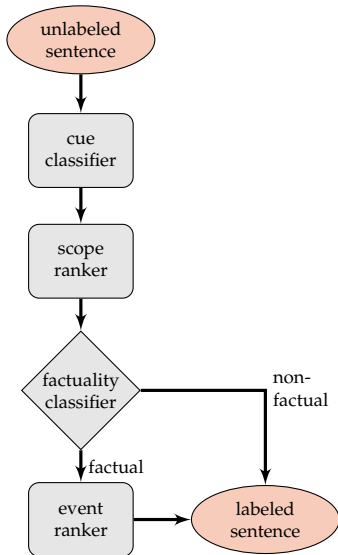


UiO₁: Constituent-Based Discriminative Ranking for Negation Resolution

Jonathon Read, Erik Velldal, Lilja Øvrelid and Stephan Oepen
University of Oslo, Department of Informatics
{jread|erikve|liljao|oe}@ifi.uio.no



UiO₁: adapted and extended from Velldal et al. (2012) to handle:

- ▶ Morphological cues
- ▶ Discontinuous scopes
- ▶ Factuality of context
- ▶ Negated events

Closed track—uses organizer-provided data only

Two submissions differ with respect to parameter tuning (including feature selection):

- I** 10-fold cross-validation on the **combined** training and development data (CDTD)
- II** using **supplied partitions** of training (CDT) and development (CDD) data

An SVM-based cue classifier **disambiguates** known cue words using features of:

- ▶ word n -grams: full form and lemma to left/right of cue
- ▶ affixal cues (e.g. *im-possible*):
 - ▶ character n -grams over base start/end ($\{possi, poss, \dots\}$, $\{sible, ible \dots\}$)
 - ▶ affix (*im*)
 - ▶ token PoS (JJ)
 - ▶ lexicon-lookup of candidate base form and character n -grams (*underlying*)

Post-processing heuristics for multi-word cues

Constituent-based hybrid system:

- ▶ “slackening” heuristics to align constituents with scope
- ▶ rules for finding scope given a cue and a syntactic tree
- ▶ data-driven ranking of candidate constituents
- ▶ heuristics for handling discontinuous scope

Scope resolution rules



RB//VP/SBAR if SBAR\WH*

RB//VP/S

RB//S

DT/NP if NP/PP

DT//SBAR if SBAR\WHADVP

DT//S

JJ//ADJPVP/S if S\VP\VB*[@lemma="be"]

JJ/NP/NP if NP\PP

JJ//NP

UH

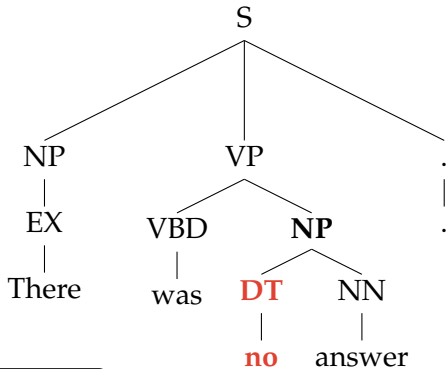
IN/PP

NN/NP//S/SBAR if SBAR\WHNP

NN/NP//S

CC/SINV

Data-driven constituent ranking

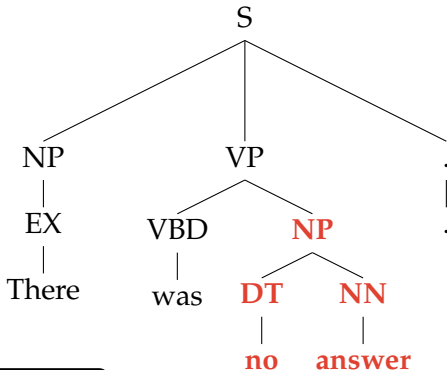


Candidates

✗ . /DT

There was no answer.

