Extracting Definitions from Court Decisions

Manfred Pinkal & Stephan Walter

Universität des Saarlandes, Saarbrücken
Computerlinguistik

Topics

- Aims:
  - Automated definition extraction from court decisions
  - Analysis and integration of extracted definitions
- Rule-based approach
- Using semantically oriented parsing technology

Outline

1. Why Definitions?
2. Why more than shallow processing?
3. Extractors and Evaluation
4. Beyond the document
5. Conclusion

Information Access for the Legal Practitioner

- Enormous amounts of text produced by courts every day
- Often available electronically
- However hardly any advanced technological support for information access (at least in Germany)

  E.g. *juris*:
  - ~27 000 decisions per year, total >8 mio docs
  - Full-text search, boolean operators
  - Fragmentary, unsystematic term index (covering only about 21000 documents)

⇒ Concept-centered access to court decisions still mainly through manually compiled (printed) commentaries

⇒ Identifying definitions is the key to enabling automatic support for this purpose

Why Definitions?

- Definitions generally regarded as valuable information nuggets
  - e.g. one established task in QA (cf. TREC)
- Definitions of particular importance in *court decisions*:
  - Backbone of legal interpretation
  - Rapidly developing body of legal knowledge (supplementary to relatively 'static' knowledge in statutes)

Normative vs. Descriptive Content in Statutes

Statutes have two different kinds of content:
Normative *and* descriptive content

Normative content:

States of affairs are assigned legal consequences

$$\forall x(Soa(x) \rightarrow \Box LConseq(x))$$

(...) the responsibility for maintaining waters shall lie with the owners of waters, the riparian owners (...)
(Section 29, Federal Water Act)

Am I responsible for the maintenance of a tubed ditch leading through my garden (built for draining my neighbour's garden)?

Descriptive content

Descriptive content defines concepts for describing the situations to be sanctioned by the statute:

*The responsibility for maintaining **waters**…*

*This Act shall apply to the following **waters**:*

1. permanently or temporarily flowing or standing waters confined within a bed

(Article 1, Federal Water Act)

⇒ Statute texts provide generic, wide coverage definitions

⇒ Body of definitions in statutes is relatively fixed (changed only by legislation)

Idealized (and naive) model: descriptive and normative content of statutes fixes decision in all possible cases

Coded law is not enough

However, reality is not as abstract: Judges have to supply further, more specific definitions when deciding concrete cases:

*...confined within a **bed***

*By a **bed of a body of water** is to be understood: the natural (...) confines of water within a cavity in the surface of the earth (cf. BVerwG, Urt. v. 31.10.1975, BVerwGE Bd. 49 S. 293, 298; Beschl. v. 17.2.1969, Buchholz 445.4 § 1 WHG Nr. 3, m.w.N.).*

*Such a **bed of a body of water** (...) can no longer be assumed if a ditch is fully tubed.*

- Judges' definitions become binding through repeated reference in other cases and commentaries
- They get elaborated with respect to the case at hand

=> More flexibility and greater fluctuation than definitions in statutes

Role and Structure of Definitions

Judges' definitions play a central role for:

- Interpretation: connecting statute and case
- Cohesion: connecting concepts
- Intertextuality: connection to precedent

But they are also special with respect to their internal structure.

Analysis of internal structure is e.g. needed to find out

- what's defined (and how)
- if definition is applicable
- what kind of definition is given

Internal Structure of a Definition

Bei einem Einfamilienreihenhaus liegt ***ein mangelhafter Schallschutz***

In a one-family row-house is an insufficient noise-insulation

in der Regel dann vor, wenn *die Haustrennwand einschalig errichtet wurde.*

as a rule then [], if the house-separating wall one-shelled built was.

(As a rule, one-family row-houses have insufficient noise insulation if their separating wall is one-shelled.)

