Sentence Boundary Detection in Decisions of the US Courts

Jaromir Savelka
Kevin D. Ashley

Intelligent Systems Program
University of Pittsburgh

jas438@pitt.edu

March 17, 2017

Presentation Overview

Sentence Boundary Detection

US Court Decisions

Data Set

Performance of Vanilla SBD Systems

Performance of Trained SBD Systems

Conditional Random Fields Model

Summary, Conclusions, Future Work

Sentence Boundary Detection (SBD)

The goal of SBD is to split a natural language text into individual sentences (i.e., identify each sentence's boundaries).

Typically, SBD is operationalized as a binary classification of a fixed number of candidate boundary points (**.**, **!**, **?**).

. . . |Because Mr**.** Lange offered information that was unknown to both competitors and to the general public, there is no reason to confront this issue in this case**.**| |The panel majority bases its reading of the federal statute on its conclusion that . . .

SBD could be a critical task in many applications such as machine translation, summarization, or information retrieval.

Approaches to SBD

- Rules – A battery of hand-crafted matching rules is applied.

IF "!" OR "?" MATCHED → MARK AS BOUND
IF "<EOL><EOL>" MATCHED → MARK AS BOUND

- Supervised Machine Learning (ML) – Given a triggering event occurs decide if it is an instance of sentence boundary.

$x_i =<$ 0:token="." , 0:isTrigger=1, -1:token="Mr" ,
-1:isAbbr=1, 1:token="Lange" , 1:isName=1 $>$
$f(x_i) \rightarrow y_i$

- Unsupervised ML – Similar to supervised ML approach but the system is trained on unlabeled data.

SBD Performance

Multiple SBD systems were reported as having an excellent performance:[1]

- 99.8% accuracy of a tree-based classifier in predicting "." as ending (or not) a sentence evaluated on Brown corpus[Riley 1989]
- 99.5% accuracy of a combination of original system based on neural nets and decision trees with existing system[Aberdeen&al. 1995] evaluated on WSJ corpus[Palmer&Hearst 1997]
- 99.75% accuracy (WSJ) and 99.64% (Brown) of a maximum entropy model in assessing ".", "!", and "?" [Reynar&Ratnaparkhi 1997]
- 99.69% (WSJ) and 99.8% (Brown) of a rule-based sentence splitter combined with a supervised POS-tagger[Mikheev 2002]
- 98.35% (WSJ) and 98.98% (Brown) of an unsupervised system based on identification of abbreviations[Kiss&Strunk 2006]

The list comes from Read et al. 2012.

SBD Performance (cont.)

	Type	Brown	CDC	Genia	WSJ	All
CoreNLP	R	87.7 (93.6)	72.1 (98.3)	98.8 (99.0)	91.3 (94.8)	89.1 (95.0)
LingPipe$_1$	R	94.9 (96.6)	87.6 (99.1)	98.3 (98.6)	97.3 (98.7)	95.2 (97.4)
LingPipe$_2$	R	93.0 (94.5)	86.3 (97.2)	**99.6** (**99.8**)	88.0 (90.9)	93.2 (95.3)
MxTerminator	S	94.7 (96.5)	97.9 (98.6)	98.3 (98.5)	97.4 (98.5)	95.8 (97.2)
OpenNLP	S	96.6 (96.6)	98.6 (98.6)	98.8 (98.8)	99.1 (99.1)	**97.4** (97.4)
Punkt	U	96.4 (96.4)	**98.7** (98.7)	99.3 (99.3)	98.3 (98.3)	97.3 (97.3)
RASP	R	**96.8** (96.8)	96.1 (99.1)	98.9 (98.9)	99.0 (99.0)	**97.4** (**97.6**)
Splitta	S	95.4 (95.4)	96.1 (96.7)	99.0 (99.0)	**99.2** (**99.2**)	96.5 (96.5)
tokenizer	R	94.9 (**96.9**)	98.6 (**99.2**)	98.6 (98.9)	97.9 (**99.2**)	96.2 (**97.6**)

	WeScience		WNB		WLB	
	A	B	A	B	A	B
CoreNLP	90.0	97.9	95.3	96.4	89.1	90.9
LingPipe$_1$	90.0	98.1	94.8	96.1	92.4	94.2
LingPipe$_2$	89.8	98.0	94.4	95.6	92.7	94.5
MxTerminator	89.5	97.2	94.7	95.9	90.3	92.2
OpenNLP	90.2	97.9	95.3	96.5	90.2	92.0
Punkt	89.9	97.7	**95.6**	96.7	92.8	94.5
RASP	**91.0**	99.1	95.4	96.6	92.8	94.6
Splitta	**91.0**	98.9	94.0	95.5	91.2	93.4
tokenizer	**91.0**	**99.2**	**95.6**	**96.8**	**93.1**	**94.9**

Experiments from Read et al. 2012

Presentation Overview

Court Decisions

The term court *decision* refers to judicial determinations.