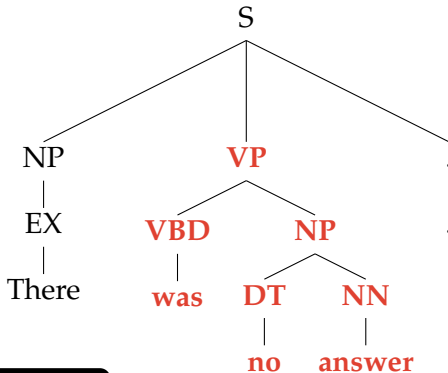
Data-driven constituent ranking



Candidates

- | | | |
|-------------------------------------|----------|-----------------------------|
| <input checked="" type="checkbox"/> | . /DT | There was <u>no</u> answer. |
| <input checked="" type="checkbox"/> | . /DT/NP | There was <u>no answer.</u> |

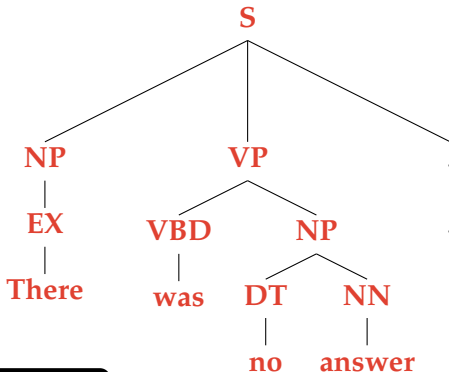
Data-driven constituent ranking



Candidates

- | | | |
|-------------------------------------|------------|------------------------------|
| <input checked="" type="checkbox"/> | ./DT | There was <u>no</u> answer. |
| <input checked="" type="checkbox"/> | ./DT/NP | There was <u>no answer</u> . |
| <input checked="" type="checkbox"/> | ./DT/NP/VP | There <u>was no answer</u> . |

Data-driven constituent ranking



Candidates

- | | | |
|-------------------------------------|--------------|------------------------------|
| <input checked="" type="checkbox"/> | ./DT | There was <u>no</u> answer. |
| <input checked="" type="checkbox"/> | ./DT/NP | There was <u>no answer</u> . |
| <input checked="" type="checkbox"/> | ./DT/NP/VP | There <u>was no answer</u> . |
| <input checked="" type="checkbox"/> | ./DT/NP/VP/S | <u>There was no answer</u> . |

Features of scopes



Features	I	II
Path from cue	•	•
Path from cue bigrams and trigrams	•	•
Path from cue to left/right boundary	•	
Path to left/right boundary		•
Path to root	•	
Punctuation to left/right	•	•
Rule prediction		•
Sibling bigrams		•
Size in tokens, relative to sentence (%)	•	•
Surface bigrams	•	•
Tree distance from cue	•	•

10.3% of scopes in the training data are **discontinuous**, e.g.

- (1) I therefore spent the day at my club and **did not** return to Baker Street until evening.
- (2) There was certainly **no** physical injury of any kind.

Post-processing heuristics:

- (1) if cue is in a conjoined phrase, remove the preceding conjunct(s) from scope
- (2) remove sentential adverbs from scope

Scope resolution results



Data set	Model	Prec	Rec	F ₁
CDTD	Baseline	98.31	33.18	49.61
	Rules	100.00	71.37	83.29
	Ranker _I	100.00	73.55	84.76
CDD	Baseline	100.00	36.31	53.28
	Rules	100.00	69.64	82.10
	Ranker _{II}	100.00	70.24	82.52
CDE	Baseline	96.47	32.93	49.10
	Rules	98.73	62.65	76.66
	Ranker _I	98.77	64.26	77.86
	Ranker _{II}	98.75	63.45	77.26

Baseline: expand scope left/right of cue until punctuation

Event annotations **connote** the factuality of contexts:

- ▶ 79.4% of scopes are **factual** (with event)
- ▶ 20.6% of scopes are **non-factual** (w/out event)

Features are variations on bag-of-words:

- ▶ centered on cue / extracted from entire sentence
- ▶ record forms, lemmas, parts-of-speech

