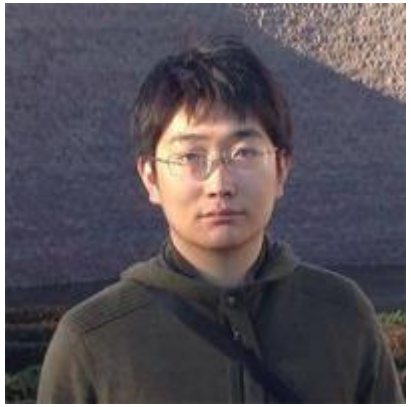
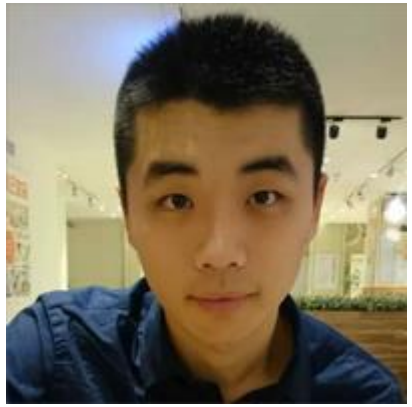
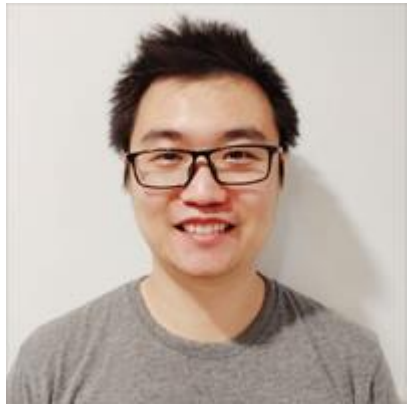
Event-Centric Natural Language Understanding


Muhao Chen Hongming Zhang Qiang Ning Manling Li Heng Ji Dan Roth

Feb 2021

AAAI Tutorials

Event-Centric Natural Language Understanding

Event-Centric Natural Language Understanding

- Human Language is used to describe and reason about events.

- We use it to
  - Describe what happened
    - (and would could have happened, or may happen)
  - Reason about Who did what to whom, and why
  - Understand what led to what? What caused what?
  - We describe and hypothesize thoughts about events, feelings, and plans

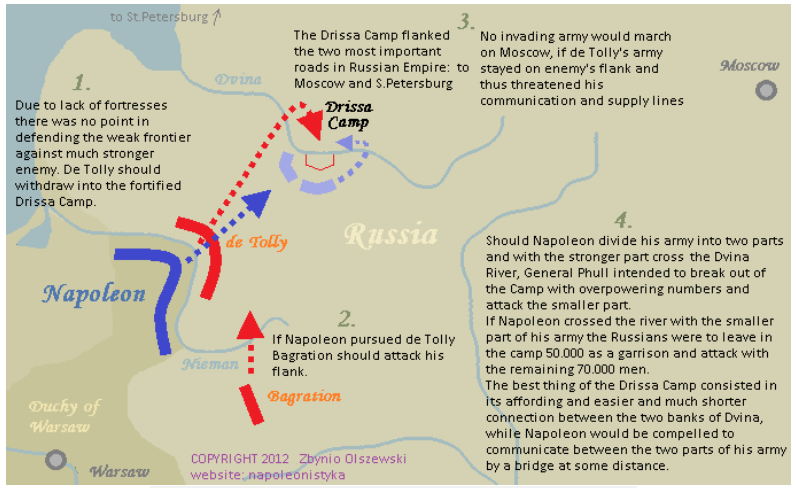
In the months leading up to the riot, Mr. Lee had helped organize a series of pro-Trump car caravans around the country, including one that temporarily blockaded a Biden campaign bus in Texas and another that briefly shut down a Hudson River bridge in the New York City suburbs. To help pay for dozens of caravans to meet at the Jan. 6 rally, he had teamed up with an online fund-raiser in Tampa, Fla., who secured money from small donors and claimed to pass out tens of thousands of dollars.

Rodgers finished 23-of-36 for 296 yards and two touchdowns. His numbers could've been even better had his receivers not dropped a couple of his passes. One dropped ball was a potential score to Allen Lazard. Despite the drop, Lazard made up for it by leading the Packers in receiving. With Davante Adams tied up with Jalen Ramsey, Lazard was able to snatch four balls for 96 yards and a touchdown. Adams still had a great game despite Ramsey's coverage, hauling in nine of his 10 targets for 66 yards and a touchdown. The score frustrated Ramsey because another defensive back was supposed to pick up Adams, who was in motion.

Event-Centric Natural Language Understanding

- We describe events in different levels of abstractions

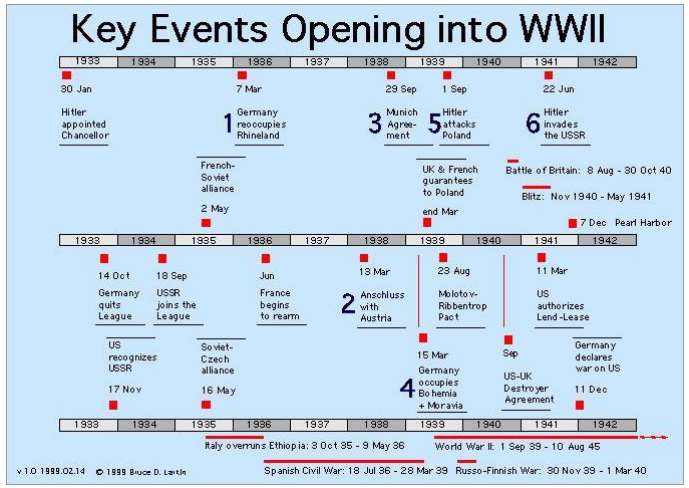
1. Turkey **forces down** Syrian plane.	4. Turkish PM says plane was carrying ammunition for Syria government.
2. Damascus sends note to Ankara over Syrian plane.	5. Last night Turkish F16s **grounded** a Syrian passenger jet.
3. Turkey Escalates Confrontation with Syria.	6. Russia angry at Turkey about Russian passengers.

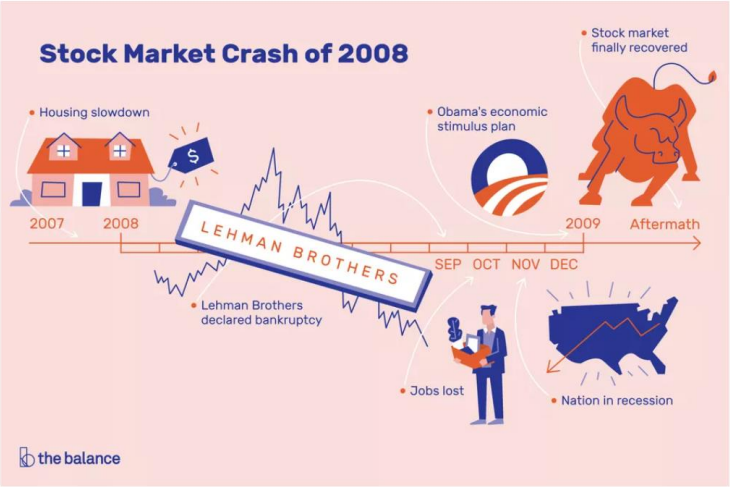


Napoleon Invades Russia

- We reason about events in multiple granularities

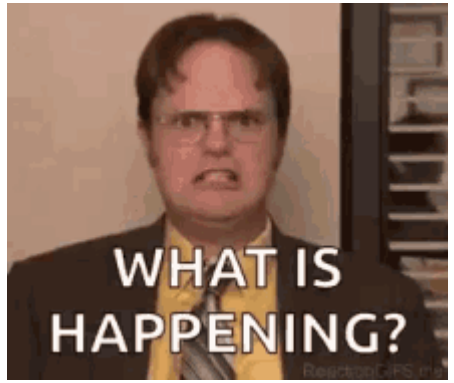
- With a range of goals in mind





Moving to Event-Centric Natural Language Processing

- Most of the work in NLP focuses on understanding "what the text says"
  - We analyze "what's written here" at the sentence level (mostly)
    - But more and more also at a document level
    - And even, rarely, at a multi-document level

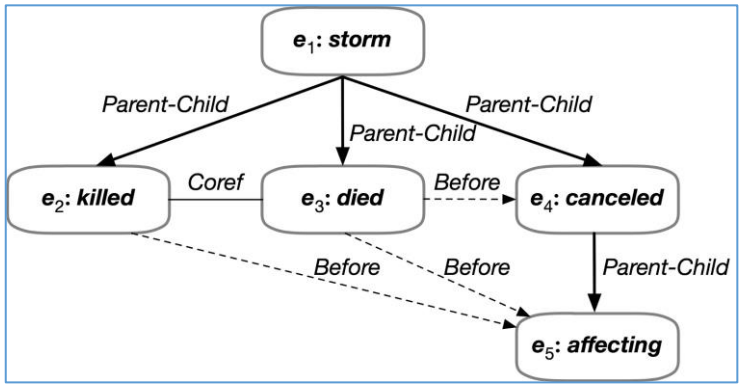


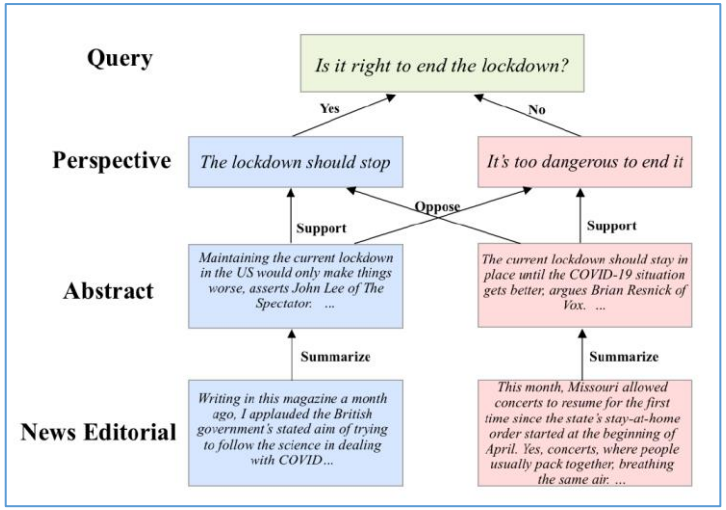
- But with the level of progress in "what the text says"
  - We can now attend to "what is happening"

Understanding "What is Happening"

- Event-Centric Natural Language Understanding and Information Extraction
  - ☐ Brings some change in foci and priorities
  - ☐ It requires "local" text understanding
  - ☐ But necessitates integrating information from multiple documents (and modalities)
    - Information aggregation and consolidation
    - Understanding multiple types of events
    - Understanding relations between events
    - Understanding time and causality
    - Acquiring and using background and commonsense knowledge
    - The ability to generalize from specific, observed, processes
    - The ability to predict – implicit events and possible future events
    - Eventually, it will also require us to identify that there are multiple perspectives
    - …
- And integrating all these into an understanding of "what is happening"

On Tuesday, there was a typhoon-strength (***e_1:storm***) in Japan. One man got (***e_2:killed***) and thousands of people were left stranded. Police said an 81-year-old man (***e_3:died***) in central Toyama when the wind blew over a shed, trapping him underneath. Later this afternoon, with the agency warning of possible tornadoes, Japan Airlines (***e_4:canceled***) 230 domestic flights, (***e_5:affecting***) 31,600 passengers.



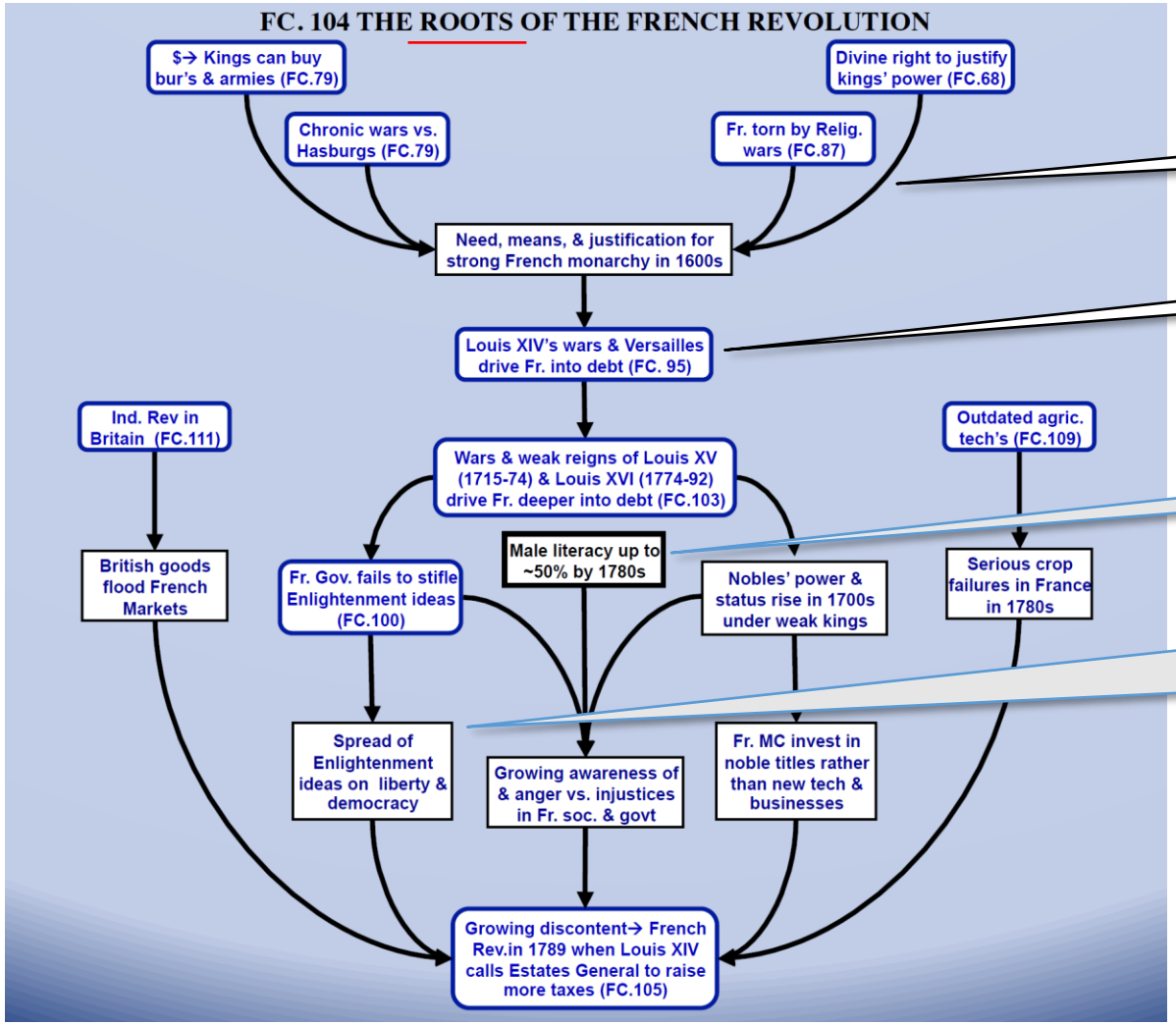


Compiling a History Book

- Ideally, at multiple levels of granularities
  - And accounting for multiple perspectives and interpretations

The history of Democracy

- One way to start is with the French revolution



Influence

Earlier Flow Charts

Earlier Flow Charts

What about the role of the rise of democracy in America?

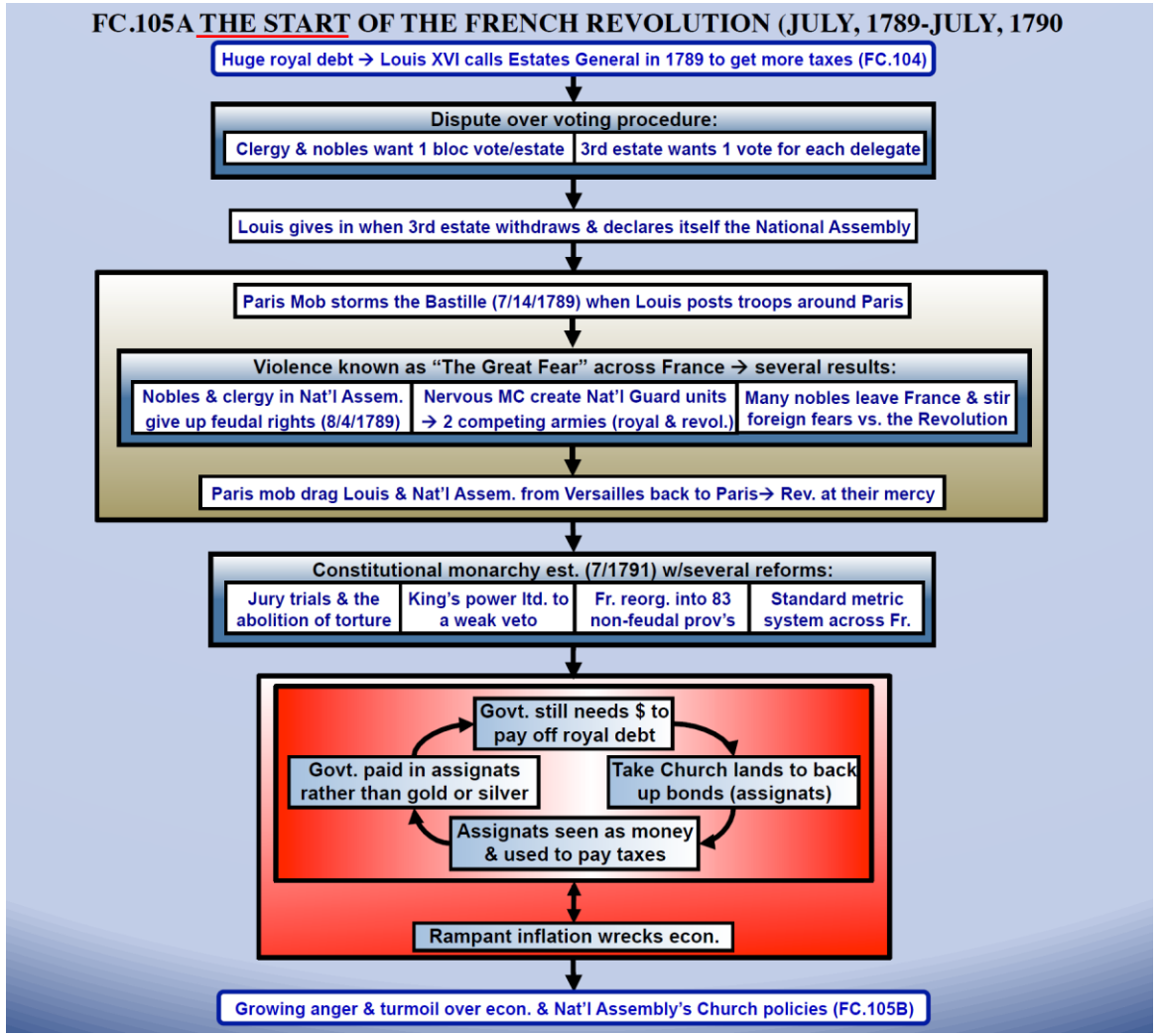
Some say that the Americans' victory over the British has been the single impact on the French revolution, and that the newly formed government in the United States became a model for the French reformers.

Credit: The Flow of History, by Chris Butler

Compiling a History Book

- Ideally, at multiple levels of granularities
  - And accounting for multiple perspectives and interpretations

The history of Democracy



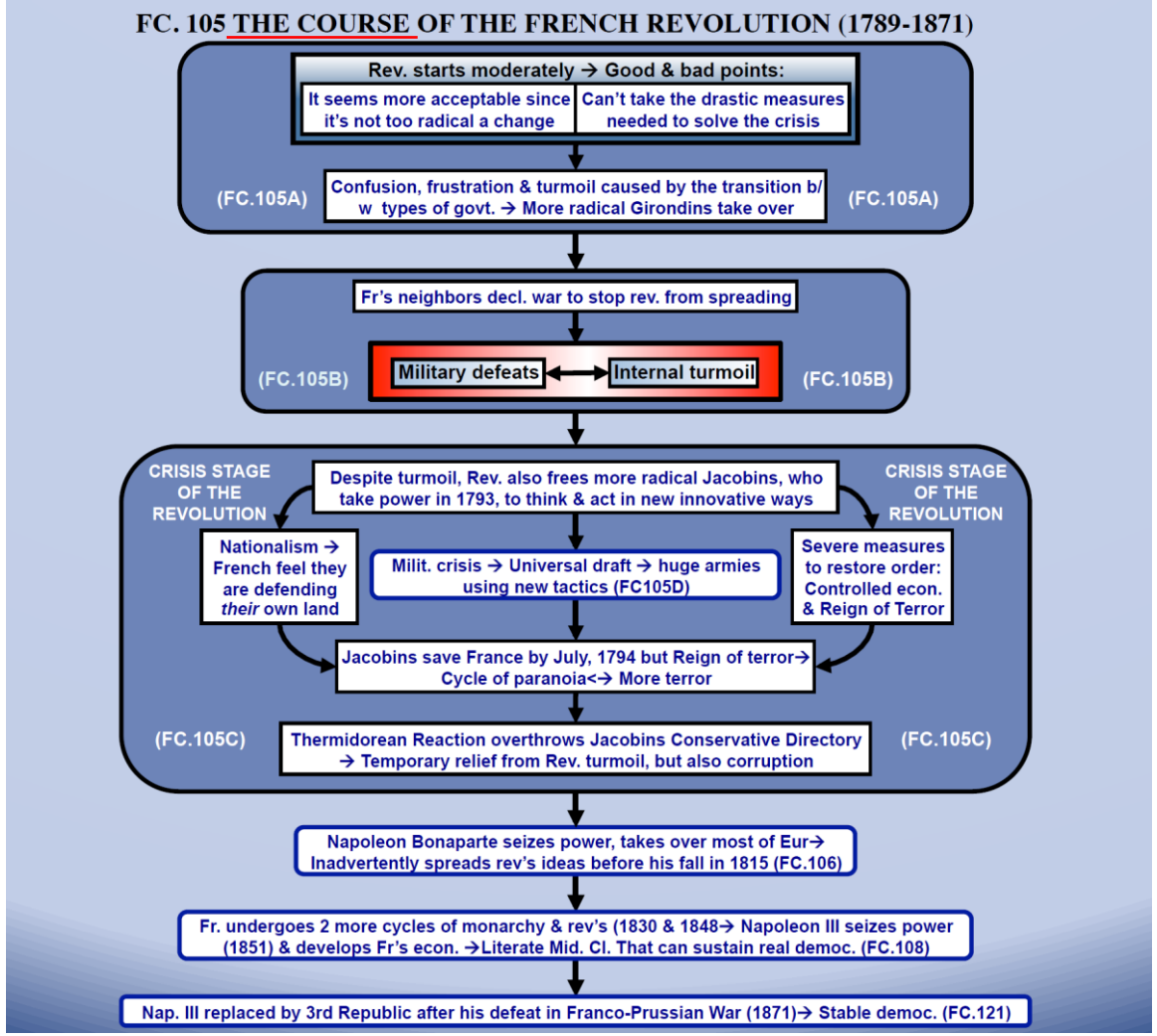
Credit: The Flow of History, by Chris Butler

7

Compiling a History Book

- Ideally, at multiple levels of granularities
  - And accounting for multiple perspectives and interpretations

The history of Democracy

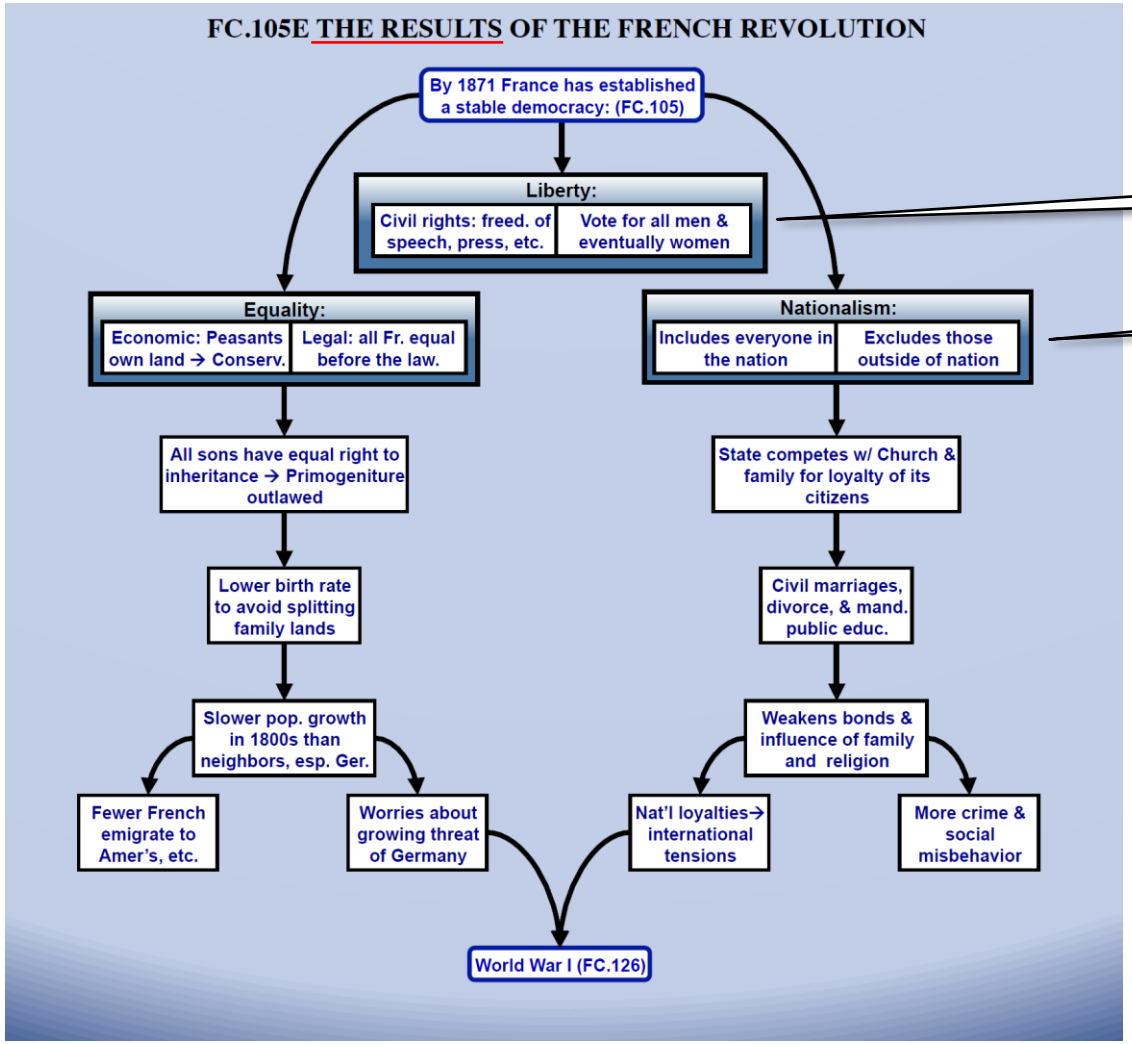


Credit: The Flow of History, by Chris Butler

Compiling a History Book

- Ideally, at multiple levels of granularities
  - And accounting for multiple perspectives and interpretations

The history of Democracy



Liberty & Equality

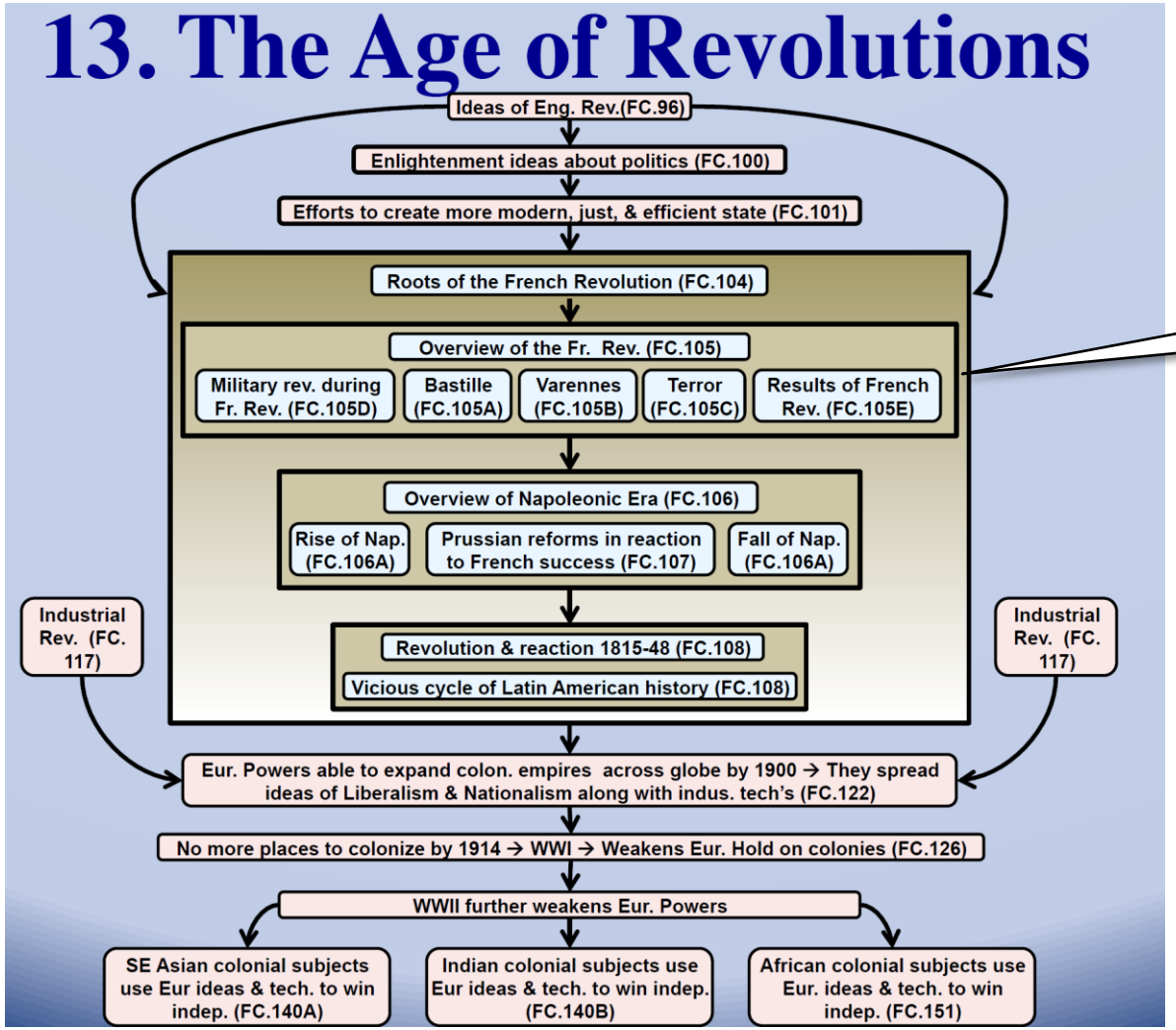
Nationalism?

Credit: The Flow of History, by Chris Butler

Compiling a History Book

- Ideally, at multiple levels of granularities
  - And accounting for multiple perspectives and interpretations

The history of Democracy



The place of the French revolution in the bigger picture

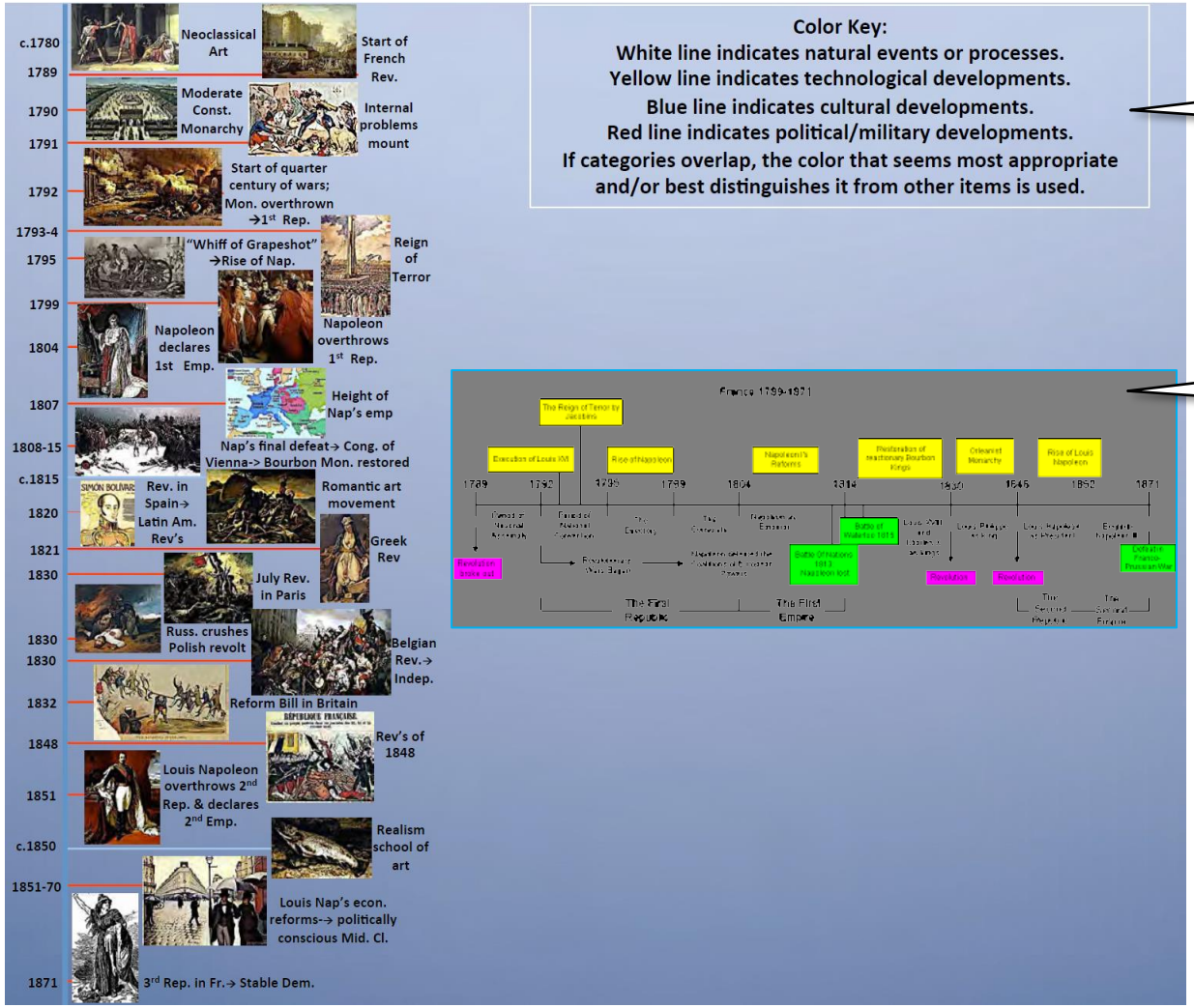
Credit: The Flow of History, by Chris Butler

Compiling a History Book – Timeline

- Ideally, at multiple levels of granularities
  - And accounting for multiple perspectives and interpretations

The history of Democracy

- One way to start is with the French revolution



Various types of processes and developments

Zoom in

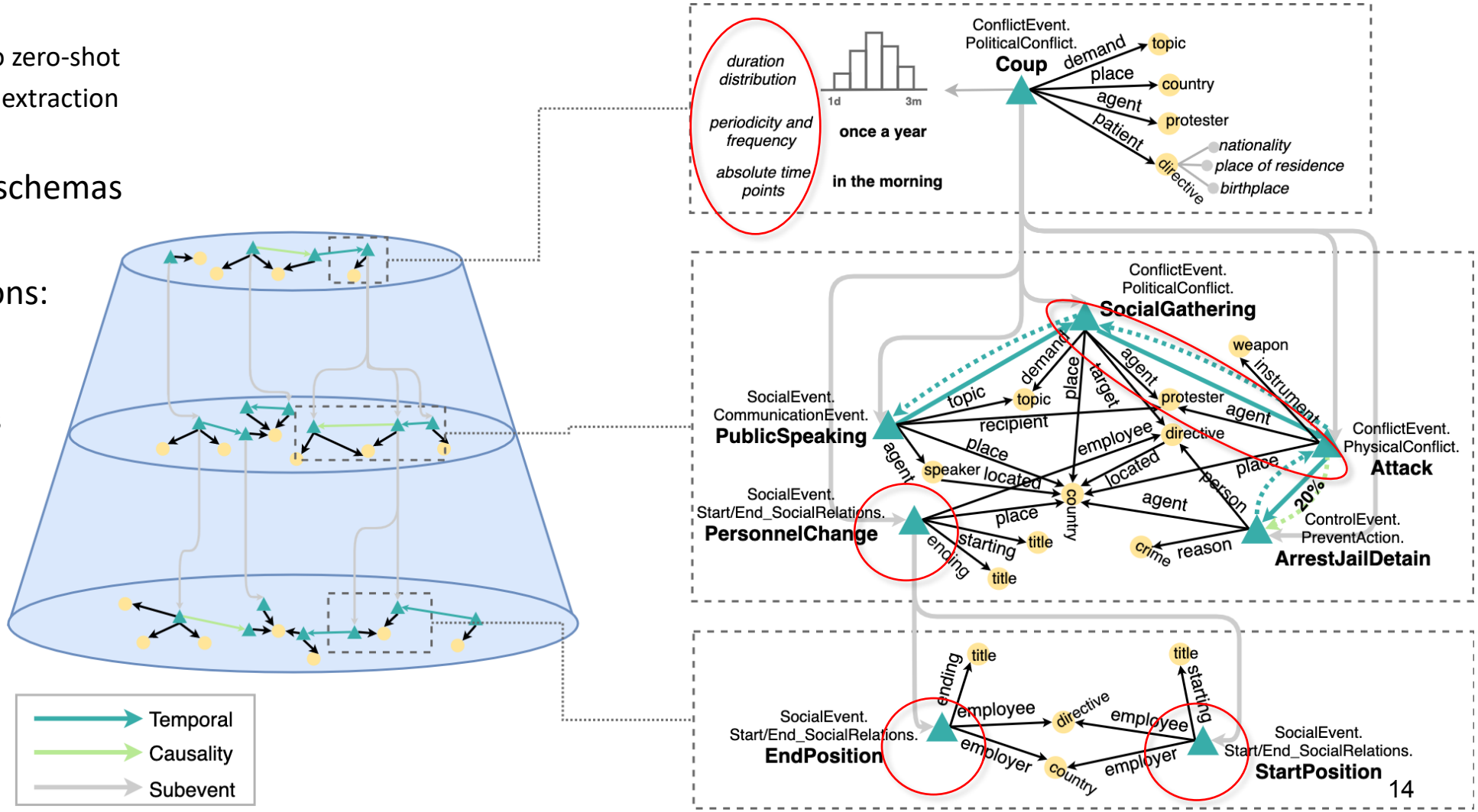
It is challenging to compile such an account given a lot of historical text, and even more so to build it from reported events, as we they occur.

Credit: The Flow of History, by Chris Butler

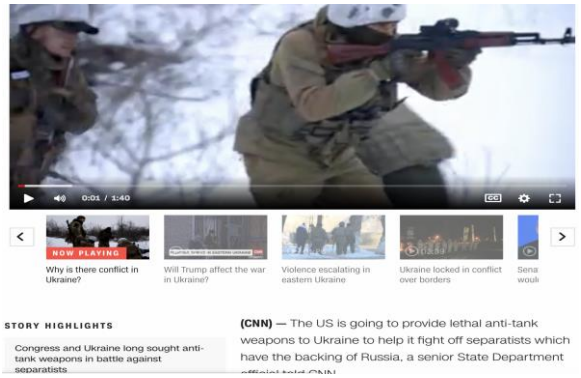
This Tutorial

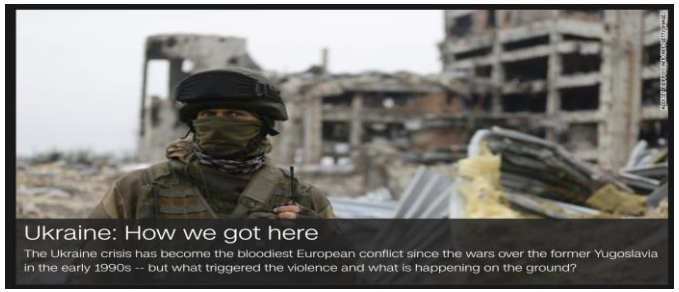
Event Extraction

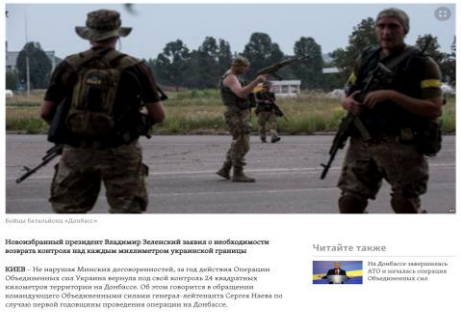
- Event Definition and Representation
- Event Extraction
  - From Supervised to zero-shot
  - Cross-media event extraction
- Hierarchies, event schemas
- Event-event Relations:
  - Coreference
  - Sub-events
  - Temporal Relations

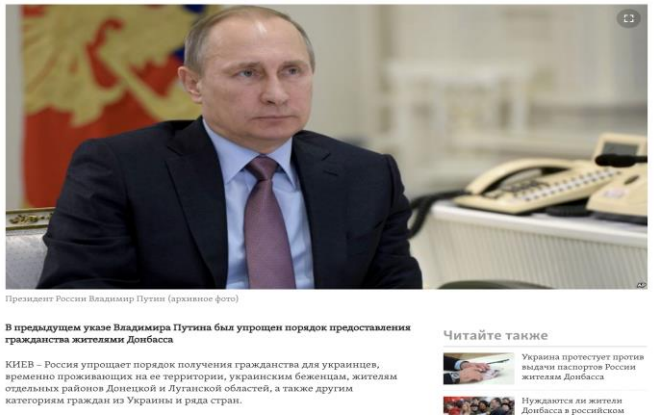


Multisource, Multimedia, Multilingual Data







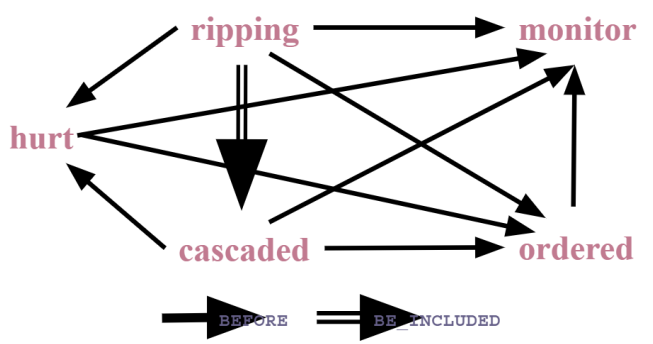


- Event extraction needs to be done across multiple sources and multiple modalities

Relations between Events

- Types of events
- Relations:
  - Coreference, Temporal, parent-child
- Incorporating knowledge constraints
  - Declarative, Statistical

*In Los Angeles that lesson was brought home Friday when tons of earth **cascaded** down a hillside, **ripping** two houses from their foundations. No one was **hurt**, but firefighters **ordered** the evacuation of nearby homes and said they'll **monitor** the shifting ground until March 23rd.*



real variable

$$\hat{I} = \arg\max_{I} \sum_{i<j} \sum_{r} f_r(ij) I_r(ij)$$

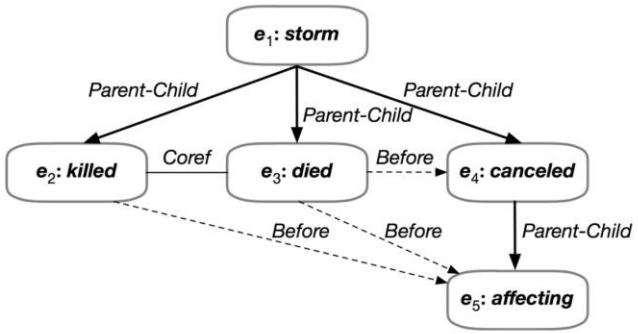
boolean variable

s.t. $\forall i, j, k$

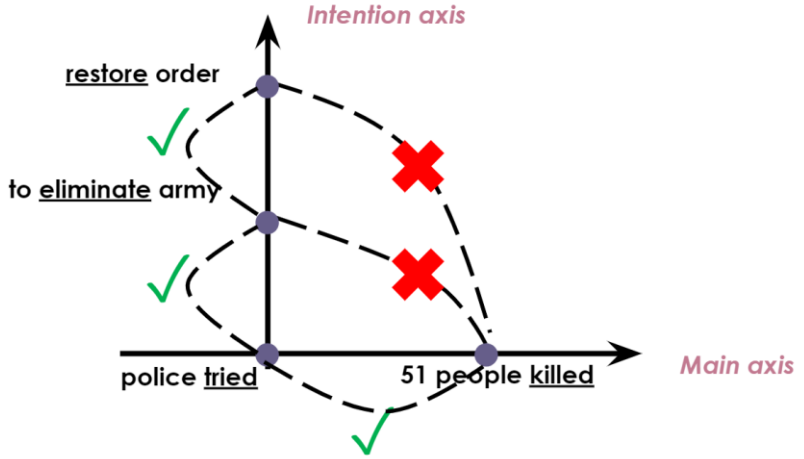
$$\sum_{r} I_r(ij) = 1, \qquad I_{r1}(ij) + I_{r2}(jk) - I_{r3}(ik) \leq 1$$

Uniqueness Transitivity (no loops)

On Tuesday, there was a typhoon-strength (e_1:***storm***) in Japan. One man got (e_2:***killed***) and thousands of people were left stranded. Police said an 81-year-old man (e_3:***died***) in central Toyama when the wind blew over a shed, trapping him underneath. Later this afternoon, with the agency warning of possible tornadoes, Japan Airlines (e_4:***canceled***) 230 domestic flights, (e_5:***affecting***) 31,600 passengers.

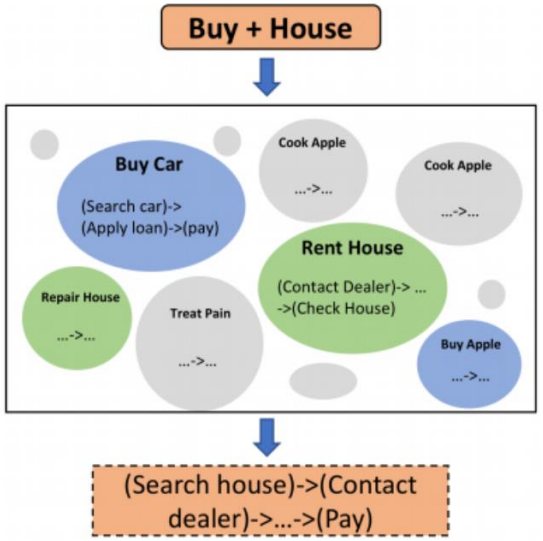


*Police **tried** to **eliminate** the pro-independence army and **restore** order. At least 51 people were **killed** in clashes between police and citizens in the troubled region.*

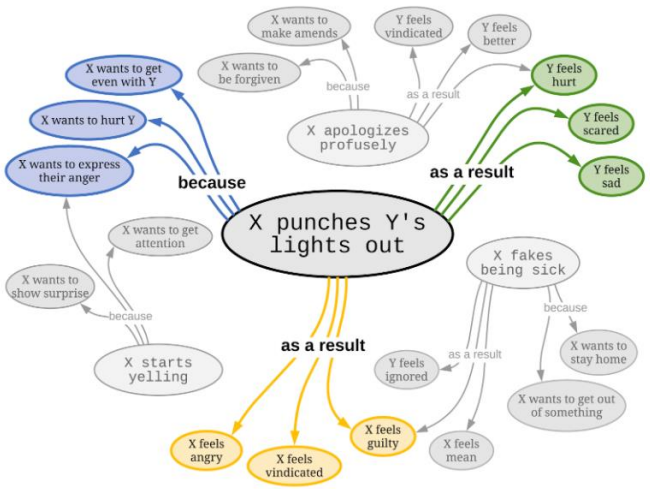


Processes

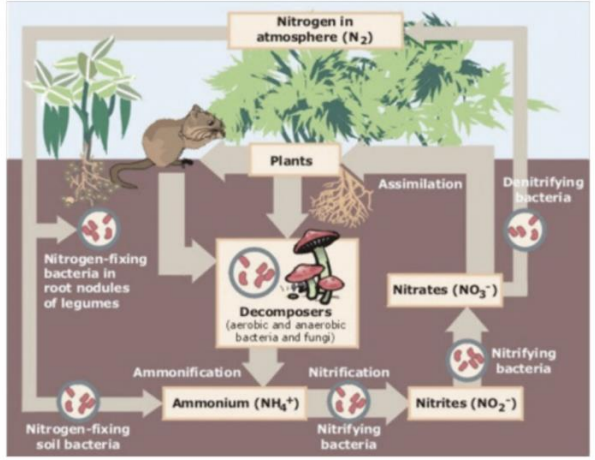
1. Event process completion



2. Event intention prediction



3. Event processes in downstream NLU tasks



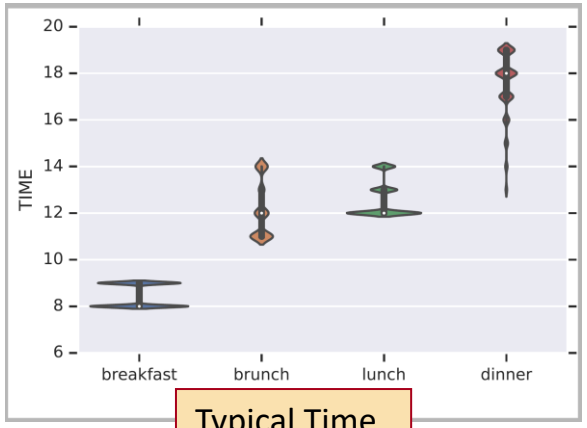
How to do this task?

What are they doing?