It includes final judgments, rulings, and inter-locutory or provisional orders made by the court.

Sometimes a distinction is made between a
decision (pronouncement of the solution or judgment in a case)
. . . and an . . .
opinion (statement of the reasons for its determination made by the court).

There are different types of decisions such as lower court decision, appellate decision, procedural decision, decision upon the merits.

legal-dictionary.thefreedictionary.com/decision

Court Decision's Characteristics

In terms of length a decision may be short (comparable to a newspaper article) or long (similar to a book).

A decision may be structured into sections and subsections preceded by a heading (possibly numbered).

A decision may contain specific constituents such as a header and a footer, footnotes, lists.

Sentences are interleaved with citations.

Sentences themselves may be extremely long, even organized as lists.

There is a high usage of sentence organizers such as ; or — and brackets (multiple types).

Quotes (possibly nested) are frequent.

Presentation Overview

Court Decision Excerpt

A long sentence with a quote (with a nested quote) organized as a list followed by citations and a short sentence with a citation.

. . . As used in the statute, "'act in furtherance of a person's right of petition or free speech under the United States or California Constitution in connection with a public issue' includes: (1) any written or oral statement or writing made before a legislative, executive, or judicial proceeding, or any other official proceeding authorized by law; (2) any written or oral statement or writing made in connection with an issue under consideration or review by a legislative, executive, or judicial body, or any other official proceeding authorized by law; (3) any written or oral statement or writing made in a place open to the public or a public forum in connection with an issue of public interest; (4) or any other conduct in furtherance of the exercise of the constitutional right of petition or the constitutional right of free speech in connection with a public issue or an issue of public interest." (§425.16, subd. (e), italics added; see Briggs v. Eden Council for Hope & Opportunity (1999) 19 Cal. 4th 1106, 1117-1118, 1123 [81 Cal.Rptr.2d 471, 969 P.2d 564] [discussing types of statements covered by anti-SLAPP statute].) The R.'s contend that plaintiffs' complaint falls within the third clause of section 425.16, subdivision (e). . . .

Court Decision Excerpt

Semicolons separate items in a list as well as independent clauses.

[O]ur family suffered: emotional distress; anxiety; sleeplessness; physical pain; insecurity; fear; pain and suffering; payment of attorneys' fees; payment of medical expenses; payment of moving expenses; payment of *1204 traveling and housing expenses to and from Los Angeles to support our business endeavors; [and] [D.C.]'s lost income. . . .

Completed assemblies must be exhaustively tested to demonstrate, to the FAA's satisfaction, that all requirements have been met; only then does the FAA certify the part for sale.

It takes RAPCO a year or two to design, and obtain approval for, a complex part; the dynamometer testing alone can cost $75,000. . . . Drawings and other manufacturing information contain warnings of RAPCO's intellectual-property rights; every employee receives a notice that the information with which he works is confidential.

Court Decision Excerpt

Informal poorly edited text may be present.

The next post, from "DAN JUSTICE," is the first to raise the rhetoric to a level that could, when considered out of context, be construed as a threat. It says "HEY [D.C.], I KNOW A GOOD *** WHEN I SEE ONE. I LIKE WHAT I SEE, LET'S GO GET SOME COFFEE. ****** im gonna kill you" and is signed "H-W student."

A sentence may span over a double line break.

. . . Section 1(4) of the Uniform Act provides:

"Trade secret" means information, including a formula, pattern, compilation, program, device, method, technique, or process . . .

Heading (possibly no triggering event)

FACTS AND PROCEDURAL HISTORY

Cyber Crime and IP Law Data Set

To evaluate performance of existing SBD systems on the decisions we created a data set of different types of US decisions.

The decisions are in the domains of cyber crime (cyber bullying, system hacking) and intellectual property protection.