- ***Definiendum*** / *Definiens*
- Connector (verb+conjunction, rel. pron, punctuation)

- Scope of application (e.g. N, if attribute is being defined)
- Modification (e.g. as a rule, typically)
- Legal field (e.g. im Umweltrecht – in environmental law)
- Citation data

Standard Methods

Standard method for definition mining:

Surface patterns + shallow processing (POS-tagging, chunking)

$\Rightarrow$Not enough for German legal text,

especially if internal structure of definitions is of interest

Difficulties for standard methods

German legal language characterized by:

- Passive constructions
- Complex, deeply embedded sentences
  => Changing word order, predicate may be split / distributed
  => Proliferation of surface patterns

- Many nominalizations
- Complex NP/PP-structure
  => term delimitation / segmentation of definitions needs to know about boundaries / dependencies within phrases

- Conscientious and meaningful use of modalities
  => Must be respected e.g. to find out if a definition is accepted or quoted+refuted

=> Deeper linguistic processing needed

A complex PP

[Bei der Umsetzung]$_0$
[der Vorgaben]$_1$
[der Gerichte]$_2$
[für eine verfassungskonforme Regelung]$_2$
[der Überführung]$_3$
[von Ansprüchen und Anwartschaften]$_4$
[aus den Zusatz- und Sonderversorgungssystemen]$_4$
[der ehemaligen DDR]$_5$
…

In implementing the requirements imposed by courts for a constitutional regulation of the transfer of claims and entitlements out of additional and special provision systems of the former GDR…

Parser

- Parser: Developed at Saarbrücken CL department, uses PREDS (Partially Resolved Dependency Structures, Braun, 2003)
  - Topological analysis (*sentence bracket* and *fields*)
  - Internal structure of topological fields: Phrase-chunking and named entity-recognition (dates, company names, citation data etc.)

- Construction of recursive partially resolved dependency structures ("PREDS")
  - Verb + prefix, auxiliaries => predicate
  - Complements => arguments
  - Adjuncts (e.g. adverbials, subclauses) => modifiers (various relations)
  - Normalization of active / passive, modalities, tense
  - Easy mapping to text spans

- Robustness: uses heuristics and underspecification for attachment
- Produces XML-Structure with accumulated linguistic information

Example PREDS

Bei einem Einfamilienreihenhaus liegt ***ein mangelhafter Schallschutz***

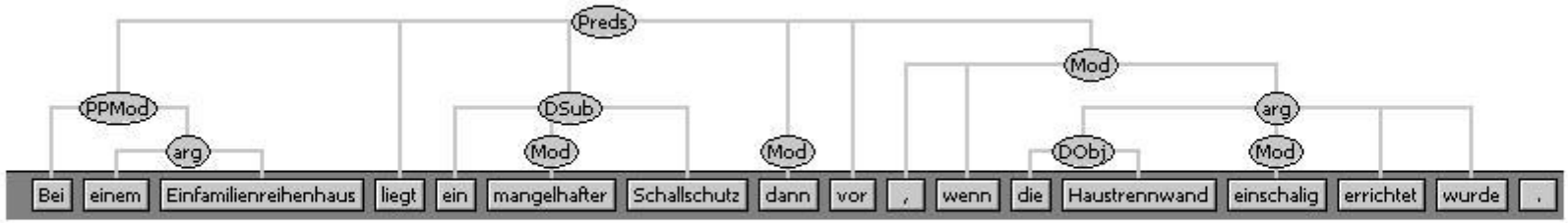
In a one-family row-house is an insufficient noise-insulation

