

BIUTEE

Bar Ilan University Textual Entailment Engine
RTE-6
TAC 2010

Asher Stern, Eyal Shnarch, Amnon Lotan, Shachar Mirkin, Lili Kolterman, Naomi Zeichner, Jonathan Berant and Ido Dagan



Presentation Outline

- **BIUTEE** Characteristics
- Knowledge
 - Lexical Graph
 - FRED
- Approximate Match
- IR Filtering
- Results

-
- Open Source

- **BIUTEE Characteristics**
- Knowledge
 - Lexical Graph
 - FRED
- Approximate Match
- IR Filtering
- Results
- Open Source

Proof System

Explicit generation of H from T

$$T = T_0 \rightarrow T_1 \rightarrow T_2 \rightarrow \dots \rightarrow T_n = H$$

- Dependency parse trees
 - Minipar

Entailment Rules

- Represented as sub-trees
 - Left hand side \rightarrow Right hand side
 - Explosion \rightarrow Blast
 - X take place at Y \rightarrow X occurs at Y

Proof System

H: A blast occurred at a hotel in Taba.

T: An explosion caused by gas took place at a Taba hotel

A **blast** caused by gas took place at a Taba hotel

A blast took place at a Taba hotel

A blast **occurred at** a Taba hotel

A blast occurred at a **hotel in Taba**.



Lexical



Lexical syntactic



Syntactic

Key property: Chaining of knowledge

X is sold by Y to Z



Y sells X to Z

Z buys X from Y

Considering Co-references



Lexical



Lexical syntactic



Syntactic

Mom of soldier killed in Iraq stages protest near Bush's ranch. The angry mother of a fallen U.S. soldier has met with advisers of President George W. Bush

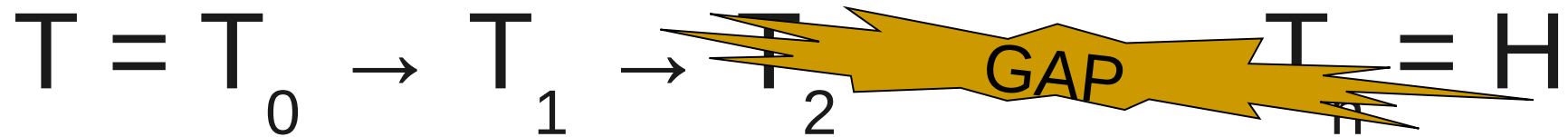
Co reference resolver

- BART as co-reference system
- Shachar Mirkin, Ido Dagan and Sebastian Padó. 2010. *Assessing the Role of Discourse References in Entailment Inference*. ACL.

Proof System flow

- Apply rules iteratively
- Generate consequents forest
- Compact forest

The inference gap problem



- Approximate match
 - To be discussed later

- BIUTEE Characteristics
- **Knowledge**
 - **Lexical Graph**
 - FRED
- Approximate Match
- IR Filtering
- Results
- Open Source

Lexical resources

Currently integrated

- WordNet
- CatVar

Future challenges – effectively integrate:

- Wikipedia
- Distributional Similarity
- Wiktionary
- Geo
- ...

- BIUTEE Characteristics
- **Knowledge**
 - Lexical Graph
 - **FRED**
- Approximate Match
- IR Filtering
- Results
- Open Source

FRED

- A lexical syntactic resource learnt from FrameNet
 - *Cure X → X's recovery*
- More accurate than other lexical-syntactic resources
- Positive results over ACE data-set

References

- Roni Ben Aharon, Idan Szpektor and Ido Dagan.
Generating entailment rules from FrameNet. In *Proceeding of the ACL 2010 Conference Short Papers*.
- www.cs.biu.ac.il/~nlp/downloads

- BIUTEE Characteristics
- Knowledge
 - Lexical Graph
 - FRED
- **Approximate Match**
- IR Filtering
- Results
- Open Source