Based on the observed phenomena presented earlier we assume:

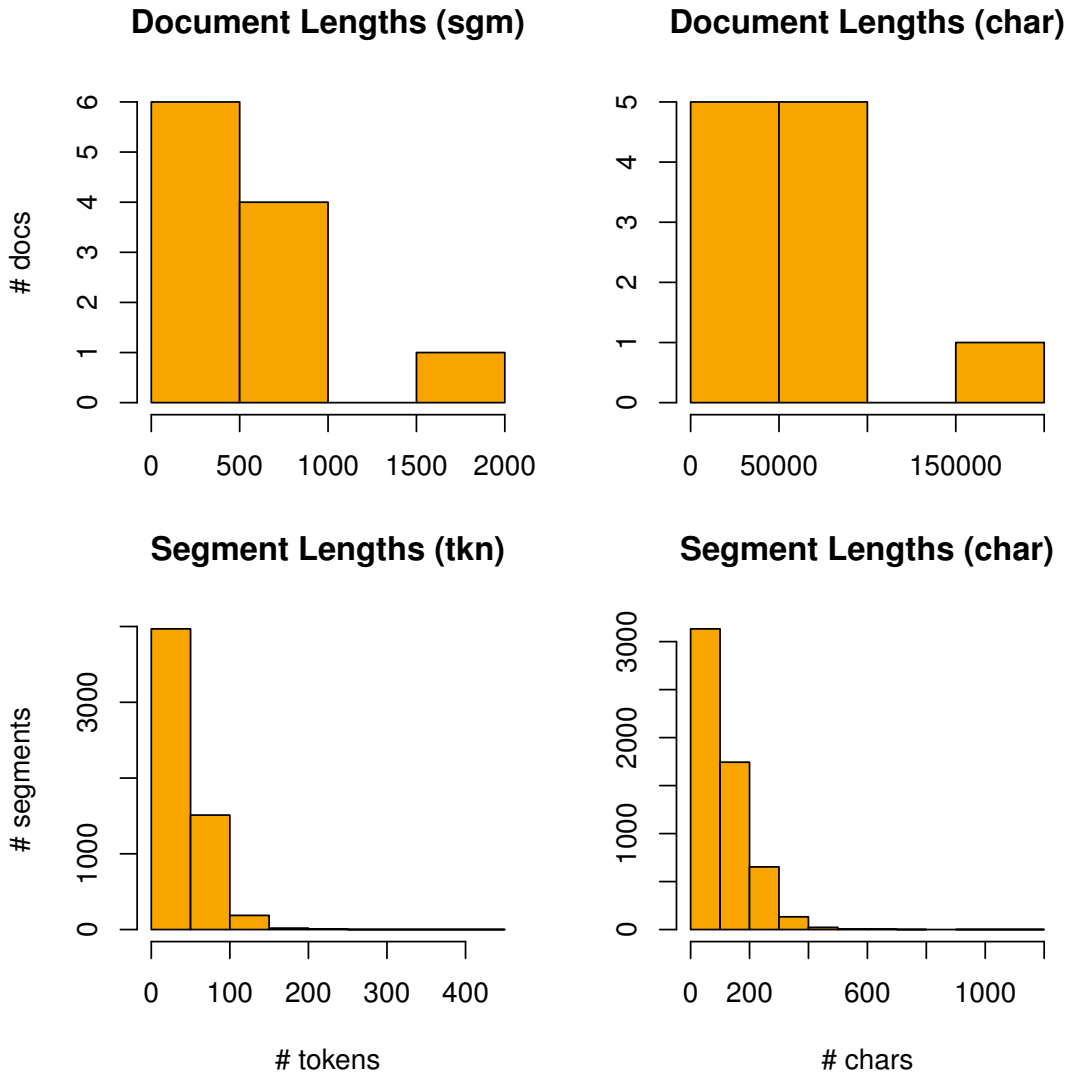
1. a definition of SBD as a binary classification of a limited number of triggering events (., !, ?) is not adequate
2. segmentation of a document into sentences (or sentence-like units) may often be done in multiple different ways

Because of #1 we do not use the idea of "triggering events" allowing possibly any token to be a boundary.