Narrative development

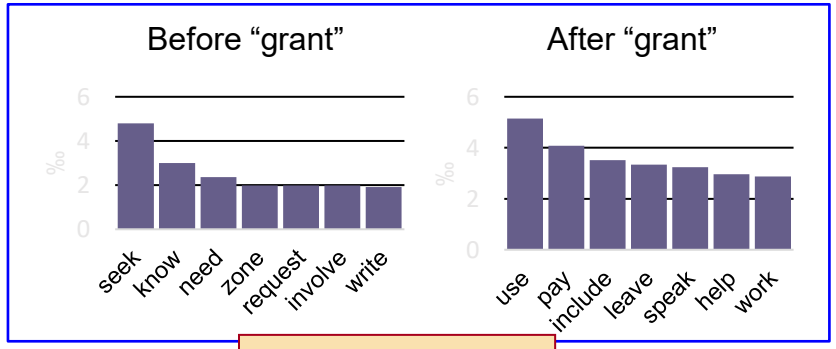
Background and Commonsense Knowledge

- Knowledge Resources that are important to facilitate reasoning about events.
- Commonsense: Event Relations
- Temporal Commonsense



Typical Time

[Elazar et al. ACL'19]



Typical Temporal Relations

[Ning et al. NAACL'18]

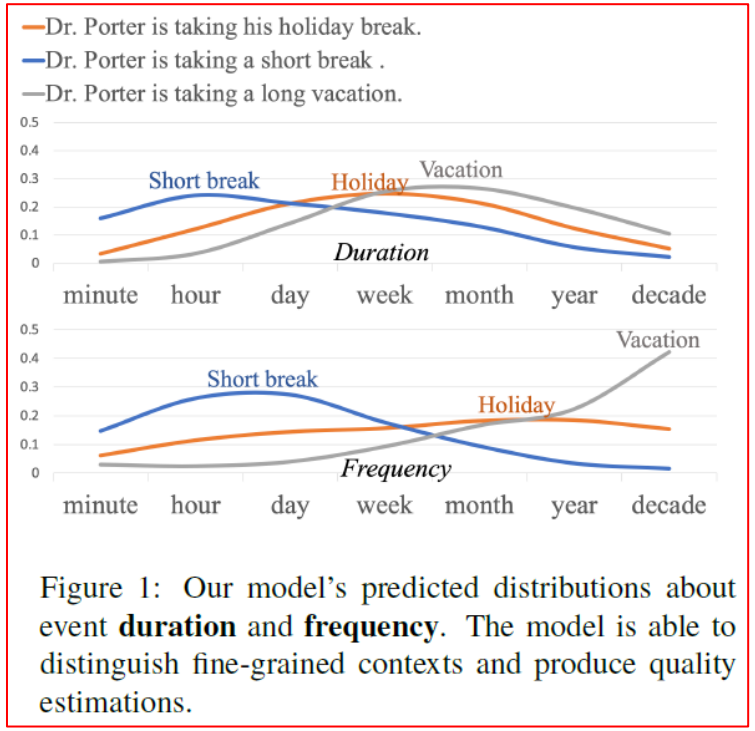
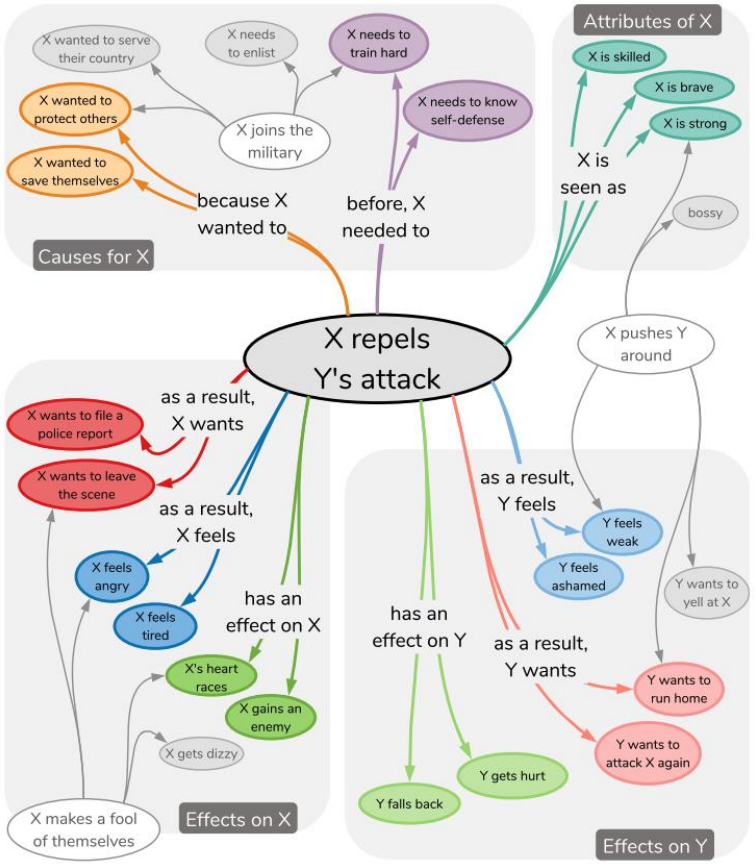


Figure 1: Our model's predicted distributions about event **duration** and **frequency**. The model is able to distinguish fine-grained contexts and produce quality estimations.

[Zhou et al. ACL'20]



Bosselut et al. ACL'19]

Tutorial Outline

- Introduction — 20 min.
  - Dan Roth
- Event-Centric Information Extraction — 40 min.
  - Heng Ji, Mingling Li
- Event-Centric Information Extraction: Relations — 30 min.
  - Qiang Ning
- Break — 30 min.
- Event-Centric Prediction: Processes — 35 min.
  - Muhao Chen
- Event-Centric Knowledge Acquisition: Commonsense — 35 min.
  - Hongming Zhang
- Conclusion and Future Work — 20 min.
  - Heng Ji, Dan Roth

Information Extraction

Event-Centric Natural Language Understanding (Part I)

Manling Li
Department of Computer Science
University of Illinois at Urbana-Champaign

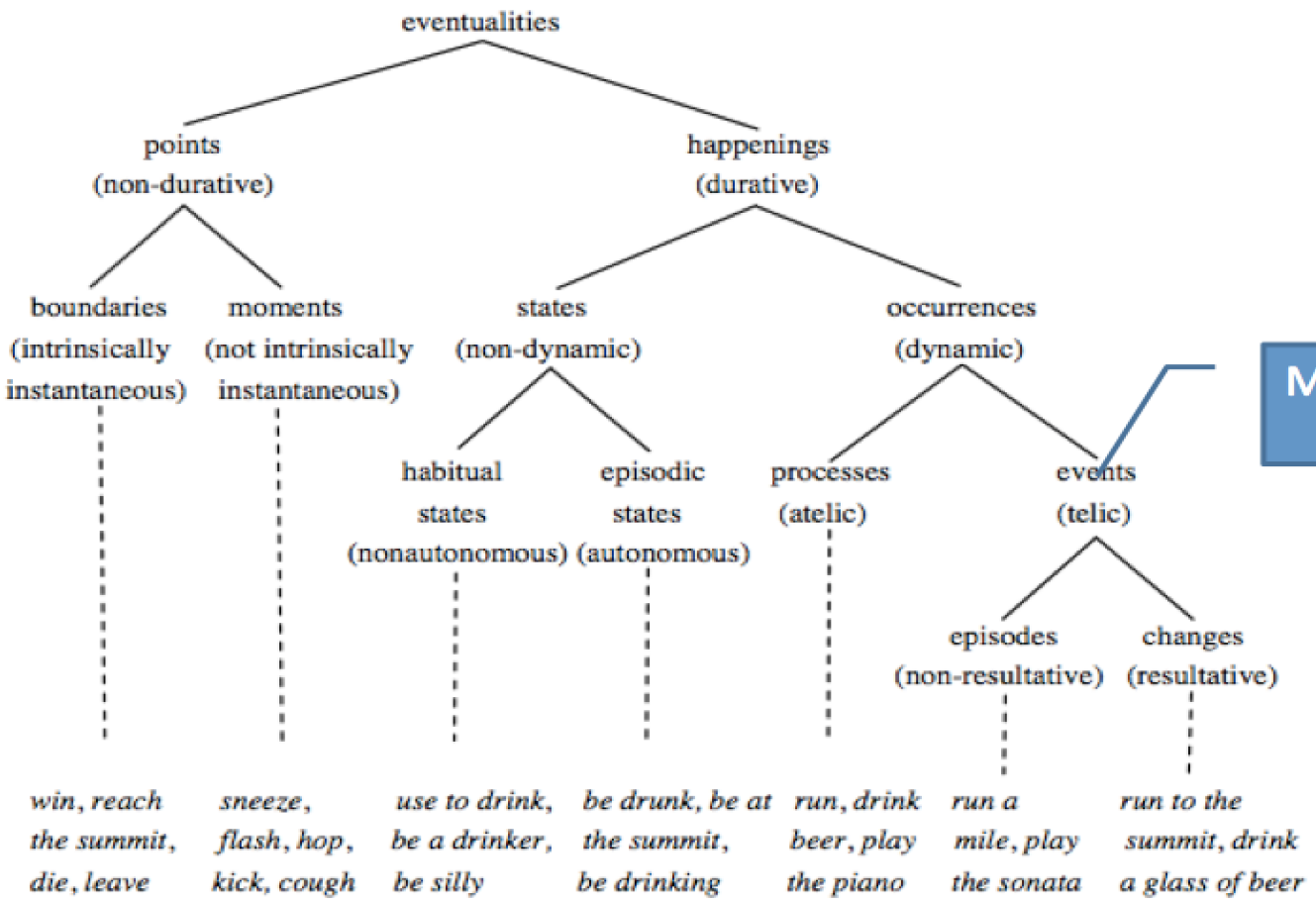
Feb 2020

AAAI Tutorials

Recent Advances in Transferable Representation Learning

What is an event?

- An Event is a specific occurrence involving participants.
- An Event is something that happens.
- An Event can frequently be described as a change of state.



Most of current NLP work focuses on this

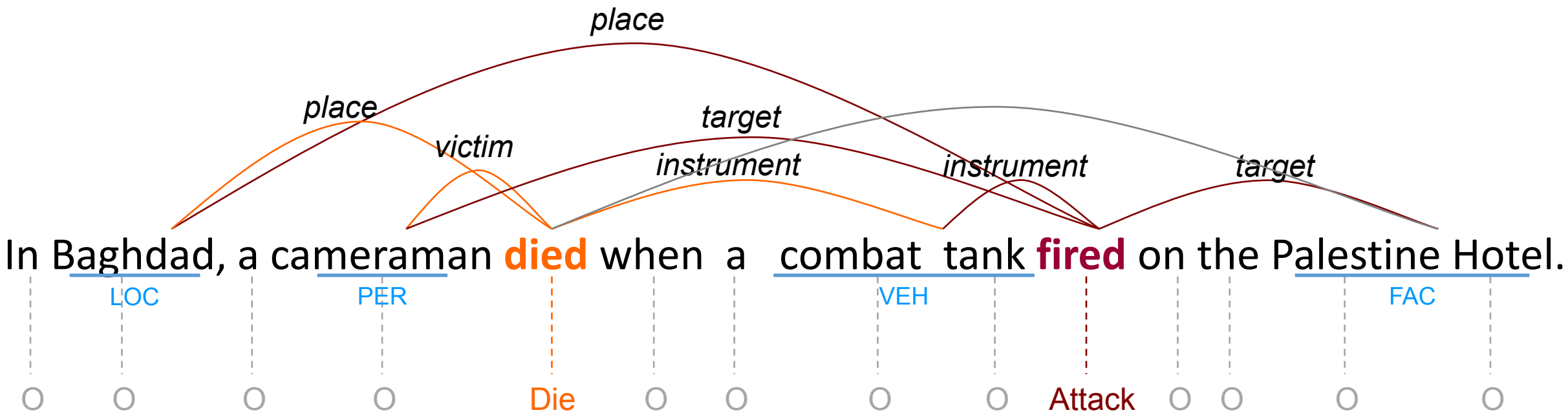
Chart from (Dölling, 2011)

Event Extraction

- Supervised Event Extraction
  - Schema-guided Event Extraction
  - Document-level Event Extraction
- Cross-domain Zero-shot Transfer for Event Extraction
- Cross-lingual Transfer for Multi-lingual Event Extraction
- Cross-media Structured Common Space for Multimedia Event Extraction

What is Information Extraction (IE)?

- Extract structured information and knowledge from unstructured data of heterogeneous data types, in various domains, genres, languages, and data modalities



- It's naturally a structure prediction task! Convert unstructured sequences to graphs

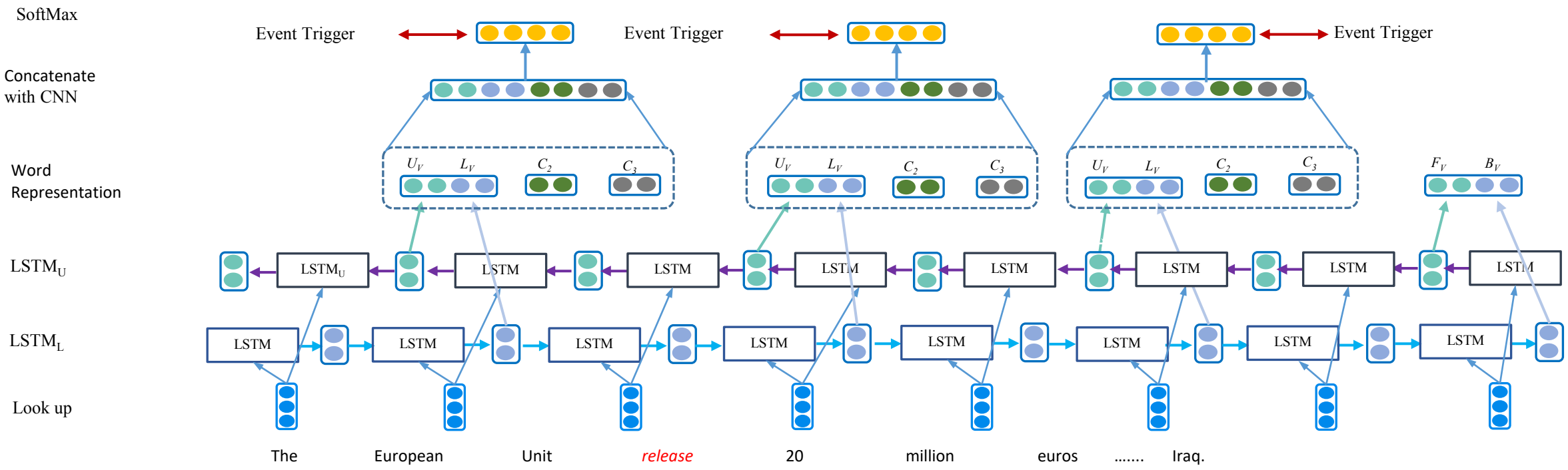
"Old" Days: Supervised Learning with Hand-crafed Features

- Trigger Labeling
  - Lexical
    - Tokens and POS tags of candidate trigger and context words
  - Dictionaries
    - Trigger list, synonym gazetteers
  - Syntactic
    - the depth of the trigger in the parse tree
    - the path from the node of the trigger to the root in the parse tree
    - the phrase structure expanded by the parent node of the trigger
    - the phrase type of the trigger
  - Entity
    - the entity type of the syntactically nearest entity to the trigger in the parse tree
    - the entity type of the physically nearest entity to the trigger in the sentence

- Argument Labeling
  - Event type and trigger
    - Trigger tokens
    - Event type and subtype
  - Entity
    - Entity type and subtype
    - Head word of the entity mention
  - Context
    - Context words of the argument candidate
  - Syntactic
    - the phrase structure expanding the parent of the trigger
    - the relative position of the entity regarding to the trigger (before or after)
    - the minimal path from the entity to the trigger
    - the shortest length from the entity to the trigger in the parse tree

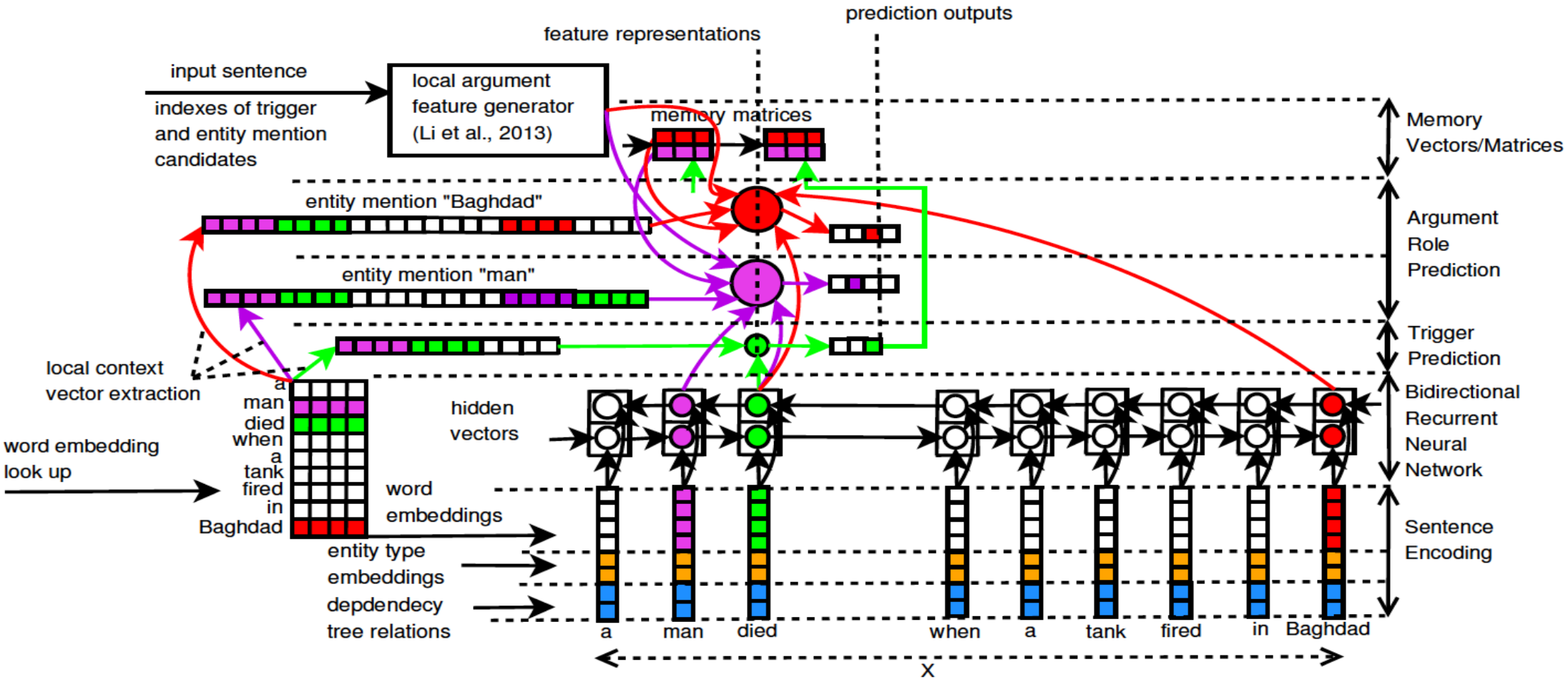
(Chen and Ji, 2009)

A More "Modern" Neural Event Extractor



- Reduce feature engineering efforts to some extent (Feng et al., 2016)
- But still rely on human annotated clean training data still fragile to noise in training data

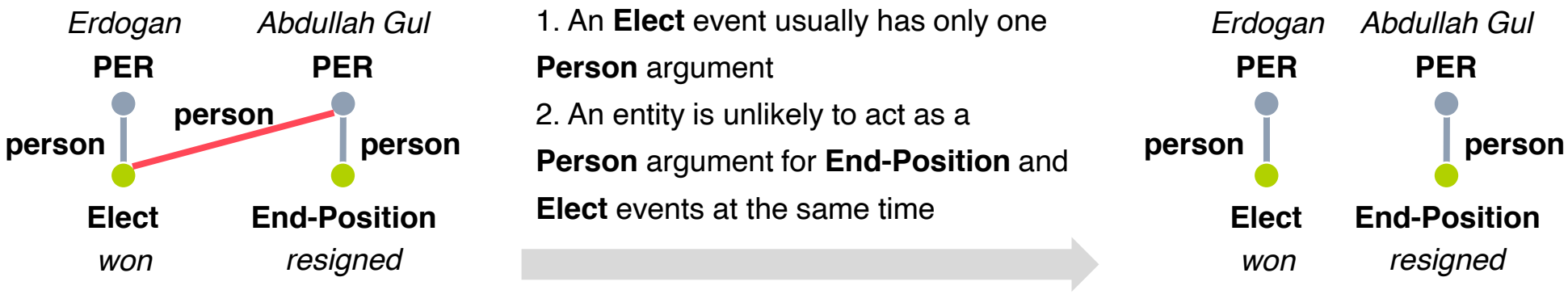
Or Put them Together…



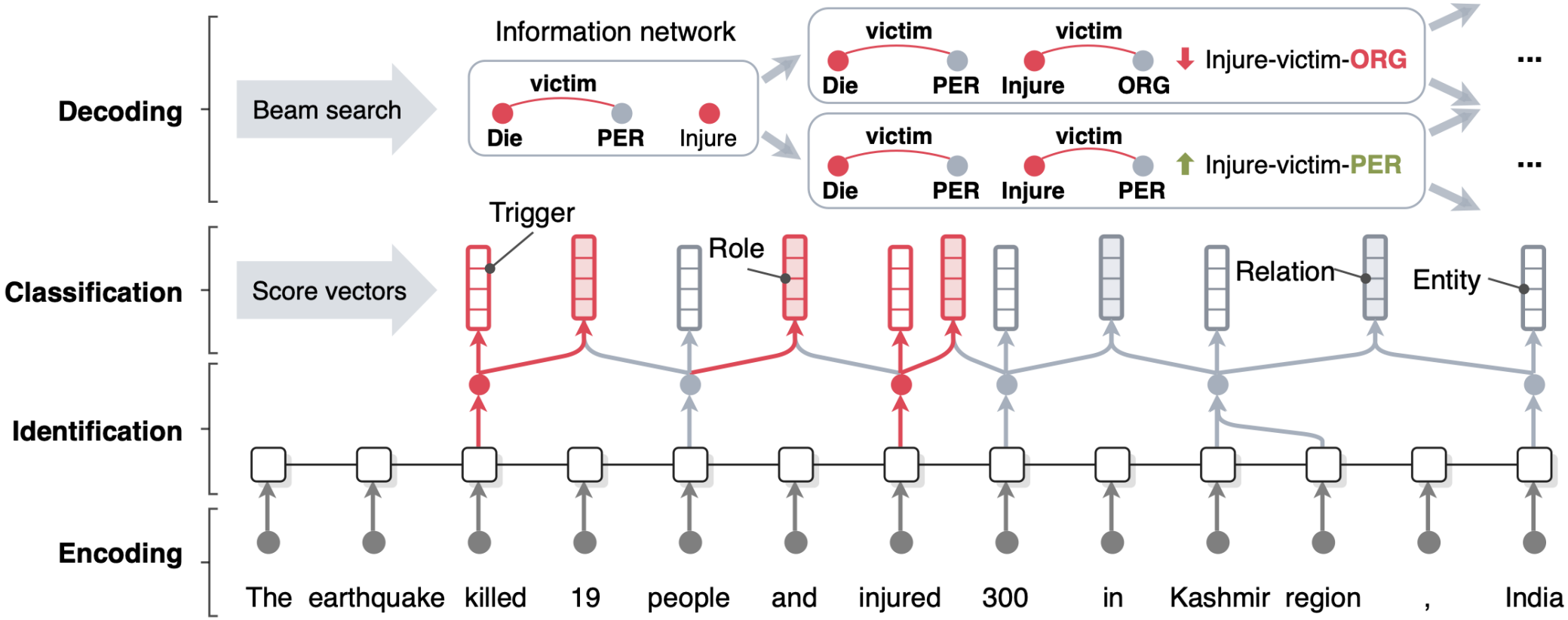
- Add symbolic features by concatenating them with embeddings (Nguyen et al., 2016)

Joint Entity, Relation and Event Extraction

- Pipelined models suffer from the error propagation problem and disallow interactions among components
- Existing neural models do not explicitly model cross-subtask and cross-instance interactions among knowledge elements
- Example: *Prime Minister **Abdullah Gul** resigned earlier Tuesday to make way for **Erdogan**, who won a parliamentary seat in by-elections Sunday.*

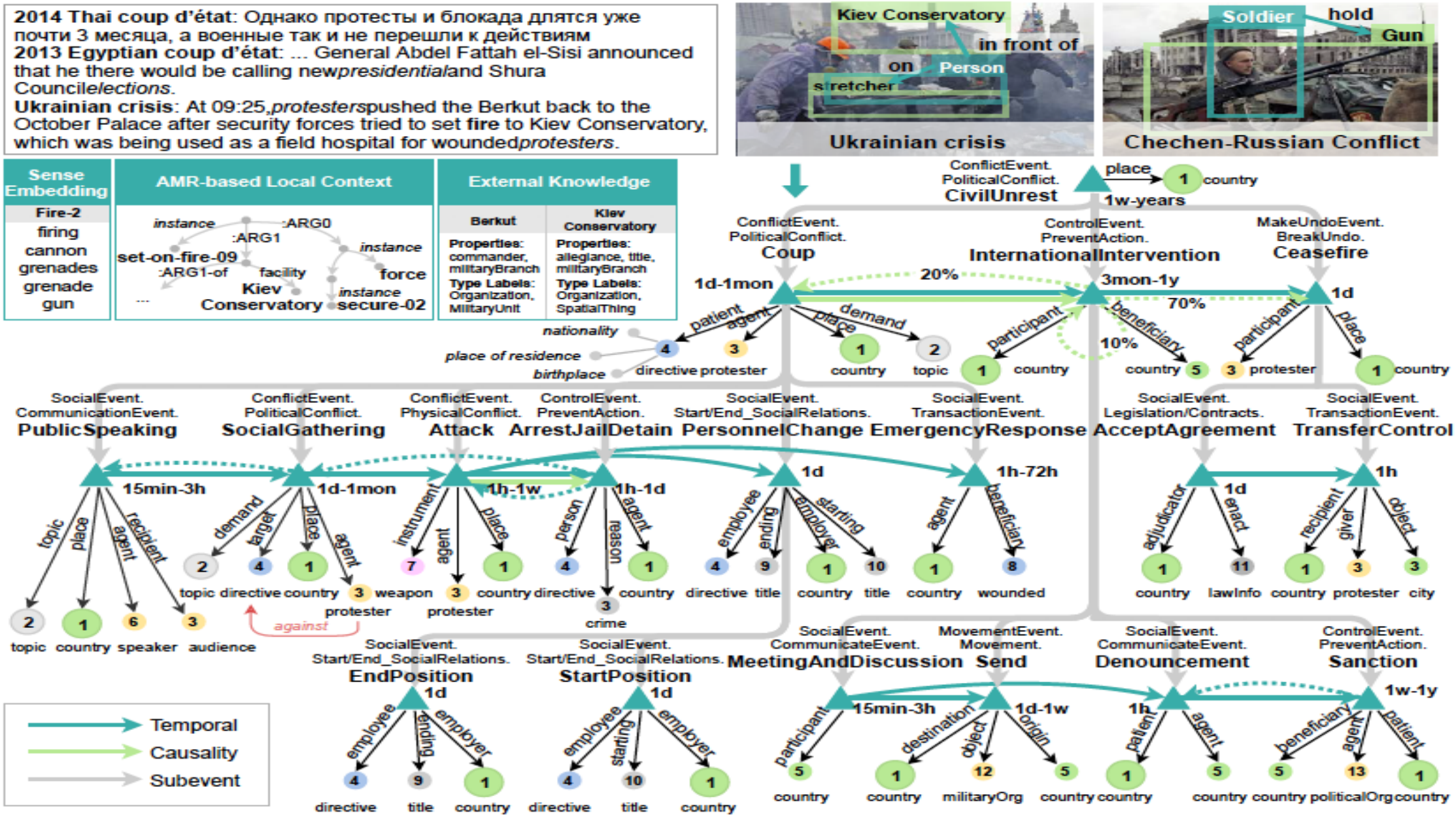


OneIE: An End-to-end Neural Model for IE (Lin et al., 2020)



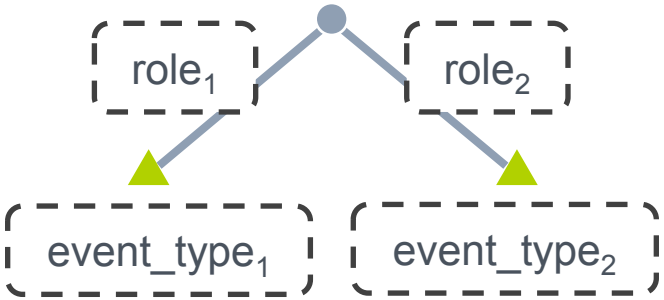
- Our OneIE framework extracts the information graph from a given sentence in four steps: encoding, identification, classification, and decoding

Move from Entity-Centric to Event-Centric NLU

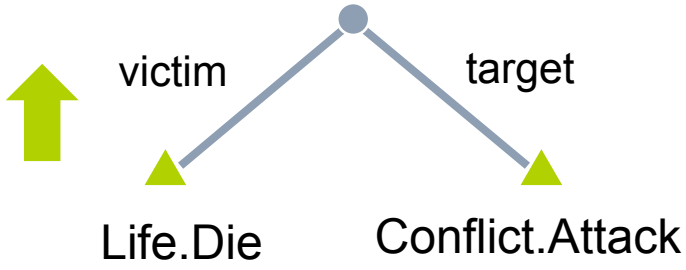


Event Schema Induction

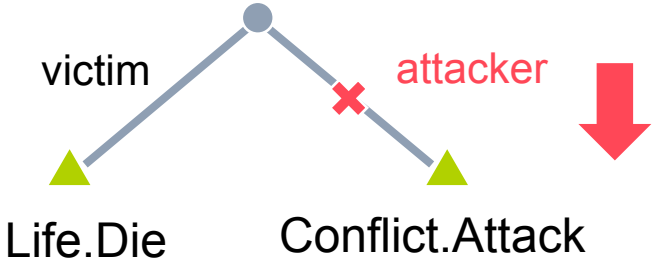
- We design a set of *global feature templates* (e.g., event_type$_1$ – role$_1$ – role$_2$ – event_type$_2$: an entity acts a role$_1$ argument for an event_type$_1$ event and a role$_2$ argument for an event_type$_2$ event in the same sentence). A more comprehensive event schema library is inducted following (Li et al, 2020).
- The model learns the *weight* of each feature during training



Template



Positive weight



Negative weight

Incorporating Global Features

- Given a graph G, we generate its global feature vector as $f(G)$, where f is a function that evaluates a certain feature and returns a scalar. For example,

$$f_i(G) = \begin{cases} 1, & G \text{ has multiple ATTACK events} \\ 0, & \text{otherwise.} \end{cases}$$

- Next, we learn a weight vector and calculate the global feature score of as the dot production of and .
- **Global score** of a graph: local graph score + global feature score:

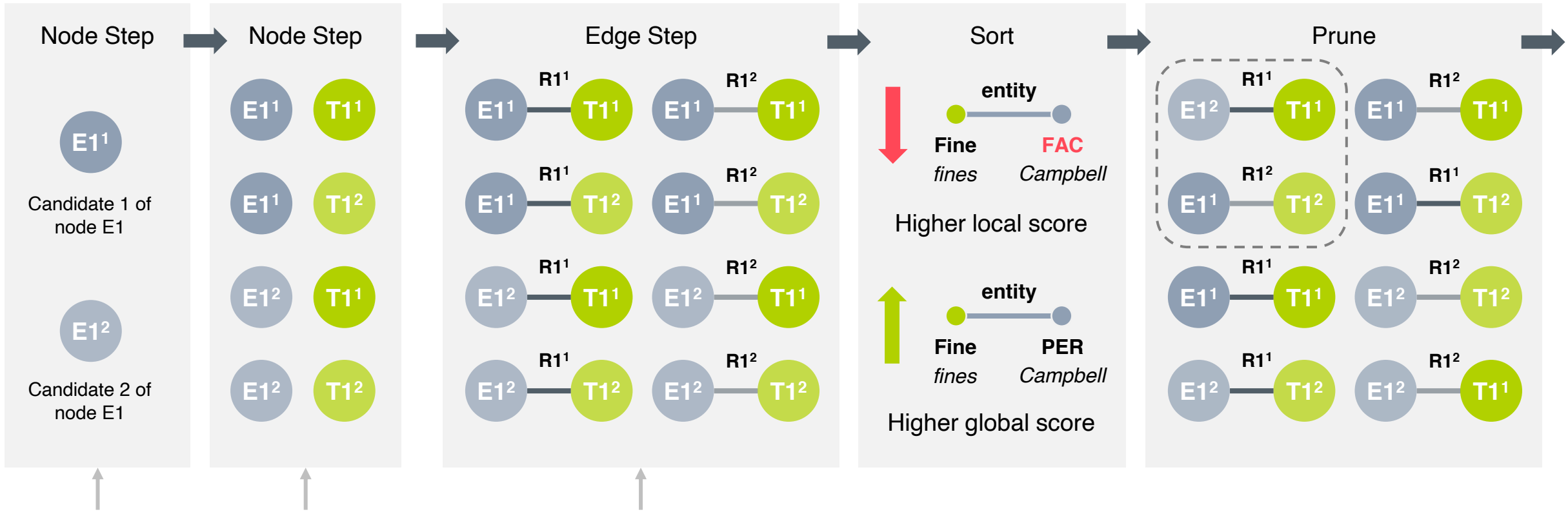
$$s(G) = s'(G) + \boldsymbol{u}\boldsymbol{f}_G$$

- We assume that the gold-standard graph for a sentence should achieve the highest global score and minimize the following loss function:

$$\mathcal{L}^{\mathrm{G}} = s(\hat{G}) - s(G)$$

Decoding

- We use beam search to decode the information graph
- Example: *He also brought a check from* ***Campbell*** *to pay the* ***fines*** *and fees.*



add node 1

Campbell: FAC, ORG

add node 2

fine: Fine, Sue

add the edge between node 1 and 2

null, entity, person, …

Experiment Results

- We conduct our experiments on ACE (Automatic Content Extraction) 2005 (F-score, %)

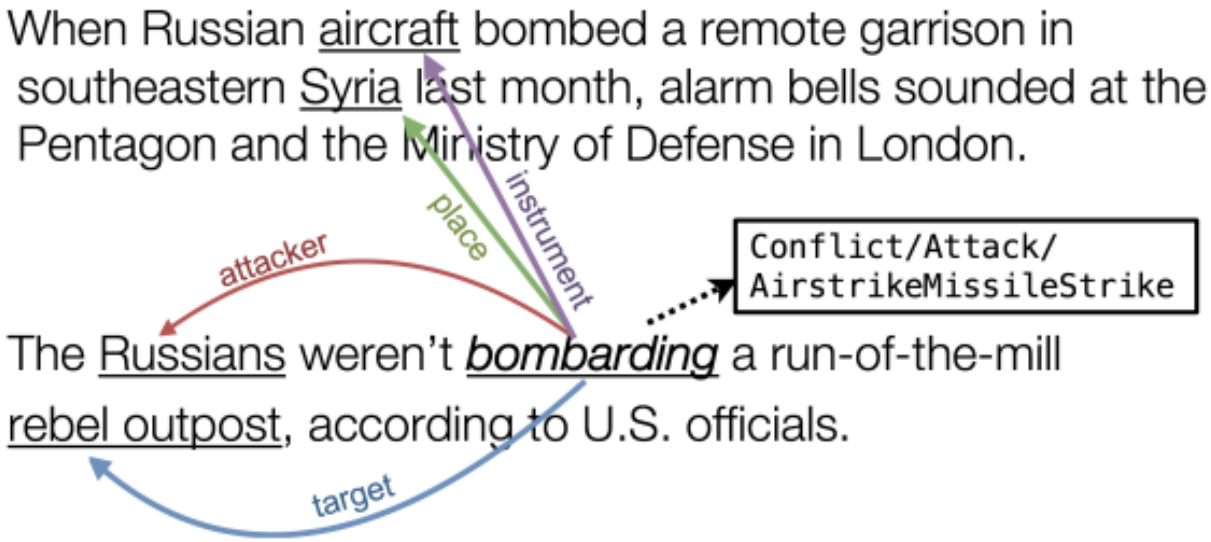
Model	ACE05-R		ACE05-E				
	Entity	Relation	Entity	Trigger Identification	Trigger Classification	Argument Identification	Argument Classification
DyGIE++	88.6	63.4	89.7	-	69.7	53.0	48.8
DyGIE++*	-	-	**90.7**	76.5	73.6	55.4	52.5
OneIE	**88.8**	**67.5**	90.2	**78.2**	**74.7**	**59.2**	**56.8**

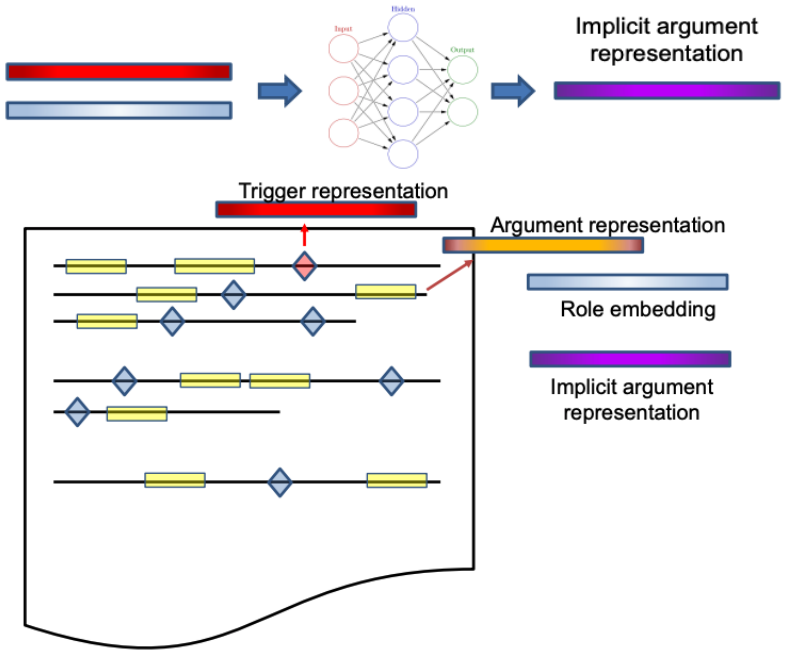
- We evaluate the portability of the proposed framework on ACE05-CN (Chinese) and ERE-ES (Spanish).

Dataset	Training	Entity	Relation	Trigger	Argument
ACE05-CN	CN	88.5	62.4	65.6	52.0
	CN+EN	89.8	62.9	67.7	53.2
ERE-ES	ES	81.3	48.1	56.8	40.3
	ES+EN	81.8	52.9	59.1	42.3

Extending from Sentence-Level to Document-Level

- Multi-Sentence Argument Linking (Ebner et al., 2020)



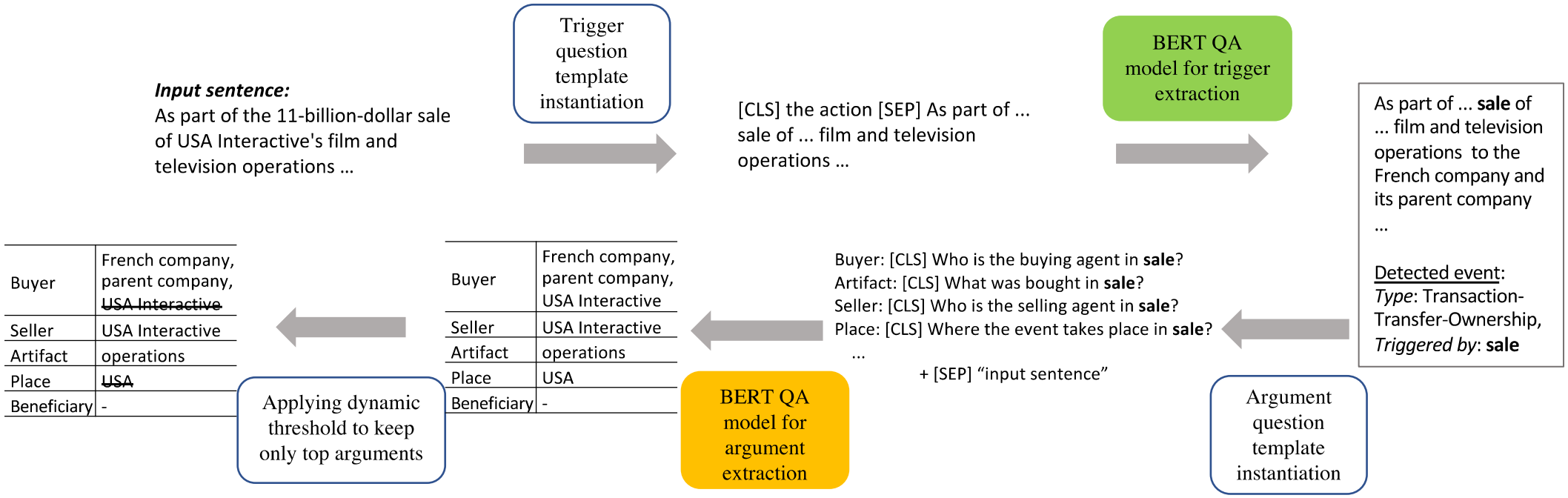


$$l(a, \tilde{a}_{e,r}) = s_{E,R}(e, r) + s_{A,R}(a, r) + s_l(a, \tilde{a}_{e,r}) + s_c(e, a), \quad a \neq \epsilon$$

- Roles are evoked by event triggers, forming implicit arguments
- Implicit arguments linked to explicit mentions in text
  - ☐ Representations: Learn span representations for each trigger span and candidate argument span
  - ☐ Prune: For each trigger, prune to top-K candidate arguments
  - ☐ Linking score: Score representations of implicit arguments with representations of explicit arguments using a decomposable scoring function

Extending from Sentence-Level to Document-Level

- Event Extraction by Answering (Almost) Natural Questions (Du and Cardie, 2020)



The input sequences for the two QA models share a standard BERT-style format

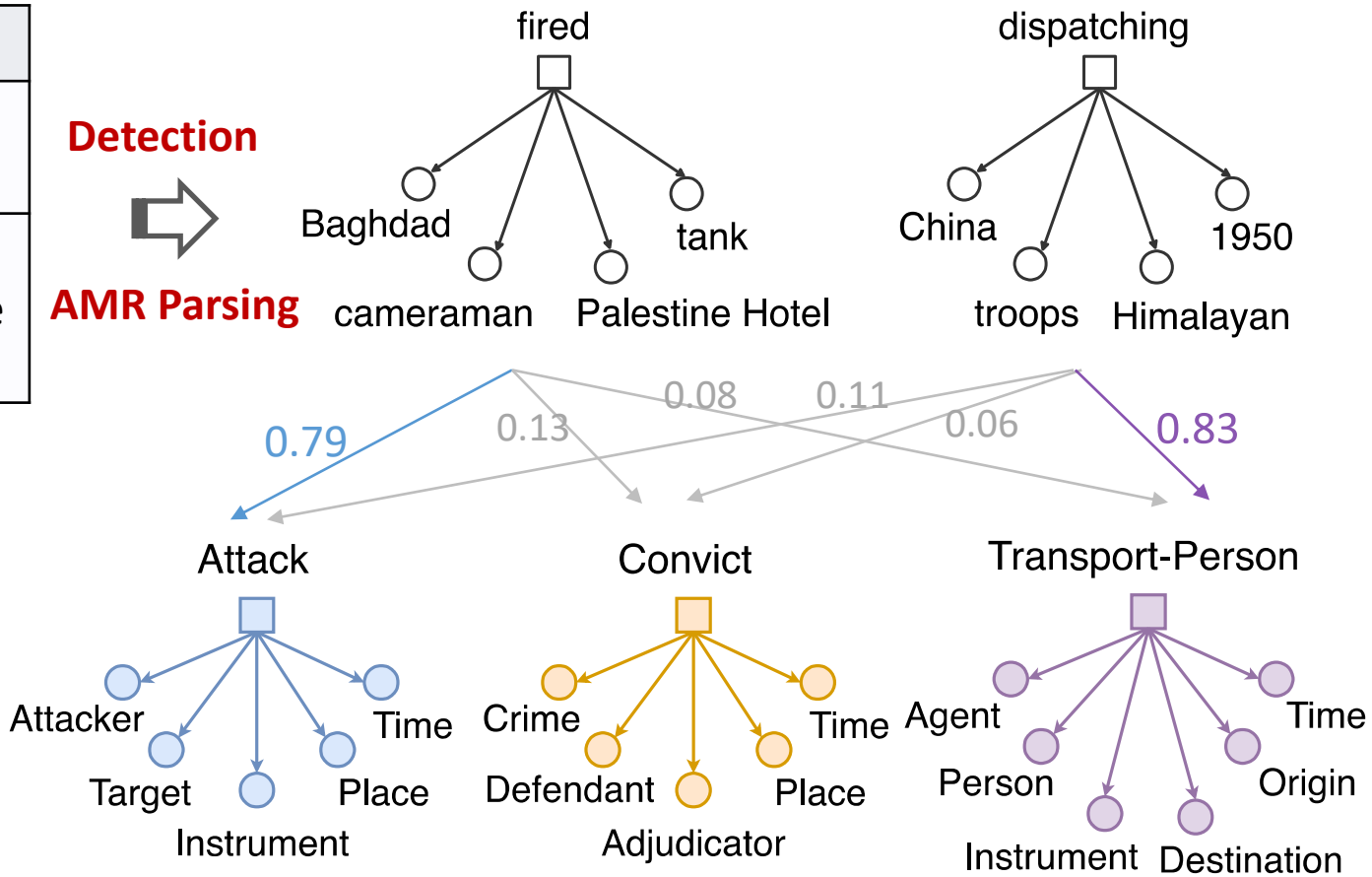
[CLS] <question> [SEP] <sentence> [SEP]

Event Extraction

- Supervised Event Extraction
  - Schema-guided Event Extraction
  - Document-level Event Extraction
- ➢ Cross-domain Zero-shot Transfer for Event Extraction
- Cross-lingual Transfer for Multi-lingual Event Extraction
- Cross-media Structured Common Space for Multimedia Event Extraction

Move to any New Types: Zero-shot Event Extraction (Huang et al., 2018)

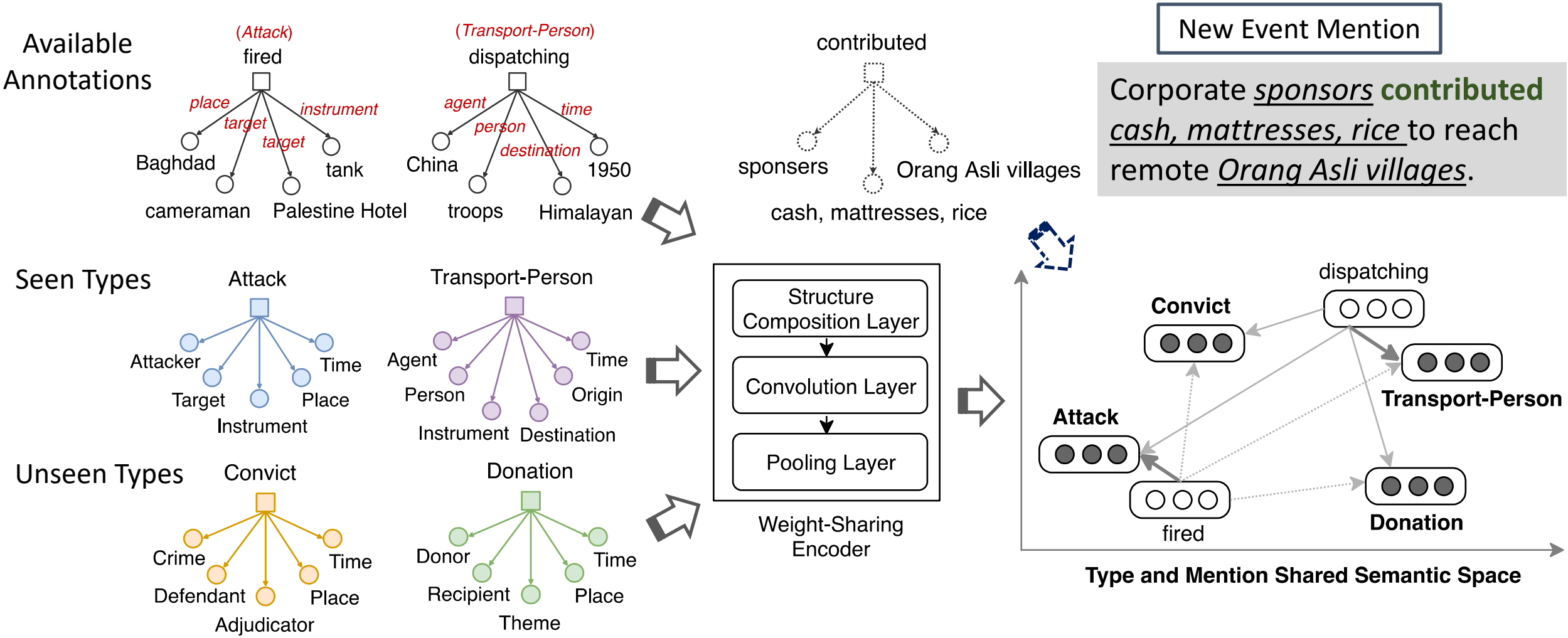
ID	Sentences
S1	In *Baghdad*, a *cameraman* **died** when a combat *tank* **fired** on the *Palestine Hotel*.
S2	The government of *China* has ruled Tibet since 1951 after **dispatching** *troops* to the *Himalayan* region in *1950*.