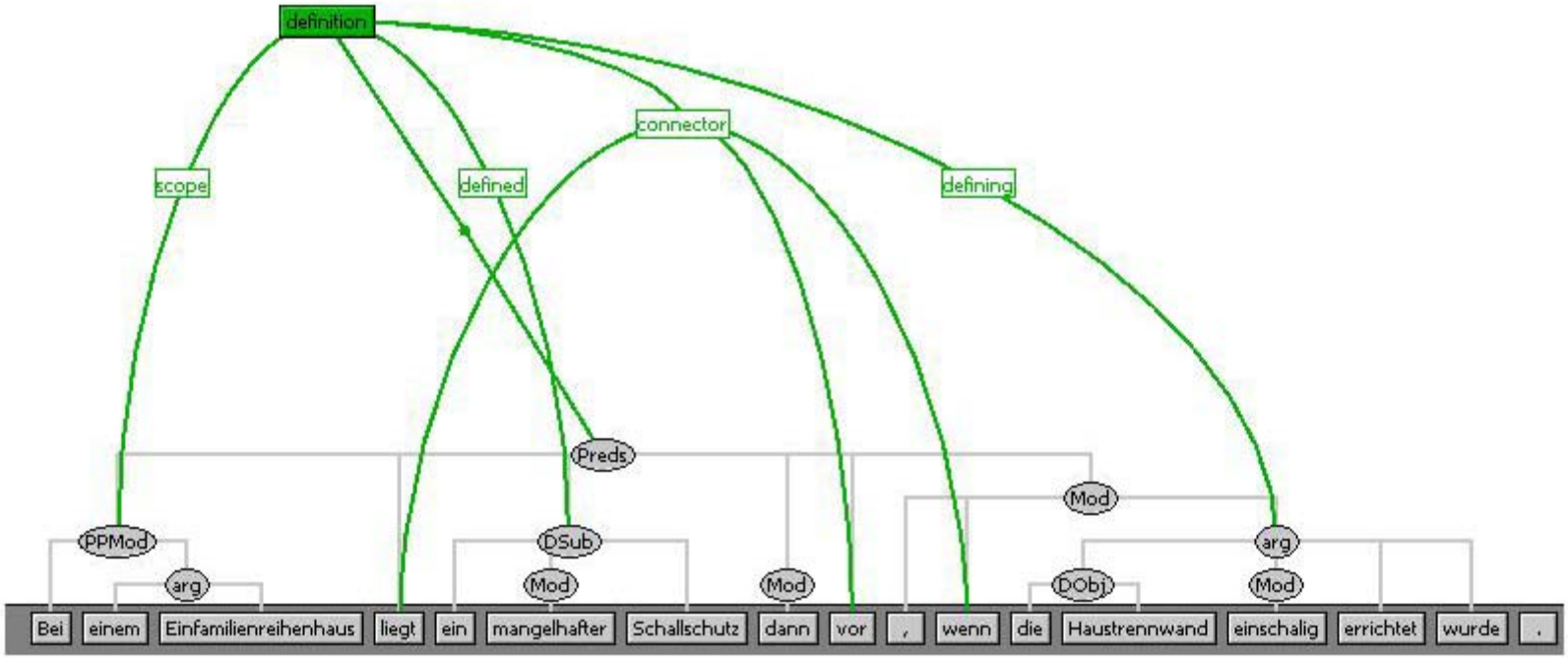
dann vor, wenn *die Haustrennwand einschalig errichtet wurde.*

then [], if the house-separating wall one-shelled built was.

(One-family row-houses have insufficient noise insulation if their separating wall is one-shelled.)

```
vorliegen[verb, sg, ind]
  -PPMod->bei[praep, dat]
            -Arg->*ein#familien#reihen#haus[noun]
  -DSub->*schall#schutz[noun]
           -Mod->mangelhaft[adj]
  -Mod->wenn[subj]
          -Arg->errichten[verb, sg, ind, pres, pass]
                  -DObj->*haus#trenn#wand[noun, defArt]
                  -Mod->einschalig[adv]
  -Mod->dann[adv]
```


Preds
Mod
PPMod
DSub
arg
Mod
Mod
DObj
arg
Mod
Bei
einem
Einfamilienreihenhaus
liegt
ein
mangelhafter
Schallschutz
dann
vor
,
wenn
die
Haustrennwand
einschalig
errichtet
wurde
.


definition
scope
defined
connector
defining
Preds
PPMod
arg
DSub
Mod
Mod
Mod
DObj
arg
Mod
Bei
einem
Einfamilienreihenhaus
liegt
ein
mangelhafter
Schallschutz
dann
vor
,
wenn
die
Haustrennwand
einschalig
errichtet
wurde
.

Extractor patterns

First study: hand annotated 40 decisions (various legal fields)

⇒ 130 definitions

Various types of lexico-syntactic indicators:

1. Explicit 'definitor'-verbs + valency frame
   ist anzunehmen*, wenn – is to be* assumed *if*
2. *'is'*-definitions (*Lexical N* is *N+RC*, *Nominalized V* is …)
3. appositive / nominal (parentheses, brackets, non-restrictive RC)
4. 'transparent' noun + support verb
   begriffliche Voraussetzung + haben/sein –
   conceptual prerequisite + have/be
5. subjunction only (e.g. sentence with *wenn / if*-subclause)
6. unmarked