Because of #2 we adapt a consistent policy of "aggressive" segmenting (i.e., if doubts exist there is a boundary)

Cyber Crime and IP Law Data Set

# of documents	11.0
# of segments	5698.0
longest document (sgm)	1792.0
shortest document (sgm)	154.0
average doc length (sgm)	518.0
longest document (char)	181009.0
shortest document (char)	16859.0
average doc length (char)	55604.3
longest segment (tkn)	430.0
shortest segment (tkn)	1.0
average sgm length (tkn)	38.9
longest segment (char)	1182.0
shortest segment (char)	1.0
average sgm length (char)	106.0



Presentation Overview

Evaluated SBD Systems

For evaluation of SBD systems' performance on the corpus of court decisions we use one system from each category:

1. We work with the SBD module from the Stanford CoreNLP toolkit[Manning&al. 2014] as an example of a system based on rules.[1]
2. To test a system based on supervised ML classifier we employ the SBD component from openNLP.[2]
3. As an example of an unsupervised system we use the punkt[Kiss&Strunk 2006] module from the NLTK toolkit.[4]

The criterion for selection of the SBD systems was the assumed wide adoption of general toolkits the SBD systems are part of.

[1] `nlp.stanford.edu/software/corenlp.shtml`
[2] `opennlp.apache.org`
[3] `nltk.org/api/nltk.tokenize.html`

Rule Based Sentence Splitter from Stanford CoreNLP

Requires a text to be already segmented into tokens.

The system is based on triggering events the presence of which is prerequisite for a boundary to be predicted.

The default events are a single "." or a sequence of "?" and "!".

The system may use information about paragraph boundaries which can be configured as either a single EOL or two consecutive EOLs.

The system may also exploit HTML or XML markup if present.

Certain patterns that may appear after a boundary are treated as parts of the preceding sentence (e.g., parenthesized expression).

`github.com/stanfordnlp/CoreNLP/blob/master/src/edu/stanford/nlp/process/WordToSentenceProcessor.java`

Supervised Sentence Splitter from OpenNLP

Based on maximum entropy model which requires a corpus annotated with sentence boundaries.

The triggering events are "**.**", "**?**", and "**!**".

Features: information about the token containing the potential boundary and about its immediate neighbours.

- the prefix
- the suffix
- the presence of particular chars in the prefix and suffix
- whether the candidate is an honorific or corporate designator
- features of the words left and right of the candidate

Ratnaparkhi 1998

Unsupervised Sentence Splitter (punkt) from NLTK

The system does not depend on any additional resources besides the corpus it is supposed to segment into sentences.

The leading idea behind the system is that the chief source of wrongly predicted boundaries are periods after abbreviations.

The system discovers abbreviations by testing the hypothesis $P(\cdot|w) = 0.99$ against the corpus.

Additionally, token length (abbreviations are short) and the presence of internal periods are taken into account.

For prediction the system uses:

- ortographic features
- collocation heuristic (collocation is evidence against split)
- frequent sentence starter heuristic (split after abbreviation)

Kiss and Strunk 2006

Evaluation

We use traditional IR measures – precision (P), recall (R), and F_1-measure (F_1).

We evaluate the SBD performance from two different perspectives:

1. boundaries
2. segments – both boundaries need to match

For each perspective we use two approaches to determine if the boundary was predicted correctly.

1. strict – boundary offsets match exactly
2. lenient – the difference between boundary offsets does not contain alphanumeric char

|Accordingly, we find that the circuit court did not abuse its discretion when it denied Mr.| |Renfrow's motion for a JNOV.|
|**|We find no merit to this issue.|

Results

	Measure	stric-B	lenient-B	strict-S	lenient-S
CoreNLP	P	$.891 \pm .057$	$.899 \pm .053$	$.654 \pm .108$	$.668 \pm .101$
	R	$.742 \pm .072$	$.748 \pm .069$	$.550 \pm .113$	$.562 \pm .109$
	F_1	$.810 \pm .060$	$.815 \pm .056$	$.597 \pm .110$	$.610 \pm .105$
openNLP	P	$.884 \pm .079$	$.891 \pm .075$	$.648 \pm .119$	$.658 \pm .113$
	R	$.731 \pm .066$	$.737 \pm .063$	$.532 \pm .110$	$.541 \pm .107$
	F_1	$.798 \pm .061$	$.804 \pm .057$	$.582 \pm .112$	$.593 \pm .107$
punkt	P	$.770 \pm .126$	$.791 \pm .107$	$.521 \pm .164$	$.544 \pm .147$
	R	$.742 \pm .068$	$.766 \pm .080$	$.496 \pm .122$	$.523 \pm .116$
	F_1	$.752 \pm .083$	$.774 \pm .068$	$.506 \pm .141$	$.530 \pm .125$

Error Analysis

Missed boundary following a unit if a triggering event is absent

B. Response to Jury Question|

Deliberate avoidance is not a standard less than knowledge;| it is simply another way that knowledge may be proven.