Hypothesis: Both event mentions and types have rich semantics and structures, which can specify their consistency and connections

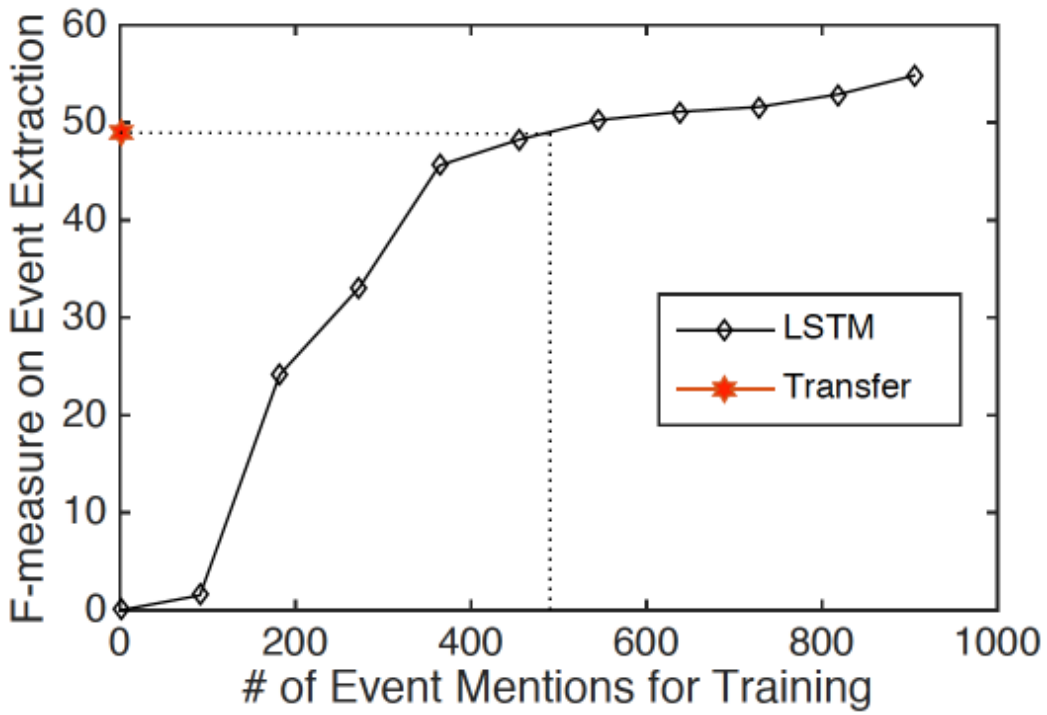
Large-Scale Target Event Ontology

Zero-shot Event Extraction



Large-Scale Target Event Ontology

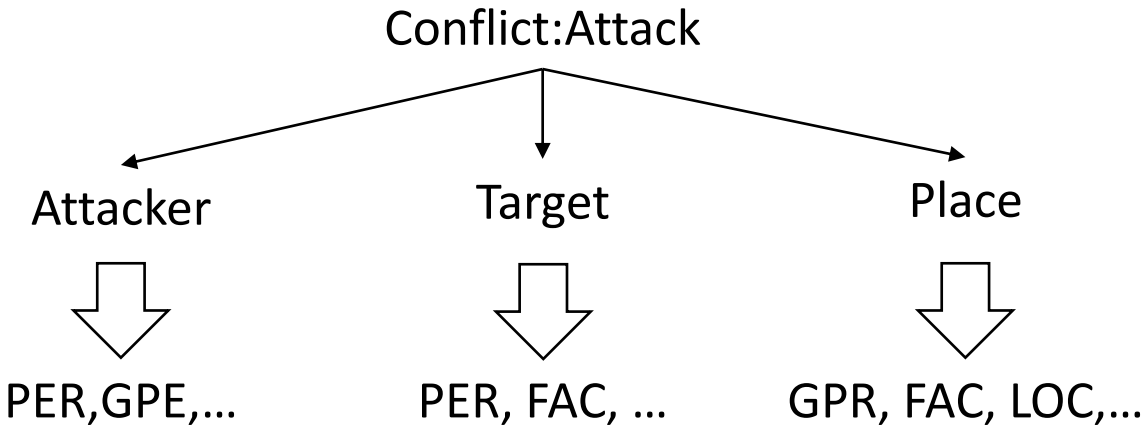
How Much Human Effort Can We Save?



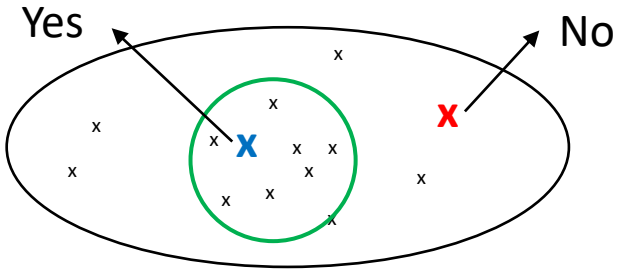
Achieved **comparable** performance as a supervised system when it's trained on **500** event mentions from **3000** sentences

- Target Event Ontology: ACE(33 types) + FrameNet (1161 frames) = 1194 types
- Seen types for training: 10 most popular ACE types
- Unseen type: 23 remaining ACE types

Label-aware Classification (Zhang et al., 2020)

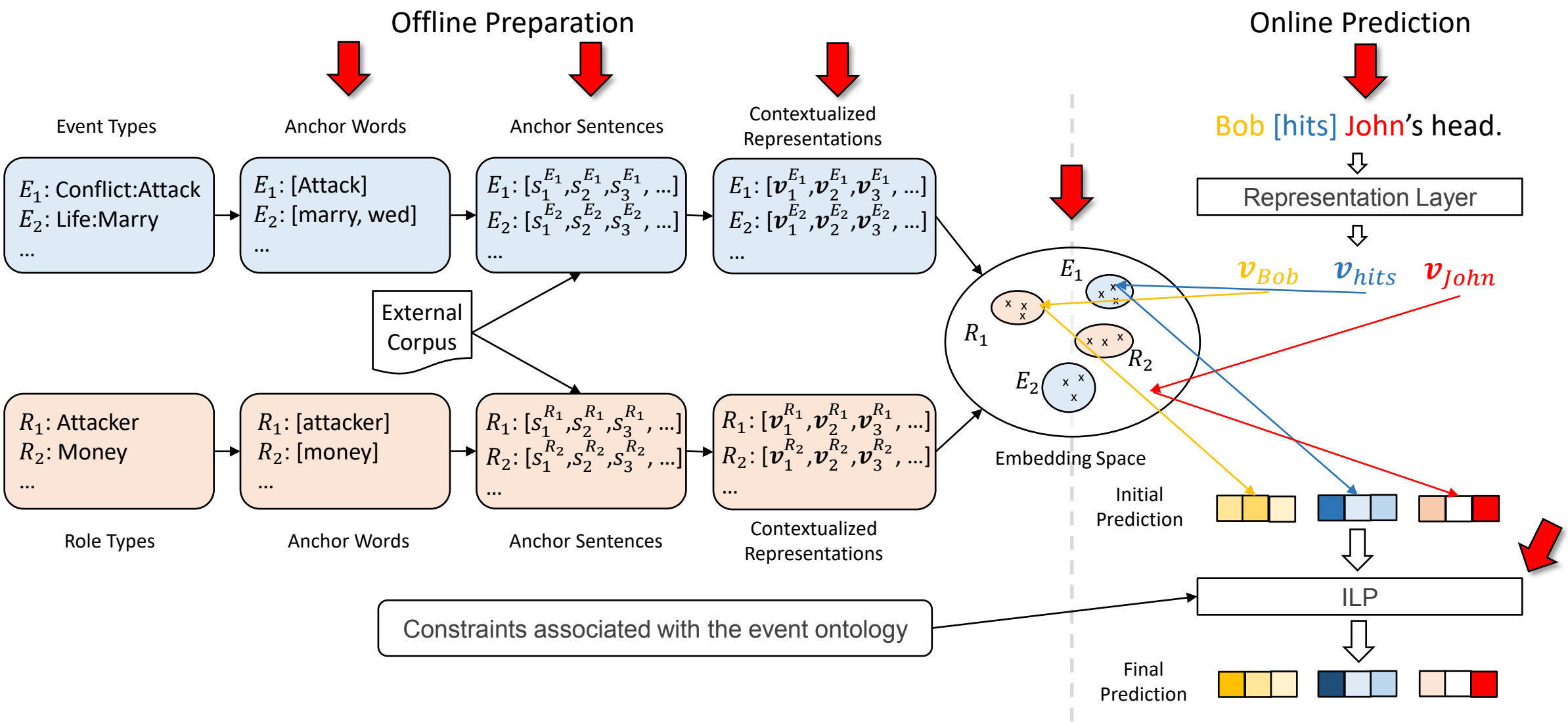


- Label semantics
  - We select “attack” as the label because we assume that it can represent the overall meaning of this event type.
- Constraints
  - “Attacker” can only be the argument of “Conflict:Attack” rather than “Life:Marry”.



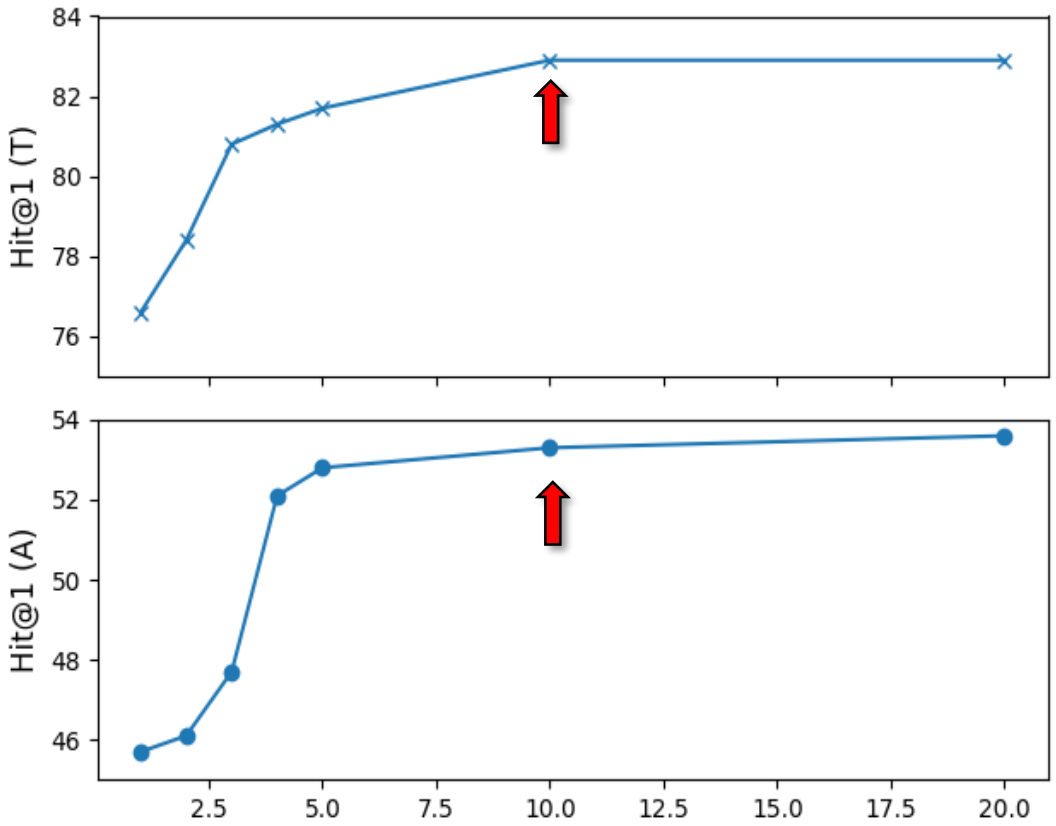
Use a cluster of contextualized embeddings to represent labels and use constraints to regularize the predictions by modeling it as an ILP problem.

The Proposed Framework



How many anchor sentences do we need?

	Model	Train types	Test types	Trig Hit@1	Trig Hit@3	Trig Hit@5	Arg Hit@1	Arg Hit@3	Arg Hit@5
Unseen types only	Frequency	0	23	9.6	27.2	42.5	25.9	63.4	80.6
	WSD	0	23	1.7	13.0	22.8	2.4	2.8	2.8
	Transfer-learning (A)	1	23	4.0	23.8	32.5	1.3	3.4	3.6
	Transfer-learning (B)	3	23	7.0	12.5	36.8	3.5	6.0	6.3
	Transfer-learning (C)	5	23	20.1	34.7	46.5	9.6	14.7	15.7
	Transfer-learning (D)	10	23	33.5	51.4	68.3	14.7	26.5	27.7
	Our Approach	0	23	**80.5**	**88.9**	**93.2**	**68.5**	**94.2**	**96.8**
Entire dataset	Frequency	0	33	28.9	53.6	62.7	13.8	33.8	51.0
	Our Approach	0	33	**82.9**	**93.1**	**96.2**	**53.6**	**87.9**	**92.4**



Ten sentences are good enough!!

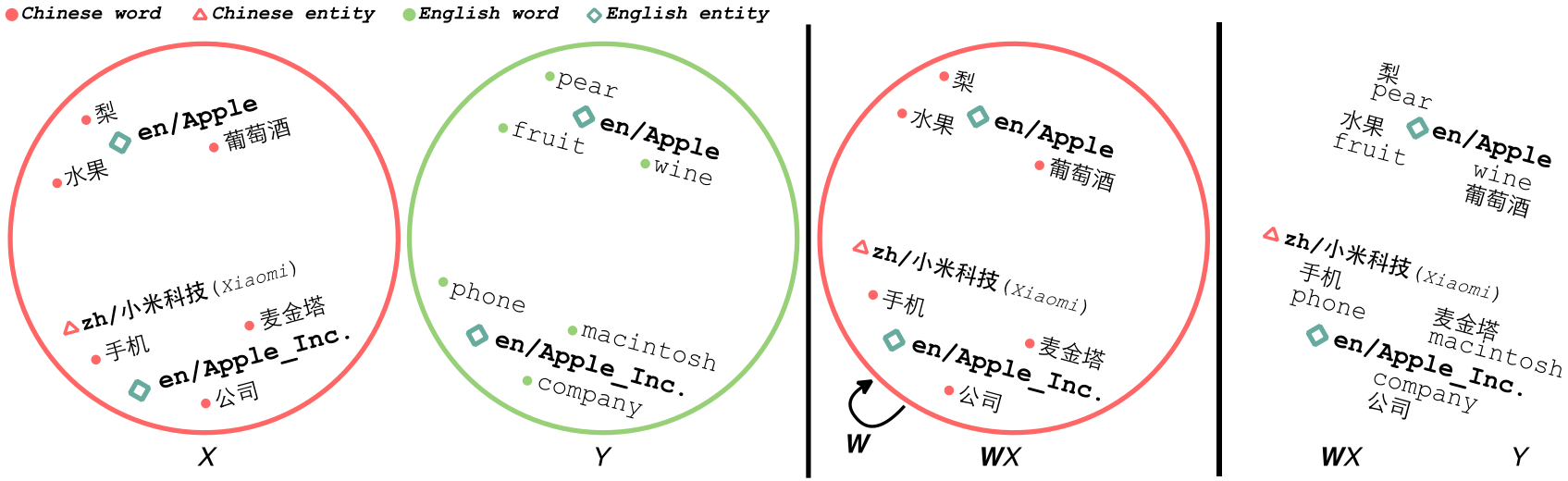
Event Extraction

- Supervised Event Extraction
  - Schema-guided Event Extraction
  - Document-level Event Extraction
- Cross-domain Zero-shot Transfer for Event Extraction
- ➢ Cross-lingual Transfer for Multi-lingual Event Extraction
- Cross-media Structured Common Space for Multimedia Event Extraction

Cross-lingual Joint Entity and Word Embedding Learning

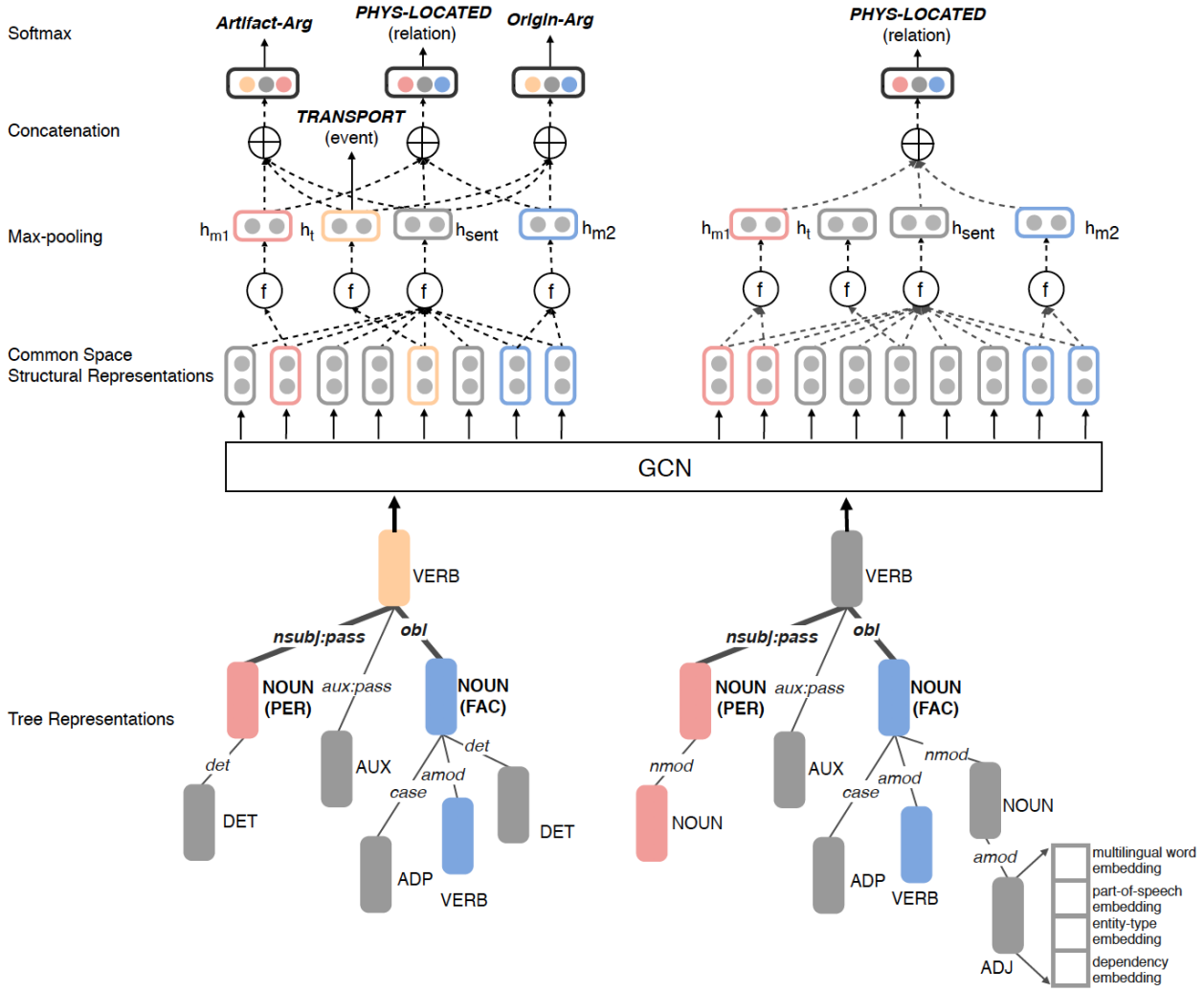
- Cross-lingual Joint Entity and Word Embedding to Improve Entity Linking and Parallel Sentence Mining (Pan et al., 2019)
  - Code-switch cross-lingual entity/word data generation

  - Use English entities as anchor points to learn a mapping (rotation matrix) W which aligns distributions in IL and English



Cross-lingual Structure Transfer Event Extraction

- Cross-lingual Structure Transfer for Relation and Event Extraction (Subburathinam et al., 2019)



The detainees were taken to a processing center

Команды врачей были замечены в упакованных отделениях скорой помощи
(teams of doctors were seen in packed emergency rooms)

Graph Convolutional Networks (GCN) Encoder

- Extend the monolingual design (Zhang et al., 2018) to cross-lingual
  - Convert a sentence with N tokens into N*N adjacency matrix *A*
  - Node: token, each edge is a directed dependency edge
- Initialization of each node's representation

$$\boldsymbol{h}_i^{(0)} = \boldsymbol{x}_i^w \oplus \boldsymbol{x}_i^p \oplus \boldsymbol{x}_i^d \oplus \boldsymbol{x}_i^e$$

Word embedding POS tag Dependency relation Entity type

- At the kth layer, derive the hidden representation of each node from the representations of its neighbors at previous layer

$$\boldsymbol{h}_i^{(k)} = \mathrm{ReLU}\left(\sum_{j=0}^{N} \frac{\boldsymbol{A}_{ij}\boldsymbol{W}^{(k)}\boldsymbol{h}_j^{(k-1)}}{d_i + b^{(k)}}\right)$$

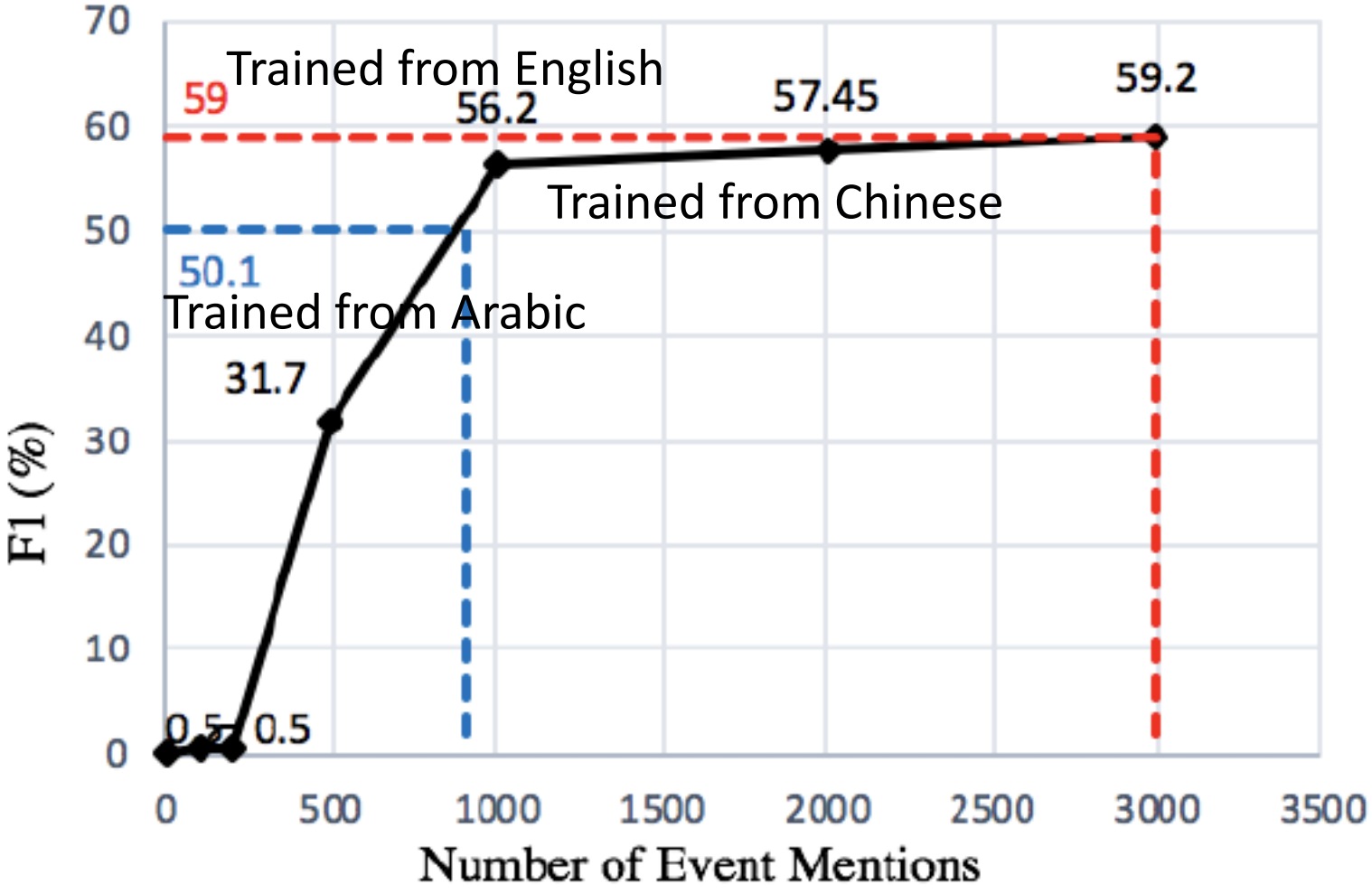
Application on Event Argument Extraction

- Task: Classify each pair of event trigger and entity mentions into one of pre-defined event argument roles or NONE
- Max-pooling over the final node representations to obtain representations for sentence, trigger and argument candidate, and concatenate them
- A softmax output layer for argument role labeling

$$L^a = \sum_{i=1}^{N} \sum_{j=1}^{L_i} y_{ij} \ \log(\sigma(\boldsymbol{U}^a \cdot [\boldsymbol{h}_i^t; \boldsymbol{h}_{ij}^s; \boldsymbol{h}_j^a]))$$

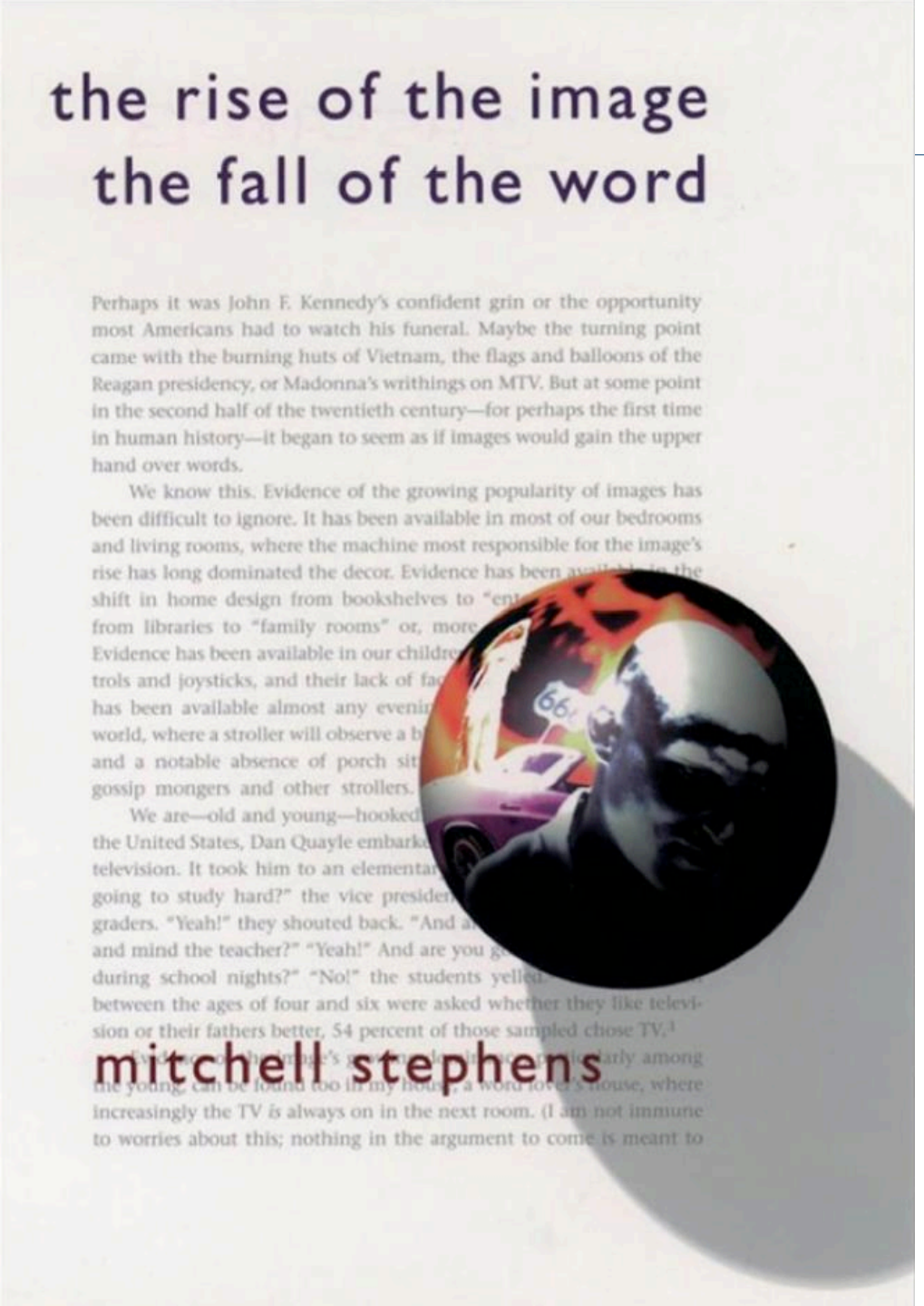
Cross-lingual Event Transfer Performance

- Chinese Event Argument Extraction (Subburathinam et al., EMNLP2019)



Event Extraction

- Supervised Event Extraction
  - Schema-guided Event Extraction
  - Document-level Event Extraction
- Cross-domain Zero-shot Transfer for Event Extraction
- Cross-lingual Transfer for Multi-lingual Event Extraction
- ➢ Cross-media Structured Common Space for Multimedia Event Extraction



Knowledge is Beyond Just Text

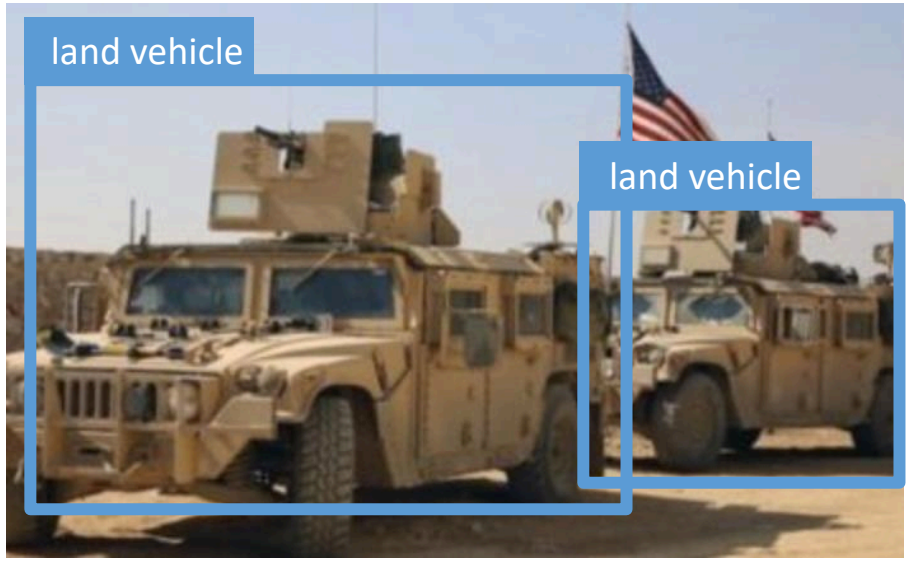
- Multimedia Event Extraction (Li et al., ACL2020)
- We produce and consume news content through multimedia, 33% of news images contain event arguments not mentioned in surrounding texts

***TransportPerson_Instrument* = stretcher**

A New Task: Multimedia Event Extraction (M^2E^2)

Input: News Article Text and Image

Last week , U.S . Secretary of State Rex Tillerson visited Ankara, the first senior administration official to visit Turkey, to try to seal a deal about the battle for Raqqa and to overcome President Recep Tayyip Erdogan's strong objections to Washington's backing of the Kurdish Democratic Union Party (PYD) militias. Turkish forces have attacked SDF forces in the past around Manbij, west of Raqqa, forcing the **United States** to **deploy** dozens of **soldiers** on the **outskirts** of the town in a mission to prevent a repeat of clashes, which risk derailing an assault on Raqqa.



Output: Events & Argument Roles

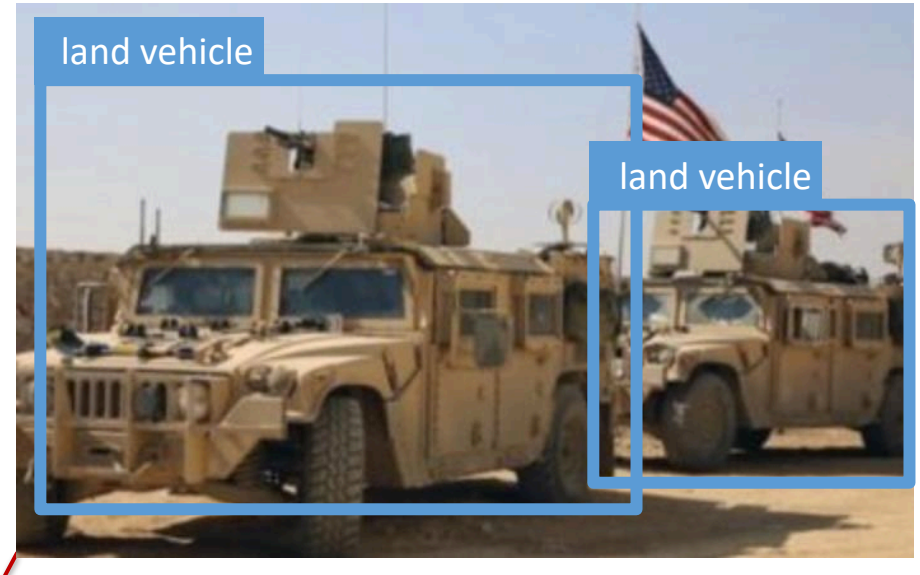
Event Type	Movement.Transport	
Event	**Text Trigger**	deploy
	Image	

Arguments	**Agent**	United States
	Destination	outskirts
	Artifact	soldiers
	Vehicle	
	Vehicle	

A New Task: Multimedia Event Extraction (M^2E^2)

Input: News Article Text and Image

Last week , U.S . Secretary of State Rex Tillerson visited Ankara, the first senior administration official to visit Turkey, to try to seal a deal about the battle for Raqqa and to overcome President Recep Tayyip Erdogan's strong objections to Washington's backing of the Kurdish Democratic Union Party (PYD) militias. Turkish forces have attacked SDF forces in the past around Manbij, west of Raqqa, forcing the **United States** to **deploy** dozens of **soldiers** on the **outskirts** of the town in a mission to prevent a repeat of clashes, which risk derailing an assault on Raqqa.



Output: Multimedia Events & Argument Roles

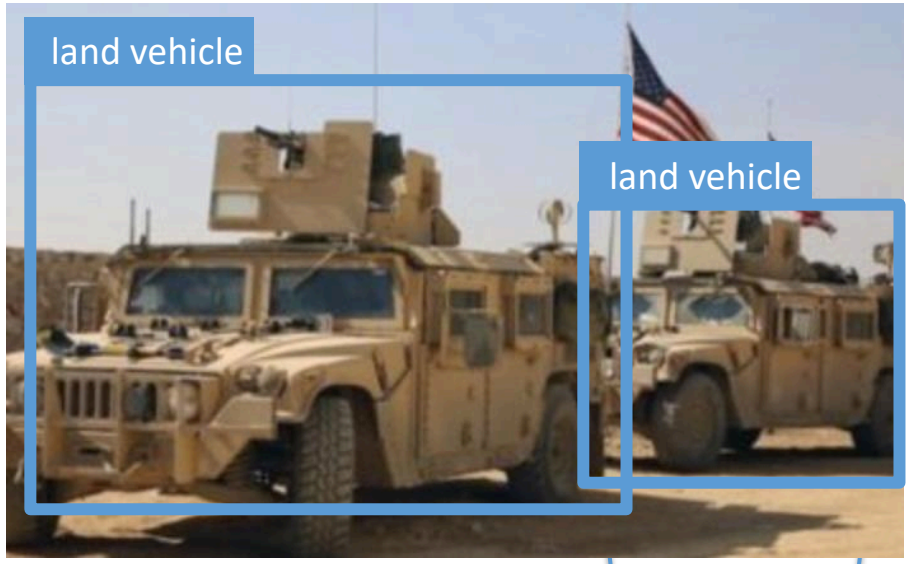
Event Type	Movement.Transport	
Event	**Text Trigger**	deploy
	Image	

Arguments	**Agent**	United States
	Destination	outskirts
	Artifact	soldiers
	Vehicle	
	Vehicle	

A New Task: Multimedia Event Extraction (M^2E^2)

Input: News Article Text and Image

Last week , U.S . Secretary of State Rex Tillerson visited Ankara, the first senior administration official to visit Turkey, to try to seal a deal about the battle for Raqqa and to overcome President Recep Tayyip Erdogan's strong objections to Washington's backing of the Kurdish Democratic Union Party (PYD) militias. Turkish forces have attacked SDF forces in the past around Manbij, west of Raqqa, forcing the **United States** to **deploy** dozens of **soldiers** on the **outskirts** of the town in a mission to prevent a repeat of clashes, which risk derailing an assault on Raqqa.



Output: Multimedia Events & Argument Roles

Event Type	Movement.Transport	
Event	**Text Trigger**	deploy
	Image	

Arguments	**Agent**	United States
	Destination	outskirts
	Artifact	soldiers
	Vehicle	
	Vehicle	

Vision vs. NLP for Event Extraction

- Vision does not study newsworthy, complex events
  - Focusing on daily life and sports (Perera et al., 2012; Chang et al., 2016; Zhang et al., 2007; Ma et al., 2017)
  - Without localizing a complete set of arguments for each event (Gu et al., 2018; Li et al., 2018; Duarte et al., 2018; Sigurdsson et al., 2016; Kato et al., 2018; Wu et al., 2019a)
- Most related: Situation Recognition (Yatskar et al., 2016)
  - Classify an image as one of 500+ FrameNet verbs
  - Identify 192 generic semantic roles via a 1-word description

CLIPPING			
ROLE	VALUE	ROLE	VALUE
AGENT	MAN	AGENT	VET
SOURCE	SHEEP	SOURCE	DOG
TOOL	SHEARS	TOOL	CLIPPER
ITEM	WOOL	ITEM	CLAW
PLACE	FIELD	PLACE	ROOM

JUMPING			
ROLE	VALUE	ROLE	VALUE
AGENT	BOY	AGENT	BEAR
SOURCE	CLIFF	SOURCE	ICEBERG
OBSTACLE	-	OBSTACLE	WATER
DESTINATION	WATER	DESTINATION	ICEBERG
PLACE	LAKE	PLACE	OUTDOOR

SPRAYING			
ROLE	VALUE	ROLE	VALUE
AGENT	MAN	AGENT	FIREMAN
SOURCE	SPRAY CAN	SOURCE	HOSE
SUBSTANCE	PAINT	SUBSTANCE	WATER
DESTINATION	WALL	DESTINATION	FIRE
PLACE	ALLEYWAY	PLACE	OUTSIDE

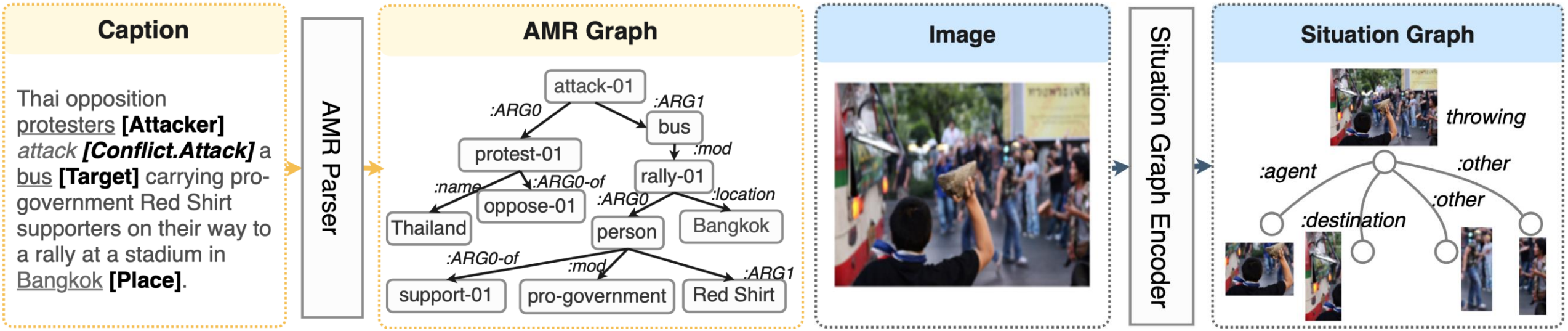
Cross-media Structured Common Space

- Treat Image/Video as a foreign language

Text	Image / Video Frame
Word	Image Region
Entity	Visual Object
Relation	Visual Relation
Entity-Relation Graph	Visual Scene Graph
Event Trigger	Visual Activity
Linguistic Structure	**Situation Graph**

Cross-media Structured Common Space

- Treat Image/Video as a foreign language
  - Represent it with a structure that is similar to AMR graph in text

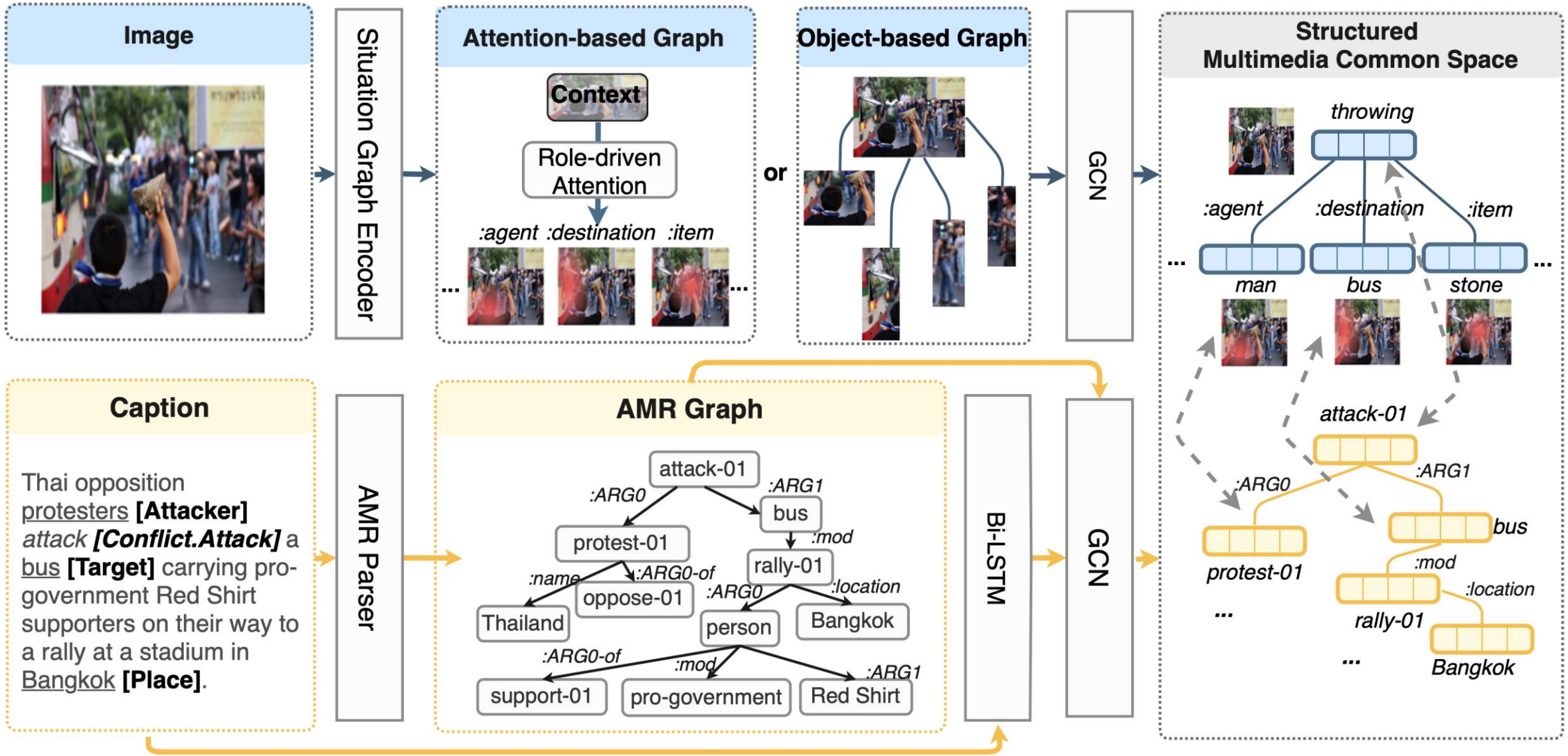


Linguistic Structure,
e.g., Dependency Tree
Abstract Meaning Representation (AMR)

Situation Graph

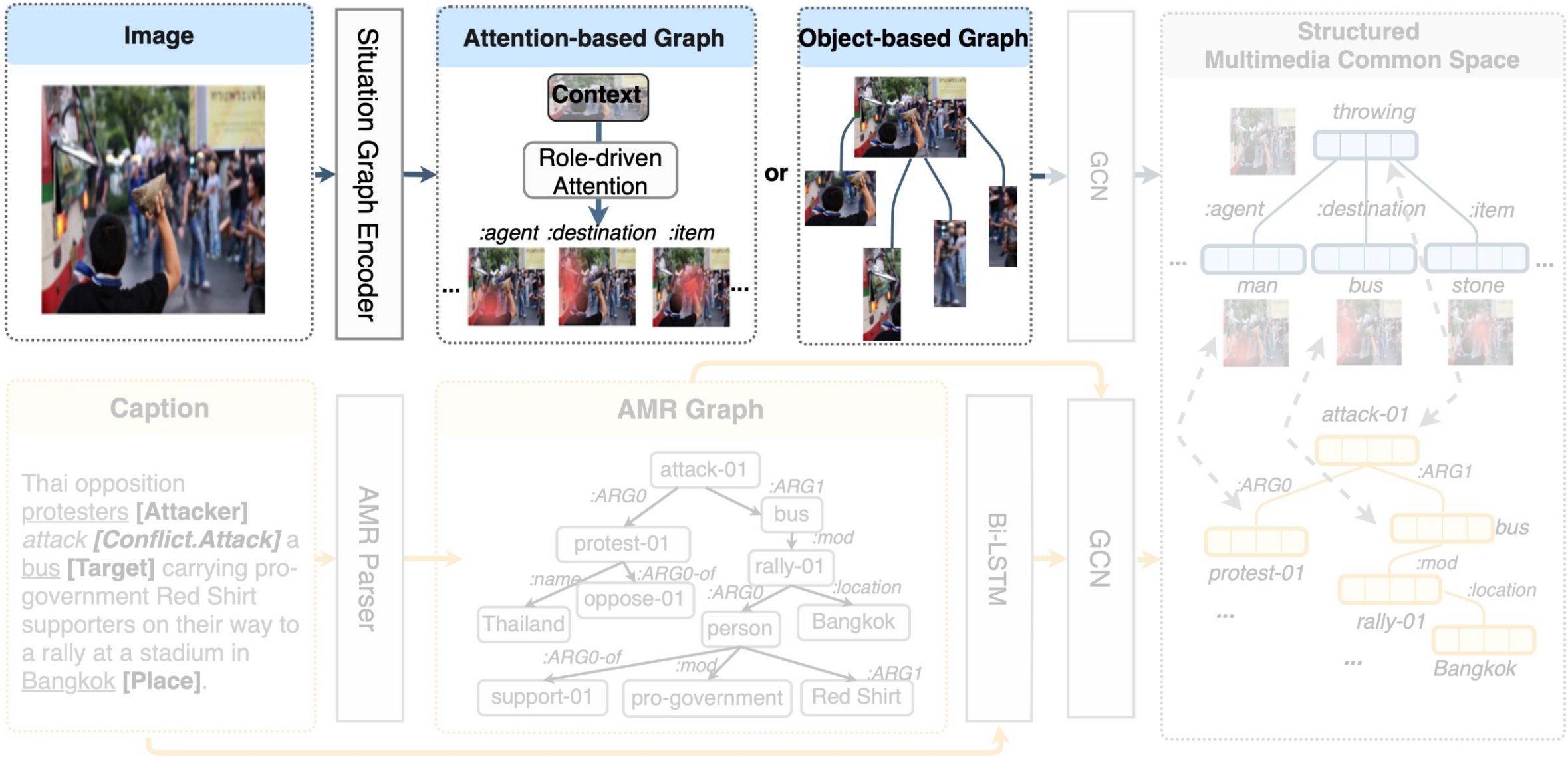
Weakly Aligned Structured Embedding (WASE)

-- Training Phase (Common Space Construction)



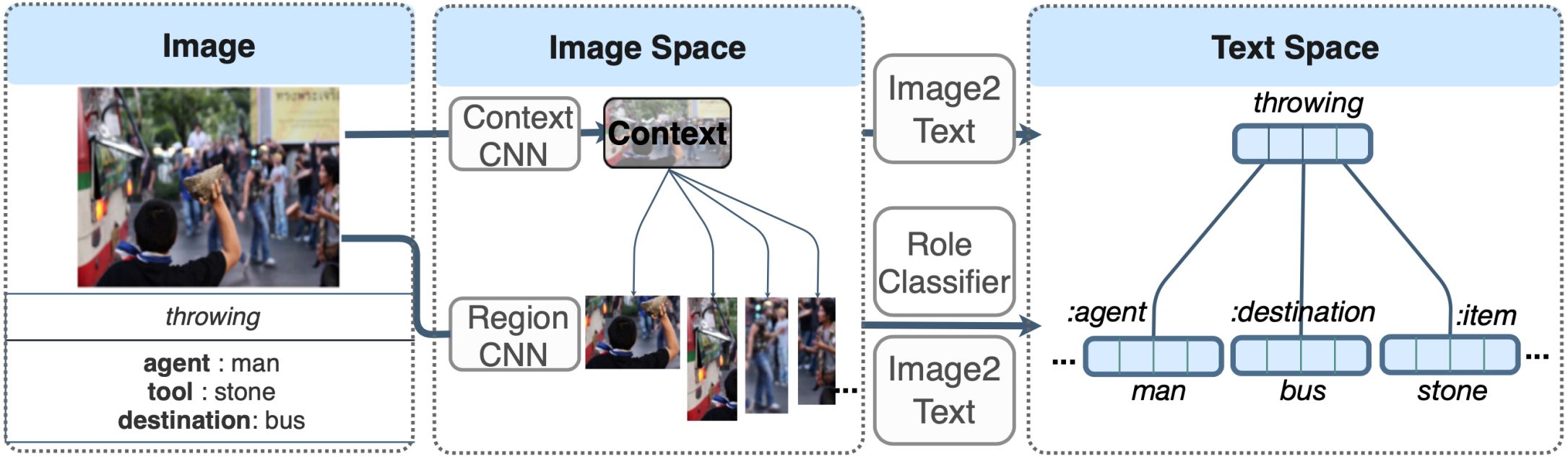
Weakly Aligned Structured Embedding (WASE)

-- Training Phase (Common Space Construction)



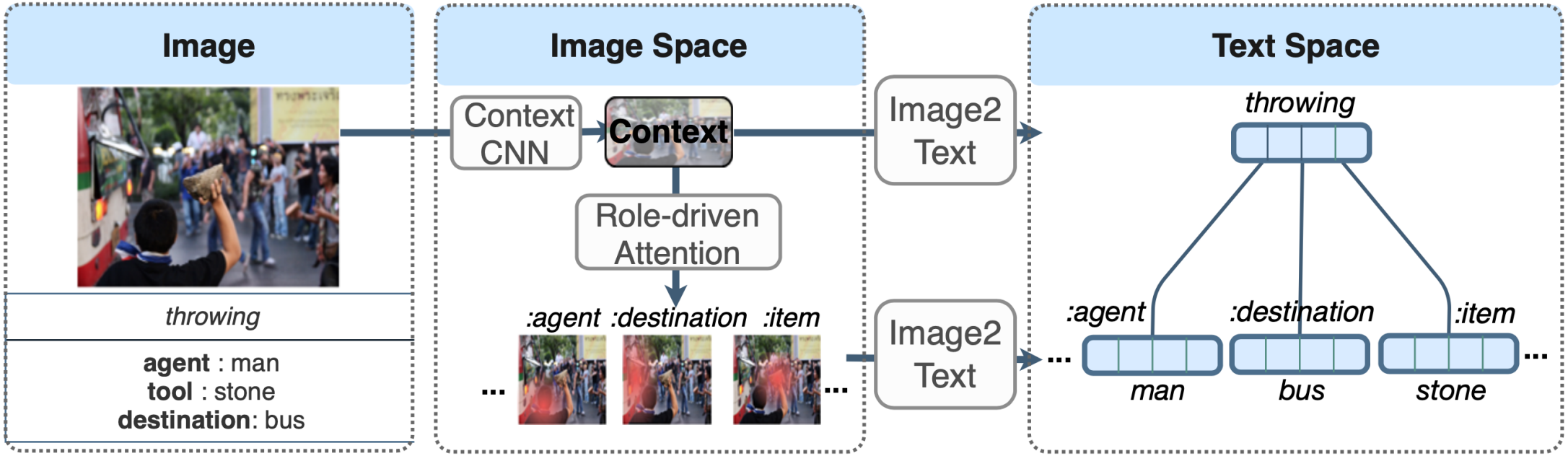
How to generate situation graph?