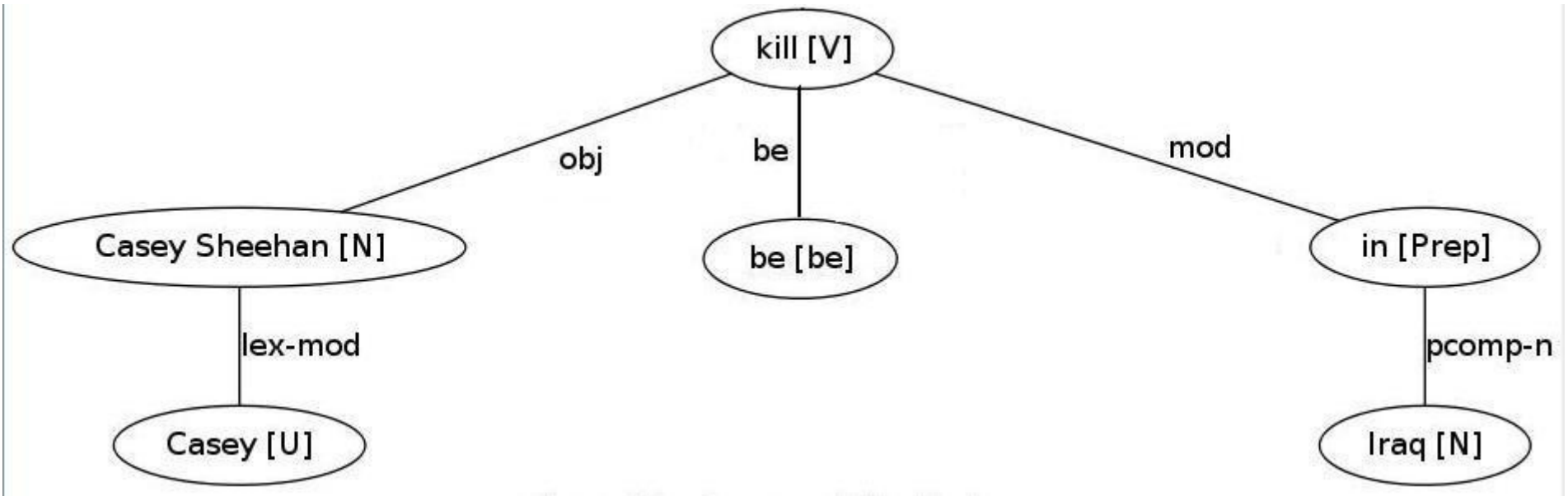
Approximate Match

- Obtaining a complete proof is very rare
- Previous years
 - Features from forest and hypothesis
 - Machine learning classifier
- Research direction: a principled model
 - Explicit model
 - Based on syntactic structure
- This year: a heuristic model
 - Mandatory relations
 - Ignorable relations

Approximate Match

- Verb argument structure
 - Arguments as bags-of-words
 - Match
 - Match main verb
 - Match arguments
 - Percentage of covered arguments reaches a predefined threshold
- Bag of words structure
 - When main predicate is "be"
 - Match
 - Percentage of word coverage reaches a predefined threshold

Approximate Match

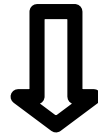


H: Casey Sheehan was killed in Iraq

Approximate Match

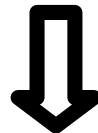
Verb argument structure

H: Casey Sheehan was killed in Iraq



<Kill (Object {Casey Sheehan} Other {Iraq})>

T: Her son, Casey, 24, was killed in Sadr City, Iraq, on April 4, 2004



<Kill (Object {Son, Her, Casey, 24}) (Other {Sadr City, Sadr, Iraq, April 4 2004, April, 4})>

- BIUTEE Characteristics
- Knowledge
 - Lexical Graph
 - FRED
- Approximate Match
- **IR Filtering**
- Results
- Open Source

IR Filtering

- Further candidates filtering
- Lucene IR system
- Hypothesis as query
- Judge only K top ranked sentences
 - Optimize value of K on development set

- BIUTEE Characteristics
- Knowledge
 - Lexical Graph
 - FRED
- Approximate Match
- IR Filtering
- **Results**
- Open Source

Results

Run #	co-ref	Knowledge resources	IR filtering	Dev F1 %	Test F1 %
BIU1	BART	✓Lexical graph ✓Syntactic rules	Top 28	40.73	37.50
BIU2			None	36.70	32.40

✓IR filtering improves results

Ablation tests

Tested (eliminated) component	Development Set		Test set	
	F1 %	Δ %	F1 %	Δ %
WordNet	39.18	+1.55	36.60	+0.90
CatVar	40.20	+0.53	36.87	+0.63
BART (co reference resolution)	41.62	-0.89	38.38	-0.88

- ✓ WordNet and CatVar have positive contribution
- ✗ Utilizing Co-reference has negative effect

RTE - KBP Validation

- What we did
 - Build forest and apply rules for all text's sentences
 - Technically, hard to process
- An alternative approach
 - Process only sub sets of the sentences
- Results:
 - F1: 16 % (Recall: 39.5%. Precision: 10%)

- BIUTEE Characteristics
- Knowledge
 - Lexical Graph
 - FRED
- Approximate Match
- IR Filtering
- Results
- **Open Source**

An Open Source Architecture

- Benefits
 - Embedded in other applications
 - New resources and ideas
- Properties supporting open architecture
 - Unified knowledge representation
 - High coding standards

Thank you!

Questions?