Missed boundaries between citations

Kolender v. Lawson, 461 U.S. 352, 357, 103 S. Ct. 1855, 75 L. Ed. 2d 903 (1983);| United States v. Lim, 444 F.3d 910, 915 (7th Cir.2006)

Wrongly predicted boundaries in citations

see United States v. X-Citement Video, Inc., 513 U.S. 64, 76-78, 115 S. Ct. 464, 130 L. Ed.| 2d 372 (1994)

Presentation Overview

Stanford CoreNLP Sentence Splitter (tuned)

Entick v. Carrington, 95 Eng.| Rep. 807 (C. P. 1765)
451 F. Supp.| 2d 71, 88 (2006).

Knotts noted the "limited use which the government made of the signals from this particular beeper," 460 U. S., at 284;| and reserved the question whether "different constitutional principles may be applicable to "dragnet-type law enforcement practices" of the type that GPS tracking made possible here, ibid.

3| Katz did not repudiate that understanding.

	Measure	stric-B	lenient-B	strict-S	lenient-S
CoreNLP	P	$.902 \pm .062$	$.904 \pm .063$	$.679 \pm .102$	$.688 \pm .107$
	R	$.752 \pm .037$	$.753 \pm .038$	$.567 \pm .079$	$.574 \pm .083$
	F_1	$.820 \pm .045$	$.821 \pm .046$	$.617 \pm .088$	$.625 \pm .093$
CoreNLP	P	$.895 \pm .043$	$.897 \pm .044$	$.722 \pm .077$	$.730 \pm .075$
(trained)	R	$.876 \pm .021$	$.877 \pm .023$	$.706 \pm .062$	$.715 \pm .060$
	F_1	$.886 \pm .031$	$.887 \pm .033$	$.714 \pm .069$	$.722 \pm .067$

openNLP Sentence Splitter (trained)

5. The Government's Hybrid Theory|

The Sealed Application does not cite the Pen/Trap Statute as authority for obtaining cell site data ...

This device delivers many different types of communication: live conversations, voice mail, pages, text messages, e-mail, alarms, internet, video, photos, dialing, signaling, etc.| The legal standard for government access depends entirely upon the type of communication involved.

	Measure	stric-B	lenient-B	strict-S	lenient-S
openNLP	P	$.898 \pm .042$	$.903 \pm .042$	$.680 \pm .097$	$.690 \pm .096$
	R	$.743 \pm .045$	$.748 \pm .042$	$.560 \pm .092$	$.568 \pm .090$
	F_1	$.813 \pm .043$	$.818 \pm .041$	$.614 \pm .094$	$.623 \pm .093$
openNLP (trained)	P	$.941 \pm .027$	$.945 \pm .031$	$.732 \pm .079$	$.740 \pm .079$
	R	$.755 \pm .042$	$.758 \pm .039$	$.585 \pm .087$	$.591 \pm .085$
	F_1	$.837 \pm .032$	$.841 \pm .032$	$.650 \pm .084$	$.657 \pm .082$

punkt Sentence Splitter (trained)

"[T]he district court retains broad discretion in deciding how to respond to a question propounded from the jury and . . . | the court has an obligation to dispel any confusion quickly and with concrete accuracy."

II. ANALYSIS|

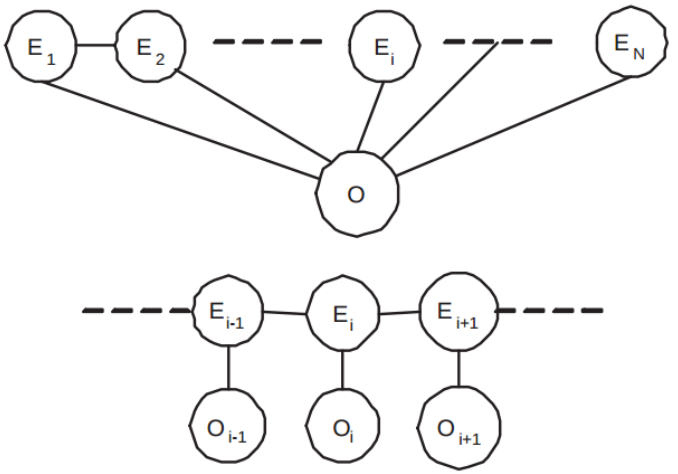
United States v. Leahy, 464 F.3d 773, 796 (7th Cir.2006);| United States v. Carrillo, 435 F.3d 767, 780 (7th Cir.2006)