- Method 1: Object-based Graph Training
  - Learn to project image to verb embedding, and object to noun
  - Learn to classify each object-image pair to a semantic role



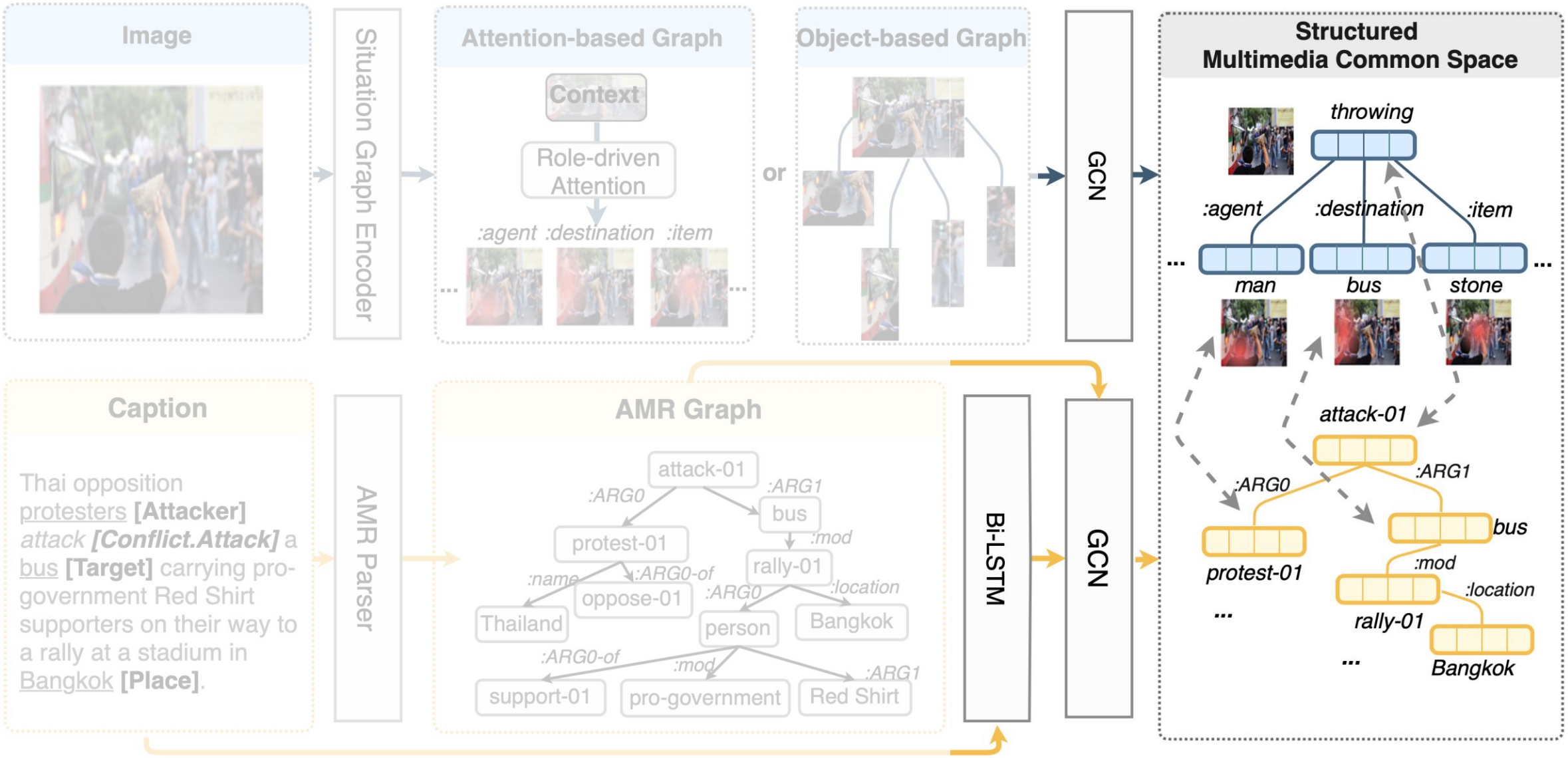
How to generate situation graph?

- Method 2: Role-driven Attention Graph
  - Learn to project image embedding to verb embedding
  - Learn a spatial attention on image for each role
  - Learn to project attended role region to noun embedding



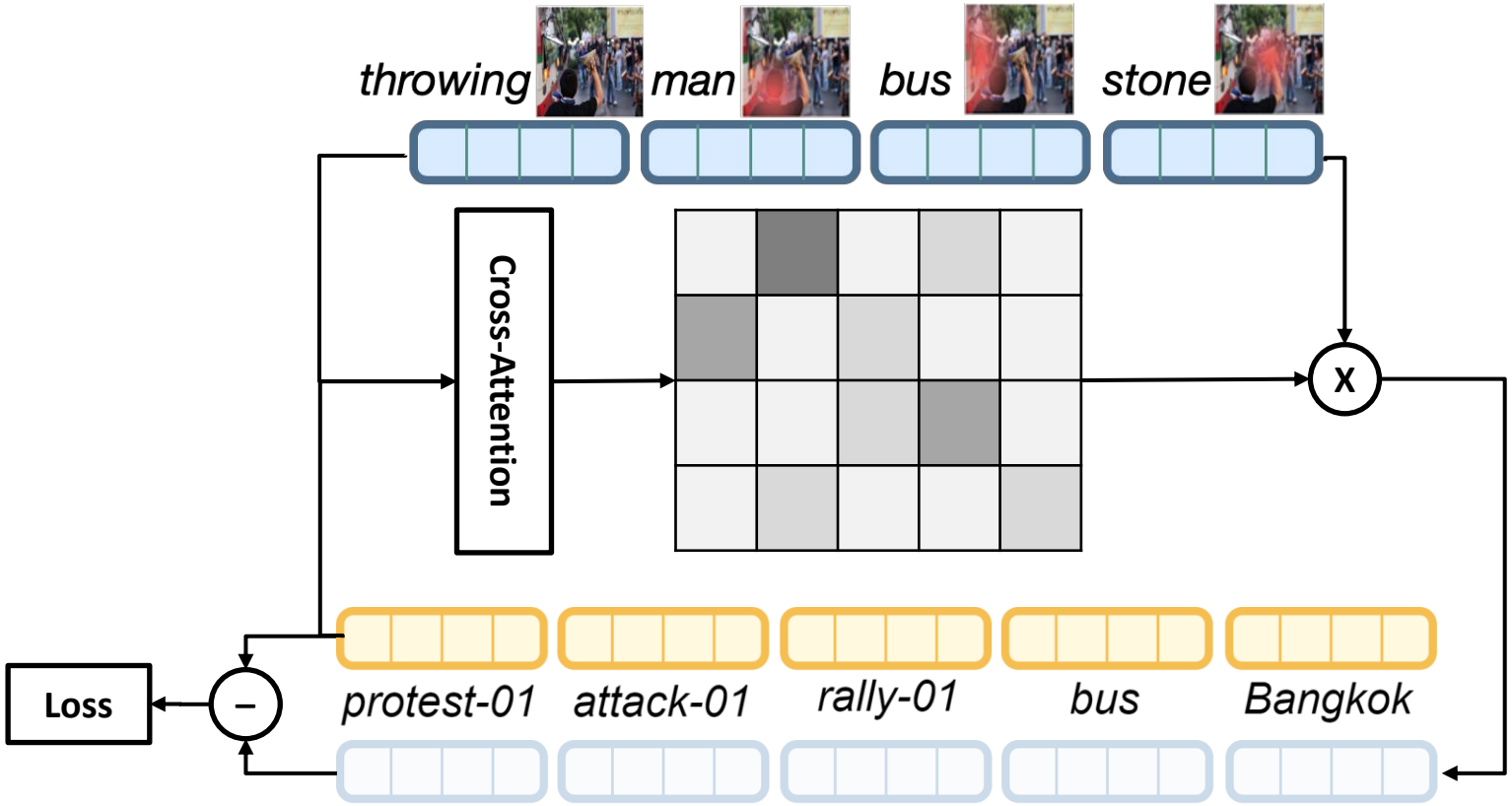
Weakly Aligned Structured Embedding (WASE)

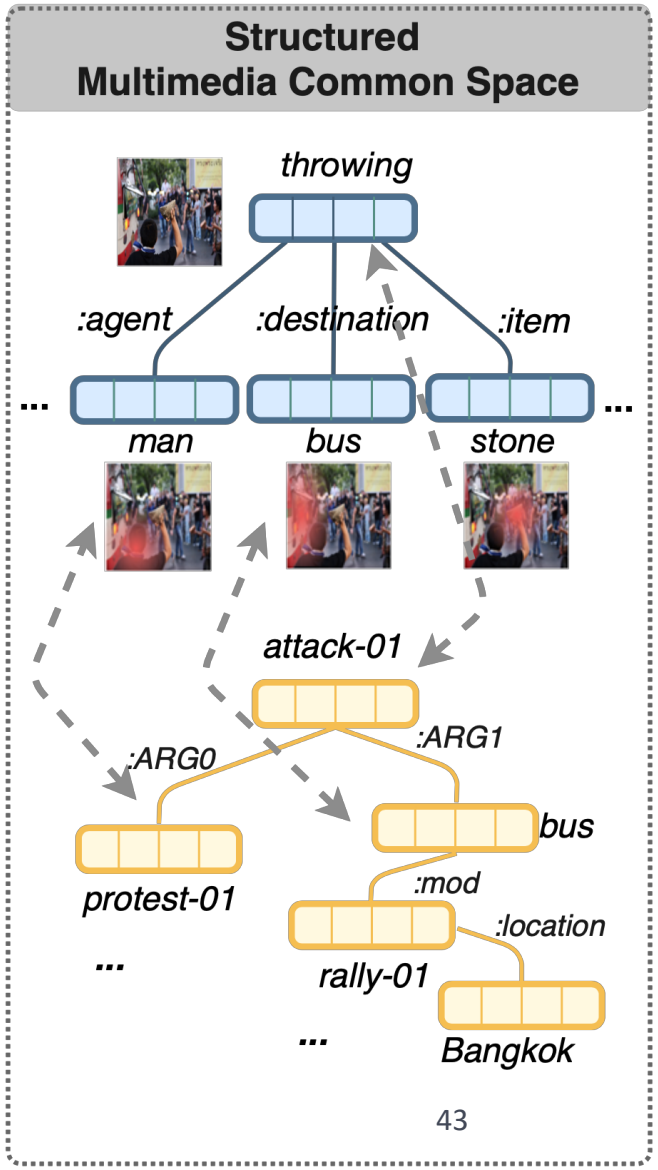
-- Training Phase (Common Space Construction)



How to align the two modalities?

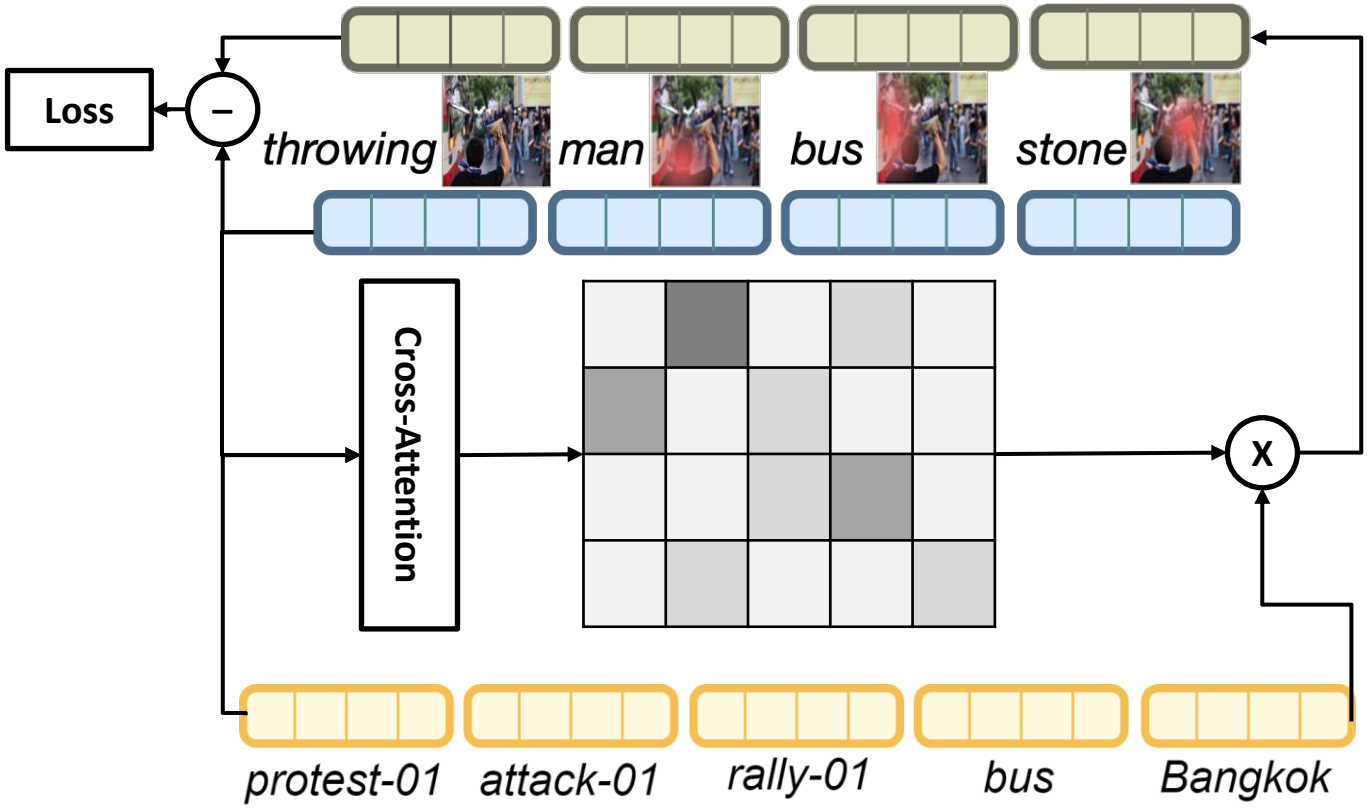
- Prior work aligns image-caption vectors by triplet loss.
- We want to align two graphs, not just single vectors.
- Ontology is shared so the nodes carry similar semantics.

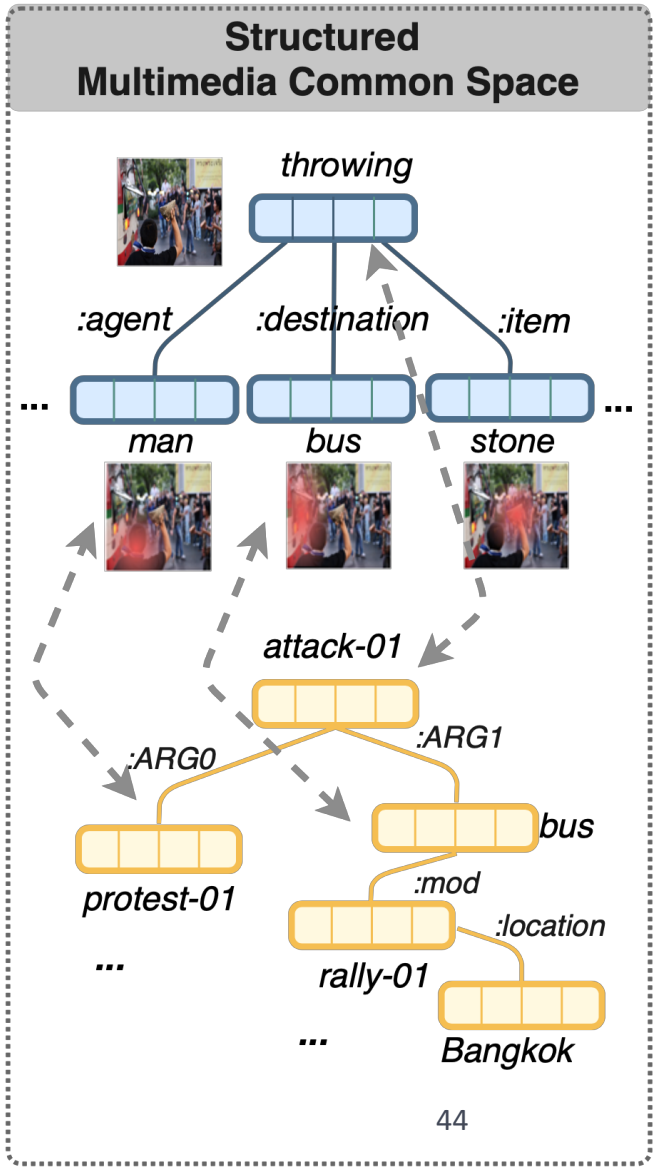




How to align the two modalities?

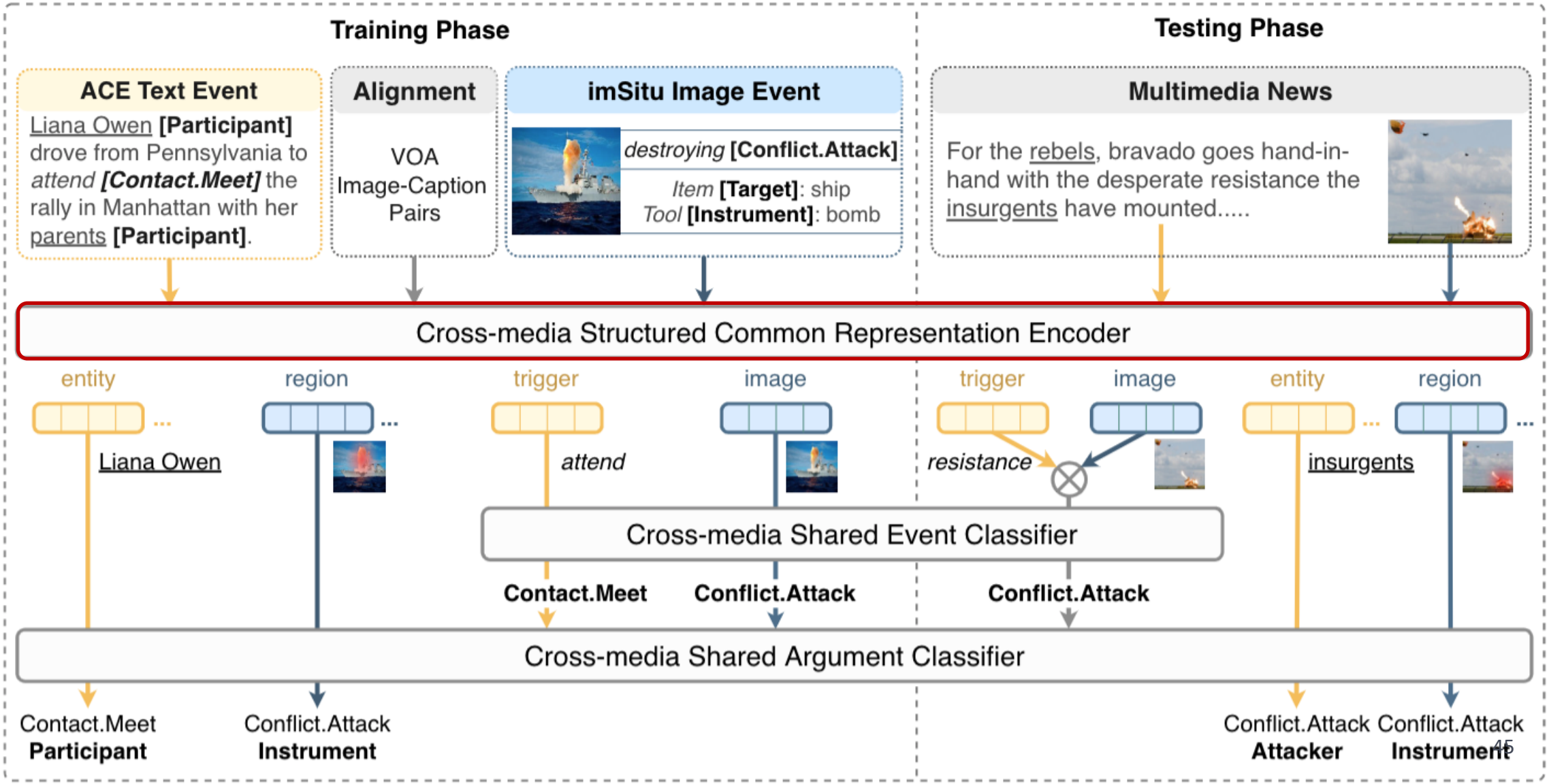
- Prior work aligns image-caption vectors by triplet loss.
- We want to align two graphs, not just single vectors.
- Ontology is shared so the nodes carry similar semantics.





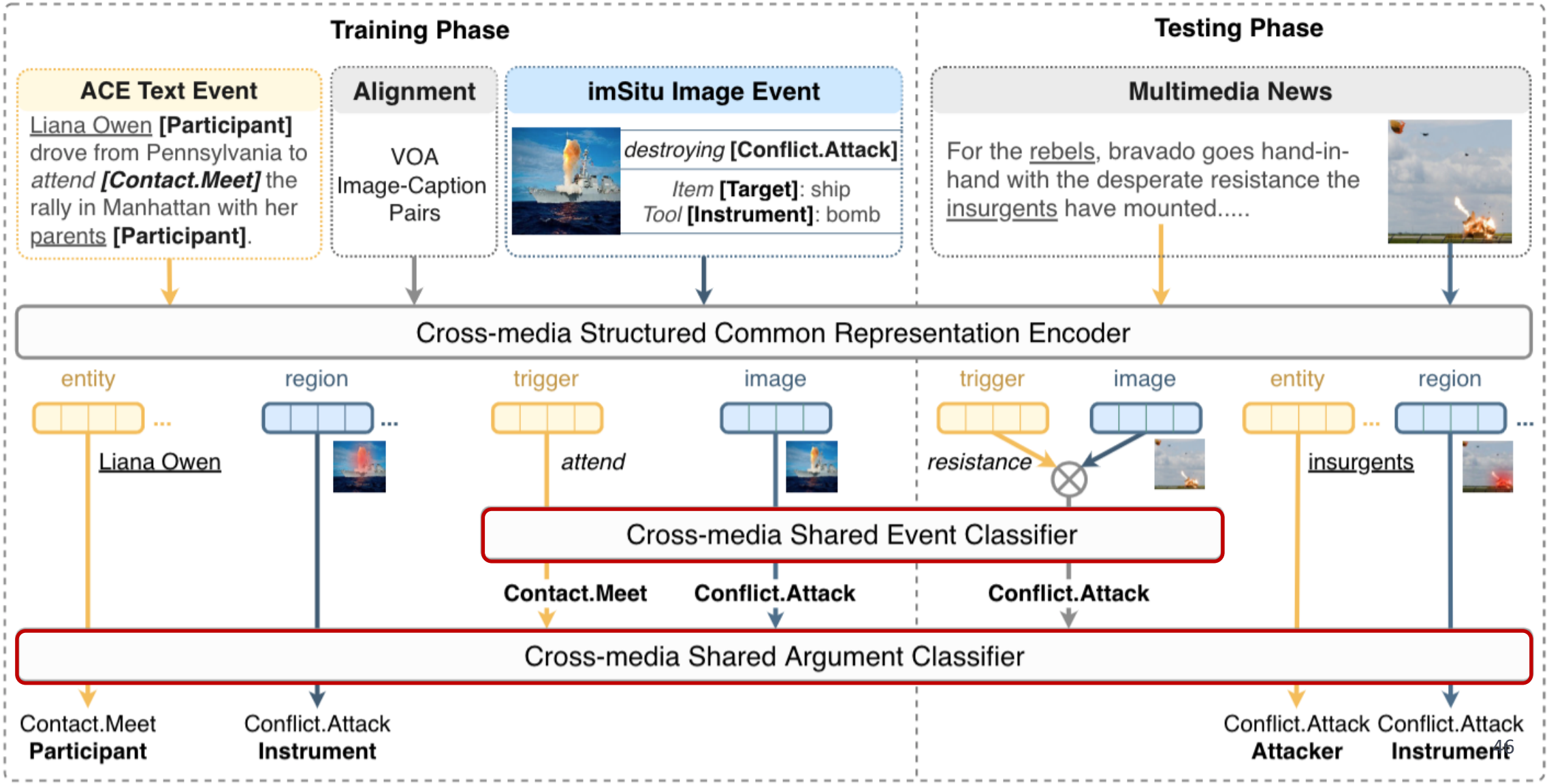
Weakly Aligned Structured Embedding (WASE)

-- Training and Testing Phase (Cross-media shared classifiers)



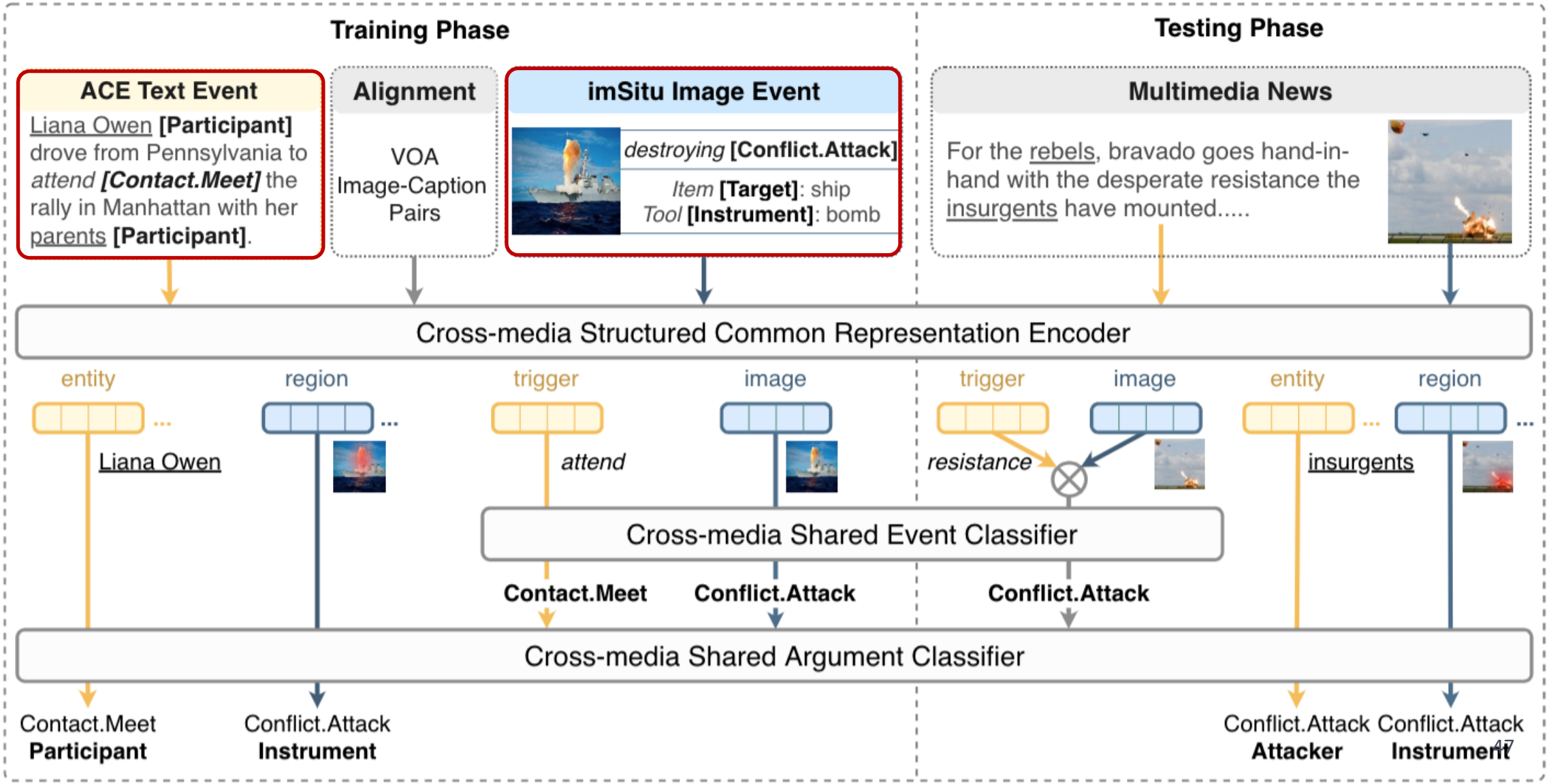
Weakly Aligned Structured Embedding (WASE)

-- Training and Testing Phase (Cross-media shared classifiers)



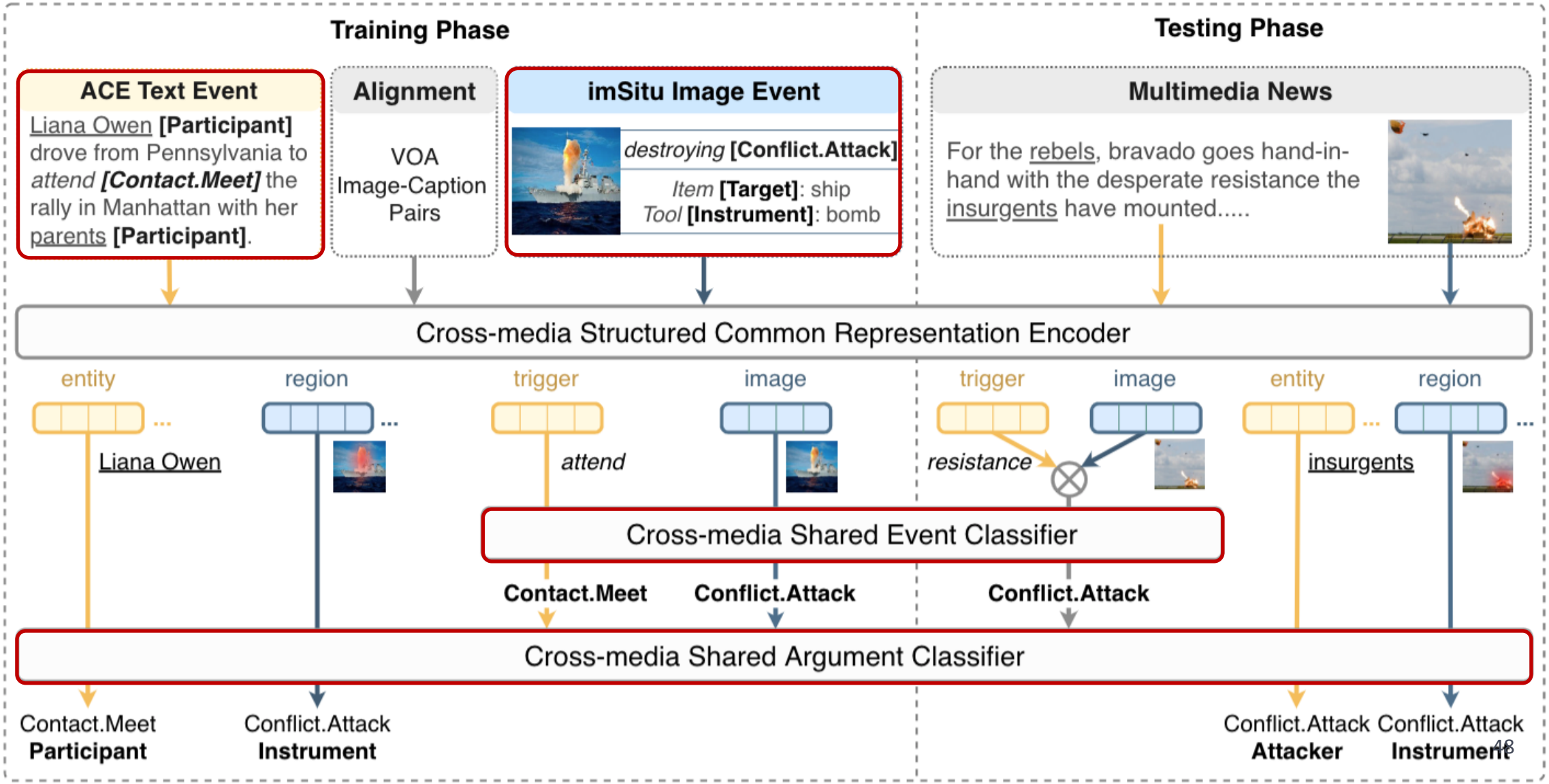
Weakly Aligned Structured Embedding (WASE)

-- Training and Testing Phase (Cross-media shared classifiers)



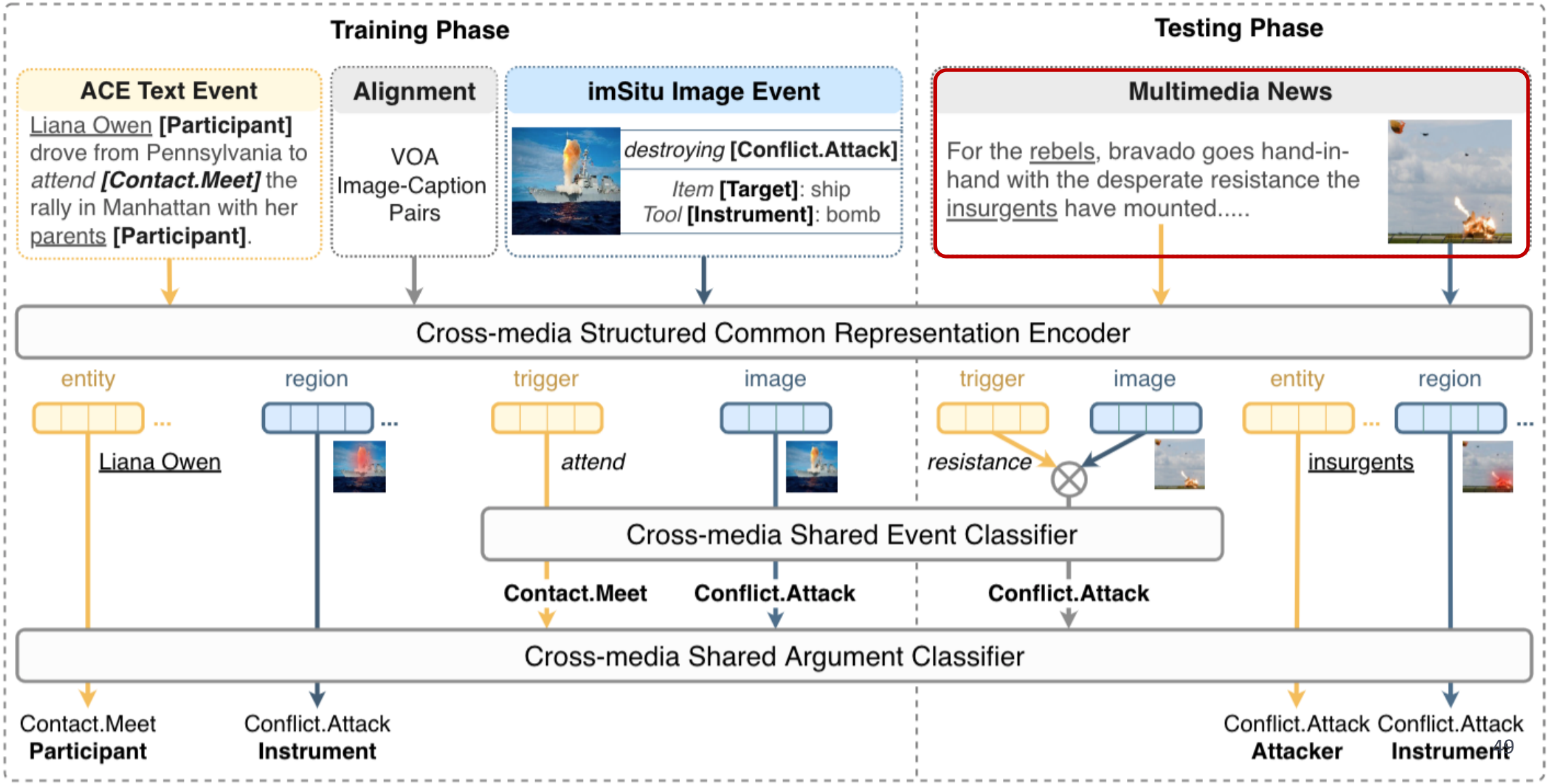
Weakly Aligned Structured Embedding (WASE)

-- Training and Testing Phase (Cross-media shared classifiers)



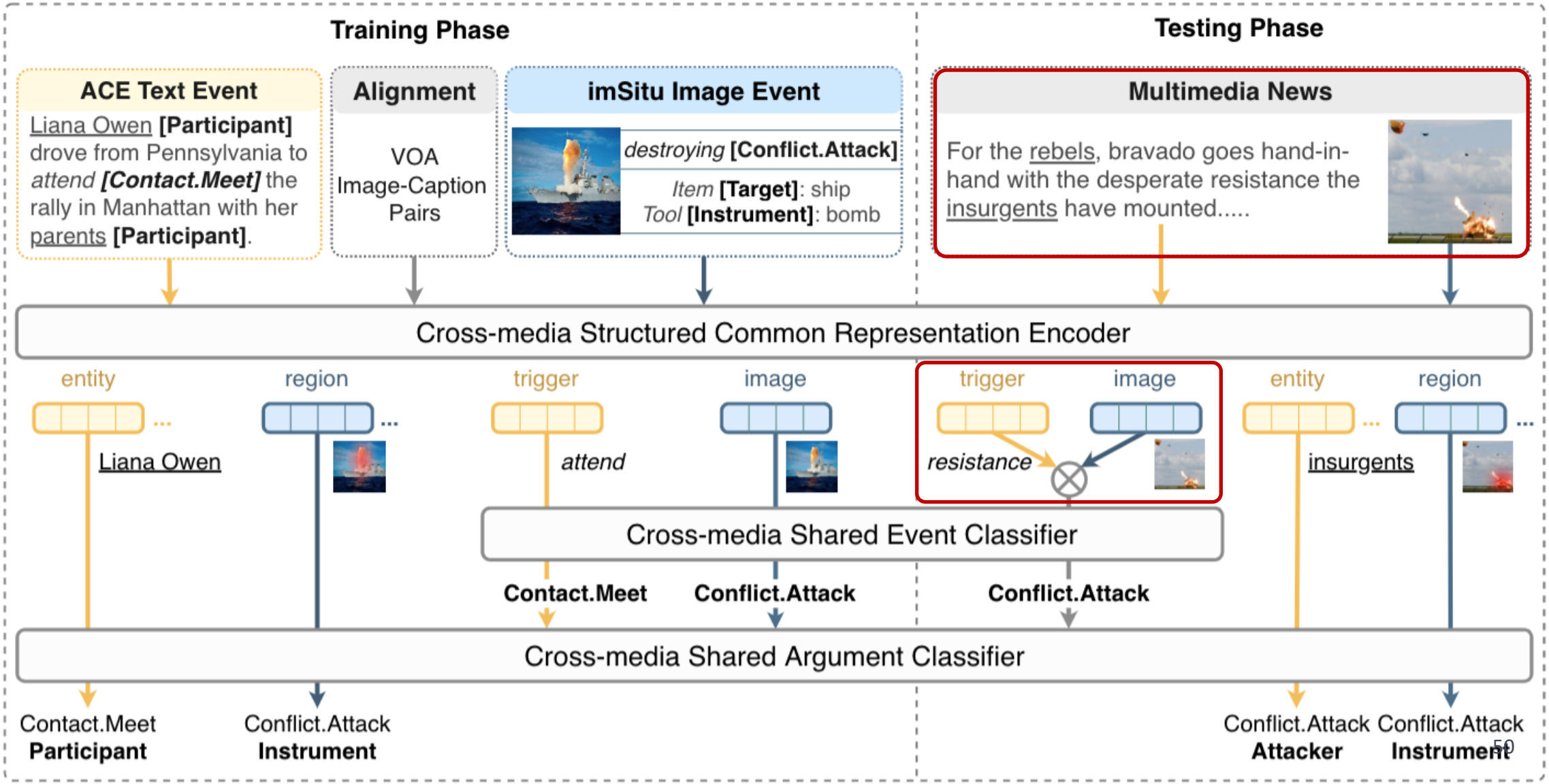
Weakly Aligned Structured Embedding (WASE)

-- Training and Testing Phase (Cross-media shared classifiers)



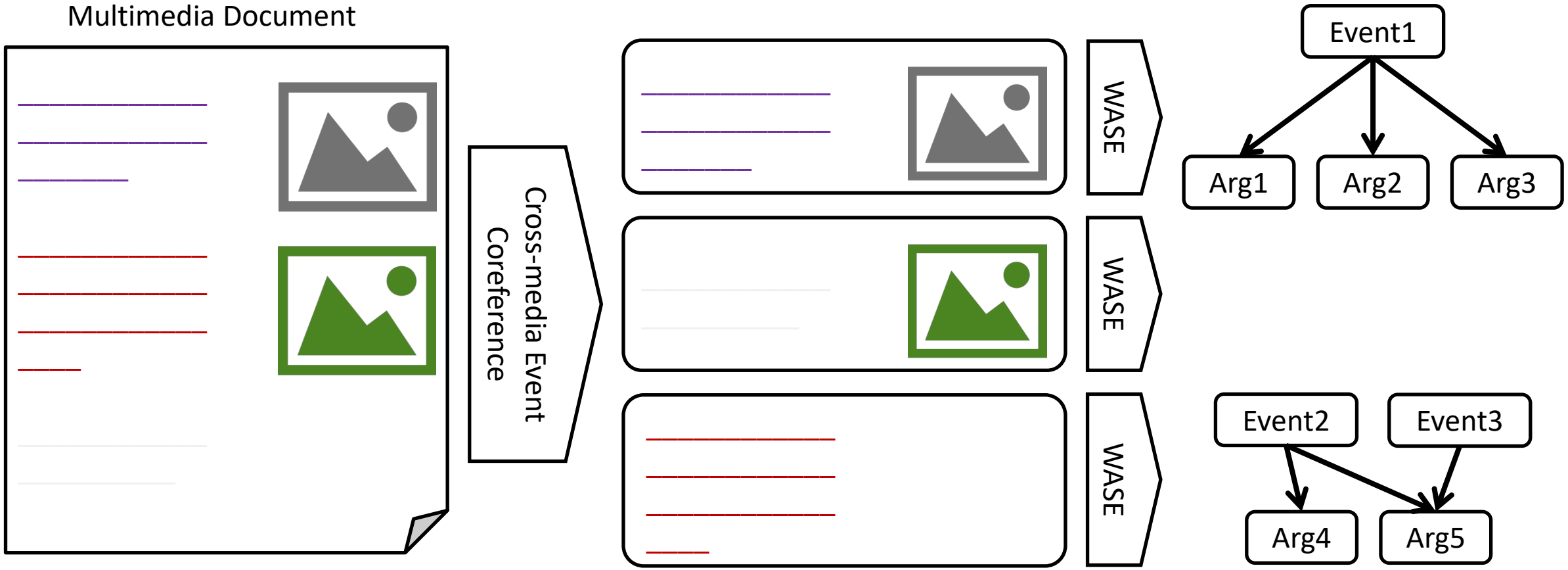
Weakly Aligned Structured Embedding (WASE)

-- Training and Testing Phase (Cross-media shared classifiers)



Weakly Aligned Structured Embedding (WASE)

-- System Diagram



Experiment Results

Training	Model	Text-Only Evaluation						Image-Only Evaluation						Multimedia Evaluation					
		Event Mention			Argument Role			Event Mention			Argument Role			Event Mention			Argument Role		
		P	R	F_1	P	R	F_1	P	R	F_1	P	R	F_1	P	R	F_1	P	R	F_1
Text	JMEE	42.5	58.2	48.7	22.9	28.3	25.3	-	-	-	-	-	-	42.1	34.6	38.1	21.1	12.6	15.8
	GAIL	43.4	53.5	47.9	23.6	29.2	26.1	-	-	-	-	-	-	44.0	32.4	37.3	22.7	12.8	16.4
	$\text{WASE}^{\mathbb{T}}$	42.3	58.4	48.2	21.4	30.1	24.9	-	-	-	-	-	-	41.2	33.1	36.7	20.1	13.0	15.7
Image	$\text{WASE}^{\mathbb{I}}_{\text{att}}$	-	-	-	-	-	-	29.7	61.9	40.1	9.1	10.2	9.6	28.3	23.0	25.4	2.9	6.1	3.8
	$\text{WASE}^{\mathbb{I}}_{\text{obj}}$	-	-	-	-	-	-	28.6	59.2	38.7	13.3	9.8	11.2	26.1	22.4	24.1	4.7	5.0	4.9
Multimedia	VSE-C	33.5	47.8	39.4	16.6	24.7	19.8	30.3	48.9	26.4	5.6	6.1	5.7	33.3	48.2	39.3	11.1	14.9	12.8
	Flat_{att}	34.2	63.2	44.4	20.1	27.1	23.1	27.1	57.3	36.7	4.3	8.9	5.8	33.9	59.8	42.2	12.9	17.6	14.9
	Flat_{obj}	38.3	57.9	46.1	21.8	26.6	24.0	26.4	55.8	35.8	9.1	6.5	7.6	34.1	56.4	42.5	16.3	15.9	16.1
	WASE_{att}	37.6	66.8	48.1	27.5	33.2	**30.1**	32.3	63.4	42.8	9.7	11.1	10.3	38.2	67.1	49.1	18.6	21.6	**19.9**
	WASE_{obj}	42.8	61.9	**50.6**	23.5	30.3	26.4	43.1	59.2	**49.9**	14.5	10.1	**11.9**	43.0	62.1	**50.8**	19.5	18.9	19.2

Experiment Results

Training	Model	Text-Only Evaluation						Image-Only Evaluation						Multimedia Evaluation					
		Event Mention			Argument Role			Event Mention			Argument Role			Event Mention			Argument Role		
		P	R	F_1	P	R	F_1	P	R	F_1	P	R	F_1	P	R	F_1	P	R	F_1
Text	JMEE	42.5	58.2	48.7	22.9	28.3	25.3	-	-	-	-	-	-	42.1	34.6	38.1	21.1	12.6	15.8
	GAIL	43.4	53.5	47.9	23.6	29.2	26.1	-	-	-	-	-	-	44.0	32.4	37.3	22.7	12.8	16.4
	$\text{WASE}^{\mathbb{T}}$	42.3	58.4	48.2	21.4	30.1	24.9	-	-	-	-	-	-	41.2	33.1	36.7	20.1	13.0	15.7
Image	$\text{WASE}^{\mathbb{I}}_{\text{att}}$	-	-	-	-	-	-	29.7	61.9	40.1	9.1	10.2	9.6	28.3	23.0	25.4	2.9	6.1	3.8
	$\text{WASE}^{\mathbb{I}}_{\text{obj}}$	-	-	-	-	-	-	28.6	59.2	38.7	13.3	9.8	11.2	26.1	22.4	24.1	4.7	5.0	4.9
Multimedia	VSE-C	33.5	47.8	39.4	16.6	24.7	19.8	30.3	48.9	26.4	5.6	6.1	5.7	33.3	48.2	39.3	11.1	14.9	12.8
	Flat_{att}	34.2	63.2	44.4	20.1	27.1	23.1	27.1	57.3	36.7	4.3	8.9	5.8	33.9	59.8	42.2	12.9	17.6	14.9
	Flat_{obj}	38.3	57.9	46.1	21.8	26.6	24.0	26.4	55.8	35.8	9.1	6.5	7.6	34.1	56.4	42.5	16.3	15.9	16.1
	WASE_{att}	37.6	66.8	48.1	27.5	33.2	**30.1**	32.3	63.4	42.8	9.7	11.1	10.3	38.2	67.1	49.1	18.6	21.6	**19.9**
	WASE_{obj}	42.8	61.9	**50.6**	23.5	30.3	26.4	43.1	59.2	**49.9**	14.5	10.1	**11.9**	43.0	62.1	**50.8**	19.5	18.9	19.2

Experiment Results

Training	Model	Text-Only Evaluation						Image-Only Evaluation						Multimedia Evaluation					
		Event Mention			Argument Role			Event Mention			Argument Role			Event Mention			Argument Role		
		P	R	F_1	P	R	F_1	P	R	F_1	P	R	F_1	P	R	F_1	P	R	F_1
Text	JMEE	42.5	58.2	48.7	22.9	28.3	25.3	-	-	-	-	-	-	42.1	34.6	38.1	21.1	12.6	15.8
	GAIL	43.4	53.5	47.9	23.6	29.2	26.1	-	-	-	-	-	-	44.0	32.4	37.3	22.7	12.8	16.4
	$\text{WASE}^{\mathbb{T}}$	42.3	58.4	48.2	21.4	30.1	24.9	-	-	-	-	-	-	41.2	33.1	36.7	20.1	13.0	15.7
Image	$\text{WASE}^{\mathbb{I}}_{\text{att}}$	-	-	-	-	-	-	29.7	61.9	40.1	9.1	10.2	9.6	28.3	23.0	25.4	2.9	6.1	3.8
	$\text{WASE}^{\mathbb{I}}_{\text{obj}}$	-	-	-	-	-	-	28.6	59.2	38.7	13.3	9.8	11.2	26.1	22.4	24.1	4.7	5.0	4.9
Multimedia	VSE-C	33.5	47.8	39.4	16.6	24.7	19.8	30.3	48.9	26.4	5.6	6.1	5.7	33.3	48.2	39.3	11.1	14.9	12.8
	Flat_{att}	34.2	63.2	44.4	20.1	27.1	23.1	27.1	57.3	36.7	4.3	8.9	5.8	33.9	59.8	42.2	12.9	17.6	14.9
	Flat_{obj}	38.3	57.9	46.1	21.8	26.6	24.0	26.4	55.8	35.8	9.1	6.5	7.6	34.1	56.4	42.5	16.3	15.9	16.1
	WASE_{att}	37.6	66.8	48.1	27.5	33.2	**30.1**	32.3	63.4	42.8	9.7	11.1	10.3	38.2	67.1	49.1	18.6	21.6	**19.9**
	WASE_{obj}	42.8	61.9	**50.6**	23.5	30.3	26.4	43.1	59.2	**49.9**	14.5	10.1	**11.9**	43.0	62.1	**50.8**	19.5	18.9	19.2

Cross-Media Coreference Accuracy

Model	P (%)	R (%)	F_1 (%)
rule_based	10.1	100	18.2
VSE	31.2	74.5	44.0
Flat_{att}	33.1	73.5	45.6
Flat_{obj}	34.3	76.4	47.3
WASE_{att}	39.5	73.5	51.5
WASE_{obj}	40.1	75.4	52.4

Compare to Single Data Modality Extraction

- Surrounding sentence helps visual event extraction.