⇒ Seed of 33 extractor patterns for type 1

⇒ Filtering: pronouns, adjectives that establish definite reference (anaphoric or specific to the concrete case): *vorgenannt – mentioned above*; *klägerisch – belonging to the plaintiff, …*

Example: Extractor Patterns

[defined] *liegt vor*, *wenn* [defining]
[defined] *vorliegt*, *wenn* [defining]
wenn [defined] muss [defining] *vorliegen*
...

```
<pattern>
 <keys>vorliegen</key>
 <frames>
  <frame id="DSub-Cond">
   <mapping id="DSub:defined_cond:defining_1"/>
  </frame>
 </frames>
 <filters>defined-anaphora, stop-adjs</filters>
</pattern>

<frame id="DSub-Cond">
 <description>KEY + DSub-Cond</description>
 <query>[@key="KEY" and INDPRES and COND and DSub]</query>
</frame>

<mapping id="DObj:defined_cond:defining_1">
    <item field="defined">DOBJ</item>
    <item field="defining">COND/arg/word</item>
    <item field="area">PPMod{PREP%bei}/arg/word</item>
</mapping>
```

Evaluation of Precision

Corpus: ~6000 decisions in environmental law (237935 sent)

- 5461 hits in 4716 sents, filtered to 1342
- 473 hits checked by two annotators

Total		
33 rules	1486 hits (1342 / 237935 sent)	
Annotator 1	Good: 211/473	(p = 44.6 %)
Annotator 2	Good: 230/473	(p = 48.6 %)
Best rules only (κ = 0.835)		
		Annotator 1
17 rules	749 hits (749 / 1342 sent)	
	Good: 176/245	(p = 71.8 %)
		Annotator 2
18 rules	764 hits (633 / 1342 sent)	
	Good: 173/230	(p = 75.2 %)

Recall

- Recall hard to assess:
  - No reference corpus with annotated definitions exists
  - Creating one is hard: Low frequency of definitions, many cases of doubt
  - Recall problems are however obvious (there must be more than this…)

- Recall problems due too:
    - Small number of patterns (only *definitor-verb*-based)
    - Insensitive filtering (stopwords same for all patterns, many hits contain anaphoric elements)
    - Technical issues (parse errors, problems with conjunctions, …)

Present Topics

- Current Corpus:
  - 15000 decisions from environmental and administrative law (~800000 sents in reasons and edited headnotes)
- Pattern induction from training set (4000 decisions):
  - search for valency frames with variable verb / definitor verbs with variable valency frame
  - bootstrapping with strongly associated nouns from definiens + definiendum

=> Currently about 200 patterns, not yet evaluated

Future Work: Beyond the document

- Structured knowledge base reflecting:
  - Relations between multiple definitions for one concept (e.g. compatible / incompatible; implied / specialisation / new area)
  - Looking into the definiens: Negative and positive conditions, extracted features for concepts
  - Hierarchy of courts / timestamp of definition
  - "Definition chains"

- Definitions as source for ontology extraction:
  - Relations directly specified by definitions (is-a, part-of)
  - Relations used in definiens

Experiment: is-a Extraction

- Setting: *is-a* extraction through N-Adj-bigrams (filtered for stopwords, ranked by co-occurrence log-likelihood)

  E.g: unsorted waste *is-a* waste

- However: Not all bigrams follow this pattern:
  - N+Adj does not denote a relevant concept :
    differenziertes Regelwerk – differentiated body of rules, vermeintliches Problem – assumed problem
  - N+Adj-concept is not a subconcept of N:
    (e.g. non compositional collocations)
    öffentliche Hand – public hand, i.e. public authorities

is-a Extraction: Baseline

LL-Ranking of all bigrams with more than 5 occurences:

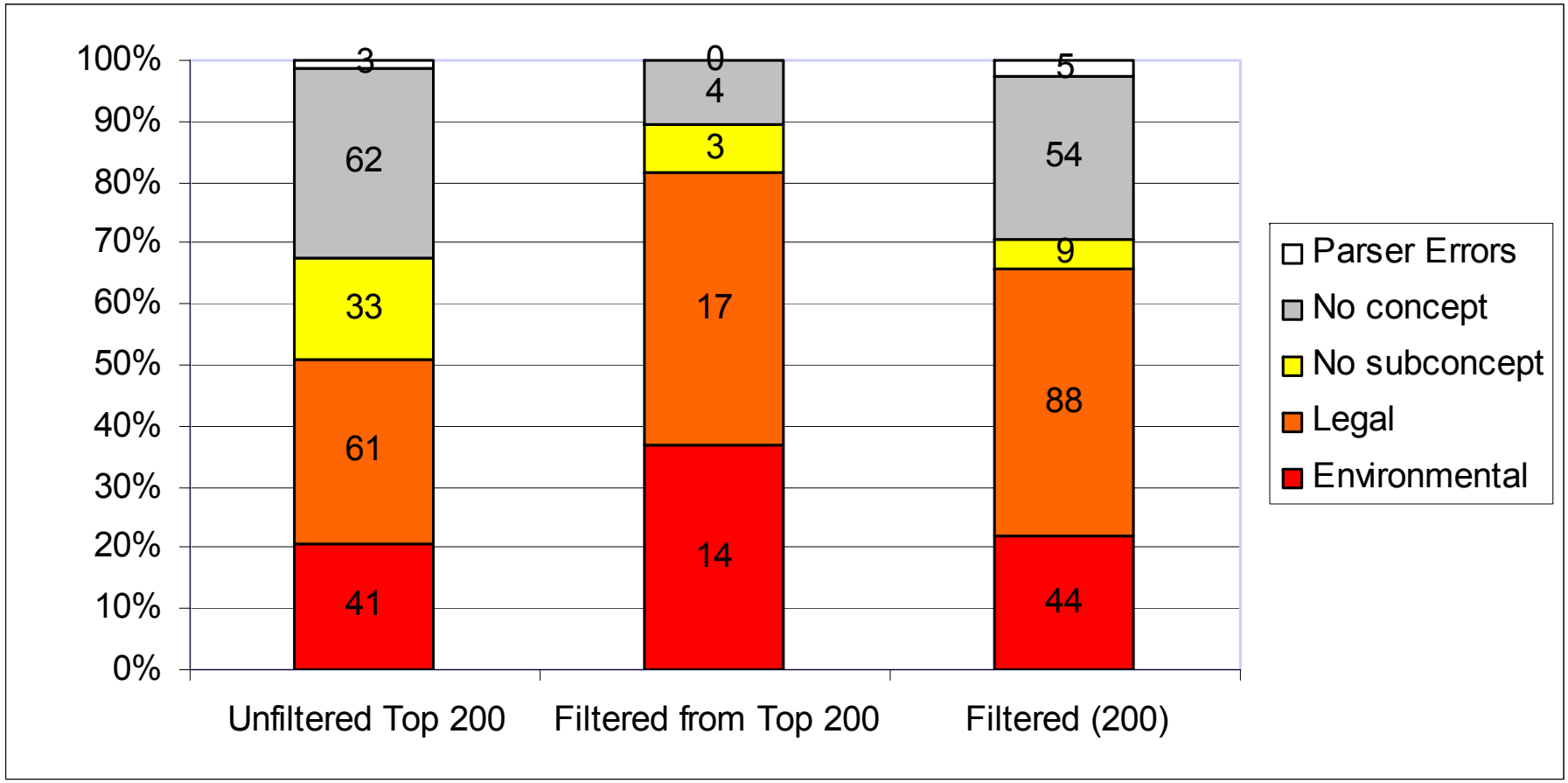
- 4371 bigrams (out of 73319) on 4320 ranks
- 46% precision on top 500
- 39% precision on ranks 3500-3600

Observation:

- Definienda are likely to contain domain terms
- Most domain terms are likely to be defined at some point

What is the effect of extracting bigrams only from definienda?