	Measure	stric-B	lenient-B	strict-S	lenient-S
punkt	P	.806 ± .083	.814 ± .089	.552 ± .119	.563 ± .123
	R	.732 ± .069	.738 ± .067	.500 ± .108	.509 ± .110
	F_1	.765 ± .062	.773 ± .064	.523 ± .110	.534 ± .113
punkt (trained)	P	.876 ± .066	.885 ± .070	.638 ± .128	.649 ± .131
	R	.726 ± .069	.733 ± .067	.530 ± .119	.538 ± .121
	F_1	.793 ± .062	.801 ± .062	.579 ± .123	.588 ± .125

Presentation Overview

Conditional Random Fields (CRF) model

A CRF is a random field model that is globally conditioned on an observation sequence O.



The states of the model correspond to event labels E.

We use a first-order CRF in our experiments (observation O_i is associated with E_i).

We use CRFsuite[1] implementation of first-order CRF.

We use the BILOU tagging scheme (B-SEN, I-SEN, L-SEN, O-SEN, U-SEN).

Features: token, preceding and following tokens, isUpper, isTitle, isDigit, isWhitespace.

Lafferty et al. 2001; Liu et al. 2005; Ozaki 2007

[1] www.chokkan.org/software/crfsuite/

CRF Results

	Measure	stric-B	lenient-B	strict-S	lenient-S
CRF	P	$.935 \pm .006$	$.937 \pm .004$	$.820 \pm .040$	$.829 \pm .034$
(trained)	R	$.922 \pm .036$	$.924 \pm .034$	$.805 \pm .068$	$.813 \pm .062$
	F_1	$.928 \pm .021$	$.930 \pm .019$	$.812 \pm .054$	$.821 \pm .048$

	P	R	F_1	support
B-SEN	.93	.82	.87	1401
I-SEN	.99	1.00	.99	51029
L-SEN	.94	.85	.89	1372
O-SEN	.93	.88	.91	1352
U-SEN	1.00	.45	.62	89
total	.98	.98	.98	55243

The judgment of the Court of Appeals for the D. C.| Circuit is affirmed.

. . . a case we have described as a monument of English freedom undoubtedly familiar to every| American statesman at the time the Constitution was adopted . . .

. . . search is not involved and resort must be had to Katz analysis;| but| there is no reason for rushing forward . . .

Presentation Overview

Summary Results of Trained Systems

	Measure	stric-B	lenient-B	strict-S	lenient-S
CoreNLP	P	$.895 \pm .043$	$.897 \pm .044$	$.722 \pm .077$	$.730 \pm .075$
(trained)	R	$.876 \pm .021$	$.877 \pm .023$	$.706 \pm .062$	$.715 \pm .060$
	F_1	$.886 \pm .031$	$.887 \pm .033$	$.714 \pm .069$	$.722 \pm .067$
openNLP	P	$.941 \pm .027$	$.945 \pm .031$	$.732 \pm .079$	$.740 \pm .079$
(trained)	R	$.755 \pm .042$	$.758 \pm .039$	$.585 \pm .087$	$.591 \pm .085$
	F_1	$.837 \pm .032$	$.841 \pm .032$	$.650 \pm .084$	$.657 \pm .082$
punkt	P	$.876 \pm .066$	$.885 \pm .070$	$.638 \pm .128$	$.649 \pm .131$
(trained)	R	$.726 \pm .069$	$.733 \pm .067$	$.530 \pm .119$	$.538 \pm .121$
	F_1	$.793 \pm .062$	$.801 \pm .062$	$.579 \pm .123$	$.588 \pm .125$
CRF	P	$.935 \pm .006$	$.937 \pm .004$	$.820 \pm .040$	$.829 \pm .034$
(trained)	R	$.922 \pm .036$	$.924 \pm .034$	$.805 \pm .068$	$.813 \pm .062$
	F_1	$.928 \pm .021$	$.930 \pm .019$	$.812 \pm .054$	$.821 \pm .048$

Conclusions

We defined a task for segmenting US court decisions into sentences and sentence-like segments (e.g., some independent clauses).

