8.7 minutes, none correct:

Notice of dishonor is excused if (i) by the terms of the instrument notice of dishonor is not necessary to enforce the obligation of a party to pay the instrument, or (ii) the party whose obligation is being enforced waived notice of dishonor.

69 parses, 25.4 minutes, none correct:

"Alteration" means (i) an unauthorized change in an instrument that purports to modify in any respect the obligation of a party, or (ii) an unauthorized addition of words or numbers or other change to an incomplete instrument relating to the obligation of a party.

Coordinated Conjunctions

37 parses, 25.5 minutes, first parse was correct:

Without dishonoring the instrument, the party to whom presentment is made may (i) return the instrument for lack of a necessary indorsement, or (ii) refuse payment or acceptance for failure of the presentment to comply with the terms of the instrument, an agreement of the parties, or other applicable law or rule.

55 parses, 66.5 minutes, none correct:

Under subsection (a) or (b), an indorsement is made in the name of a payee if (i) it is made in a name substantially similar to that of the payee or (ii) the instrument, whether or not indorsed, is deposited in a depository bank to an account in a name substantially similar to that of the payee.

71 parses, 37.2 minutes, none correct:

"Traveler's check" means an instrument that (i) is payable on demand, (ii) is drawn on or payable at or through a bank, (iii) is designated by the term "traveler's check" or by a substantially similar term, and (iv) requires, as a condition to payment, a counter-signature by a person whose specimen signature appears on the instrument.

Prepositional Phrases

100 parses, 31.2 minutes.

First parse was correct: VP(perform) --> VB(perform) NPB(duties) PP(through)

A party may perform its duties through a delegate.

100 parses, 13.2 minutes.

Correct on phrase: NP(acceptance) --> NNP(acceptance) PP(of) PP(by)

Acceptance of a delegation of duties by the assignee constitutes a promise to perform those duties.

100 parses, 60.4 minutes.

Never correct: VP(explained) --> VBN(explained) PP(by) PP(without)

Terms in a record may be explained by evidence of course of performance, course of dealing, or usage of trade without a preliminary determination.

Prepositional Phrases

100 parses, 144.6 minutes.

Never correct: NP(affixing) --> NPB(affixing) PP(of) PP(to) VP(does)

The affixing of a seal to a record evidencing a contract for sale or an offer to buy or sell goods does not constitute the record a sealed instrument.

100 parses, 103.4 minutes.

First parse was correct: VP(made) --> VBN(made) PP(by) PP(to)

Any remedial promise made by the seller to the immediate buyer creates an obligation that the promise will be performed upon the happening of the specified event.

100 parses, 90.1 minutes.

Never correct on phrase: NP(expression) --> NPB(expression) PP(of) PP(in)

A definite and seasonable expression of acceptance in a record operates as an acceptance even if it contains terms additional to or different from the offer.

Conclusion

We need a tagged legal corpus ...