People celebrate Supreme Court ruling on Same Sex Marriage in front of the Supreme Court in Washington.

- Image helps textual event extraction.

Iraqi security forces *search* ***[Justice.Arrest]*** a civilian in the city of Mosul.

Why Does Vision Help NLP?

- Various triggers and context can be coherent in visual space.
- Cross-media Common space pushes scattered sentences towards the visual cluster.

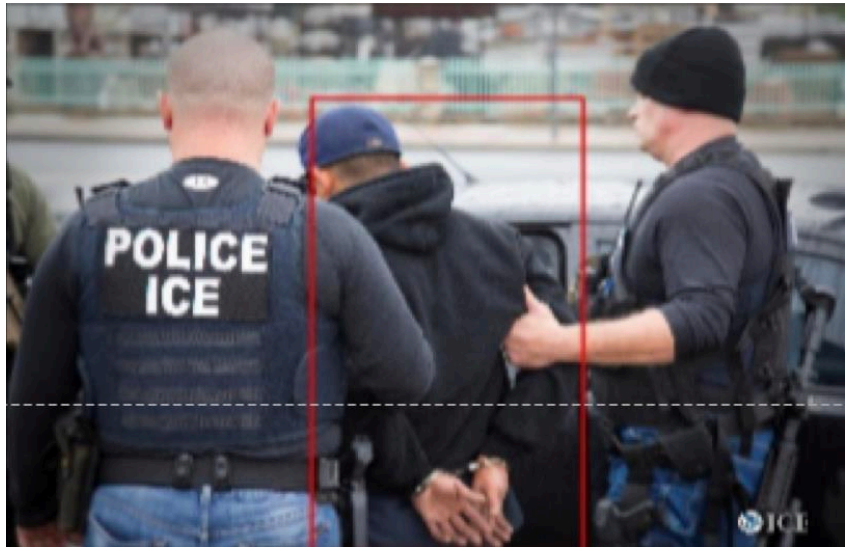
Berlin police tweeted that six people were arrested after a joint operation with the Berlin's prosecutor's office.

The man in Kosovo is an ethnic Albanian arrested south of the capital, Pristina.

He was asleep in a suburban Seattle house last week morning when immigration agents showed up to arrest his father.

But shortly after the round table began, Marko Djuric, head of the Serbian government office on Kosovo, was detained by police.

Compare to Cross-media Flat Representation



Model	Event Type	Argument Role	
Flat	Justice.ArrestJail	Agent =	man
Ours	Justice.ArrestJail	Entity =	man

Model	Event Type	Argument Role	
Flat	Movement.Transport	Artifact =	none
Ours	Movement.Transport	Artifact =	man

Summary of Event Extraction Methods

IE Methods		Supervised Learning	Bootstrapping	Distant Supervision	Open IE/ Zero-shot	Schema/ Discovery
Approach Overview		Learn rules or supervised model from labeled data	Send seeds to extract common patterns from unlabeled data	Project large database entries into unlabeled data to obtain annotations	Open-domain IE based on syntactic patterns	Automatically discover scenarios, event types and templates
Requirement of labeled data		Large unstructured labeled data	Small seeds	Large seeds	Small unstructured labeled data	Little labeled data
Quality	Precision	High	Moderate	Low	Moderate	Moderate
	Recall	High	Difficult to measure	Moderate	Low	Moderate
Portability		Poor	Moderate	Moderate	Good	Good
Scalability		Poor	Moderate	Moderate	Good	Good

Part II: Event-event Relation Extraction

Qiang Ning
Alexa AI, Amazon

Feb 2020

AAAI Tutorials

Recent Advances in Transferable Representation Learning

Events are not isolated…

- …and there are various types of relationships between two events
  - Coreference relations
  - Temporal relations
  - Parent-child relations
  - Causal relations
  - …

AAAI-21 is held virtually due to the pandemic. Its attendees are thus giving remote presentations of their research.

AAAI-21 is [held virtually]$_1$ due to [the pandemic]$_2$. [Its]$_3$ attendees are thus [giving remote presentations]$_4$ of [their research]$_5$.

AAAI-21 is [held virtually]$_1$ due to [the pandemic]$_2$. [Its]$_3$ attendees are thus [giving remote presentations]$_4$ of [their research]$_5$.

[the pandemic]$_2$ CAUSES [held virtually]$_1$

AAAI-21 is [held virtually]$_1$ due to [the pandemic]$_2$. [Its]$_3$ attendees are thus [giving remote presentations]$_4$ of [their research]$_5$.

[held virtually]$_1$ CAUSES [giving remote presentations]$_4$

AAAI-21 is [held virtually]$_1$ due to [the pandemic]$_2$. [Its]$_3$ attendees are thus [giving remote presentations]$_4$ of [their research]$_5$.

[Its]$_3$ REFERS to the conference being [held virtually]$_1$

AAAI-21 is [held virtually]$_1$ due to [the pandemic]$_2$. [Its]$_3$ attendees are thus [giving remote presentations]$_4$ of [their research]$_5$.

[giving remote presentations]$_4$ is a SUBEVENT of [Its]$_3$ (i.e., AAAI)

AAAI-21 is [held virtually]$_1$ due to [the pandemic]$_2$. [Its]$_3$ attendees are thus [giving remote presentations]$_4$ of [their research]$_5$.

[held virtually]$_1$ HAPPENS DURING [the pandemic]$_2$

AAAI-21 is [held virtually]$_1$ due to [the pandemic]$_2$. [Its]$_3$ attendees are thus [giving remote presentations]$_4$ of [their research]$_5$.

[their research]$_5$ HAPPENS BEFORE [giving remote presentations]$_4$

Events are not isolated…

- …and there are various types of relationships between two events
  - Coreference relations
  - Temporal relations
  - Parent-child relations
  - Causal relations
  - …

[Its]$_3$ REFERS to the conference being [held virtually]$_1$
[held virtually]$_1$ HAPPENS DURING [the pandemic]$_2$
[their research]$_5$ HAPPENS BEFORE [giving remote presentations]$_4$
[giving remote presentations]$_4$ is a SUBEVENT of [Its]$_3$ (i.e., AAAI)
[the pandemic]$_2$ CAUSES [held virtually]$_1$
[held virtually]$_1$ CAUSES [giving remote presentations]$_4$
…

- These event-event relationships are important for understanding stories.
  - We can tell a different story with the same set of events but with different relationships (see example next).

[held virtually]$_1$ CAUSES [giving remote presentations]$_4$

[their research]$_5$ HAPPENS BEFORE [giving remote presentations]$_4$

AAAI-21 is [held virtually]$_1$ due to [the pandemic]$_2$. [Its]$_3$ attendees are thus [giving remote presentations]$_4$ of [their research]$_5$.

[held virtually]$_1$ CAUSES [giving remote presentations]$_4$

[their research]$_5$ HAPPENS BEFORE [giving remote presentations]$_4$

AAAI-21 is [held virtually]$_1$ due to [the pandemic]$_2$. [Its]$_3$ attendees are thus [giving remote presentations]$_4$ of [their research]$_5$.

[giving remote presentations]$_4$ CAUSES [held virtually]$_1$

[their research]$_5$ HAPPENS DURING [the pandemic]$_2$

AAAI-21 is [held virtually]$_1$ because it has received many requests to [give remote presentations]$_4$. Many have also reported unexpected delays in [their research]$_5$ during [the pandemic]$_2$.

General Problem Statement

- Given
  - a piece of text
  - the head phrases of two events
- Extract the relationship(s) between this event pair
  - most works focus on one type of relationship, e.g., only predicting coreference relations, or only predicting temporal relations.
  - some also attempts to predict multiple types at the same time.
- Evaluated by
  - precision and recall on all relations

General Problem Statement (cont'd)

- Given
  - a piece of text (often long enough to contain multiple events)
  - the head phrases of ~~two~~ many events
- Extract the relationship(s) between ~~this~~ all event pairs
  - most works focus on one type of relationship, e.g., only predicting coreference relations, or only predicting temporal relations.
  - some also attempts to predict multiple types at the same time.
  - people start to consider multiple events and their relations jointly
- Evaluated by
  - precision and recall on all relations
  - metrics that consider global coherency (B^3, MUC, temporal awareness, etc.)

Cai & Strube, 2010. UzZaman & Allen, 2011. Moosavi & Strube, 2016.

General Problem Statement (cont'd)

- Given
  - a piece of text (often long enough to contain multiple events)
  - ~~the head phrases of two many events~~
- Extract the relationship(s) between ~~this~~ all event pairs
  - most works focus on one type of relationship, e.g., only predicting coreference relations, or only predicting temporal relations.
  - some also attempts to predict multiple types at the same time.
  - people start to consider multiple events and their relations jointly
- Evaluated by
  - precision and recall on all relations
  - metrics that consider global coherency (B^3, MUC, temporal awareness, etc.)

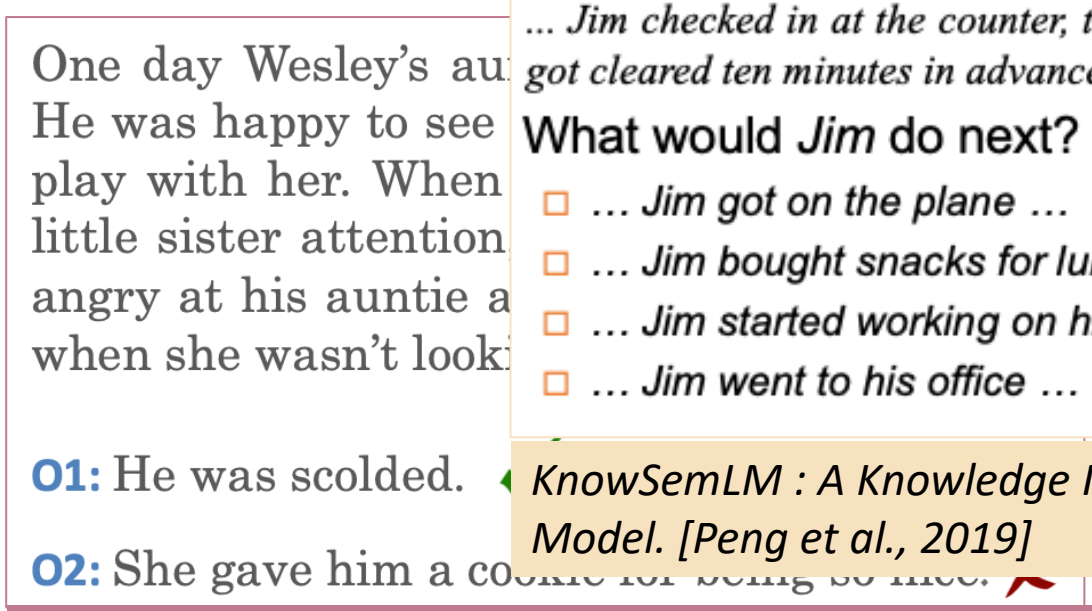
General Problem Statement (cont'd)

- Given
  - a piece of text (often long enough to contain multiple events)
  - ~~the head phrases of two many events~~

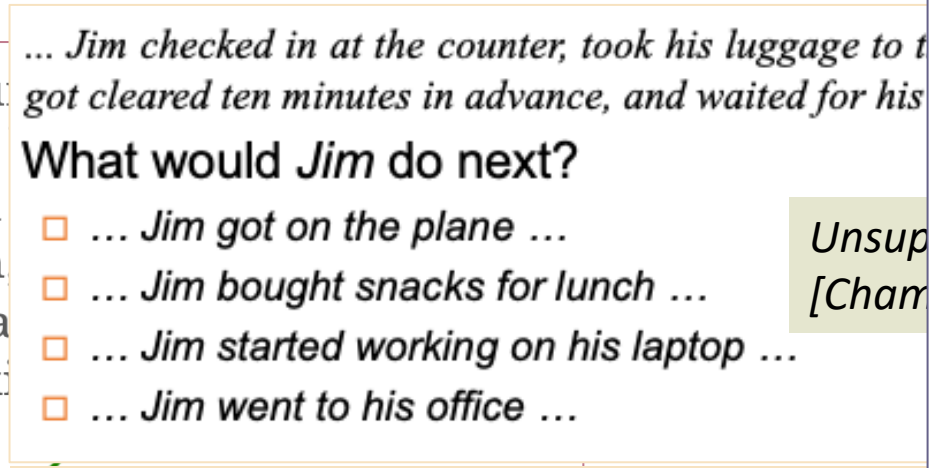
- Extract the events and the relationship(s) between ~~this~~ all event pairs
  - most works focus on one type of relationship, e.g., only predicting coreference relations, or only predicting temporal relations.
  - some also attempts to predict multiple types at the same time.
  - people start to consider multiple events and their relations jointly
  - joint extraction of events and relations
- Evaluated by
  - precision and recall on all relations
  - metrics that consider global coherency (B^3, MUC, temporal awareness, etc.)
  - end-to-end metrics that consider event extraction errors

General Problem Statement (cont'd)

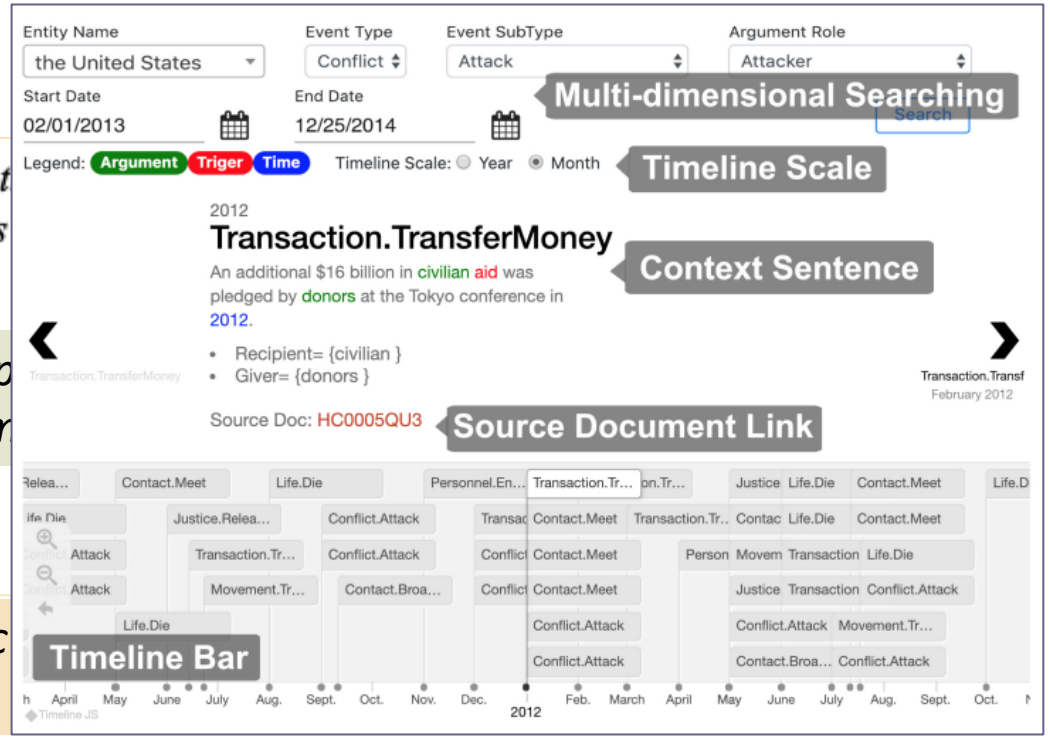
- This part only covers event-event relationships.
- StoryCloze, script learning, schema induction, timeline construction, etc. can also be viewed as tackling relationships among multiple events, but will be covered in later sections of this tutorial.



Story Comprehension for Predic3ng What Happens Next [Chaturvedi et al., 2017]



KnowSemLM : A Knowledge Infused Semantic Model. [Peng et al., 2019]

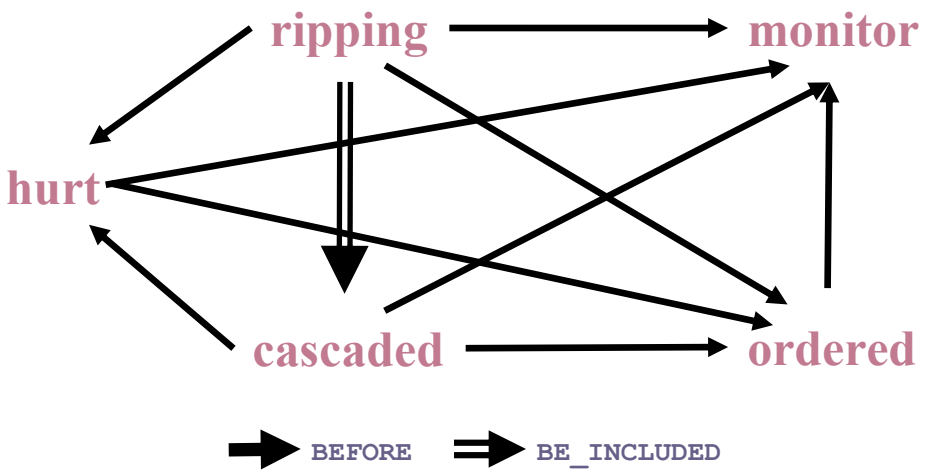


Multilingual Entity, Relation, Event and Human Value Extraction [Li et al., 2019]

Challenge 1

- Events are inter-related due to the transitive property of relations
  - Coreference: If A == B, B == C, then A == C.
  - Temporality: If A before B, B before C, then A before C.
  - Parent-child: If A contains B, B contains C, then A contains C.
  - Causality: If A leads to B, B leads to C, then A leads to C.*

*In Los Angeles that lesson was brought home Friday when tons of earth **cascaded** down a hillside, **ripping** two houses from their foundations. No one was **hurt**, but firefighters **ordered** the evacuation of nearby homes and said they'll **monitor** the shifting ground until March 23rd.*



A Structured Learning Approach to Temporal Relation Extraction. Ning et al., 2017.

Challenge 2

- Different types of relations are also inter-related
  - Coreference vs other relationships: If event A is a coreference of event B, then other relationships of A must be the same with those of B.
  - Parent-child relationship vs temporal relationship: If A is the parent of B, then the time span of A must include that of B.
  - Causal relationship vs temporal relationship: Physically, a cause should be temporally before its effect

AAAI-21 is [held virtually]$_1$ due to [the pandemic]$_2$.

[the pandemic]$_2$ CAUSES [held virtually]$_1$

|
V

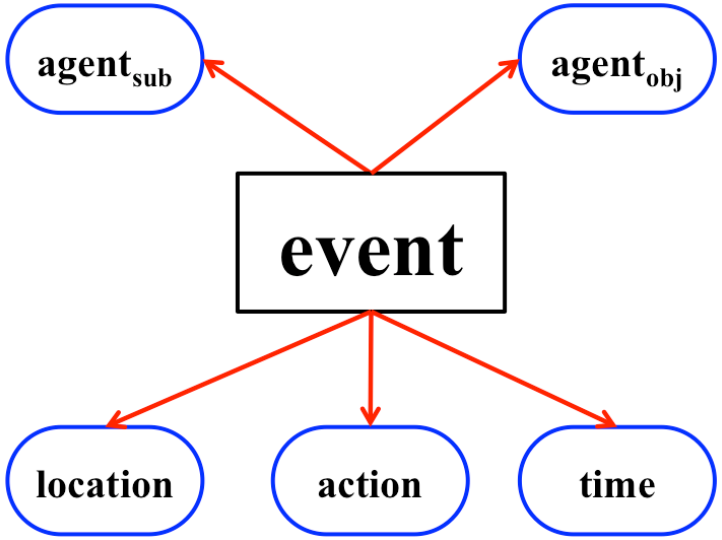
[the pandemic]$_2$ HAPPENS BEFORE [held virtually]$_1$

Joint Reasoning for Temporal and Causal Relations. Ning et al., ACL2018.

Challenge 3

- Event itself is a complex concept, with many components, and can have different modalities
  - which often leads to many difficult cases when designing relation formalisms

One event has many components



Event Detection and Co-reference with Minimal Supervision. Peng et al., 2016.

Events in different modes

The lion had a large meal and slept for 24 hours.

[Negated] The lion didn't sleep after having a large meal.

[Uncertain] The lion may have had a large meal before sleeping.

[Hypothetical] If the lion has a large meal, it will sleep for 24 hours.

[Repetitive] The lion used to sleep for 24 hours after having large meals.

[Generic] After having a large meal, lions may sleep longer.

TORQUE: A Reading Comprehension Dataset of Temporal Ordering Questions. Ning et al., 2020.

Researchers [went]$_1$ to New York to [give presentations]$_2$ at AAAI in 2020.

- *To [give presentations]$_2$ is the cause of [went]$_1$*
- *But, [give presentations]$_2$ happened after [went]$_1$*

Shouldn't the cause happen before the effect?

He used to take a [walk]$_1$ after [dinner]$_2$.

He took a [walk]$_1$ after [dinner]$_2$ today.

[walk]$_1$ happens after [dinner]$_2$ in both sentences.

But, are they the same relationship?

He used to take a [walk]$_1$ after [dinner]$_2$.

He took a [walk]$_1$ after [dinner]$_2$ today.

He used to take a [walk]$_1$ after [dinner]$_2$, but today he took a [walk]$_3$ beforehand.

What's their relationship?

Challenges and How to Handle Them

- Events are inter-related due to transitive property of relations
- Different types of relations are also inter-related
- Event itself is a complex concept, with many components, and can have different modalities

- "Joint" – taking into consideration the structural constraints among multiple events, cross multiple relation types, and event properties and extraction.

How to Handle Them

Multiple relation types

Coreference vs other relationships
Causal relationship vs temporal relationship
Parent-child relationship vs temporal relationship
…

How do we define events?
How do we jointly extract events and relations?

Event properties and extraction

Multiple events

Coreference: If A == B, B == C, then A == C.
Temporality: If A before B, B before C, then A before C.
*Causality: If A leads to B, B leads to C, then A leads to C.**
Parent-child: If A contains B, B contains C, then A contains C.

A Non-exhaustive Overview

Multiple relation types

- *T: Temporal*
- *C: Causal*
- *E: Coreferential*
- *P: Parent-child*

T, C: Mirza COLING'16

T, C: Ning ACL'18, NAACL'18, EMNLP'19

T, C: Mostafazadeh 2016

T, P, E: Wang EMNLP'20

E, P: Zhou ACL'20

T: Denis IJCAI'11

T: Do EMNLP'12

T: Chambers TACL'14

T: Ning EMNLP'17

T: Han CoNLL'19

E: Bagg MUC'98

E: Chen GMNLP'09

Multiple events

T: UzZaman SEM'13

P: Glavas LREC'14

Event properties and extraction

E: Cybulska RANLP'13

C: Do EMNLP'11

T: Han EMNLP'19

T: Vashishtha ACL'19

T: Ning EMNLP'20

*E: Ji ACL'08

E: Naughton PhD'09

*E: Liao ACL'10

E: Peng EMNLP'16

A Non-exhaustive Overview

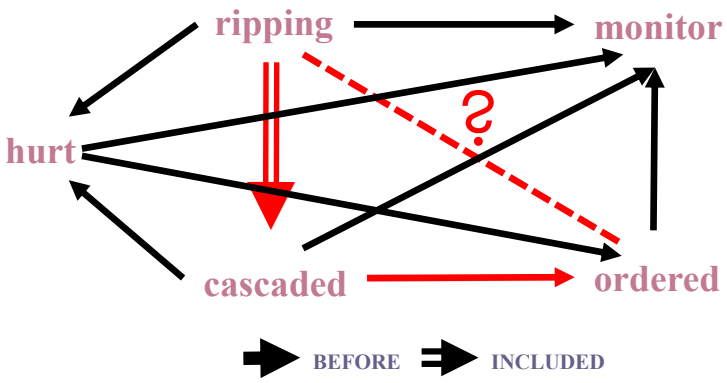
The general methodology:

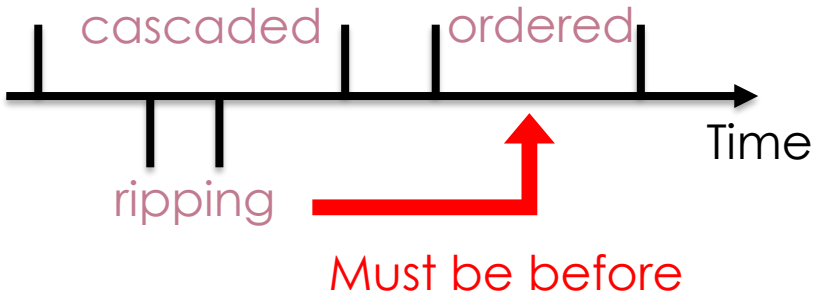
- Find structures in data/task
- Enforce (strictly/loosely) the structure
  - in inference
  - in learning
- Investigate the underlying linguistic formalism

Example: Enforce Temporal Transitive Structure

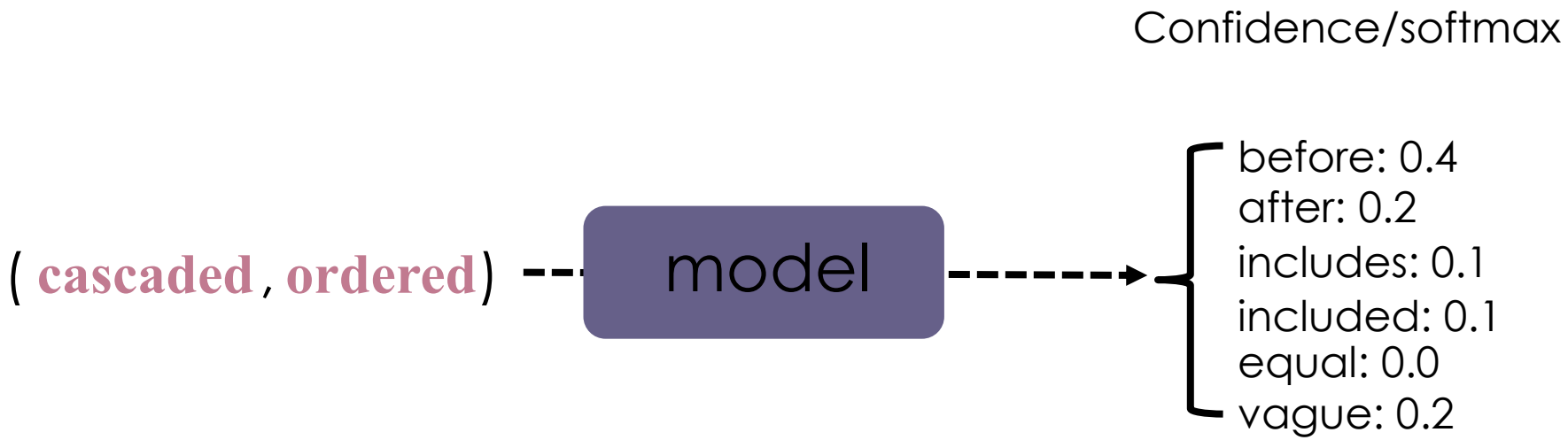
Due to transitivity, temporal relations are not independent

Global inference: respect these transitive constraints in inference



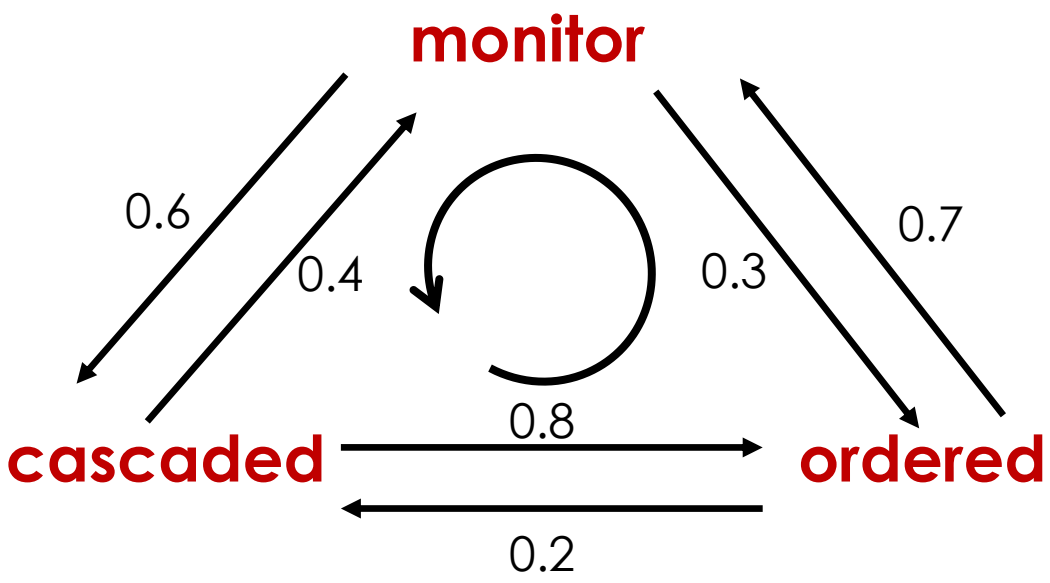


Assume a model is already trained



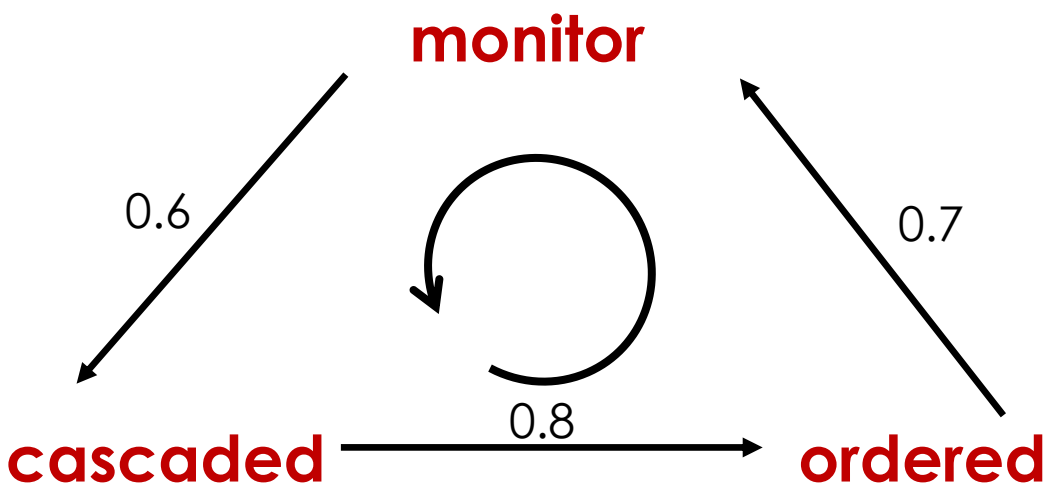
Based on these confidence scores, we need to solve for the final temporal graph.

Global Inference (A Toy Example)



Time cannot be a loop!

Global Inference (A Toy Example)



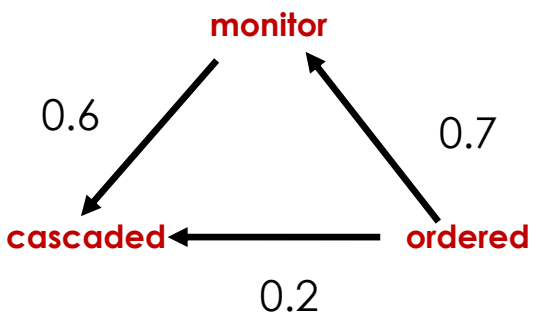
Local inference:

Mani et al., 2006
Chambers et al., 2007
Bethard et al., 2007

We should not only select the assignment with the best score, but also avoid loops

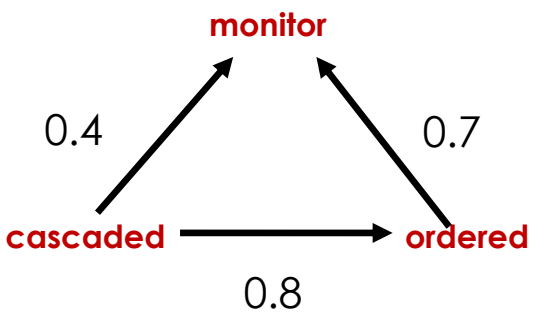
Global Inference (A Toy Example)

Option 1



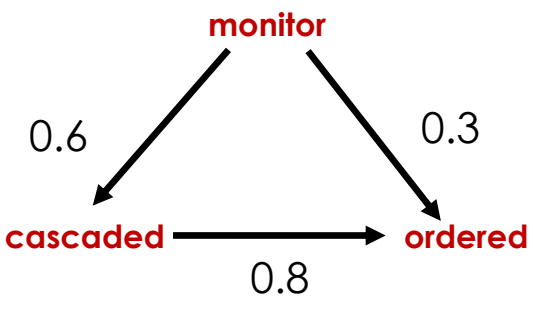
0.6+0.2+0.7=1.5

Option 2



0.4+0.8+0.7=1.9

Option 3

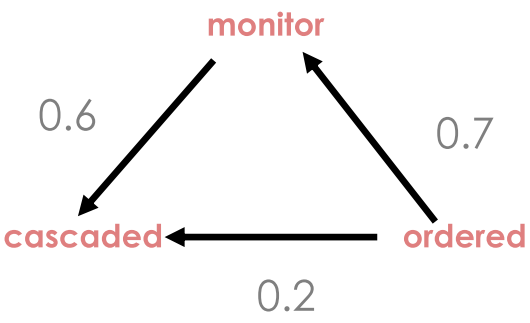


0.6+0.3+0.8=1.7

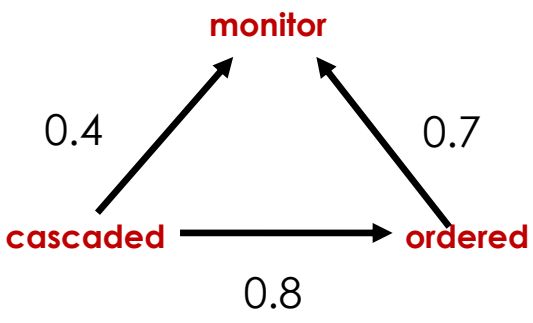
We should not only select the assignment with the best score, but also avoid loops

Global Inference (A Toy Example)

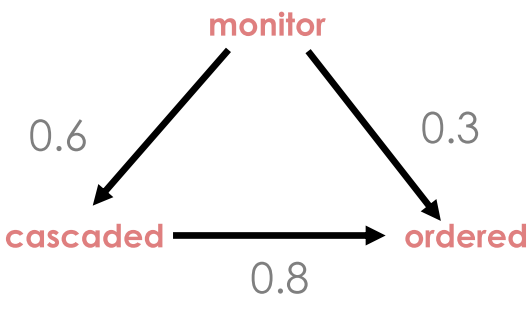
Option 2



0.6+0.2+0.7=1.5



0.4+0.8+0.7=1.9

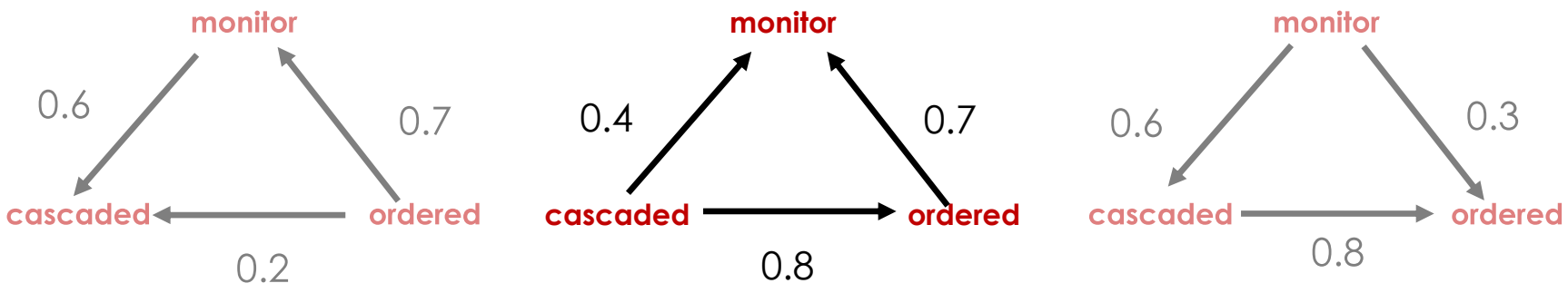


0.6+0.3+0.8=1.7

We should not only select the assignment with the best score, but also avoid loops

Global Inference (A Toy Example)

Option 2



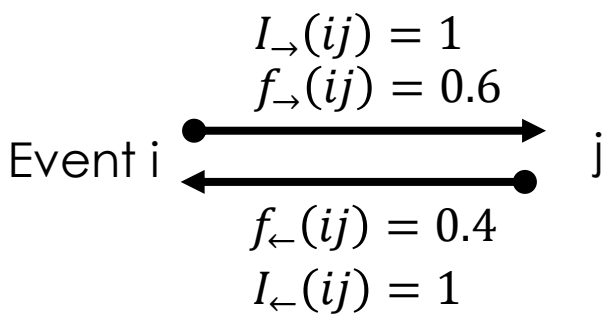
0.6+0.2+0.7=1.5 0.4+0.8+0.7=1.9 0.6+0.3+0.8=1.7

This "global inference" procedure is often formulated as an integer linear programming (ILP) problem.

A Linear Programming Formulation for Global Inference in Natural Language Tasks. Roth & Yih, CoNLL2004.

Global Inference via ILP

Integer Linear Programming (ILP)



real variable

$$\hat{I} = arg\max_{I} \sum_{i<j} \sum_{r} f_r(ij) \boxed{I_r(ij)}$$

boolean variable

s.t. $\forall i, j, k$

$$\sum_{r} I_r(ij) = 1, \qquad I_{r1}(ij) + I_{r2}(jk) - I_{r3}(ik) \leq 1$$

Uniqueness

Transitivity (no loops)

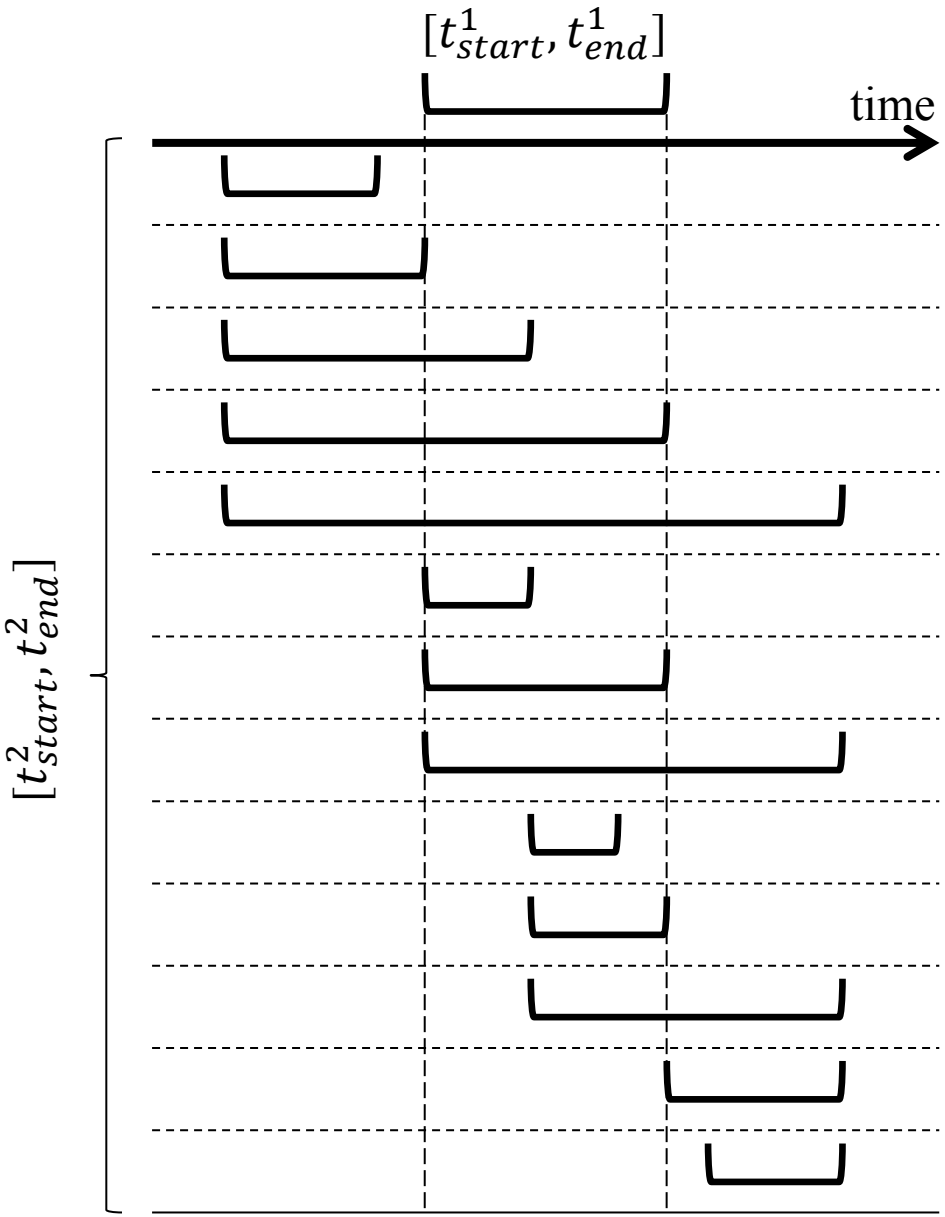
We're maximizing the score of an entire graph **while enforcing transitivity constraints**.

Global inference for temporal relation extraction: Bramsen et al., 2006. Chambers & Jurafsky, 2008. Denis & Muller, 2011. Do et al., 2012. Chambers et al., 2014. Mirza & Tonelli, 2016. Ning et al., 2017. Han et al., 2019.

The Key Step

- How do we understand $I_{r1}(ij) + I_{r2}(jk) - I_{r3}(ik) \leq 1$?
- Recall I_r are binary variables.
- If both $I_{r1}(ij) = I_{r2}(jk) = 1$, then $I_{r3}(ik)$ must be 1 due to this constraint.
- Otherwise, $I_{r3}(ik)$ is not constrained.

Constraints for Temporal Relations



Constraints for Temporal Relations

Relation between Event1 and Event2 — Relation between Event2 and Event3 — Relation between Event1 and Event3

No.	r_1	r_2	$\text{Trans}(r_1, r_2)$
1	r	r	r
2	r	**s**	r
3	r_1	r_2	$\overline{\text{Trans}(\bar{r}_2, \bar{r}_1)}$
4	**b**	**i**	**b, i, v**
5	**b**	**ii**	**b, ii, v**
6	**b**	**v**	**b, i, ii, v**
7	**a**	**i**	**a, i, v**
8	**a**	**ii**	**a, ii, v**
9	**a**	**v**	**a, i, ii ,v**
10	**i**	**v**	**b, a, i, v**
11	**ii**	**v**	**b, a, ii, v**

- Relation labels
  - b: before
  - a: after
  - i: including
  - ii: included
  - s: simultaneously
  - v: vague

Joint Reasoning for Temporal and Causal Relations. Ning et al., ACL2018.

The Key Step

- How do we understand $I_{r1}(ij) + I_{r2}(jk) - I_{r3}(ik) \leq 1$?
- Recall I_r are binary variables.
- If both $I_{r1}(ij) = I_{r2}(jk) = 1$, then $I_{r3}(ik)$ must be 1 due to the constraint.
- Otherwise, $I_{r3}(ik)$ is not constrained.

- What if r_3 has multiple choices?
- A small extension: $I_{r1}(ij) + I_{r2}(jk) - \sum_{r_3} I_{r3}(ik) \leq 1$

- What if we want to enforce constraints across different relation types, e.g., temporal & causal?

Joint Reasoning for Temporal and Causal Relations. Ning et al., ACL2018.

Constraints for Multiple Relation Types

Temporal only

- $\hat{I} = arg\max_{I} \sum_{i<j} \sum_{r} f_r(ij) I_r(ij)$

  s.t. $\forall i, j, k$

  $$\sum_{r} I_r(ij) = 1,$$

  $$I_{r1}(ij) + I_{r2}(jk) - I_{r3}(ik) \leq 1$$

→

Temporal & Causal

- $\hat{I} = arg\max_{I} \sum_{i<j} (\sum_{r} f_r(ij) I_r(ij) + \sum_{c} h_c(ij) J_c(ij))$

  s.t. $\forall i, j, k$

  $$\sum_{r} I_r(ij) = 1,$$

  $$I_{r1}(ij) + I_{r2}(jk) - I_{r3}(ik) \leq 1$$

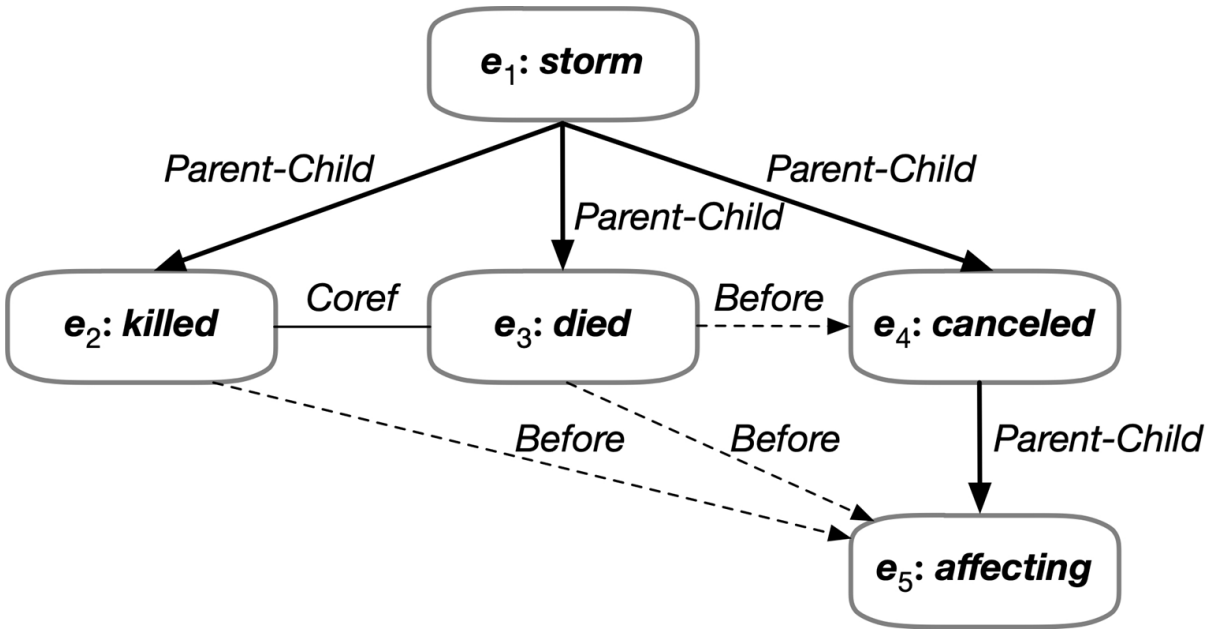
  $$J_{causes}(ij) \leq I_{before}(ij)$$

Joint Reasoning for Temporal and Causal Relations. Ning et al., ACL2018.

Constraints for Temporal, Parent-child, and Coreference

- Temporal Relations
- Subevent Relations
- Event Coreference

On Tuesday, there was a typhoon-strength (***e***$_1$***:storm***) in Japan. One man got (***e***$_2$***:killed***) and thousands of people were left stranded. Police said an 81-year-old man (***e***$_3$***:died***) in central Toyama when the wind blew over a shed, trapping him underneath. Later this afternoon, with the agency warning of possible tornadoes, Japan Airlines (***e***$_4$***:canceled***) 230 domestic flights, (***e***$_5$***:affecting***) 31,600 passengers.



Constrained Learning for Event-Event Relation Extraction. Wang et al., EMNLP2020.