We evaluated selected SBD systems on their performance with respect to this task on a data set we have created.

On the basis of analyzing the degraded performance we tuned the systems towards the corpus which led to an improved performance.

We trained a CRF model that out performed the SBD systems.

For future work we would like to:

- improve the performance of the CRF model even further
- test the model on US court decisions from other domains and decisions from other countries written in English

Thank you!

Questions, comments and suggestions are welcome now
or any time at jas438@pitt.edu.

References I

Aberdeen, J., Burger, J., Day, D., Hirschman, L., Robinson, P., & Vilain, M. (1995, November). MITRE: description of the Alembic system used for MUC-6. In *Proceedings of the 6th conference on Message understanding* (pp. 141–155). Association for Computational Linguistics.

Kiss, Tibor, and Jan Strunk. "Unsupervised multilingual sentence boundary detection." Computational Linguistics 32.4 (2006): 485-525.

John Lafferty, Andrew McCallum, and Fernando Pereira. "Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data". Proceedings of the 18th International Conference on Machine Learning. 282–289. 2001.

Liu, Y., Stolcke, A., Shriberg, E., & Harper, M. (2005, June). Using conditional random fields for sentence boundary detection in speech. In Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics (pp. 451-458). Association for Computational Linguistics.

Manning, Christopher D., Surdeanu, Mihai, Bauer, John, Finkel, Jenny, Bethard, Steven J., and McClosky, David. 2014. The Stanford CoreNLP Natural Language Processing Toolkit In Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations, pp. 55-60.

References II

- Mikheev, Andrei. “Periods, capitalized words, etc.” Computational Linguistics 28.3 (2002): 289-318.
- Okazaki, Naoaki. “CRFsuite: a fast implementation of conditional random fields (CRFs).” (2007).
- Palmer, David D., and Marti A. Hearst. “Adaptive multilingual sentence boundary disambiguation.” Computational Linguistics 23.2 (1997): 241-267.
- Ratnaparkhi, Adwait. Maximum entropy models for natural language ambiguity resolution. Diss. University of Pennsylvania, 1998.
- Reynar, Jeffrey C., and Adwait Ratnaparkhi. “A maximum entropy approach to identifying sentence boundaries.” Proceedings of the fifth conference on Applied natural language processing. Association for Computational Linguistics, 1997.
- Read, J., Dridan, R., Oepen, S., & Solberg, L. J. (2012). Sentence boundary detection: A long solved problem?. COLING (Posters), 12, 985-994.
- Riley, Michael D. 1989. Some applications of tree-based modelling to speech and language. In Proceedings of the workshop on Speech and Natural Language (HLT ’89). Association for Computational Linguistics, Stroudsburg, PA, USA, 339-352.

References III

- *CRFsuite: A fast implementation of Conditional Random Fields (CRFs)* [online]. Accessed on 2017-03-14 at `http://www.chokkan.org/software/crfsuite/`
- *GATE: general architecture for text engineering* [online]. Accessed on 2017-03-14 at `http://gate.ac.uk`
- *LingPipe* [online]. Accessed on 2017-03-14 at `http://alias-i.com/lingpipe/`
- *MxTerminator* [online]. Accessed on 2017-03-14 at `ftp://ftp.cis.upenn.edu/pub/adwait/jmx/`
- *openNLP* [online]. Accessed on 2017-03-14 at `http://opennlp.apache.org/`
- *NLTK 3.0*: punkt module [online]. Accessed on 2017-03-14 at `http://nltk.org/api/nltk.tokenize.html`
- *iLexIR*: RASP [online]. Accessed on 2017-03-14 at `https://www.ilexir.co.uk/nlp-applications/index.html`
- *splitta* [online]. Accessed on 2017-03-14 at `https://code.google.com/archive/p/splitta/`
- *Stanford CoreNLP* – a suite of core NLP tools [online]. Accessed on 2017-03-14 at `http://nlp.stanford.edu/software/corenlp.shtml`

References IV

tokenizer [online]. Not accessible on 2017-03-14 at `http://www.cis.uni-muenchen.de/~wastl/misc/`