Evaluation


100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
3
62
33
61
41
0
4
3
17
14
5
54
9
88
44
Unfiltered Top 200
Filtered from Top 200
Filtered (200)
Parser Errors
No concept
No subconcept
Legal
Environmental

Evaluation II

Gain in precision:

51 % good hits in top 200 filtered vs.
66 % in top 200 filtered

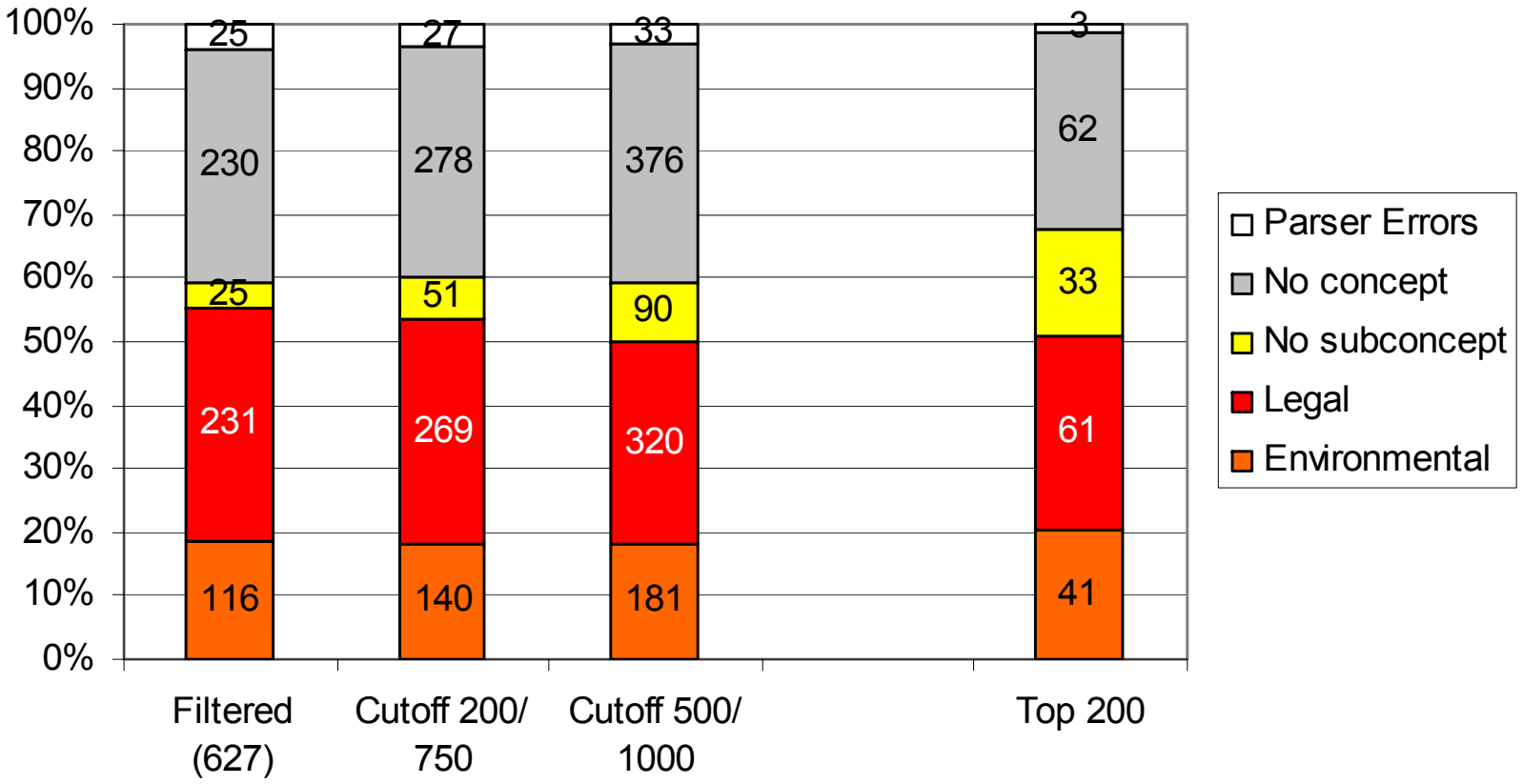
Enormous loss in recall:

Total of 227 bigrams left after filtering (out of 4371)

Solutions:

- Improve recall of definition extraction (better/more extractors)
- Combine top ranks of unfiltered ranking with lower ranks from filtered one (may even use ones with n<5)

Results of Combined Method



Conclusion

- Definitions from court decisions contain valuable knowledge for the legal practitioner
- Extraction and analysis requires relatively deep linguistic processing
- Precise extraction is possible. Recall is a problem, but there's still hope…
- Extracted definitions can be integrated in various ways to form new resources