Constraints for Temporal, Parent-child, and Coreference

Relation between Event2 and Event3

Relation between Event1 and Event3

α \ β	PC	CP	CR	NR	BF	AF	EQ	VG
PC	PC, ¬AF	–	PC, ¬AF	¬CP, ¬CR	BF , ¬CP, ¬CR	–	BF , ¬CP, ¬CR	–
CP	–	CP, ¬BF	CP, ¬BF	¬PC, ¬CR	–	AF, ¬PC, ¬CR	AF, ¬PC, ¬CR	–
CR	PC, ¬AF	CP, ¬BF	CR, EQ	NR	BF , ¬CP, ¬CR	AF, ¬PC, ¬CR	EQ	VG
NR	¬CP, ¬CR	¬PC, ¬CR	NR	–	–	–	–	–
BF	BF , ¬CP, ¬CR	–	BF , ¬CP, ¬CR	–	BF , ¬CP, ¬CR	–	BF , ¬CP, ¬CR	¬AF, ¬EQ
AF	–	AF, ¬PC, ¬CR	AF, ¬PC, ¬CR	–	–	AF, ¬PC, ¬CR	AF, ¬PC, ¬CR	¬BF , ¬EQ
EQ	¬AF	¬BF	EQ	–	BF , ¬CP, ¬CR	AF, ¬PC, ¬CR	EQ	VG, ¬CR
VG	–	–	VG, ¬CR	–	¬AF, ¬EQ	¬BF , ¬EQ	VG	–

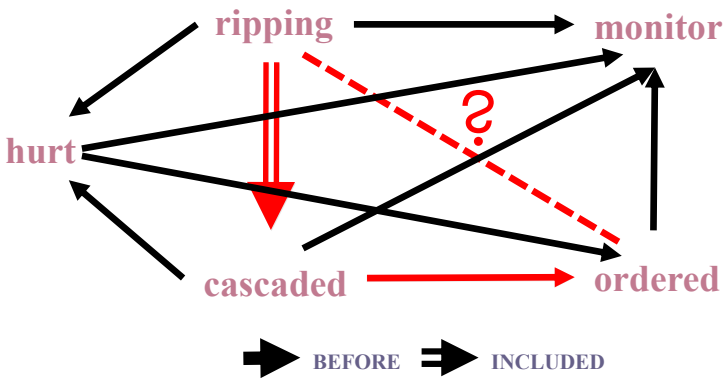
Relation between Event1 and Event2

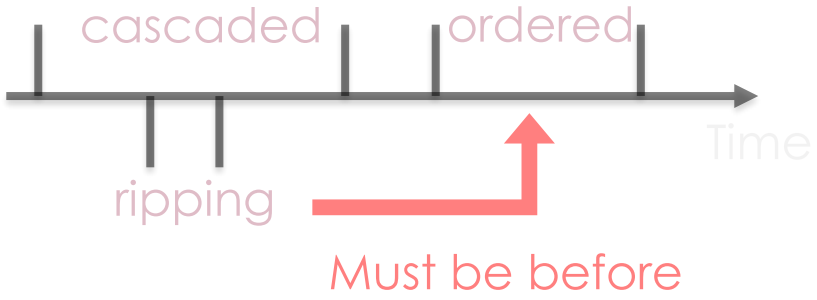
Constrained Learning for Event-Event Relation Extraction. Wang et al., EMNLP2020.

But how do we train the model?

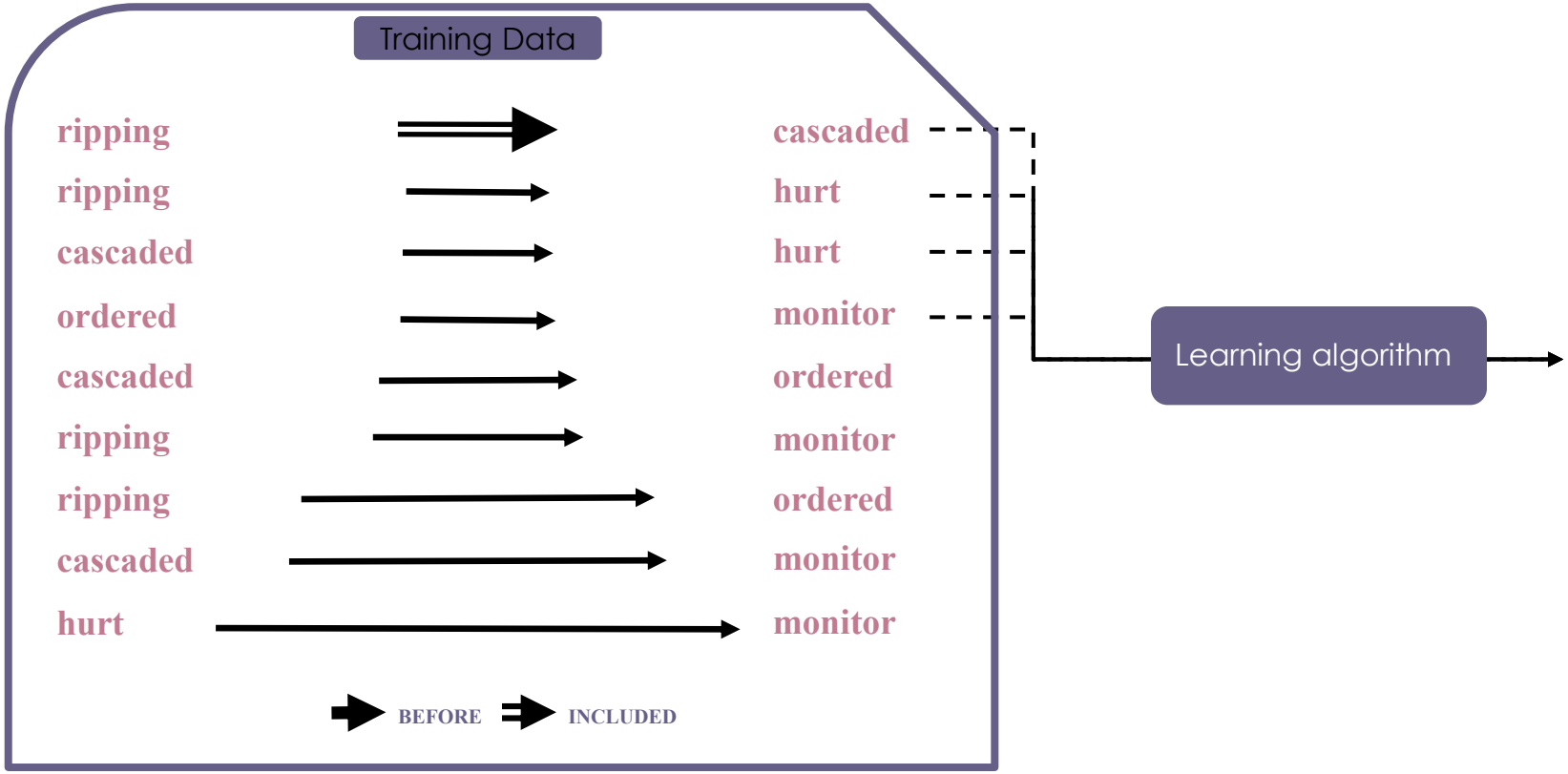
Due to transitivity, temporal relations are not independent

Existing methods: global inference **with local learning**





Local learning



Local learning is not sufficient

tons of earth **cascaded** down a hillside,

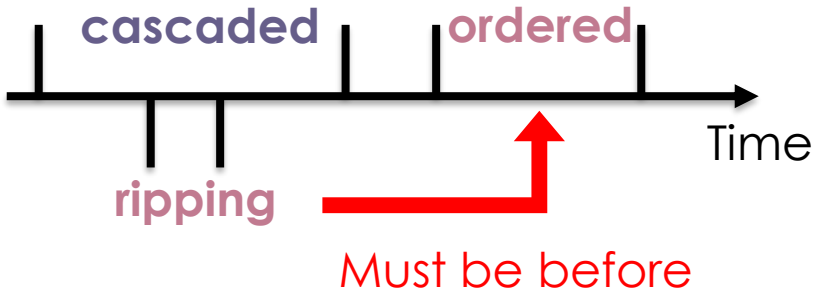
…**ripping** two houses…firefighters **ordered** the evacuation of nearby homes…

Q: (**ripping**, **ordered**)=? (difficult even for humans)

Annotation says "before", if we update the parameters to fit it,

Then it leads to <u>overfitting</u>

Overfitting is mitigated.



Structured learning

Structured learning

Standard Perceptron

For each (x, y)

$\hat{y} = sgn(w^T x)$

If $y \neq \hat{y}$

Update w

- (x, y): feature and label for **a single pair of events**
- **Unaware** of decisions in other pairs

Structured Perceptron

For each (X, Y)

$\hat{Y}$ =“solution to ILP”

If $Y \neq \hat{Y}$

Update W

- (X, Y): features and labels from **the entire graph**
- **Aware** of other pairs thanks to the global inference in-between

Discriminative training methods for hidden markov models: Theory and experiments with perceptron algorithms. Collins, ACL2002.

Structured learning

$$L = L_A + \lambda_S L_S + \lambda_C L_C$$

Fidelity to annotations

$$L_A = \sum_{e_1, e_2 \in \mathcal{E}_D} -w_r \log r_{(e_1, e_2)}$$

Symmetry constraints

$$L_S = \sum_{e_1, e_2 \in \mathcal{E}, \alpha \in \mathcal{R}_S} |\log \alpha_{(e_1, e_2)} - \log \bar{\alpha}_{(e_2, e_1)}|$$

Transitivity constraints

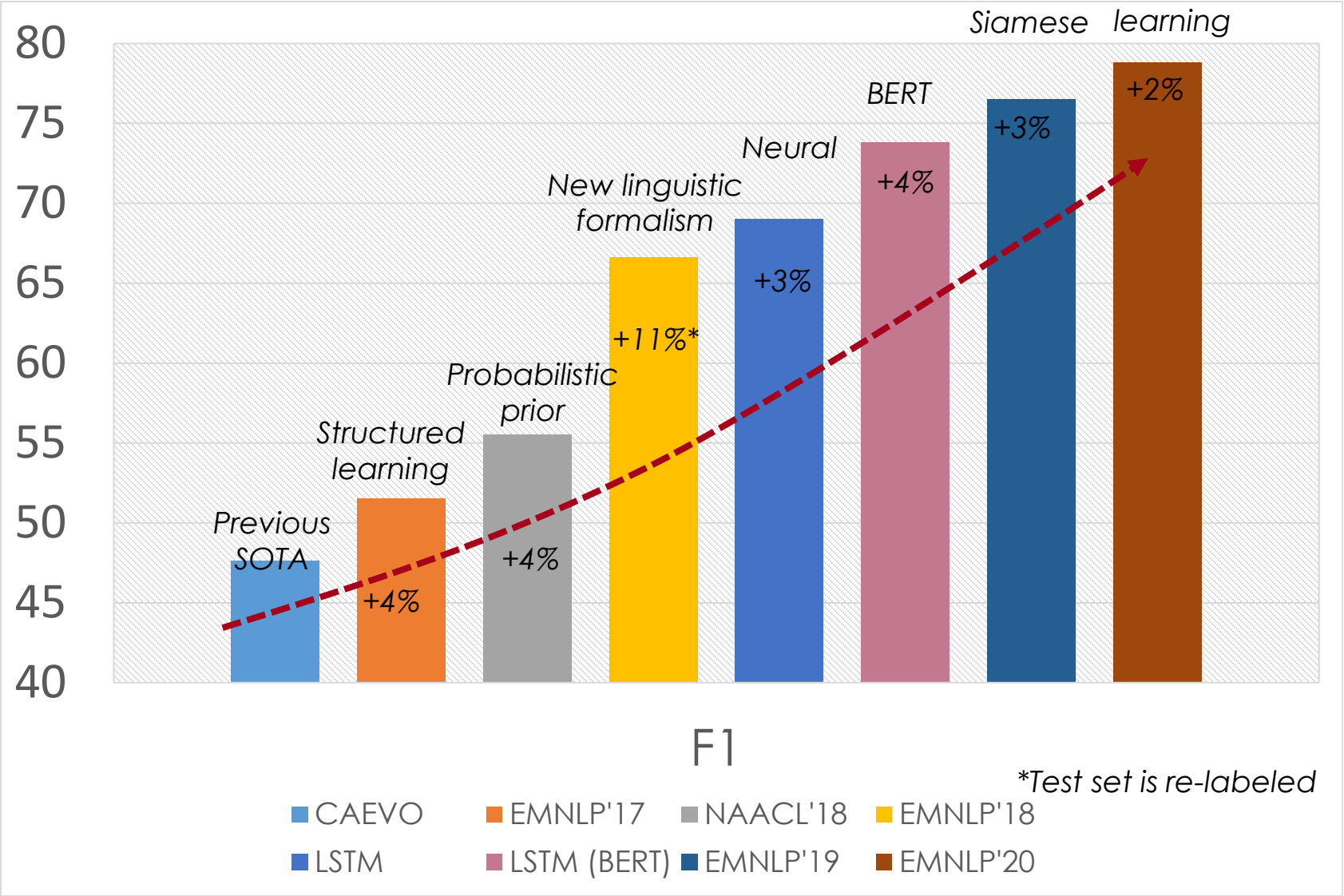
$$L_C = \sum_{\substack{e_1, e_2, e_3 \in \mathcal{E}_D, \\ \alpha, \beta \in \mathcal{R}, \gamma \in \mathrm{De}(\alpha, \beta)}} |L_{t_1}| + \sum_{\substack{e_1, e_2, e_3 \in \mathcal{E}_D, \\ \alpha, \beta \in \mathcal{R}, \delta \notin \mathrm{De}(\alpha, \beta)}} |L_{t_2}|$$

$$L_{t_1} = \log \alpha_{(e_1, e_2)} + \log \beta_{(e_2, e_3)} - \log \gamma_{(e_1, e_3)}$$

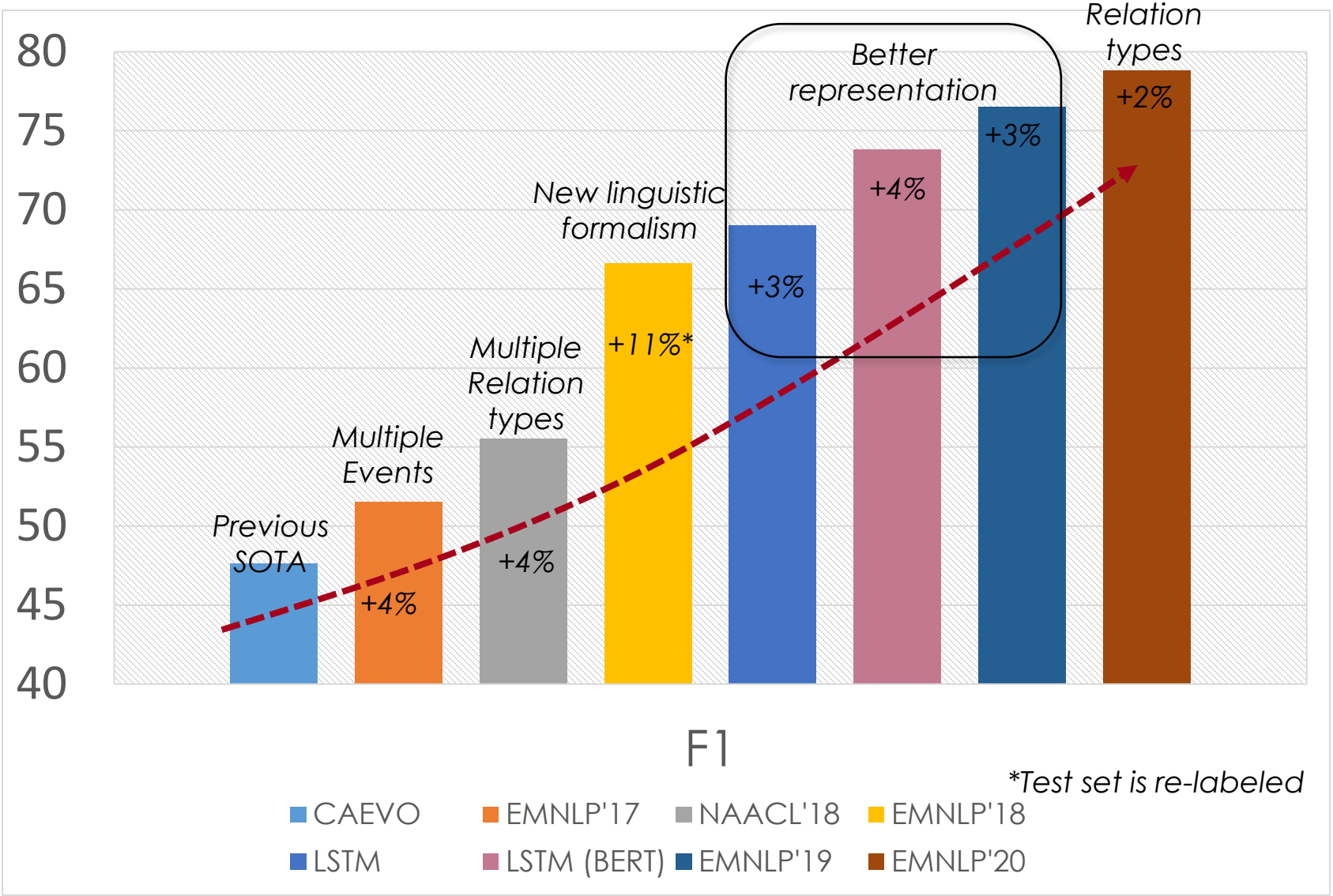
$$L_{t_2} = \log \alpha_{(e_1, e_2)} + \log \beta_{(e_2, e_3)} - \log(1 - \delta_{(e_1, e_3)})$$

Constrained Learning for Event-Event Relation Extraction. Wang et al., EMNLP2020.

Temporal relation extraction in recent years

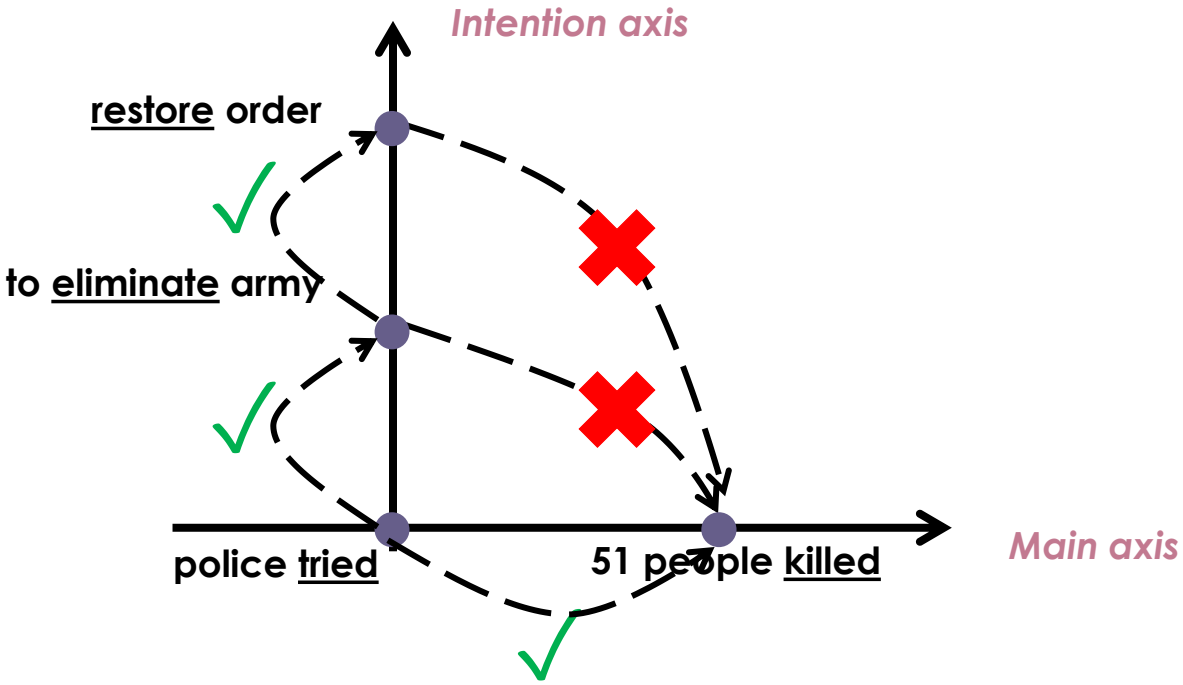


Temporal relation extraction in recent years



Study on Linguistic Formalisms

- Time is one-dimensional physically.
- But, **multiple time axes may exist in natural language** (*Ning et al., 2018*)
  - *Police **tried** to **eliminate** the pro-independence army and **restore** order. At least 51 people were **killed** in clashes between police and citizens in the troubled region.*

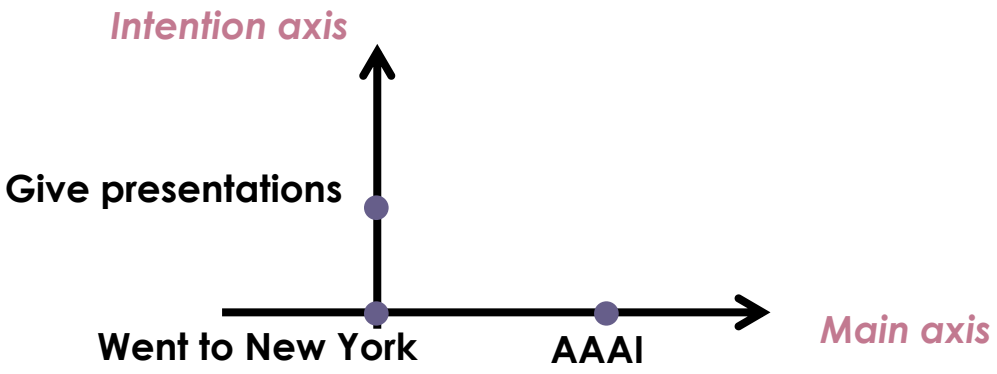


MATRES: A Multi-Axis Annotation Scheme for Event Temporal Relations. Ning et al., ACL2018.

Researchers [went]$_1$ to New York to [give presentations]$_2$ at AAAI in 2020.

- *To [give presentations]$_2$ is the cause of [went]$_1$*
- *But, [give presentations]$_2$ happened after [went]$_1$*

Shouldn't the cause happen before the effect?



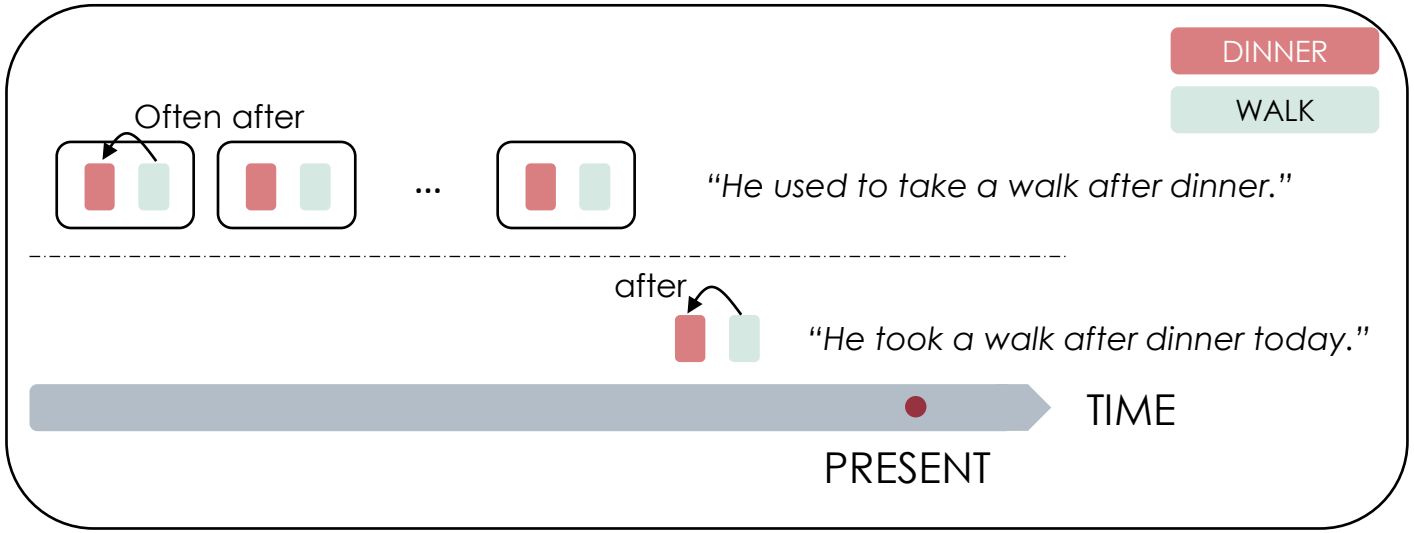
MATRES: A Multi-Axis Annotation Scheme for Event Temporal Relations. Ning et al., ACL2018.

He used to take a [walk]$_1$ after [dinner]$_2$.

He took a [walk]$_1$ after [dinner]$_2$ today.

[walk]$_1$ happens after [dinner]$_2$ in both sentences.

But, are they the same relationship?



This can be easily distinguished by the two questions below:

Q1: What did he often do after dinner?

Q2: What did he do after dinner today?

TORQUE: A Reading Comprehension Dataset of Temporal Ordering Questions. Ning et al., EMNLP2020.

TORQUE

Heavy snow is causing disruption to transport across the UK, with heavy rainfall bringing flooding to the south-west of England. Rescuers searching for a woman trapped in a landslide at her home said they had found a body.

Q1: What event has already finished?
A: searching trapped landslide said found
Q2: What event has begun but has not finished?
A: snow causing disruption rainfall bringing flooding
Q3: What will happen in the future?
A: No answers.

Hard-coded questions

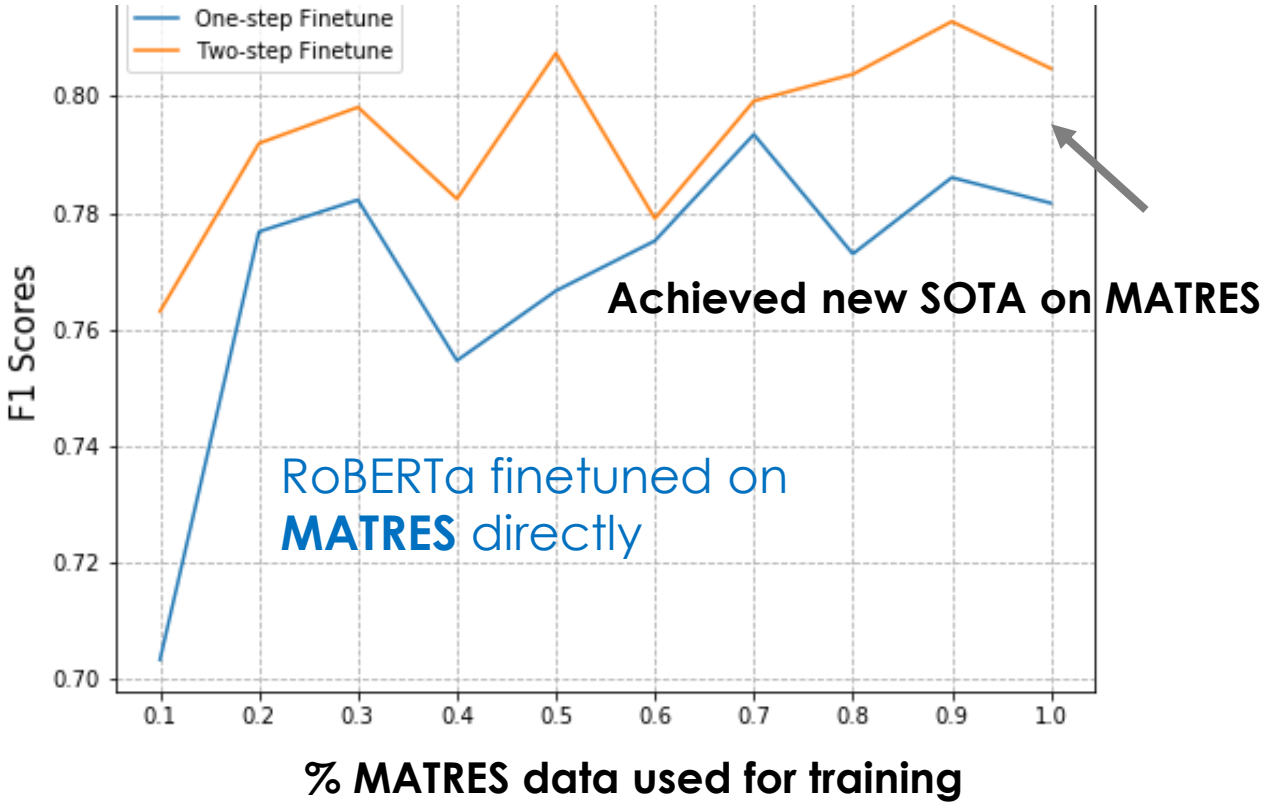
Q4: What happened before a woman was trapped?
A: landslide
Q5: What had started before a woman was trapped?
A: snow rainfall landslide
Q6: What happened while a woman was trapped?
A: searching
Q7: What happened after a woman was trapped?
A: searching said found

Group of contrast questions

Q8: What happened at about the same time as the snow?
A: rainfall
Q9: What happened after the snow started?
A: causing disruption bringing flooding searching trapped landslide said found
Q10: What happened before the snow started?
A: No answers.

Group of contrast questions

RoBERTa finetuned on **TORQUE** first and then on **MATRES**



TORQUE: A Reading Comprehension Dataset of Temporal Ordering Questions. Ning et al., EMNLP2020.

Summary

- Relations between events are important for story understanding.
- Event relation extraction is difficult because
  - Each type of relation forms a complex structure
  - Different types of relations also influences each other
  - Event formalisms are naturally difficult to define
- A key word in existing works is “JOINT”
  - Find event structures
  - Enforce these structures in inference and/or in learning
- But, the more important problem often lies in “how should we define these relations?”, or more fundamentally, “what is an event?”.

References

1. Algorithms for scoring coreference chains. Bagga & Baldwin, 1998.
2. Discriminative training methods for hidden markov models: Theory and experiments with perceptron algorithms. Collins, 2002.
3. A Linear Programming Formulation for Global Inference in Natural Language Tasks. Roth & Yih, 2004.
4. Inducing temporal graphs. Bramsen et al., 2006.
5. Refining Event Extraction through Cross-document Inference. Ji & Grishman, 2008.
6. Unsupervised Learning of Narrative Event Chains. Chambers & Jurafsky, 2008.
7. Jointly combining implicit constraints improves temporal ordering. Chambers & Jurafsky, 2008.
8. Sentence Level Event Detection and Coreference Resolution. Naughton, 2009.
9. Graph-based event coreference resolution. Chen & Ji, 2009.
10. Using document level cross-event inference to improve event extraction. Liao & Grishman, 2010.
11. Evaluation Metrics For End-to-End Coreference Resolution Systems. Cai & Strube, 2010.
12. Predicting globally-coherent temporal structures from texts via endpoint inference and graph decomposition. Denis & Muller, 2011.
13. Minimally supervised event causality identification. Do et al., 2011.
14. Temporal evaluation. UzZaman & Allen, 2011.
15. Joint inference for event timeline construction. Do et al., 2012.
16. Semantic Relations between Events and their Time, Locations and Participants for Event Coreference Resolution. Cybulska & Vossen, 2013.
17. TEMPEVAL-3: Evaluating time expressions, events, and temporal relations. UzZaman et al., 2013.
18. HiEve: A Corpus for Extracting Event Hierarchies from News Stories. Glavas et al., 2014.
19. CATENA: CAusal and TEmporal relation extraction from NAtural language texts. Mirza & Tonelli, 2016.
20. Event Detection and Co-reference with Minimal Supervision. Peng et al., 2016.
21. Which Coreference Evaluation Metric Do You Trust? A Proposal for a Link-based Entity Aware Metric. Moosavi & Strube, 2016.
22. Story Comprehension for Predic3ng What Happens Next. Chaturvedi et al., 2017.
23. A Structured Learning Approach to Temporal Relation Extraction. Ning et al., 2017.
24. Joint Reasoning for Temporal and Causal Relations. Ning et al., 2018.
25. A Multi-Axis Annotation Scheme for Event Temporal Relations. Ning et al., 2018.
26. Improving Temporal Relation Extraction with a Globally Acquired Statistical Resource. Ning et al., 2018.
27. KnowSemLM: A Knowledge Infused Semantic Language Model. Peng et al., 2019.
28. Multilingual Entity, Relation, Event and Human Value Extraction. Li et al., 2019.
29. Joint Event and Temporal Relation Extraction with Shared Representations and Structured Prediction. Han et al., 2019.
30. An Improved Neural Baseline for Temporal Relation Extraction. Ning et al., 2019.
31. Fine-Grained Temporal Relation Extraction. Vashishtha et al., 2019.
32. Constrained Learning for Event-Event Relation Extraction. Wang et al., 2020.
33. TORQUE: A Reading Comprehension Dataset of Temporal Ordering Questions. Ning et al., 2020.


THANKS
AI
NLP
Graph2Vec
Structure
Time
Hilbert
Semantics
Crowdsourcing
Fourier
Ranking
Vision
CRF
CommonSense
Embedding
BiLSTM
NER
BERT
ILP
Inference
Summarization
Biomedicine

Understanding Event Processes in Natural Language

Event-Centric Natural Language Understanding (Part III)

Muhao Chen
Viterbi School of Engineering
University of Southern California

Feb 2021

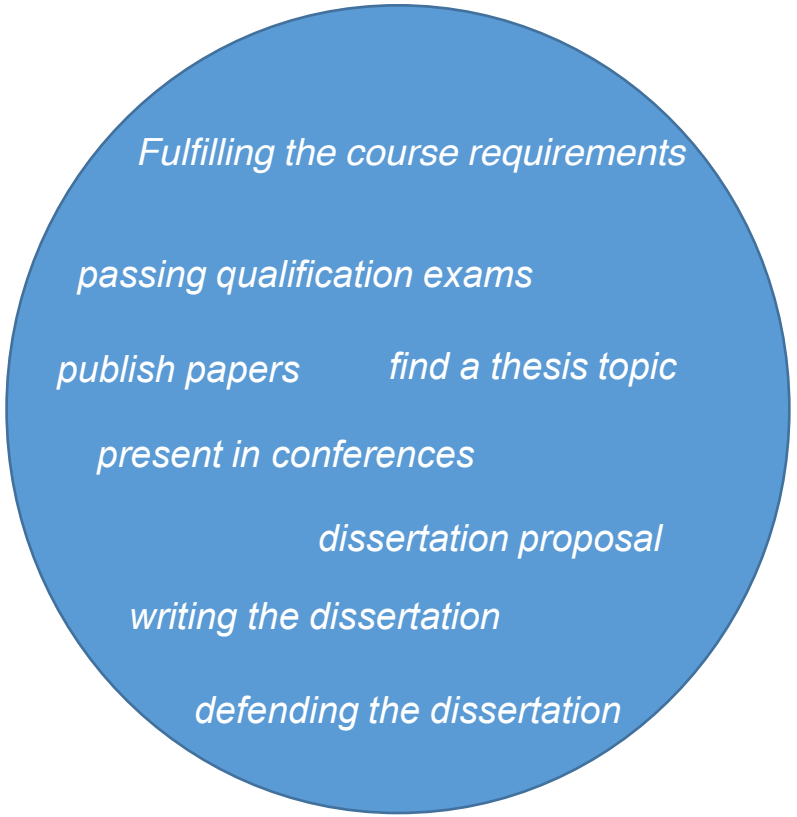
AAAI Tutorials

Event-Centric Natural Language Understanding

How Do Machines Understand the *Evolution* of Events?

Human Language Always Communicates About Events

Earning a PhD in Computer Science typically takes around 5 years. It first involves ***fulfilling the course requirements*** and ***passing qualification exams***. Then within several years, the student is expected to ***find a thesis topic***, ***publish several papers*** about the topic and ***present them in conferences***. The last one or two years are often about ***completing the dissertation proposal***, ***writing*** and ***defending the dissertation***.



Event-centric IE has helped machines in

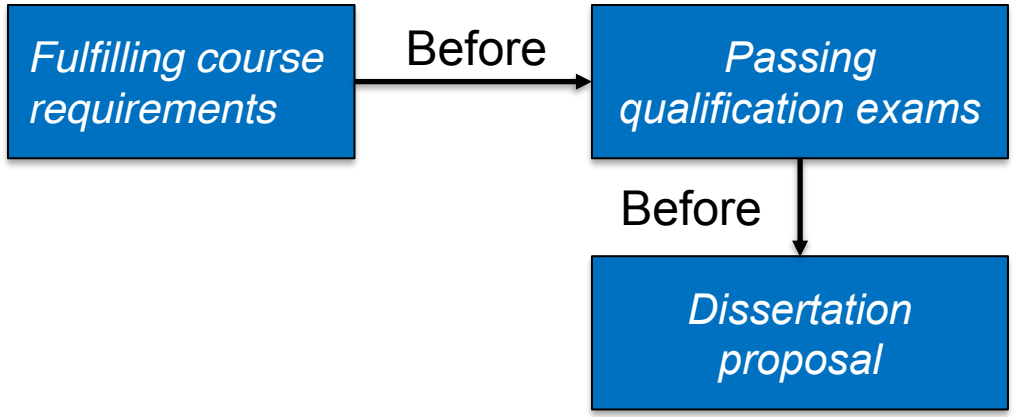
- Recognition and typing events
- Inducing the relations between two events
- Identifying the equivalence of events

Event Process Understanding And Prediction

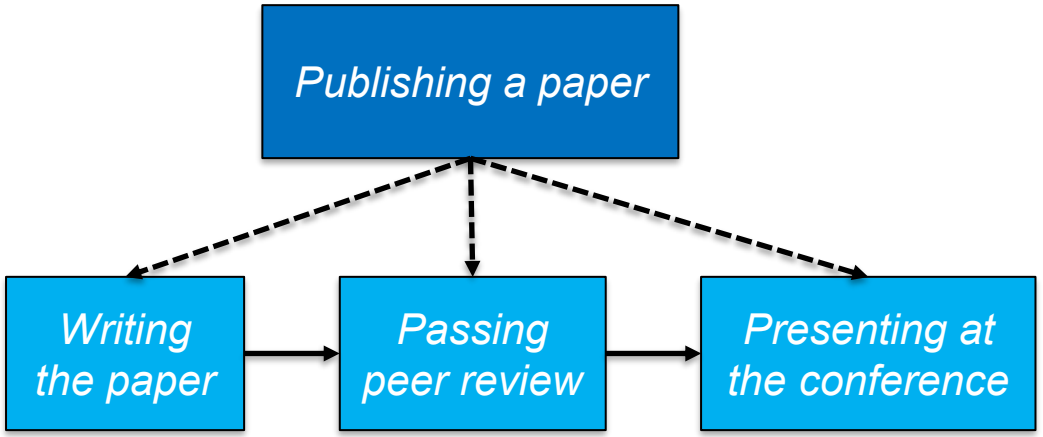
Extraction only is not enough.

Events are **NOT simple, static** predicates.

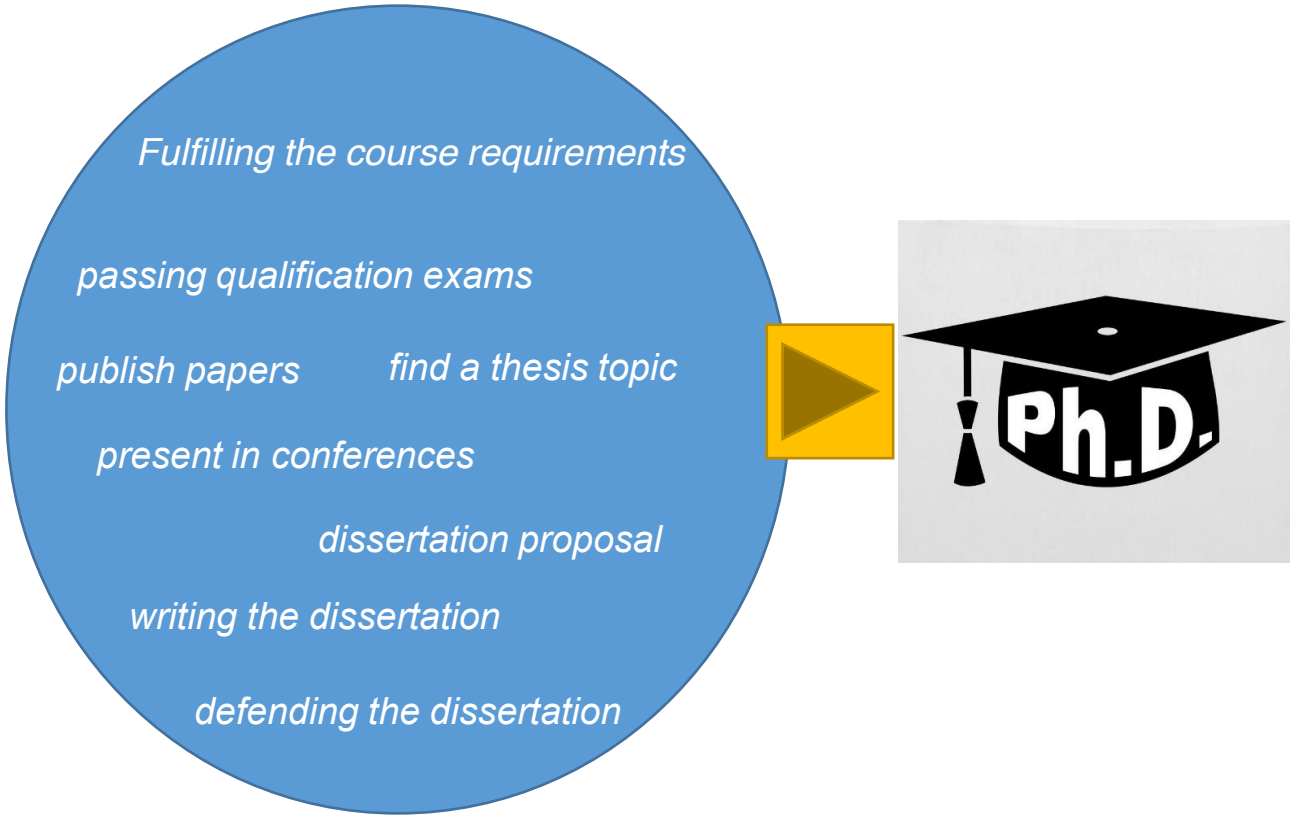
They evolve,



are described in different granularities,



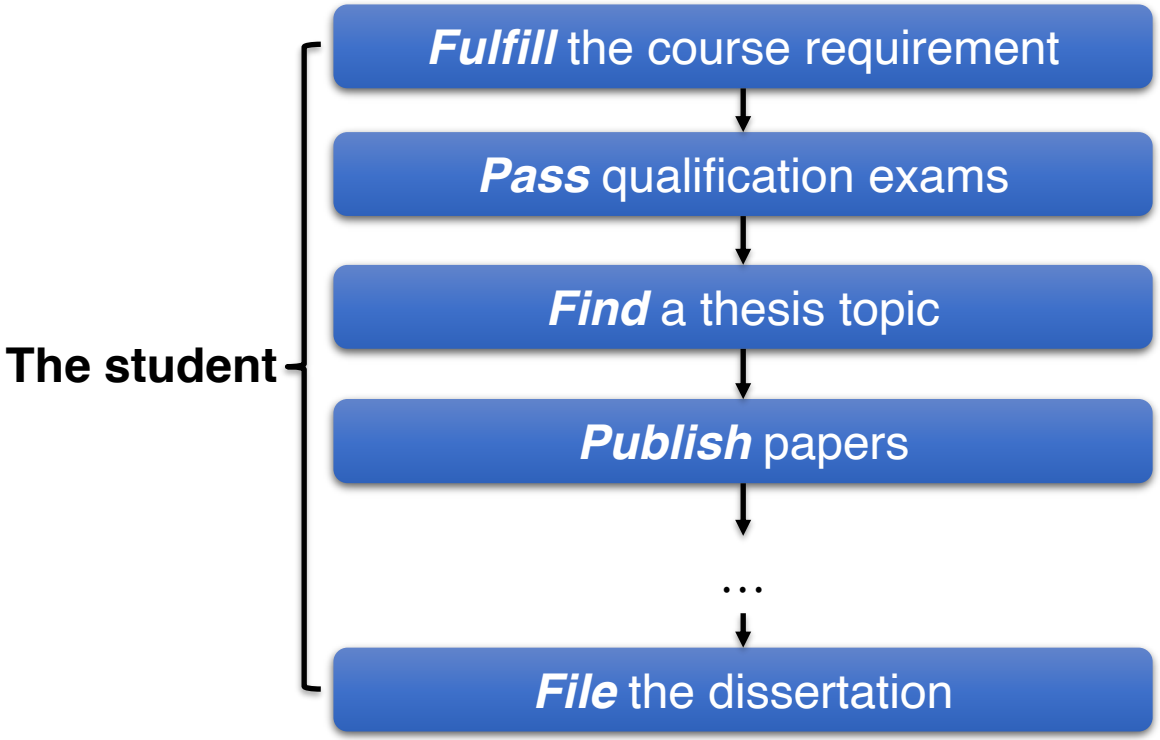
and are always directed by specific intents or central goals [Zacks et al. *Nature Neuroscience*, 2001]



Event Process Understanding And Prediction

An event process (or event chain)

- Partially ordered events that are centered around common protagonists [Chambers et al., ACL-08]



Prediction problems on event processes

Event process completion

- What happens next?

Intention prediction

- What is the goal of "*digging a hole, putting some seeds in the hole* and *filling it with soil*"?

Membership prediction

- What are the steps of "buying a car"?

Salience prediction

- Is *defending the dissertation* more important than *doing an internship*?

Event Processes Are Essential to Downstream NLU Tasks

Narrative prediction

One day Wesley's auntie came over to visit. He was happy to see her, because he liked to play with her. When she started to give his little sister attention, he got **jealous**. He got **angry** at his auntie and **bit** his sister's hand when she wasn't looking.

Then what might happen?

O1: He was **scolded**. ✓

O2: She **gave him a cookie** for being so nice. ✗

jealous → angry → bit → scolded ✓

bit → get a cookie ✗

Machine comprehension

Water is split, providing a source of electrons and protons (hydrogen ions, H^+) and giving off O_2 as a by-product. ***Light absorbed*** by chlorophyll drives a ***transfer of the electrons and hydrogen ions*** from water to an acceptor called $NADP^+$.

What can the splitting of water lead to?

