Raytheon
BBN Technologies

TAC 2017

Jay DeYoung

Yee Seng Chan, Chinnu Pittapally, Hannah Provenza

Ryan Gabbard*, Marjorie Freedman*

*now at USC ISI

Document Level Event Extraction

- Argument Assertions e.g. *(Contact.Meet, Place, Pittsburgh, Actual)*
  1. Logistic regression to identify (1) event –focused terms and (2) roles/arguments for events
     - Two argument classifiers: one that depends on event-focused terms, the second relies of just identifying a role in the argument context
  2. Identify a canonical string for the argument using
     - SERIF within document coreference
     - SERIF time normalization
  3. ERE-based trained classifier for distinguishing ACTUAL/GENERIC
     - Syntactic rules for identifying past/negated as OTHER
  4. Joint optimization using system confidence of 1-3
  5. World-knowledge based inference using event structure
- Within document event frame creation
  – Sieve-based system that relies on argument overlap, argument conflict, and syntactic links between arguments and event-focused terms

2017 Updates

Raytheon
BBN Technologies

- Incorporated additional training data
  - More Rich ERE
  - Event Nugget Training
  - BBN-developed targeted training
- Incorporated additional event types
  - Contact.Broadcast
  - Contact.Contact
  - Transaction.Transaction

Challenges with Contact.Broadcast

Raytheon BBN Technologies

- Rich ERE only marks the first mention of a Contact.Broadcast, subsequent mentions are ignored
  - Unmarked RichERE text is ambiguous between
    - Negative example for Contact.Broadcast
    - 2nd, 3rd, 4th,…. positive example of a Contact.Broadcast event
- System trained exclusively with targeted training on EAL dry-run data
  - Many false alarms that seem like annotation errors
  - Contact.Broadcast annotation agreement may be low enough to interfere with measuring system performance

Targeted Training (1)

Raytheon BBN Technologies

- Core challenge of EAL task is sparsity of training data
  - Many annotated documents
  - Few positive examples of events
- Develop targeted event annotation using human intuitions about event contexts
  - Ask annotator to find useful examples
  - Let annotator skip hard examples
- Annotation process
  - Annotator asked to come up with a list of likely event-related phrases
    - Nuggets OR other words likely to be associated with an event
  - Annotator searches & then marks ~10 examples per-term
    - Only marks sentences with one event mention (and may skip confusing sentences)
    - Marks all words that could be considered an event trigger
    - Marks arguments
  - Annotator asked to mark negative examples in the surrounding context (e.g. sentence N-1 does not contain a Contact.Meet event)
  - Annotator revises list to include additional event words
- Resulting annotation is
  - Dense in events
  - Likely to contain multiple syntactic contexts for arguments <-> triggers
  - For polysemous triggers, likely to contain positives and negatives

Targeted Training (2)

Raytheon
BBN Technologies

- 2015: Annotated ~5.8K positive & 6.4K negative sentences
  - Each sentence for a single event type
    - 4-8 hours per event type for all event types
    - Additional annotation for a few event types where we observed poor system performance
  - 2015 TAC system used only trigger annotation
    - ~12% relative improvement on *argument score* for system (BBN1 vs BBN2)
      - Arg F1: BBN2 35.5
      - Arg F1: BBN1 38.0 *(rank 1)*
- 2016: Additional annotation for new event types

	P	R	F1
No spannotator	26.3	26	26.2
Target:Trigger	26.1	26	26.1
Target:Trigger+Arg	28.1	26.2	27.1

2016 Dry Run Data: All Event Types

Context Embeddings (2015)

