Automatic Committed Belief Tagging

presented by Chris Hidey

Columbia University

February 26, 2015

Overview

1. Background
2. Corpus
3. Methodology
4. Results
5. Critique

Background

Cognitive State

- **Belief**
- Desire
- Intention

Background

Cognitive State

- **Belief**
- Desire
- Intention

Examples

- Desire and ¬ Belief

  I know John won't be here but I wouldn't mind if he were

Background

Cognitive State

- **Belief**
- Desire
- Intention

Examples

- Desire and ¬ Belief

  I know John won't be here but I wouldn't mind if he were

- Belief and ¬ Belief

  #John won't be here but nevertheless I think he may be here.

Corpus

10,000 tokens annotated for belief

Verbal Propositions

- Verb

  Republican leader Bill Frist **said** *the Senate was* **hijacked**.

Corpus

10,000 tokens annotated for belief

Verbal Propositions

- Verb

  Republican leader Bill Frist **said** *the Senate was* **hijacked**.

- Nominal/Adjectival/Prepositional predicate

  Republican leader Bill Frist **said** *the Senate was* **useless**.

Corpus

Belief Tags

- Committed belief (CB)

  GM has laid off workers.

Corpus

Belief Tags

- Committed belief (CB)

  GM has laid off workers.

- Non-committed belief (NCB)

  GM may lay off workers.

Corpus

Belief Tags

- Committed belief (CB)

  GM has laid off workers.

- Non-committed belief (NCB)

  GM may lay off workers.

- Not applicable (NA)

  Some wish GM would lay off workers.

Corpus

Belief Tags

- Committed belief (CB)

  GM has laid off workers.

- Non-committed belief (NCB)

  GM may lay off workers.

- Not applicable (NA)

  Some wish GM would lay off workers.

- Other (O)

Corpus

Belief Tags

- Committed belief (CB)

  GM has laid off workers.

- Non-committed belief (NCB)

  GM may lay off workers.

- Not applicable (NA)

  Some wish GM would lay off workers.

- Other (O)

Subtasks

- Identifying propositions (binary classification)
- Tagging propositions (3-way classification)

Methodology

Inference

- Joint Inference
  - Support Vector Machine (SVM) Chunker
  - Conditional Random Field (CRF)
- Pipeline

Methodology

Inference

- Joint Inference
  - Support Vector Machine (SVM) Chunker
  - Conditional Random Field (CRF)
- Pipeline

Feature Engineering

- Lexical
- Syntactic

Lexical Features

Derived without parsing

Informative

- numeric
- POS
- regular/modal/auxiliary

Uninformative

- lemma
- stem

Syntactic Features

Derived from dependency parse

Informative

- infinitive form
- reporting ancestor
- parent POS
- child
  - perfect tense
  - wh-word
  - auxilary/modal

Uninformative

- parent lemma/stem
- supertags

Experiments

SVM			
Kernel:	quadratic	*Features*	*F1-score*
Slack:	$c = 0.5$	Lexical	56.9
Context Width:	2	Lexical and Syntactic	**64.0**

Experiments

SVM			
Kernel:	quadratic	*Features*	*F1-score*
Slack:	$c = 0.5$	Lexical	56.9
Context Width:	2	Lexical and Syntactic	**64.0**

CRF			
Order:	1	*Features*	*F1-score*
Gaussian Variance:	1	Lexical	49.6
		Lexical and Syntactic	59.0

Experiments

SVM			
Kernel:	quadratic	*Features*	*F1-score*
Slack:	$c = 0.5$	Lexical	56.9
Context Width:	2	Lexical and Syntactic	**64.0**

CRF			
Order:	1	*Features*	*F1-score*
Gaussian Variance:	1	Lexical	49.6
		Lexical and Syntactic	59.0

Pipeline			
SVM binary tagger		*Features*	*F1-score*
Context Width:	2	Lexical and Syntactic	46.1
SVM 3-way classifier			
trained on gold data			

Critique

- Clarity?
  - Well-written and understandable
  - More explanation of features and parameters
  - Confusion matrix or error analysis

Critique

- Clarity?
  - Well-written and understandable
  - More explanation of features and parameters
  - Confusion matrix or error analysis
- Methodology/Evaluation?
  - Additional baselines
  - Training/testing
  - Statistical significance

Critique

- Clarity?
  - Well-written and understandable
  - More explanation of features and parameters
  - Confusion matrix or error analysis
- Methodology/Evaluation?
  - Additional baselines
  - Training/testing
  - Statistical significance
- Originality?

Critique

- Clarity?
  - Well-written and understandable
  - More explanation of features and parameters
  - Confusion matrix or error analysis
- Methodology/Evaluation?
  - Additional baselines
  - Training/testing
  - Statistical significance
- Originality?
- Reproducibility?
  - Depends on code and data availability
  - Parameter tuning
  - Training/evaluation/testing

Critique

- Clarity?
  - Well-written and understandable
  - More explanation of features and parameters
  - Confusion matrix or error analysis
- Methodology/Evaluation?
  - Additional baselines
  - Training/testing
  - Statistical significance
- Originality?
- Reproducibility?
  - Depends on code and data availability
  - Parameter tuning
  - Training/evaluation/testing
- Impact?

Critique

- Positives?
  - Well-defined problem, high IAA
  - Not too domain-specific
  - Comparison of pipeline versus joint inference

Critique

- Positives?
  - Well-defined problem, high IAA
  - Not too domain-specific
  - Comparison of pipeline versus joint inference
- Negatives?
  - Requires dependency parsing
  - Language-specific