A: Light absorption

B: Transfer of ions

Light absorbed

↓

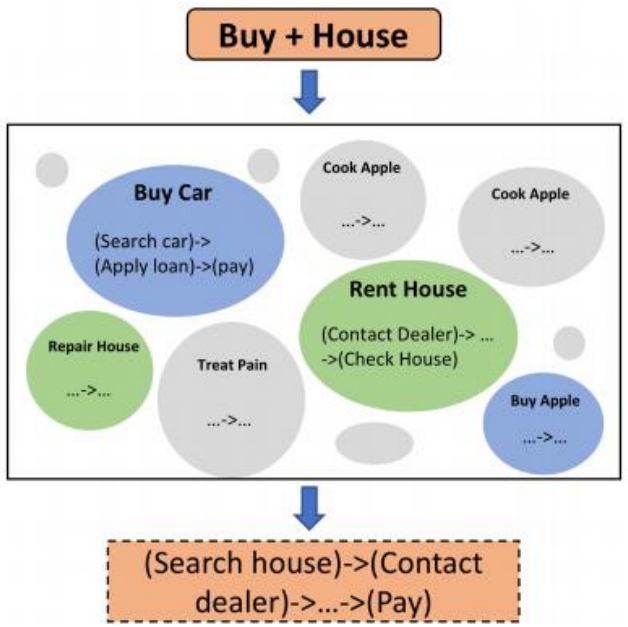
Water is split

↓

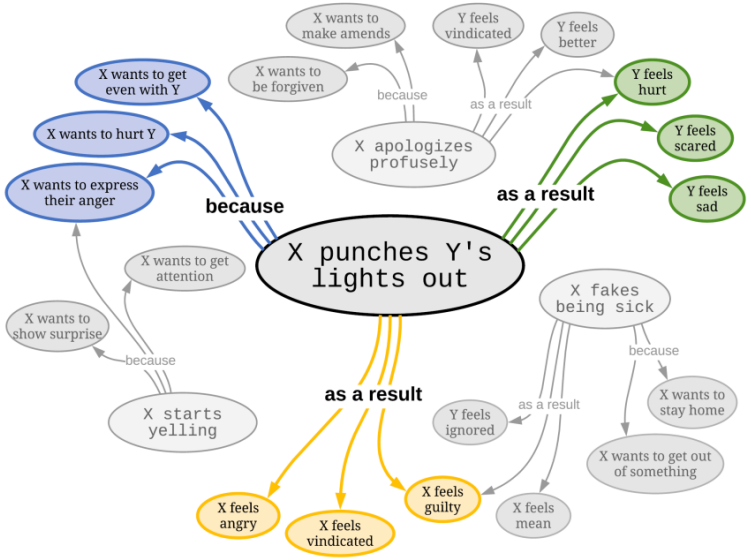
Transfer of the electrons and hydrogen ions

Agenda

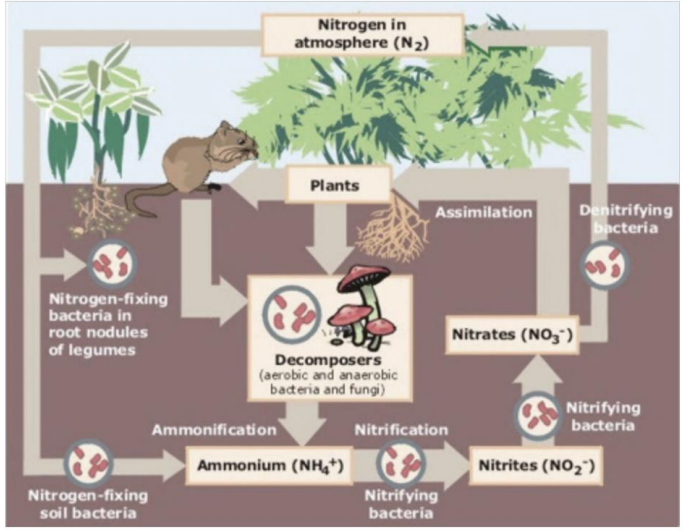
1. Event process completion



2. Event intention prediction



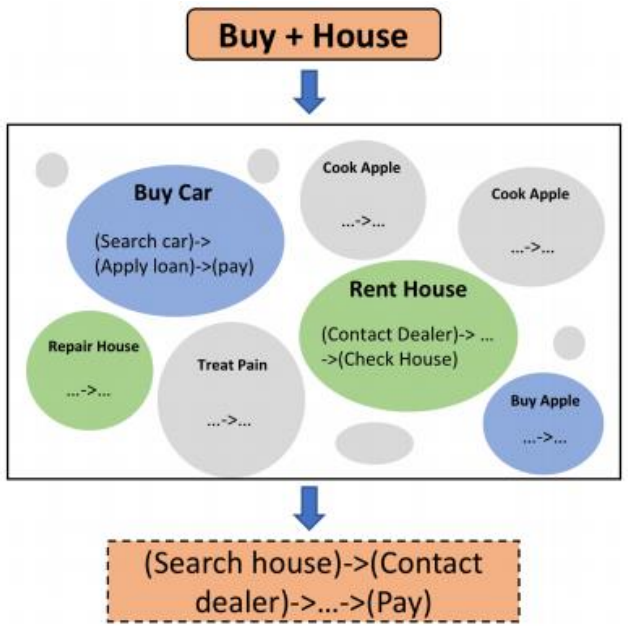
3. Event processes in downstream NLU tasks



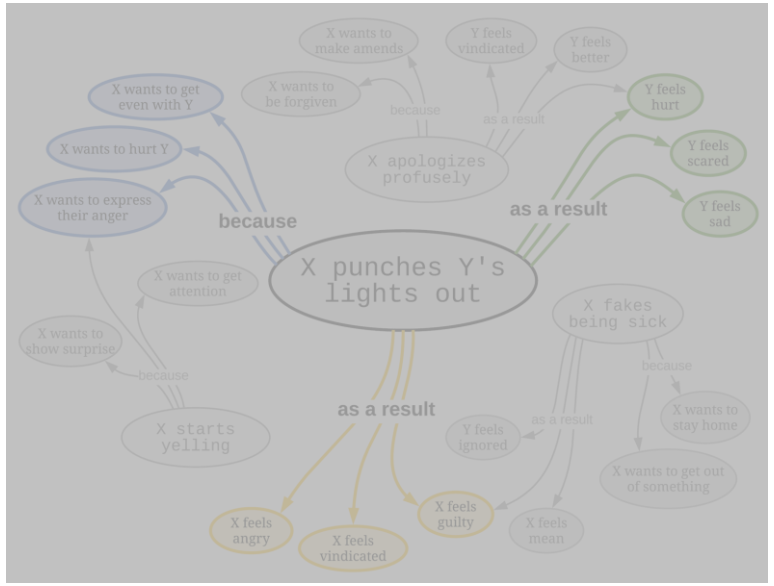
4. Open Research Directions

Agenda

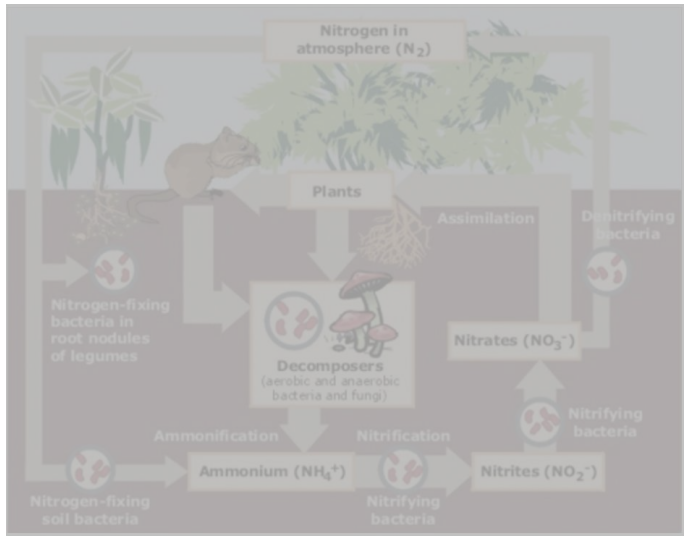
1. Event process completion



2. Event intention prediction



3. Event processes in downstream NLU tasks



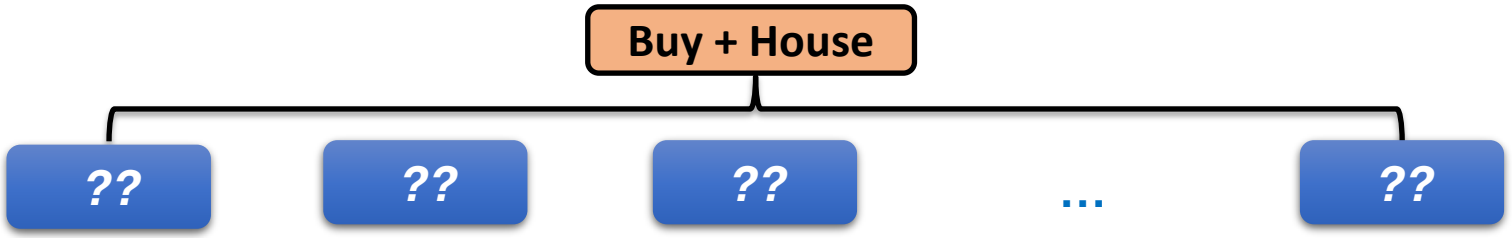
4. Open Research Directions

Event Process Completion

Two forms of process prediction

1. Predicting steps of the process

2. Inducing the entire process from scratch.



Event Process Completion

Chambers and Jurafsky. Unsupervised Learning of Narrative Event Chains. ACL-08

Unsupervised event process completion can be done using corpus statistics

- Capturing the co-occurrence of events using pointwise mutual information

$$pmi(e(w,d), e(v,g))$$

- The next most likely forthcoming event can be found by maximizing the accumulated PMI

$$\max_{j:0<j<m} \sum_{i=0}^{n} pmi(e_i, f_j)$$

(*n*: #events in the process; *m*: #events in the vocabulary.

Known events:
(pleaded subj), (admits subj), (convicted obj)

Likely Events:

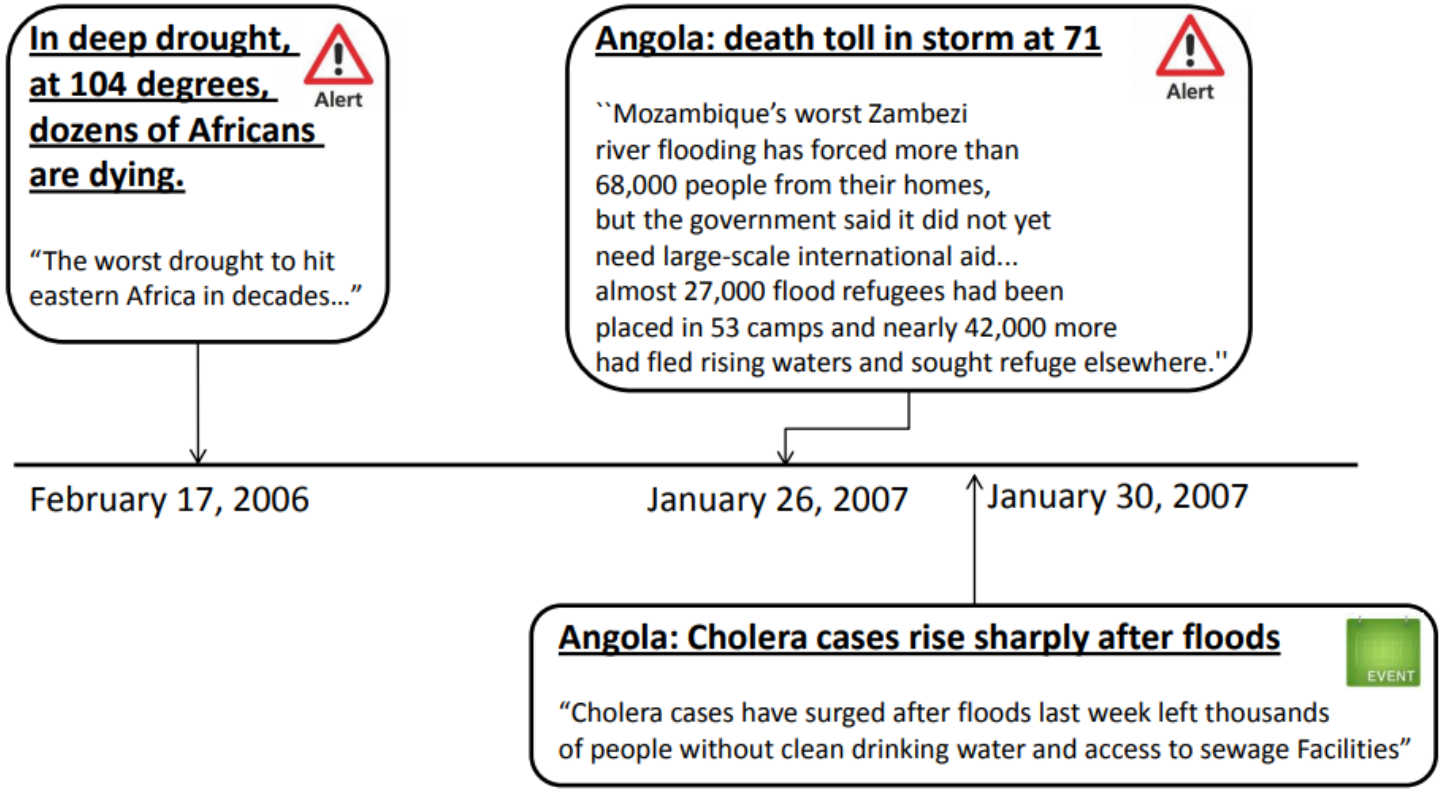
sentenced obj	0.89	indicted obj	0.74
paroled obj	0.76	fined obj	0.73
fired obj	0.75	denied subj	0.73

Improves narrative cloze tests (36% improvement on NYT Narrative Cloze).

Event Process Completion

Radinsky and Horvitz. Mining the Web to Predict Future Events. WSDM, 2013

Extension of the event chain model on multiple **dated** and **topically cohesive** documents.



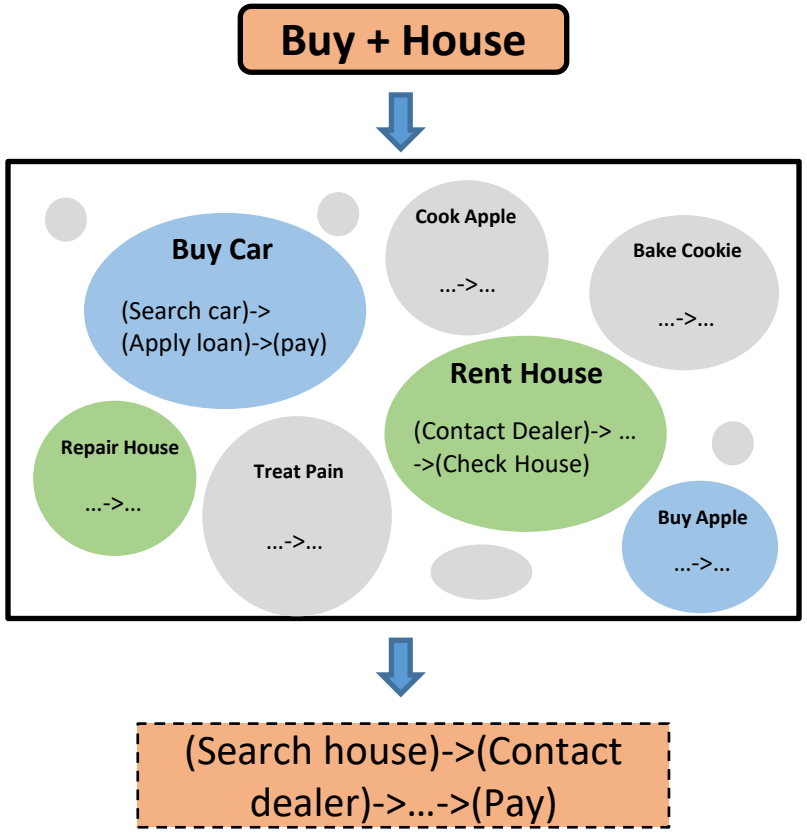
Maximum-entropy-based event co-occurence model.

The likelihood of cholera rising is predicted high after a drought followed by storms in Angola (*based on corpus statistics*).

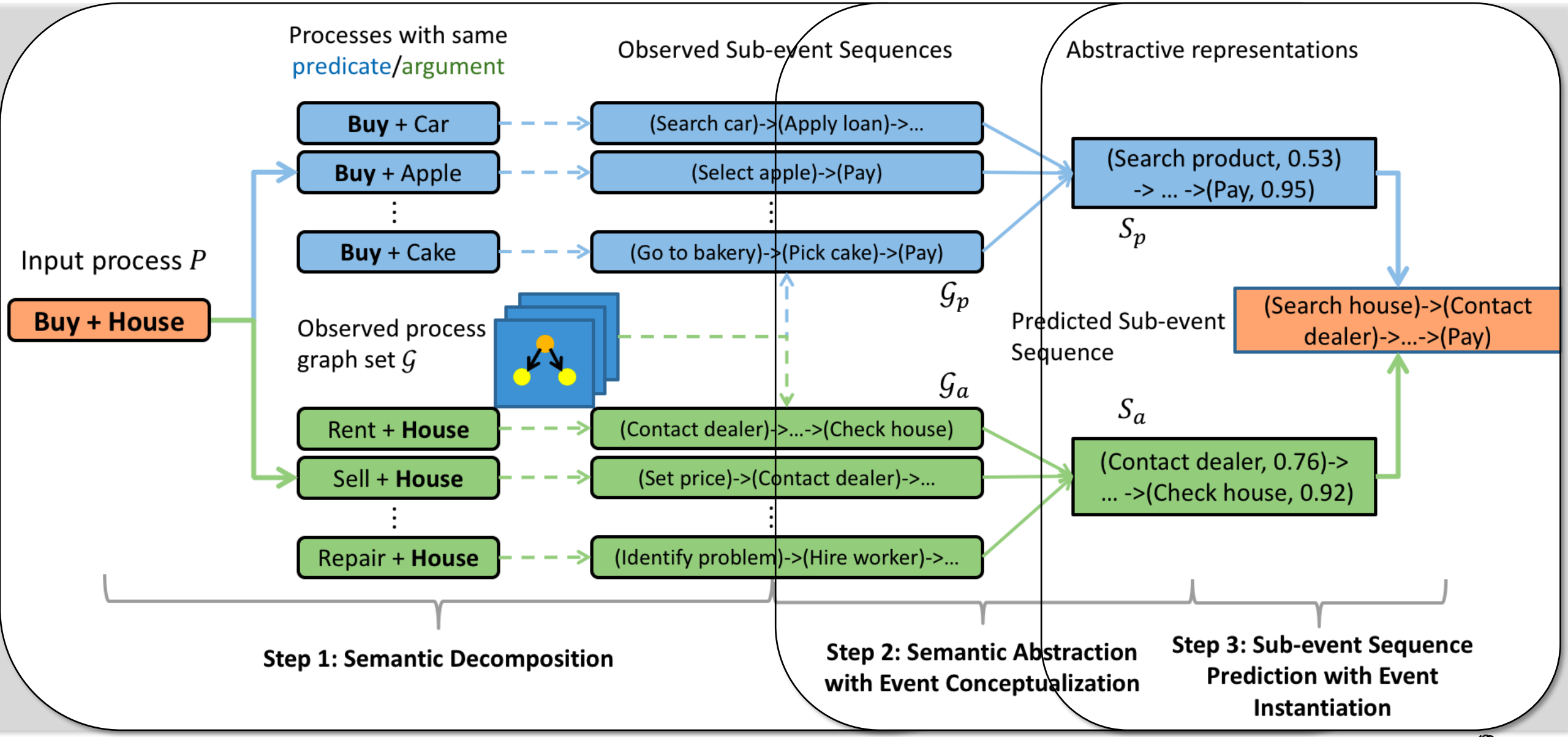
Analogous Event Process Induction

Can we perform *de novo* process induction?

Zhang, et al. Analogous Process Structure Induction for Sub-event Sequence Prediction. EMNLP, 2020



Analogous Event Process Induction



Evaluation Based on wikiHow Event Processes

Model	String Match		Hypernym Allowed	
	E-ROUGE1	E-ROUGE2	E-ROUGE1	E-ROUGE2
Random	2.9165	0.4664	23.5873	8.1089
Seq2seq (GloVe)	5.0323	1.4965	27.8710	13.0946
Seq2seq (RoBERTa)	4.5455	0.4831	28.0032	12.8502
Top one similar process (Jaccard)	8.8589	5.1000	28.6548	14.6231
Top one similar process (GloVe)	9.8797	5.1452	29.4203	13.6001
Top one similar process (RoBERTa)	9.2599	4.7390	30.6599	15.8417
Analogous Process Structure Induction (APSI)	**14.8013**	**6.6045**	**36.1648**	**19.2418**
Human	29.0189	15.2542	50.4647	29.4423

(a) Basic Setting (for each sub-event, we only predict and evaluate the verb)

Model	String Match		Hypernym Allowed	
	E-ROUGE1	E-ROUGE2	E-ROUGE1	E-ROUGE2
Random	0.0000	0.0000	0.5104	0.0903
Seq2seq (GloVe)	0.1935	0.0534	0.9677	0.1069
Seq2seq (RoBERTa)	0.4870	0.0000	1.7857	0.2899
Top one similar process (Jaccard)	0.6562	0.2257	2.4797	0.5867
Top one similar process (GloVe)	0.8750	0.2106	2.8801	0.7372
Top one similar process (RoBERTa)	0.9479	0.3009	3.2811	0.9929
Analogous Process Structure Induction (APSI)	**3.4988**	**0.4513**	**6.1611**	**1.1885**
Human	11.6351	5.5905	18.0034	8.2695

(b) Advanced Setting (for each sub-event, we predict and evaluate all words)

Quantitative results

Process Name: **Treat Pain**

References: ('learn cause'->'identify symptom'->'see doctor')
('identify cause'->'learn injury'->'recognize symptom'->`recognize symptom')

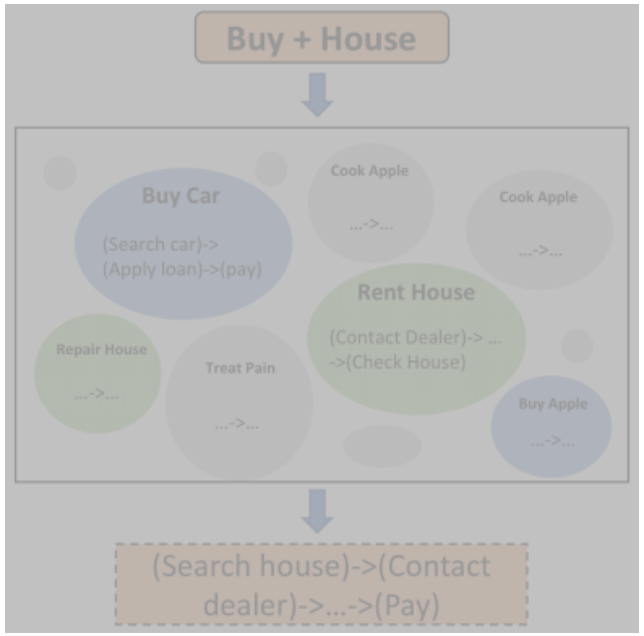
APSI Prediction: ('Identify symptom'->'see doctor'->'recognize symptom'->'take supplement')

Qualitative results

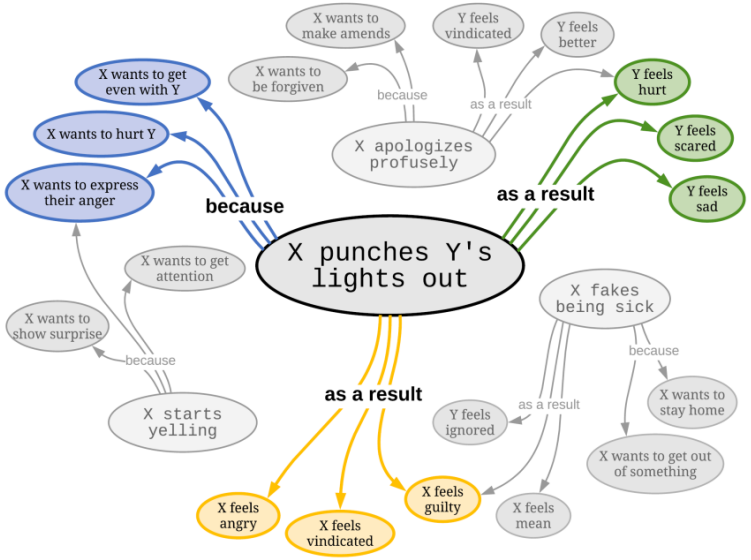
Resources are available at https://cogcomp.seas.upenn.edu/page/publication_view/910

Agenda

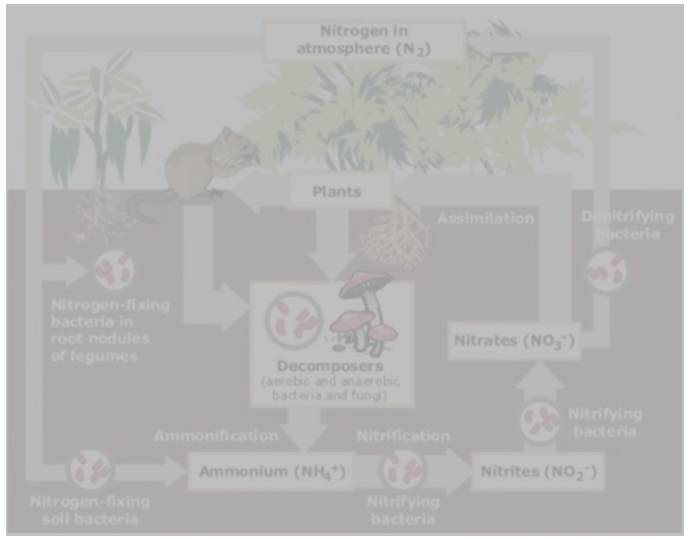
1. Event process completion



2. Event intention prediction



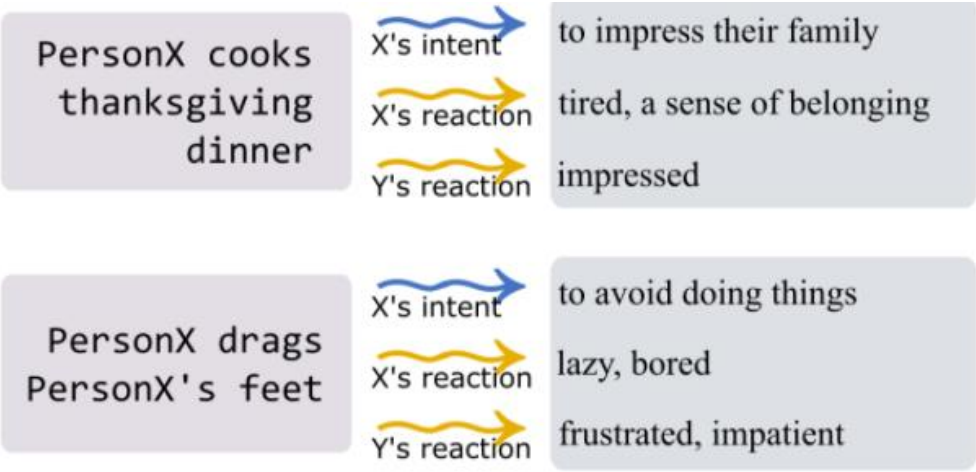
3. Event processes in downstream NLU tasks

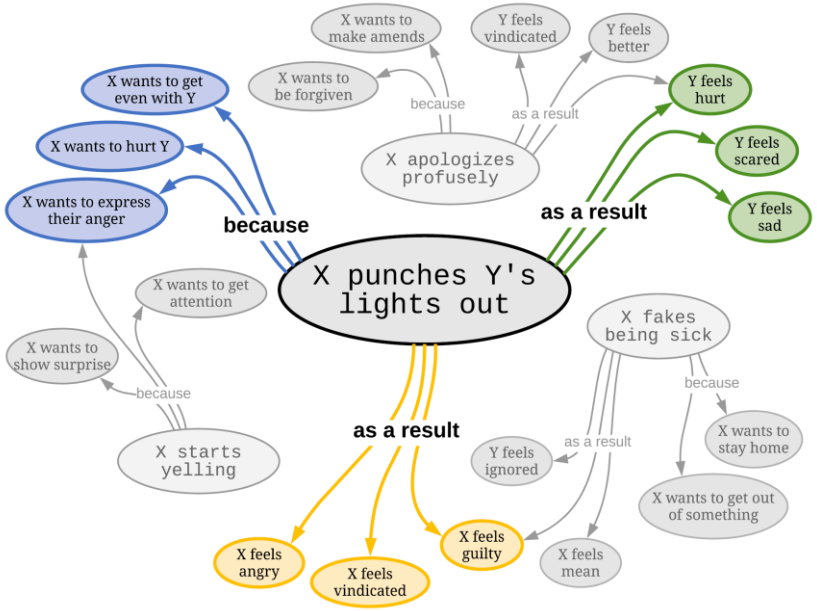


4. Open Research Directions

Intention Prediction for Events

People can easily anticipate the intents and possible reactions of participants in an event.





A commonsense-aware system should also perform such prediction.

Event2Mind – A learning system that understands stereotypical intents and reactions to events (Rashkin et al. ACL-18)

Event2Mind

Is developed based on large crowdsourced corpora:

- 25,000 events
- Free-form descriptions of their intents and reactions

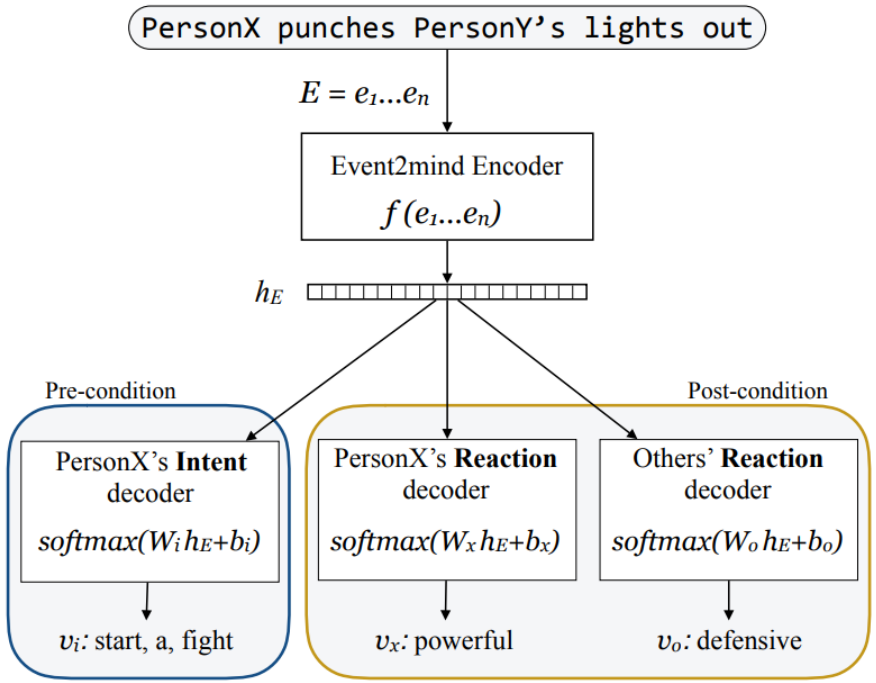
Performs Seq2NGram generation:

PersonX cooks steak

PersonX's intent: ["steak", "to kill their hunger", "to make dinner for the family", "to eat steak"]

PersonX's reaction: ["excited", "accomplished", "proud", "full"]

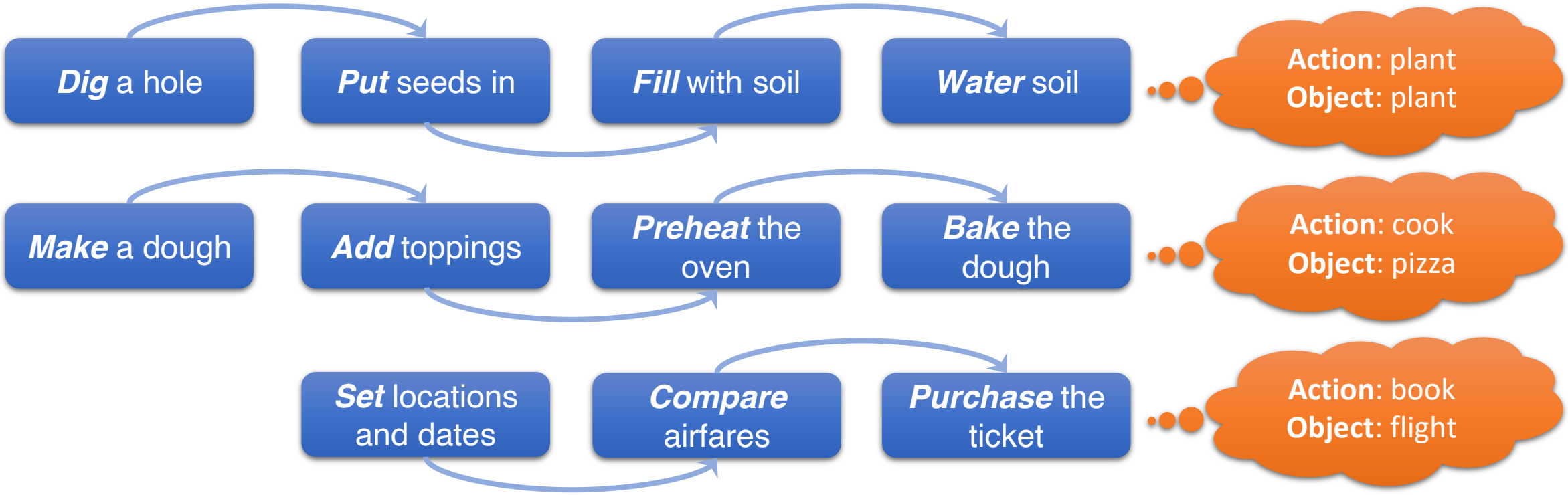
Other people's reaction: ["none", "happy", "person x cooked well."]



More follow-ups of Event2Mind

- ATOMIC: An Atlas of Machine Commonsense for If-Then Reasoning (Sap+ AAAI 2019)
- COMET: Commonsense Transformers for Automatic Knowledge Graph Construction (Bosselut+, ACL-19)

Intention Prediction for Event Processes



Event processes are directed by the **central goal**, or the **intention** of its performer [Zacks+, Nature Neuroscience 2001].

- Inherent to human's common sense.
- Missing from current computational methods.
- Important to machine commonsense reasoning, summarization, schema induction, etc.

A New Task: Multi-axis Event Process Typing

Chen et al. “What are you trying to do?” Semantic Typing of Event Processes. CoNLL-2020

A new (cognitively motivated) **semantic typing task** for understanding event processes in natural language. Two **type axes**:

- What ***action*** the event process seeks to take? (**action type**)
- What type of **object**(s) it should affect? (**object type**)

This research also contributes with

- A **large dataset** of typed event processes (>60k processes)
- A **hybrid learning framework** for event process typing based on **indirect supervision**

A Large Event Process Typing Dataset

A large dataset of typed event processes from wikiHow

- 60,277 event processes with free-form labels of action and object types

A challenging typing system

- Diversity: 1,336 action types and 10,441 object types (in free froms)
- Few-shot cases: 85.9% labels appear less than 10 times, (~half 1-shot).
- External labels: in 91.2% (84.2%) processes, the action (object) type label does not appear in the process body.

A non-trivial learning problem with **ultra fine-grained** and **extremely few-shot** labels.

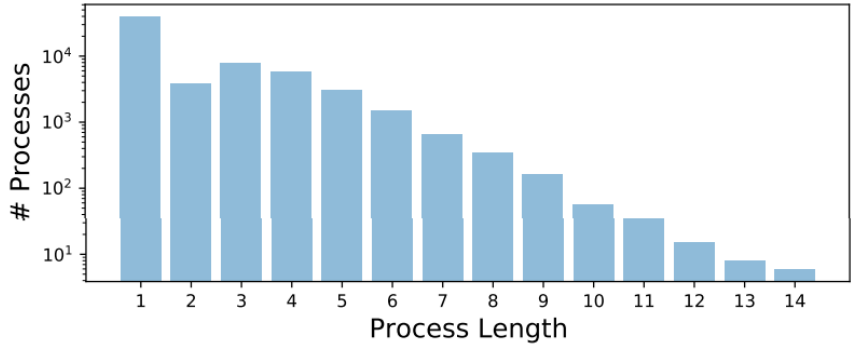


Figure 2: Distribution of process lengths.

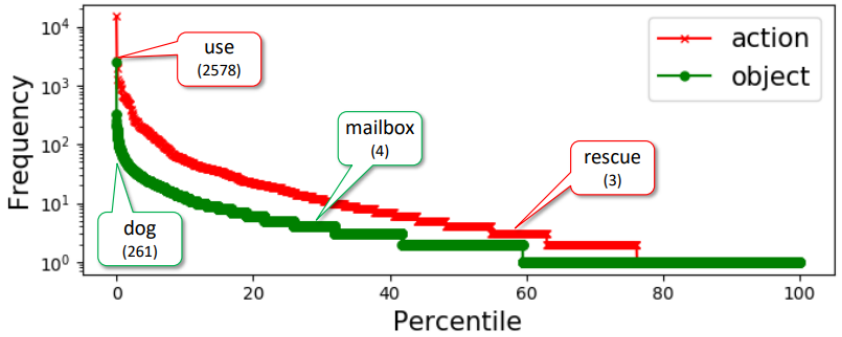


Figure 3: Distribution of actions and objects. Number of frequencies are shown in the brackets.

Indirect Supervision from Gloss Knowledge

An event process

Directly inference
(Difficult)

Make
Cocktail

Labels

Indirect inference
(Much Easier)

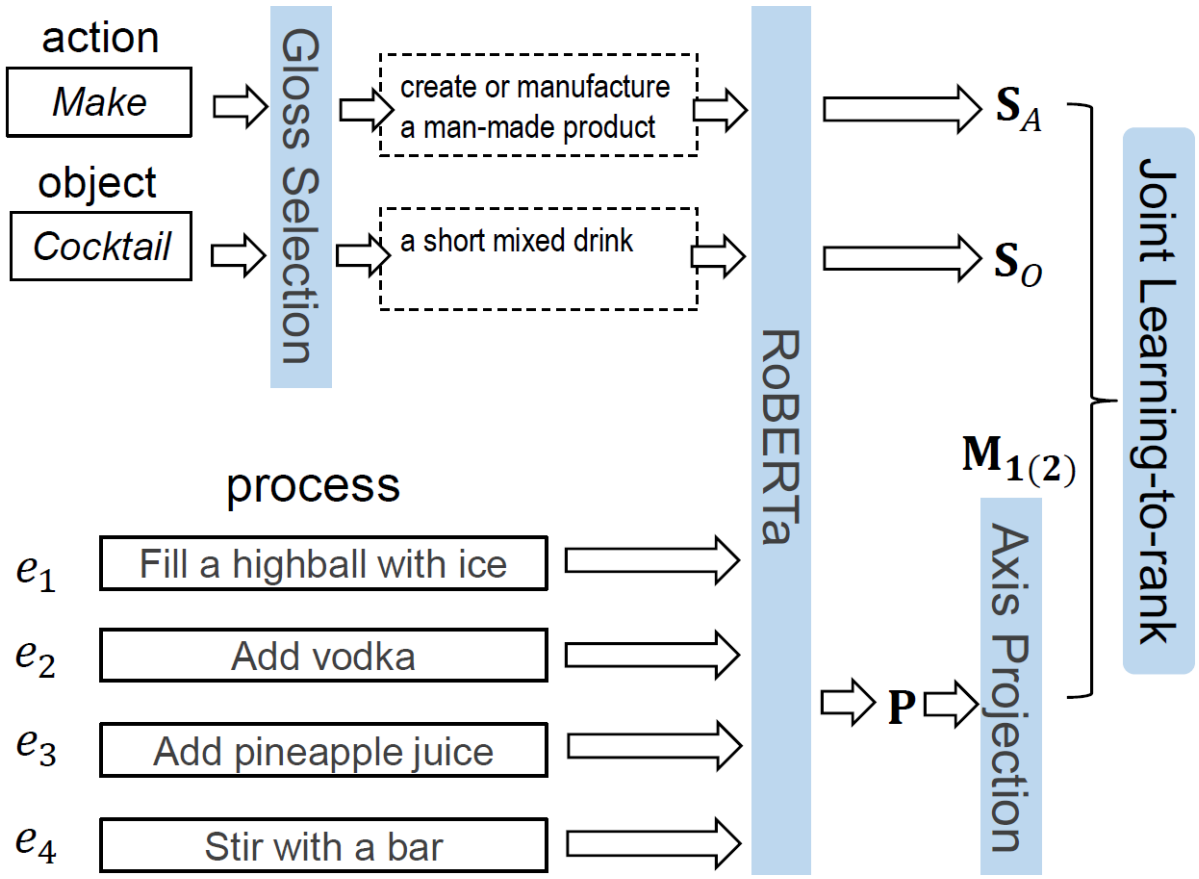
Make	create or manufacture a man-made product
Cocktail	a short, mixed drink

Label glosses (from WordNet)

Why using label glosses?

- Semantically richer than labels themselves
- Capturing the association of a process-gloss pair (two sequences) is much easier
- Jump-starting few-shot label representations (and benefiting with fairer prediction)

Indirect Supervision from Gloss Knowledge



How to represent the process?

- RoBERTa encodes concatenated event contents (VERB and ARG1).

How to represent a label?

- The same RoBERTa encodes the label gloss

Which gloss for a polysemous label?

- WSD [Hadiwinoto+, EMNLP-19]
- MFS (Most frequent sense)

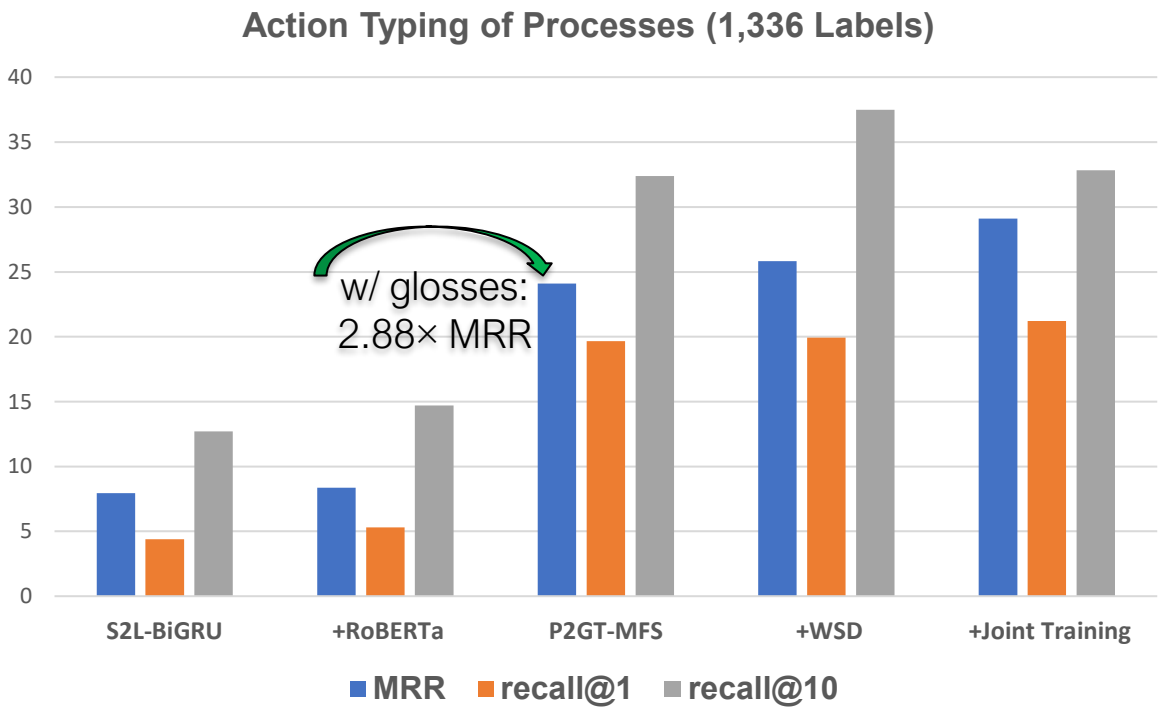
Learning objective?

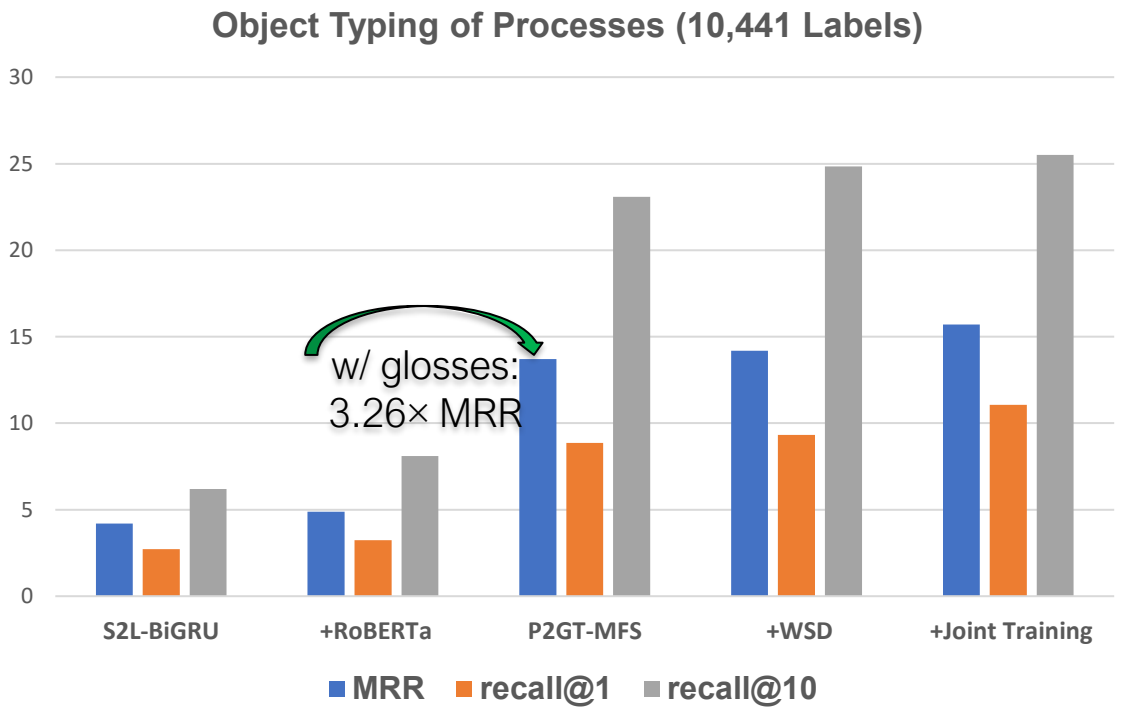
- Joint **learning-to-rank** for both type axes (different projection)

Inference?

- Ranking all glosses for all labels in the vocab

Results





- Gloss knowledge brings along the most improvement (2.88~3.26 folds of MRR)
- Joint training indicates the effectiveness of leveraging complementary supervision signals
- Sense selection (WSD) leads to lesser improvement (predominant senses are representative enough)

Case Study

Event processes	Predictions
Make explosive materials ⇒ Obtain a container ⇒ Obtain shrapnel ⇒ Install a trigger	*A*: **detonate**, **assemble**, blacken *O*: **grenade**, **blaster**, **mine**
Go to DMV ⇒ Take photos ⇒ Take vision test ⇒ Take permit test ⇒ Take road test	*A*: **obtain**, *verify*, explore *O*: **license**, check, visa
Ignore order ⇒ Enter area ⇒ Enforce blockade ⇒ Force to retreat from area	*A*: **conquer**, *disarm*, **invade** *O*: *barrier*, **soldier**, **fortress**
Capture two opposition posts ⇒ Kill many fighters ⇒ Destroy three armed trucks ⇒ Confiscate artillery guns	*A*: *kill*, **demolish**, **fight** *O*: *melee*, **conflict**, **stronghold**
Cooperate with the counsel investigation ⇒ Open his remarks ⇒ Apologize many times ⇒ Try to restore public trust	*A*: *respond*, disagree, **accept** *O*: *apology*, *disagreement*, **slander**
Travel in a presidential motorcade ⇒ Be shot once in the back ⇒ Be taken to hospital ⇒ Be pronounced dead	*A*: *survive*, **die**, tackle *O*: **assassin**, crash, *roadkill*
Give advance notice ⇒ Give notice ⇒ Issue dividends	*A*: **honor**, **pay**, **reward** *O*: *finance*, **equity**, **subsidy**
Target quotes ⇒ Target shares quotes ⇒ Ask to clarify offer ⇒ Challenge to merge agreement ⇒ Challenge to merge businesses	*A*: **compare**, **maximize**, **negotiate** *O*: *prospectus*, **quote**, **settlement**
Clean windows ⇒ Buy plants ⇒ Hang pictures ⇒ Paint walls ⇒ Carpet floors	*A*: **redecorate**, **decorate**, *refurbish* *O*: **room**, **bedroom**, *makeover*

Table 3: Case study for typing event processes in the news domain. The predictions are given by Joint P2GT-WSD trained on our full dataset. Each case is given top 3 predictions on both axes, whereof reasonably correct ones are boldfaced, and relevant ones are italic. Few-shot labels appearing up to 10 times in our dataset are in blue.

System Demonstration

A web demonstration of our prototype system is running at https://cogcomp.seas.upenn.edu/page/demo_view/step

Examples

Decoration

Event process (choose an example or write the subevents of a process separated by '@' to get its intention)

clean windows @ buy plants @ paint walls @ hang pictures @ carpet floors @ reorganize furniture

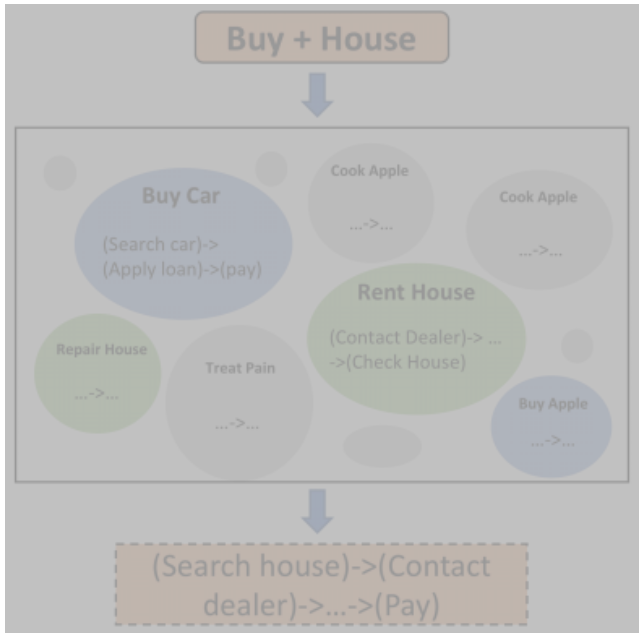
Get intention >

redecorate room

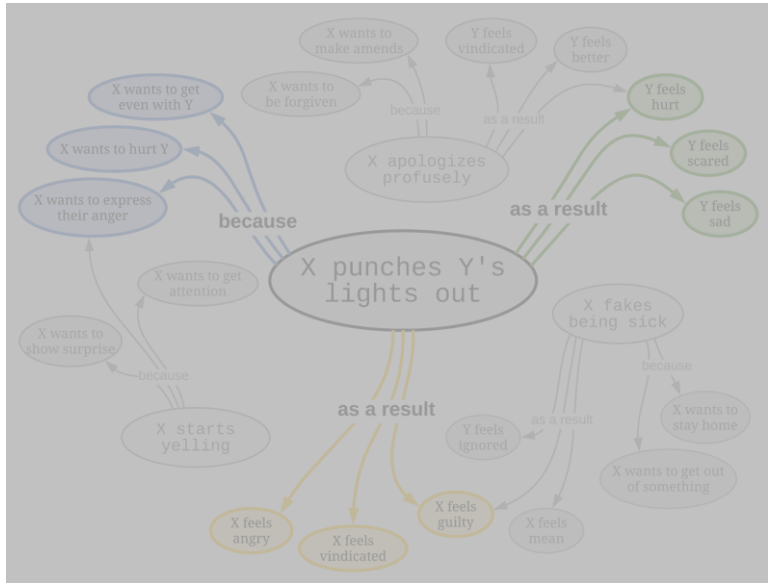
Cosine similarity	Action	Object	Cosine similarity
0.678	redecorate	room	0.623
0.650	stage	atmosphere	0.599
0.500	brighten	mosaic	0.589
0.427	preoccupy	suite	0.574
0.418	furnish	interior	0.573

Agenda

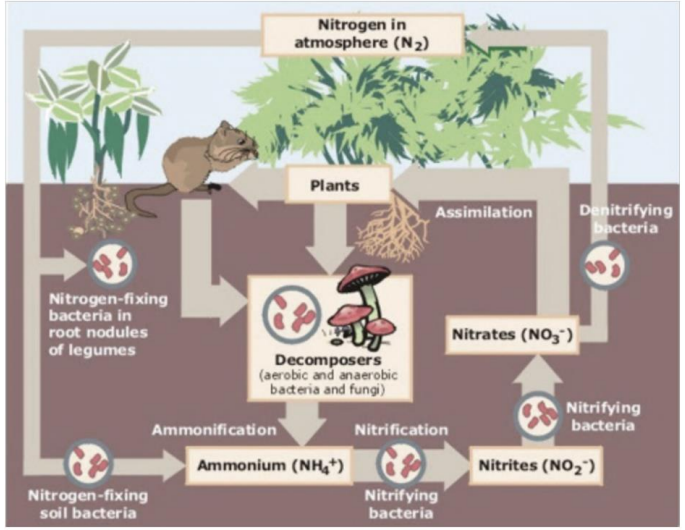
1. Event process completion



2. Event intention prediction



3. Event processes in downstream NLU tasks



4. Open Research Directions

Narrative Prediction

The ROC Story Narrative Cloze Test [Mostafazadeh+, NAACL 2016]**:**

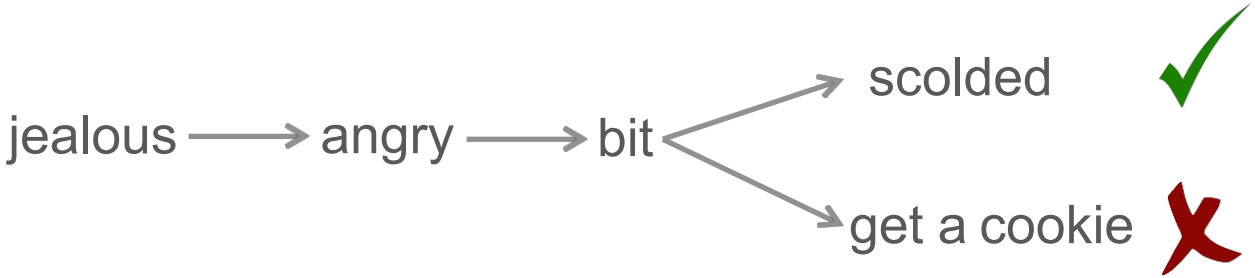
> One day Wesley's auntie came over to visit. He was happy to see her, because he liked to play with her. When she started to give his little sister attention, he got **jealous**. He got **angry** at his auntie and **bit** his sister's hand when she wasn't looking.