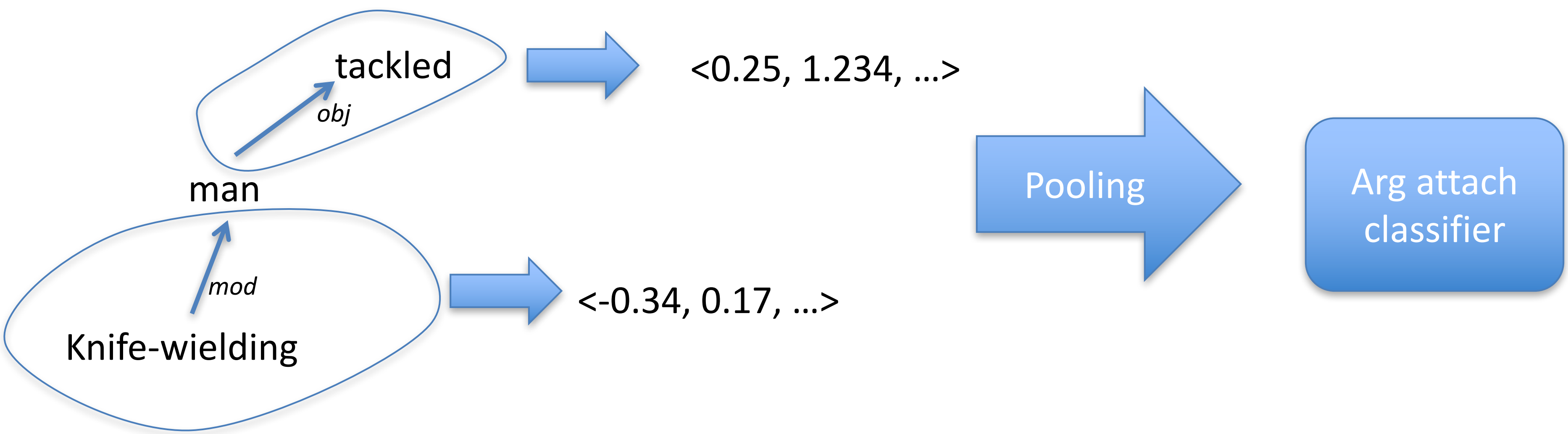
Raytheon BBN Technologies

- Event arguments can often be distant from event triggers
- But often the argument context is informative
  - *The* ***knife-wielding*** *<u>man</u> was tackled by a bystander, but only after three people were severely injured in the <u>attack</u>.*
  - *<u>Acme Inc.</u>***’s creditors*** *were disappointed by Friday’s <u>bankruptcy</u> filing.*
- We would like to learn informative argument contexts which never appear in our supervised training data based on those which do

Context Embeddings: AA (2015)

Raytheon
BBN Technologies

- We trained dense vector representations of the normalized dependency trees contexts of words on Gigaword(s) using a variant of the skip-gram model due to (Levy & Goldberg, ‘14)
- We include this representation in our AA model



Context Embeddings: AA (2015)

Raytheon
BBN Technologies

- Internal development tests on KBP-2014 EA newswire eval corpus (English)
  - Embeddings improve on 2014's best system (BBN1), scored using 2014 EA scorer
- 2015's BBN1 used context embeddings, 2015's BBN3 did not
  - ~10% relative improvement from context embeddings
- Context embeddings used in all languages in 2017

CROSS DOC EVENT FRAME COREFERENCE

Cross Document Event Coreference

Raytheon
BBN Technologies

- Task: Identify coreferent event frames across corpus

Event-1

MEET GID: M1	Role	Fillers
	ENTITY	• EU heads of government • Ahmet Davutoglu
	LOCATION	Brussels
	DATE	11-29-2015

MEET GID: M1	Role	Fillers
	ENTITY	• Turkey • 28 EU member states • the presidents of European Council…
	LOCATION	Brussels
	DATE	11-29-2015

Event-2

MEET GID: M2	Role	Fillers
	ENTITY	• Mehment Simsek • EU
	LOCATION	Brussels
	DATE	12-14-2015

MEET GID: M2	Role	Fillers
	DATE	12-14-2015

- System can (and probably needs to) use
  - Information that is available in the event frames
  - Information directly derived from the document
  - Information provided by other automatic processes
    - Cross-document entity coreference (EDL)
    - Event nuggets and their context
    - Discovered topics
    - …

Challenges

Raytheon BBN Technologies

MEET	Role	Fillers
	ENTITY	• Turkey ~~• 28 EU member states~~ ~~• the presidents of European Council…~~ ***protesters***
	LOCATION	~~Brussels~~ ***Istanbul***
	DATE	11-29-2015

- Imperfect automatic event-frame detection
  - Top performing 2015 system:
    - Precision: 36.8
    - Recall: 39.2
    - Linking F1: 23.3
- Event-frames represent a snapshot of what goes into a knowledge-base, *not* all of the information necessary for coreference decision
  - Marjorie Freedman and Jason Duncan both attended 3 distinct meetings 09-29-2016
- Event nuggets do not provide same discrimination as entity names
  - Nuggets for the 09-29-2016 meetings would be: *attend* or *telecon*
- Currently, no frame-level exhaustive training data
  - Small number of assessments from pilot
  - Even when training data exists, it is likely to be small in quantity

BBN Approach: Overview

Raytheon
BBN Technologies

- Pipeline of decisions
  - Find arguments *(previous section)*
  - Link arguments into per-document event frames *(previous section)*
  - Cluster event-frames across the corpus using event-type (and role) specific intuitions

BBN Approach: Argument Specific Intuitions

- Define per-role equivalence
  - TIME: Year, month, and day (if available) relying on SERIF's Timex normalization
  - PLACE: Containment of GeoNames' Admin districts
  - AGENT/ENTITY/etc.:
    - For named entities, AWAKE cross-document coreference
    - Ignore non-named entities (e.g*. 7 soldiers*, *the crowd*)
- Event Frame Coreference heuristics include
  - Specific roles that must be matched (e.g. TIME or PLACE)
  - Minimum number of arguments that must be matched (e.g. at least three arguments)
  - Maximum number of
    - Documents in which an event can be mentioned
    - Distinct arguments in an event

Thanks!