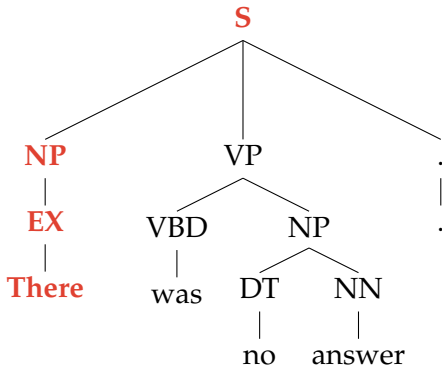
Factuality detection results



Data set	Model	Prec	Rec	F_1	Acc
CDE	Baseline	69.48	100.00	81.99	69.48
	Classifier _I	77.73	95.91	85.86	78.31

Baseline: choose majority class (factual)

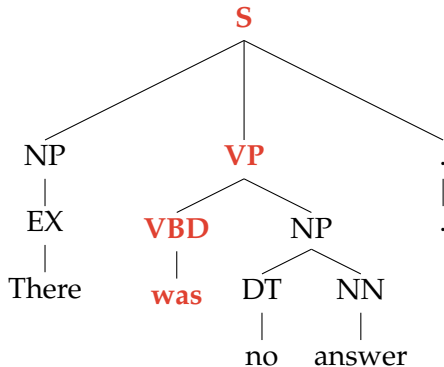
Data-driven event ranking



Candidates

☒ . /EX/NP/S There was **no** answer.

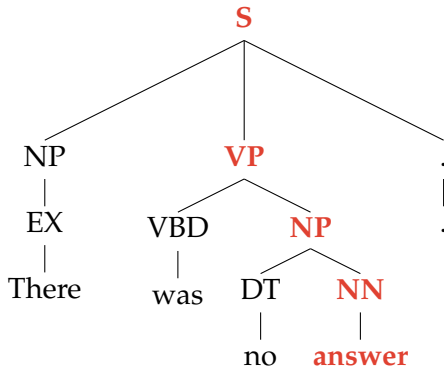
Data-driven event ranking



Candidates

- ☒ . /EX/NP/S There was **no** answer.
- ☒ . /VBD/VP/S There was **no** answer.

Data-driven event ranking



Candidates

- ☒ . /EX/NP/S There was **no** answer.
- ☒ . /VBD/VP/S There was **no** answer.
- ☒ . /NN/NP/VP/S There was **no** answer.

Features of events



Feature type	I	II
Contains affixal cue	•	
Following lemma		•
Lemma	•	•
Path to scope constituent	•	•
Path to scope constituent bigrams	•	•
Part-of-speech	•	•
Position in scope	•	•
Preceding lemma	•	•
Preceding part-of-speech	•	•
Token distance from cue	•	•

Held-out evaluation



	Submission I			Submission II		
	Prec	Rec	F ₁	Prec	Rec	F ₁
Cues	91.42	92.80	92.10	89.17	93.56	91.31
Scopes	87.43	61.45	72.17	83.89	60.64	70.39
Scope Tokens	81.99	88.81	85.26	75.87	90.08	82.37
Events	60.50	72.89	66.12	60.58	75.00	67.02
Full negation	83.45	43.94	57.57	79.87	45.08	57.63

Cue detection:

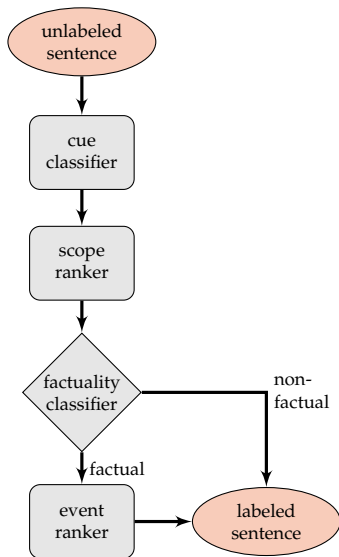
- ▶ most errors involve affixal cues
- ▶ multi-word cues not covered, e.g. *no more*

Scope resolution:

- ▶ parse errors (coordination)
- ▶ discourse-level adverbials (*in the second place*)

Event detection:

- ▶ most errors involve multi-word events
- ▶ others occur when event is the main verb



UiO₁—a hybrid approach to negation:

- ▶ *n*-grams for cue disambiguation
- ▶ constituent-based scope ranking
- ▶ bag-of-words for factuality classification
- ▶ parse tree-based event ranking

Thanks!