Then what might happen?

O1: He was **scolded**.
O2: She **gave him a cookie** for being so nice.

Chaturvedi, et al (EMNLP, 2017) train a language model that captures three types of sequential features:

1. **Event sequences** in 20 years of NYT data



2. Sentiment trajectories
3. Sentential topical consistency

Features	Accuracy
All	74.4%
Event-sequence	71.6%
Sentiment	64.5%
Topic	55.2%

Event sequences are most important.

Machine Reading Comprehension

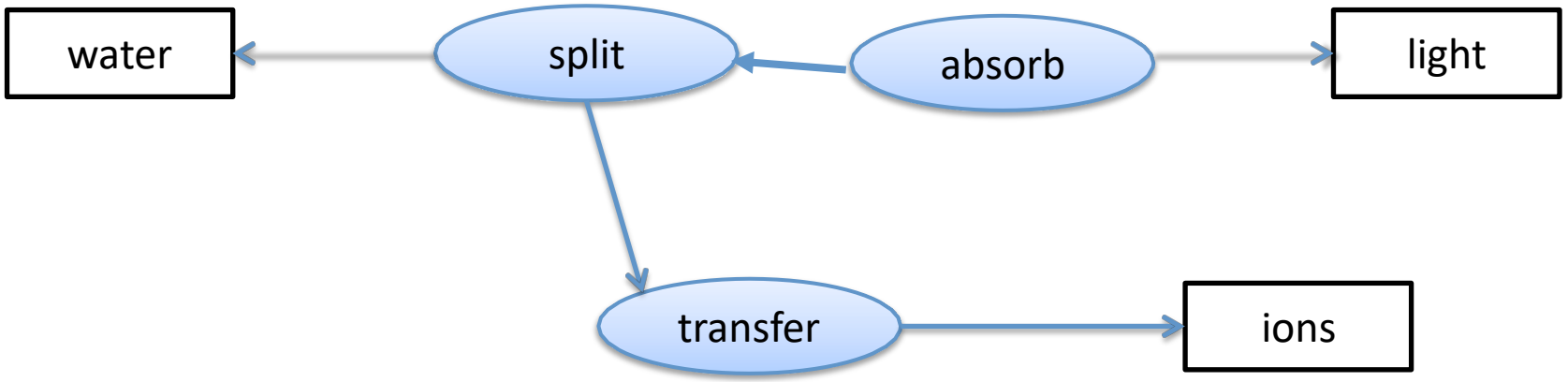
QA based on articles in biology

Water is split, providing a source of electrons and protons (hydrogen ions, H^+) and giving off O_2 as a by-product. ***Light absorbed*** by chlorophyll drives a ***transfer of the electrons and hydrogen ions*** from water to an acceptor called $NADP^+$.

What can the splitting of water lead to?

A: Light absorption **B:** Transfer of ions

1. Extracting events and event-event relations from articles



2. Matching questions and candidate answers with extracted event processes

Berant, et al. Modeling Biological Processes for Reading Comprehension. EMNLP, 2014 (Best Paper Award)

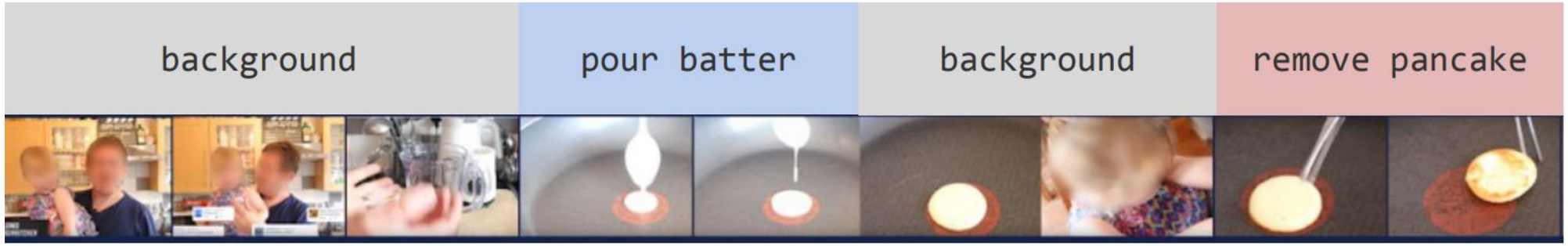
Video Segmentation

Events in a process as anchors of video segments.

wikiHow process: make pancakes: {add egg, add flour, …, pour batter, remove pancake}

Video segments:

Video:

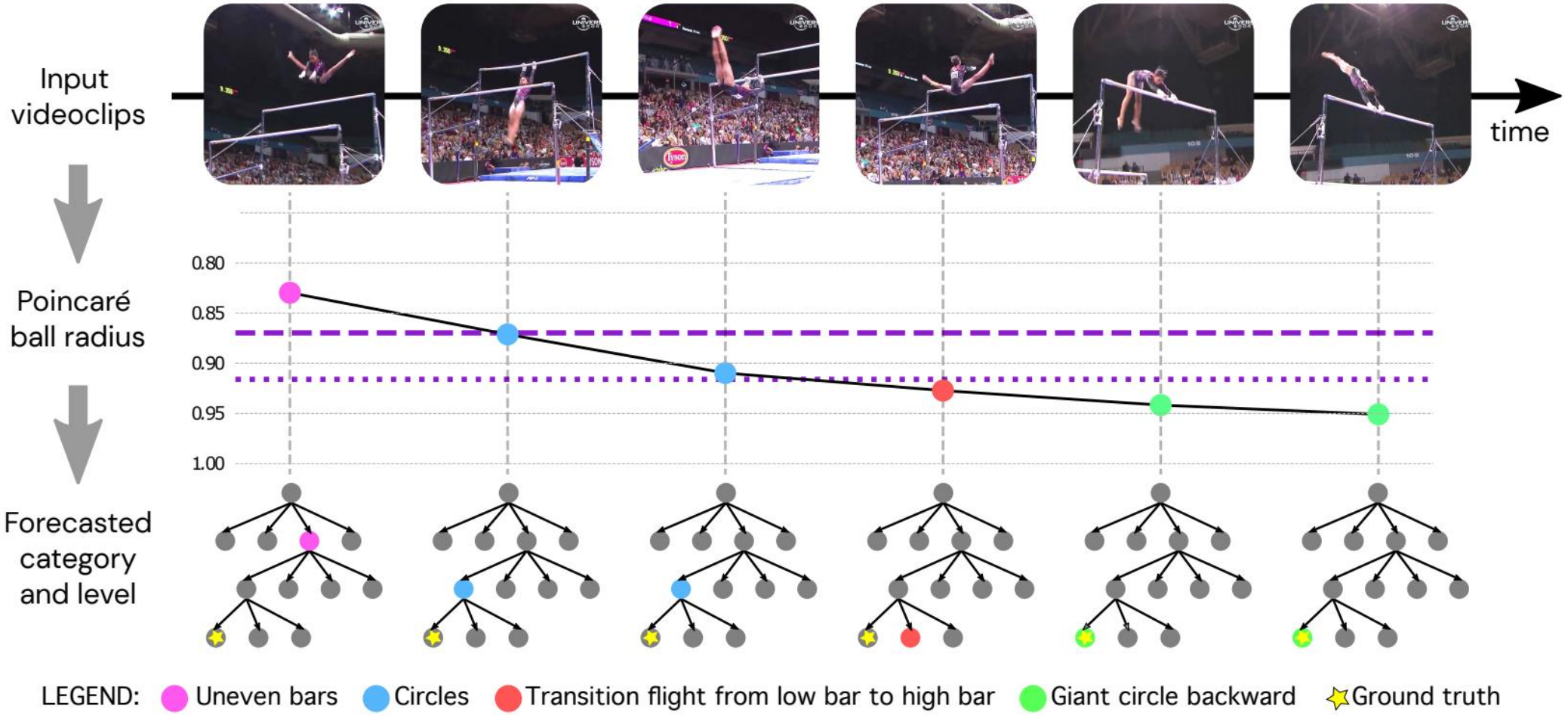


Video narration: *hey folks here welcome to my kitchen … pour a nice-sized amount … change the angle to show … and take it out*

Alignment learning between video narration and wikiHow event processes help action segmentation in videos.

Zhukov et al. Cross-task weakly supervised learning from instructional videos. CVPR 2019
Fried et al. Learning to Segment Actions from Observation and Narration. ACL 2020

Future Event Prediction in Videos

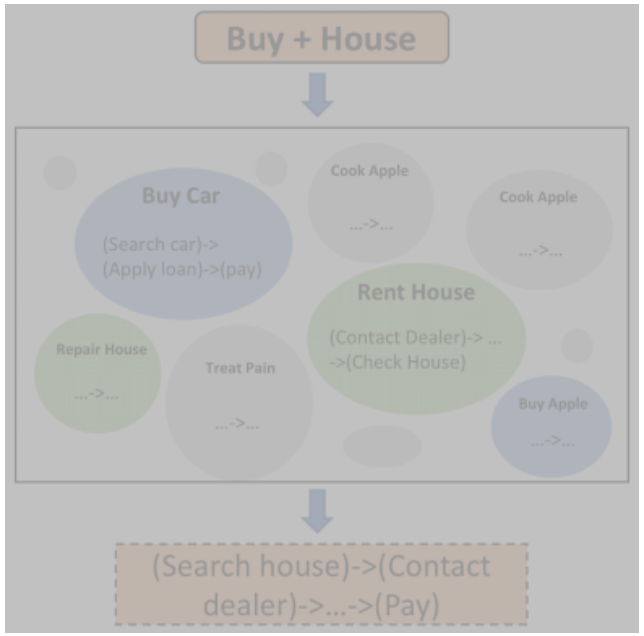


Hyperbolic embeddings model hierarchies of possible event evolution processes in videos.

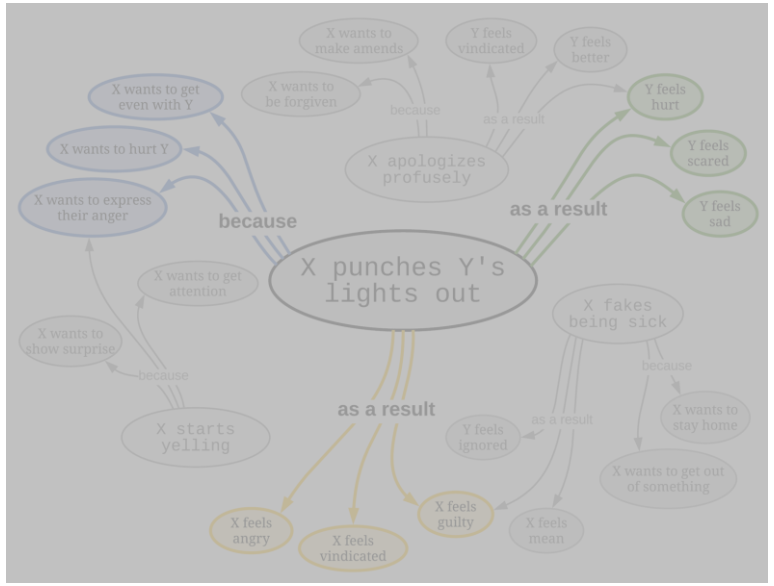
Surís et al. Learning the Predictability of the Future. arXiv:2101.01600, 2021

Agenda

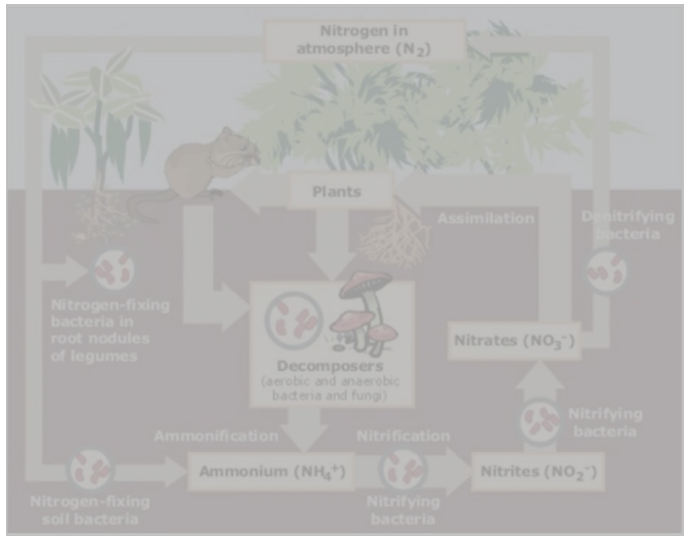
1. Event process completion



2. Event intention prediction



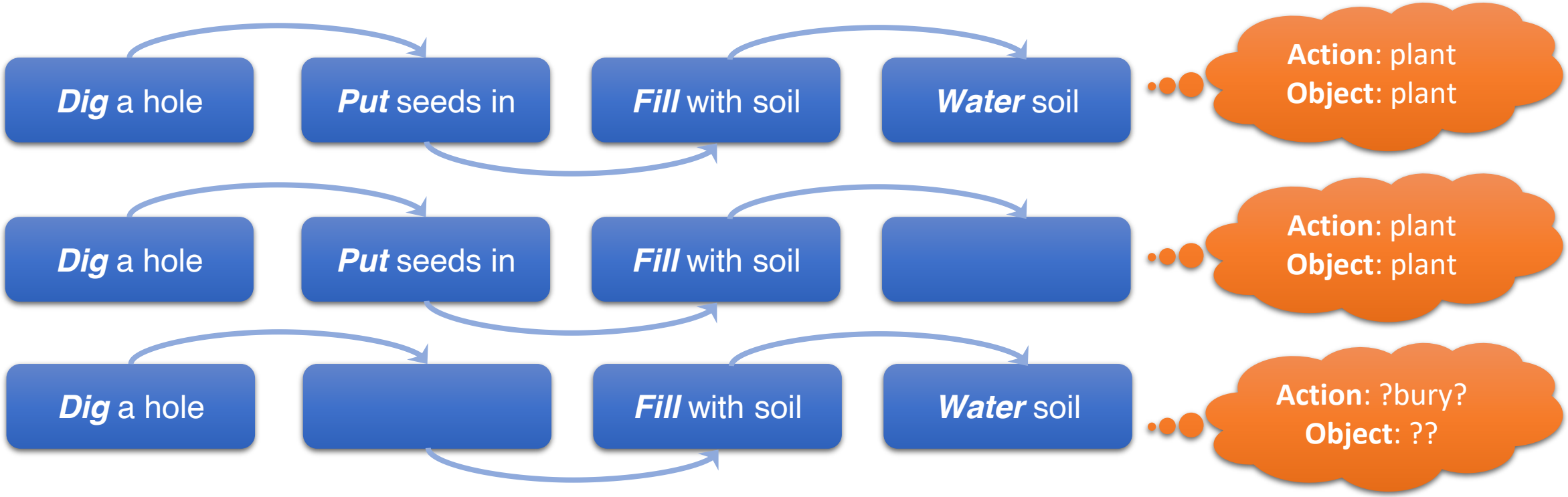
3. Event processes in downstream NLU tasks



4. Open Research Directions

Salience Detection in Event Processes

Events in a process are not equally important



Defending your dissertation is essential; **Doing a TAShip** is less important; **Doing an internship** is optional…

Is there a way to automatically identify salient events in a process?
Would that help downstream tasks such as abstractive summarization?

Reasoning About Ordering

Identifying the order of member events in a process is an unresolved challenge

Heavy snow is causing disruption to transport across the UK, with heavy rainfall bringing flooding to the south-west of England. Rescuers searching for a woman trapped in a landslide at her home in Looe, Cornwall, said they had found a body.

Q1: What events have already finished?
A: searching trapped landslide said found
Q2: What events have begun but has not finished?
A: snow causing disruption rainfall bringing flooding
Q3: What will happen in the future?
A: No answers.
warm-up

Q4: What happened before a woman was trapped?
A: landslide
Q5: What had started before a woman was trapped?
A: snow rainfall landslide
Q6: What happened while a woman was trapped?
A: searching
Q7: What happened after a woman was trapped?
A: searching said found
User-provided

Ning, et al. TORQUE: A Reading Comprehension Dataset of Temporal Ordering Questions. EMNLP, 2020

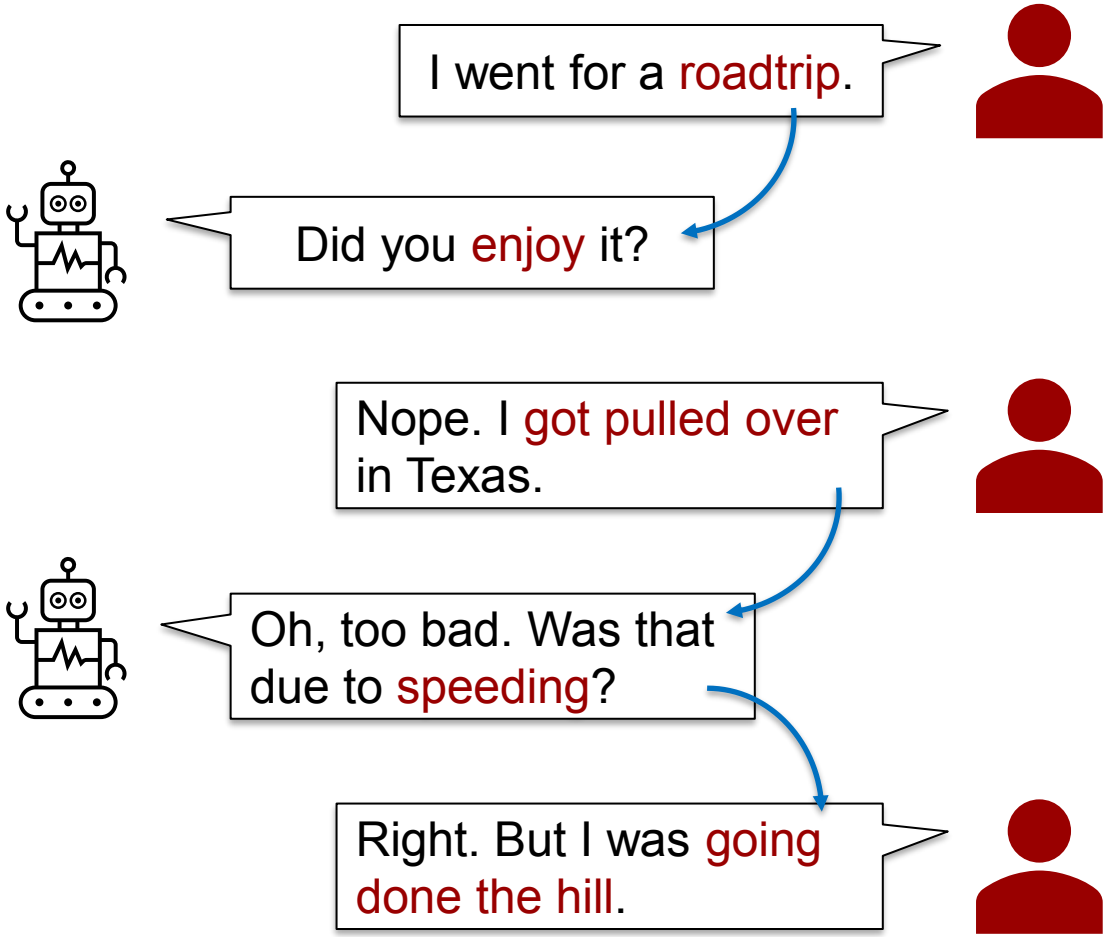
- 3.2k news snippets with 21k human-generated questions querying temporal relationships

Lyu, et al. Reasoning about Goals, Steps, and Temporal Ordering with WikiHow. EMNLP, 2020

- A wikiHow-based testbed about event ordering (and more)

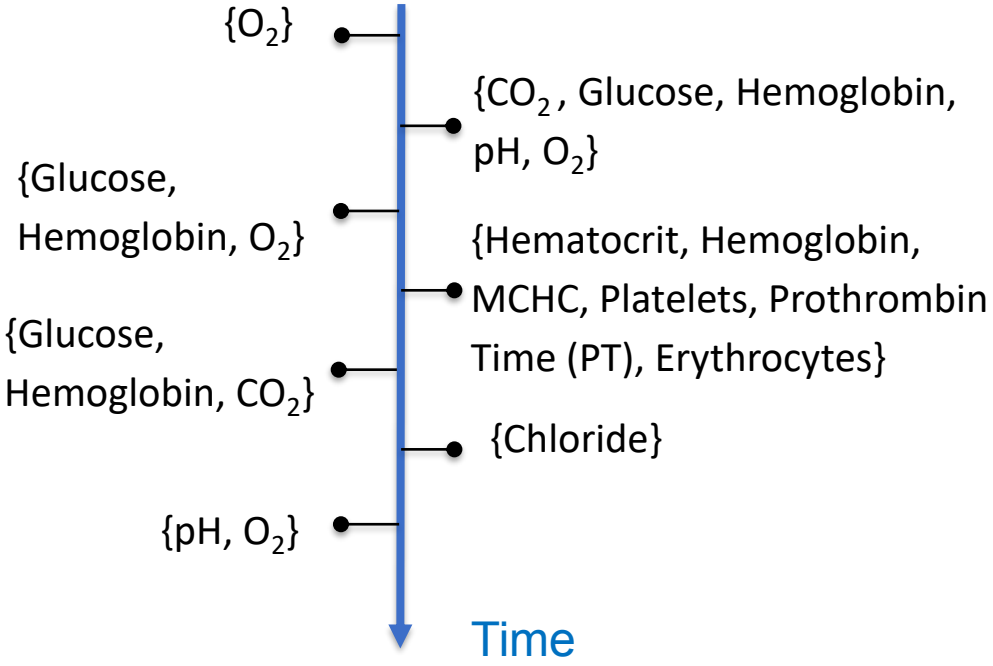
More Tasks

Chatbots



Can event processes improve the consistency of utterance generation/retrieval?

Understanding clinical event processes



Diagnostic prediction (Zhang et al. AIME-20), phenotype prediction, …

- Transfer learning can be important (naturally lack of data)
- Structured prediction can be important (dependency of phenotypes, disease labels)

References

- Zack, et al. Human brain activity time-locked to perceptual event boundaries. Nature neuroscience, 4(6):651–655. 2001
- Chambers and Jurafsky. Unsupervised learning of narrative event chains. ACL, 2008
- Radinsky and Horvitz. Mining the Web to Predict Future Events. WSDM, 2013
- Berant, et al. Modeling Biological Processes for Reading Comprehension. EMNLP, 2014
- Chaturvedi, et al. Story comprehension for predicting what happens next. EMNLP, 2017
- Rashkin, et al. Event2Mind: Commonsense Inference on Events, Intents, and Reactions. ACL, 2018
- Liu, et al. Automatic event salience identification. EMNLP, 2018
- Zhukov et al. Cross-task weakly supervised learning from instructional videos. CVPR, 2019
- Zhang, et al. Analogous Process Structure Induction for Sub-event Sequence Prediction. EMNLP, 2020
- Chen, et al. “What are you trying to do?” Semantic typing of event processes. CoNLL, 2020
- Ning et al. TORQUE: A Reading Comprehension Dataset of Temporal Ordering Questions. EMNLP, 2020
- Lyu, et al. Reasoning about Goals, Steps, and Temporal Ordering with WikiHow. EMNLP, 2020
- Jindai, et al. Is Killed More Significant than Fled? A Contextual Model for Salient Event Detection. COLING, 2020
- Fried, et al. Learning to Segment Actions from Observation and Narration. ACL, 2020
- Zhang, et al. Diagnostic Prediction with Sequence-of-sets Representation Learning for Clinical Events. AIME, 2020
- Surís, et al. Learning the Predictability of the Future. arXiv:2101.01600, 2021

Thank You

Event and Commonsense

Event-centric Natural Language Understanding (Part IV)

Hongming Zhang

HKUST/UPenn

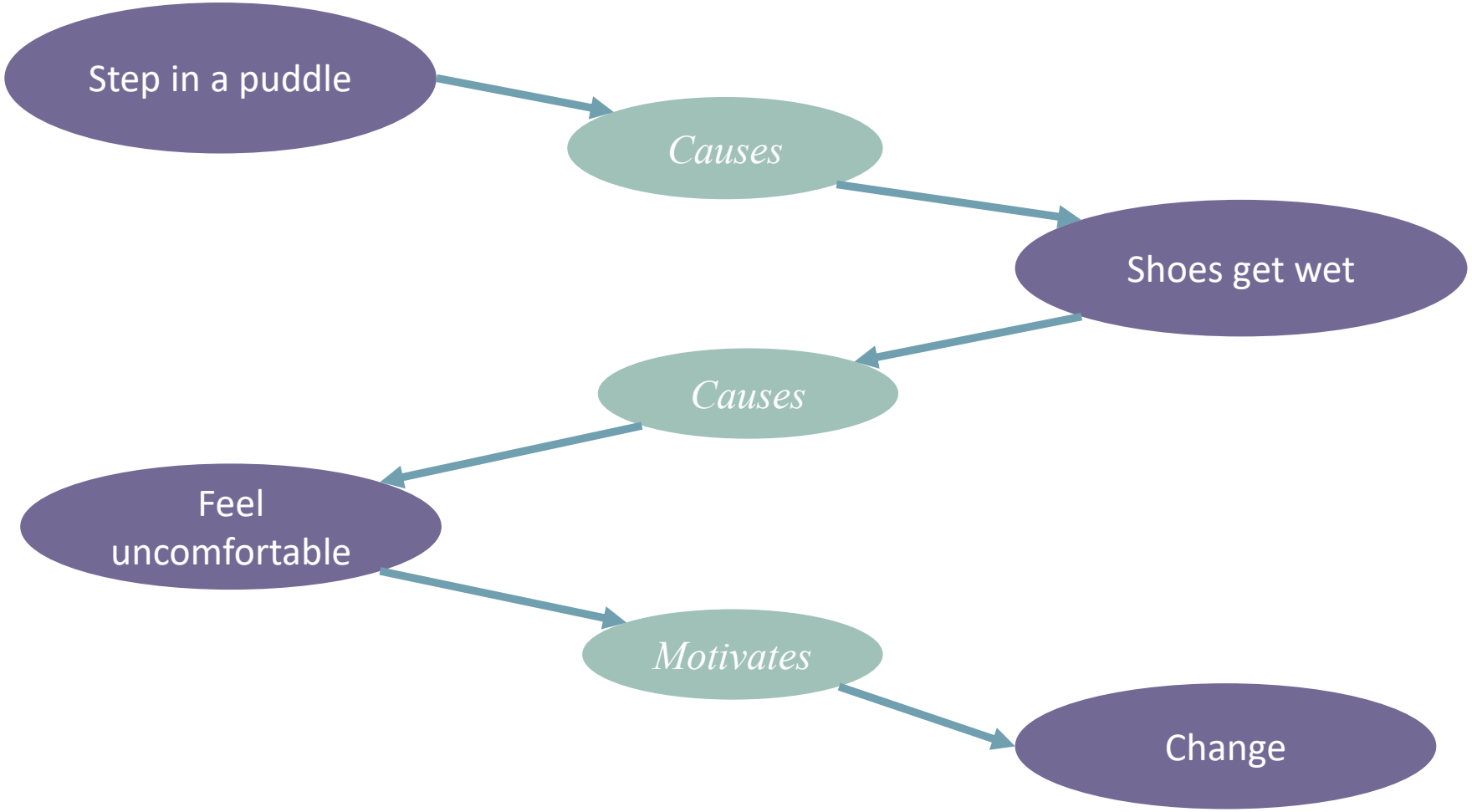
Feb 2021

AAAI Tutorials

Event-centric Natural Language Understanding

Commonsense is crucial for NLU

Example: John stepped in a puddle and had to go home to change.



Outline

- Understanding Commonsense from the Angle of Events
- Instance-level Event Knowledge Acquisition
  - Human Annotation
  - Automatic Event Knowledge Extraction
  - Language Modeling
- Schema-level Event Knowledge Acquisition
- Conclusion

Commonsense Knowledge

- Modern Definition of Commonsense Knowledge (Liu & Singh, 2004)
  - "While to the average person the term 'commonsense' is regarded as synonymous with 'good judgement'"
  - "the AI community it is used in a technical sense to refer to the millions of basic facts and understandings possessed by most people."
  - "Commonsense is about preference and not always true"
    - If you forget someone's birthday, they may be unhappy with you.
    - But if your friends understand that you are busy, he will not by angry.

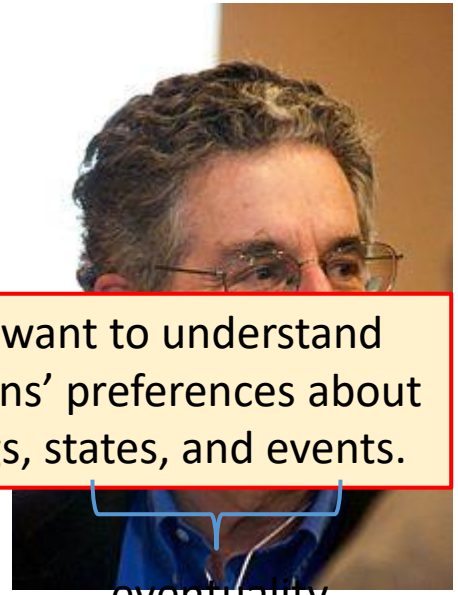
Unlike factual knowledge, they are not inevitably true.

Commonsense is about preference.

Hugo Liu and Push Singh, ConceptNet - a practical commonsense reasoning tool-kit, BTTJ, 2004

What kinds of preference?

- Semantic meaning in our language can be described as "a finite set of mental primitives and a finite set of mental combination." (Jackendoff, 1990)
- The primitive units of semantic meanings include
  - Thing (or entity)
    - cat
  - State
    - The cat is cute.
    - The cat is smiling.
  - Event
    - The cat is running.



We want to understand humans' preferences about things, states, and events.

eventuality

States describe things.

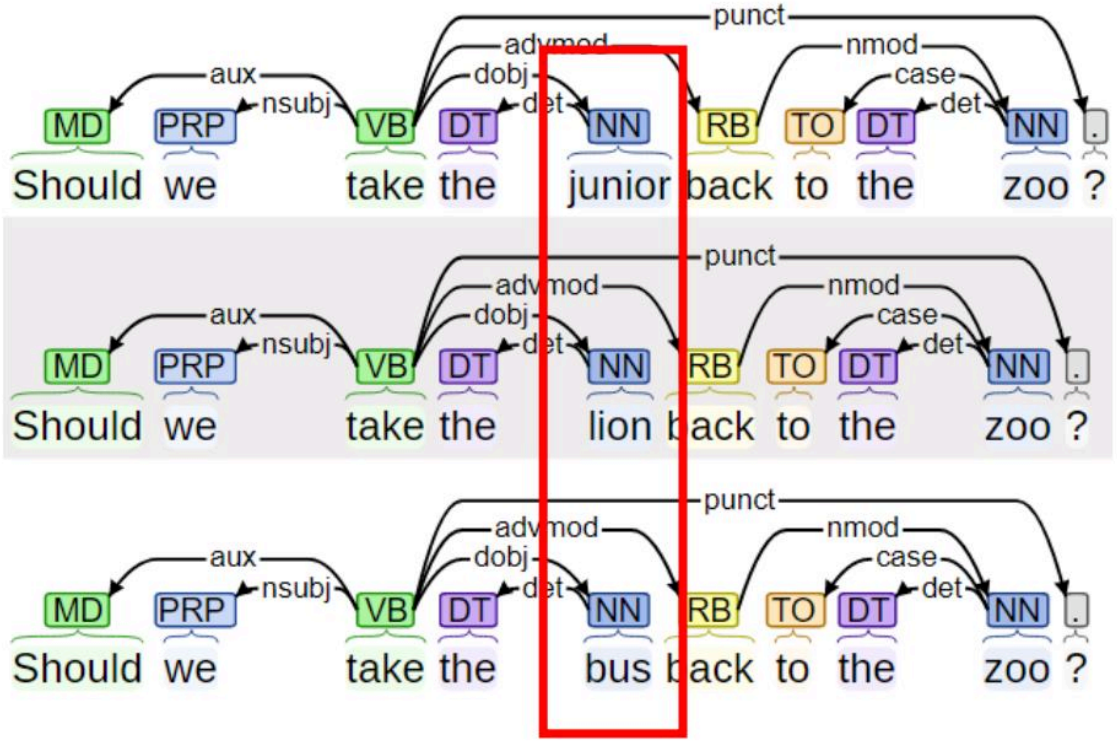
Events describe the changing of states.

Ray Jackendoff Semantic Structures. Cambridge, Massachusetts: MIT Press. 1990

How to represent the preference?

- The lower bound of a semantic theory (Katz and Fodor, 1963)
  - Linguistic description – grammar = semantics
  - Understanding language needs both "the speaker's knowledge of his language and his knowledge about world" (Katz and Fodor, 1963)

It is so dangerous!!!



When the grammar is controlled, the selection we made can reflect our understanding about the world.

Katz, J. J., & Fodor, J. A. (1963). The structure of a semantic theory. Language, 39(2), 170–210.

Selectional Preference

- Selectional Preference (Resnik, 1993)
  - A relaxation of selectional restrictions (Katz and Fodor, 1963) and is often used as syntactic features (Chomsky, 1965).
  - Applied to IsA hierarchy in WordNet and verb-object relations.
  - With this formulation, we can easily use the frequency/plausibility scores of different combinations to reflect humans' preference.

  - Examples:
    - ("Cat" -IsA- "Animal") > ("Cat" -IsA- "Plant")
    - ("eat" -dobj- "food") > ("eat" -dobj- "rock")

Katz, J. J., & Fodor, J. A. (1963). The structure of a semantic theory. Language, 39(2), 170–210.
Noam Chomsky. 1965. Aspects of the Theory of Syntax. MIT Press, Cambridge, MA.
Philip Resnik. 1993. Selection and information: A class-based approach to lexical relationships. Ph.D. thesis, University of Pennsylvania.

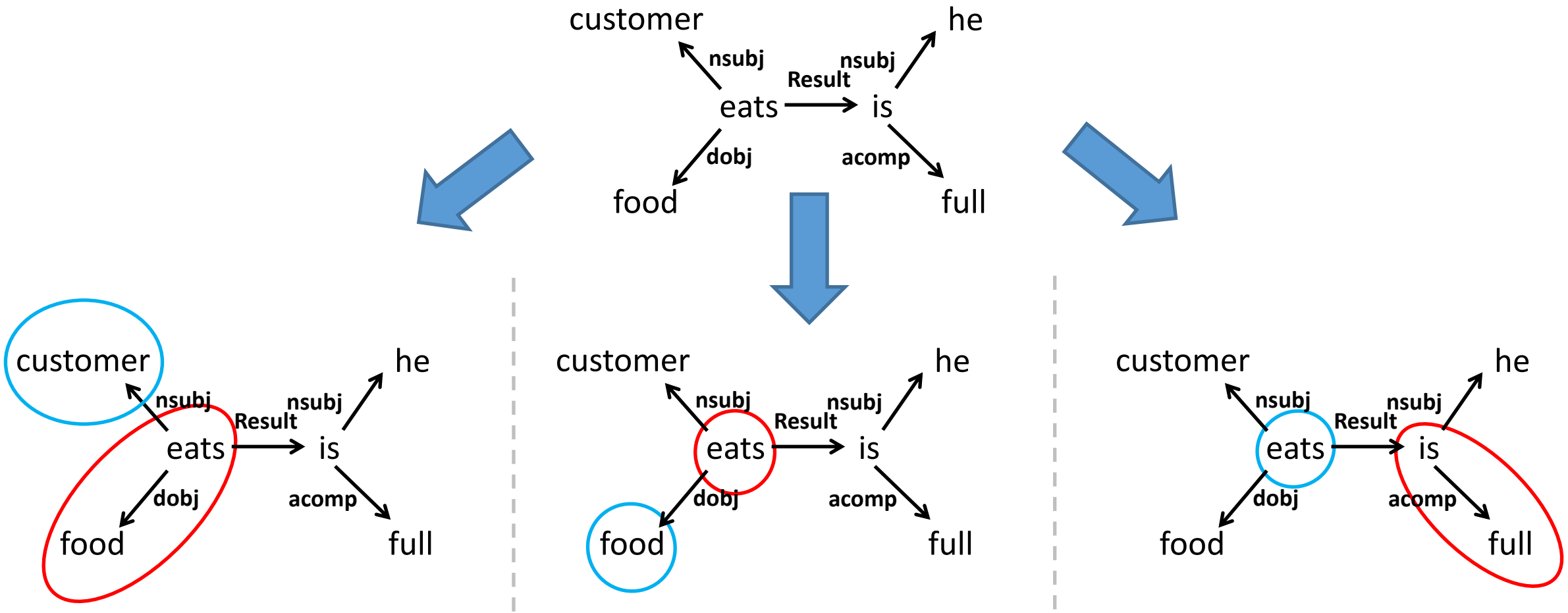
Higher-order Selectional Preference

- First-order
  - dobj: (“eat”–>dobj->“food”) > (“eat”->dobj->“house”)
  - Nsubj: (“sing”->nsubj->“singer”) > (“sing”->nsubj->“house”)
  - …
- Second-order (Zhang et al., 2019)
  - Nsubj-amod / dobj-amod
  - (“eat”->nsubj->“[SUB]”->amod->“hungry”) > (“eat”->dobj>“[OBJ]”->amod->“hungry”)
- Higher-order
  - (“I eat dinner”->Causes->“I am full”) > (“I eat dinner”->Causes->“I am hungry”)

Commonsense can be represented by the higher-order selectional preference over eventualities.

Hongming Zhang, Hantian Ding, and Yangqiu Song. SP-10K: A Large-Scale Evaluation Set for Selectional Preference Acquisition. ACL 2019.

Transferability from event knowledge to Commonsense



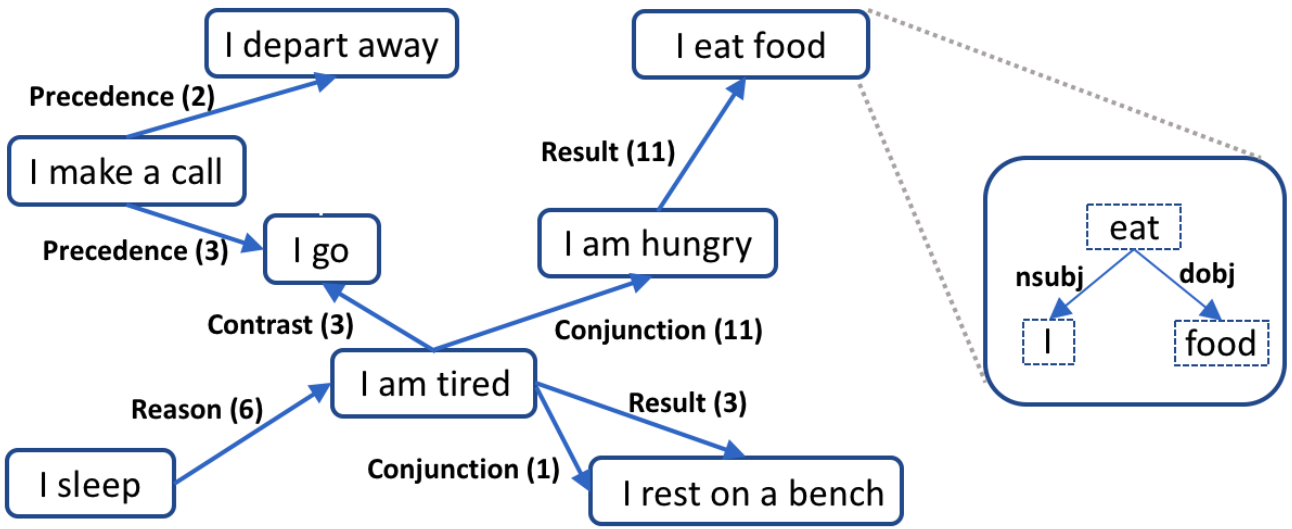
('customer'-*CapableOf*-'eat food')

('food'-*ReceivesAction*-'eat')

('eat'-*Causes*-'be full')

Hongming Zhang, Daniel Khashabi, Yangqiu Song, and Dan Roth. TransOMCS: From Linguistic Graphs to Commonsense Knowledge. IJCAI 2020.

Transferability from event knowledge to Commonsense



Event-centric KG

"human" *CapableOf*

1. stand
2. think
3. die
4. learn
5. make mistake
6. lie
7. typically have 🤔
8. create society
9. have cell
10. create life

"love" *Causes*

1. be friendly
2. be happy
3. pain
4. marriage
5. be quaint 🤔
6. be unhappy
7. be allergic 🤔
8. be desperate
9. be apart
10. be silly

Human-defined commonsense

Outline

- Understanding Commonsense from the Angle of Events
- Instance-level Event Knowledge Acquisition
  - Human Annotation
  - Automatic Event Knowledge Extraction
  - Language Modeling
- Schema-level Event Knowledge Acquisition
- Conclusion

Event-centric KBs

	# Events	# Event relation	# Relation Types
FrameNet (Baker et al., 1998)	27,691	1,709	7
ACE (Aguilar et al., 2014)	3,290	0	0
PropBank (Palmer et al., 2005)	112,917	0	0
NomBank (Meyers et al., 2004)	114,576	0	0
TimeBank (Pustejovsky et al., 2003)	7,571	8,242	1
ConceptNet (Liu and Singh, 2004)	74,989	116,097	4
Event2Mind (Smith et al., 2018)	24,716	57,097	3
ProPora (Dalvi et al., 2018)	2,406	16,269	1
ATOMIC (Sap et al., 2019)	309,515	877,108	9
ATOMIC 2020* (Hwang et al., 2020)	-	165,164	4

Pro: High quality
Con: Expensive; Small Scale; Limited relation types

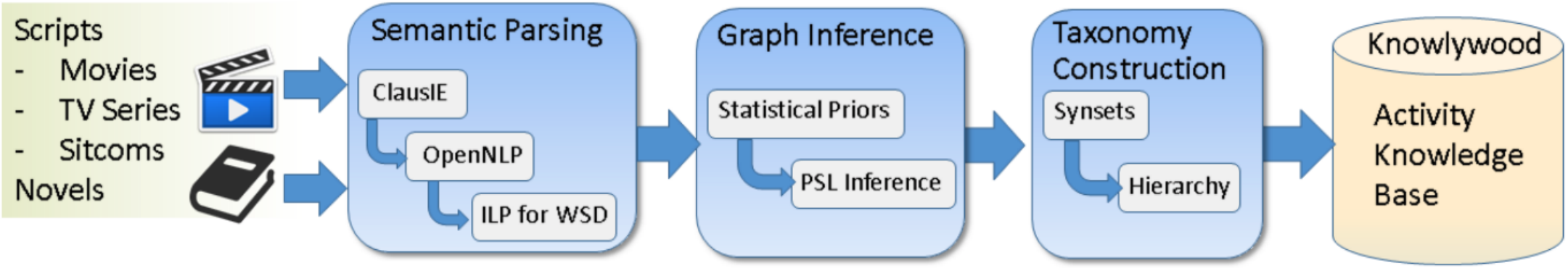
*For ATOMIC 2020, we only count the unique edges and ignore the edges it inherits from other KBs.

Outline

- Understanding Commonsense from the Angle of Events
- Instance-level Event Knowledge Acquisition
  - Human Annotation
  - Automatic Event Knowledge Extraction
  - Language Modeling
- Schema-level Event Knowledge Acquisition
- Conclusion

Knowlywood (Tandon et al., 2015)

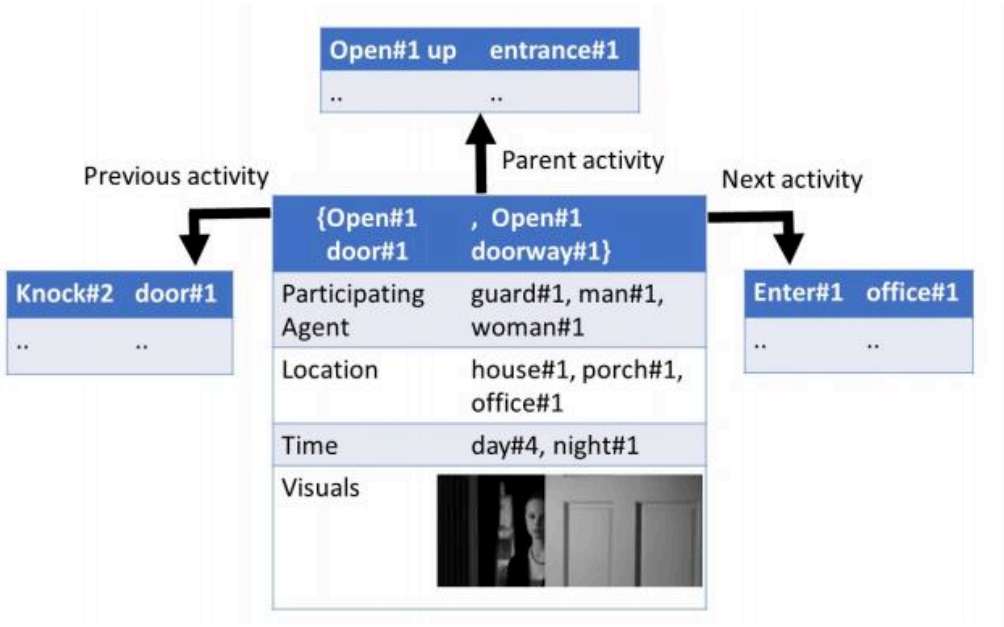
- KG Format
  - Node: Verb + Object
  - Edge: Temporal Relation
- Resource
  - 560 movie scripts
- Extraction Methodology



Niket Tandon, Gerard de Melo, Abir De, and Gerhard Weikum. 2015. Knowlywood: Mining Activity Knowledge From Hollywood Narratives. CIKM 2015.

Knowlywood

- Example



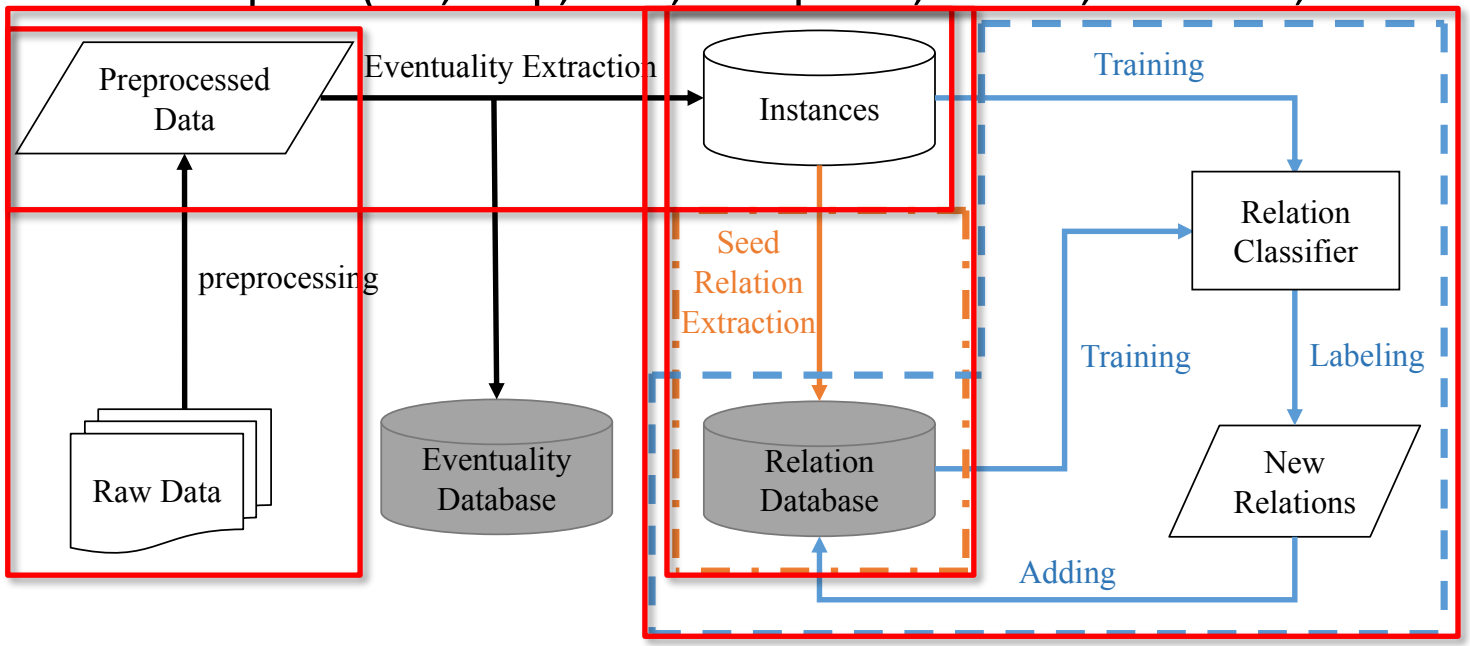
"Knock door"->"open up entrance"->"enter office"

- Quantity

Source	#Input Scripts	#Scenes	#Unique Activities	Parent	Participant	Prev	Next	Loc.	Time	Avg.
Movie scripts	560	148,296	244,789	0.87	0.86	0.78	0.85	0.79	0.79	0.84
TV series	290	886,724	565,394	0.89	0.85	0.81	0.84	0.82	0.84	0.86
Sitcoms	179	286,266	200,550	0.88	0.85	0.81	0.87	0.81	0.83	0.87
Novels	103	383,795	137,365	0.84	0.84	0.78	0.88	0.85	0.72	0.84
Crowdsrc.	25	3,701	9,575	0.82	0.91	0.91	0.87	0.74	0.40	0.86
Knowlywood	1,157	1,708,782	964,758	0.87	0.86	0.84	0.85	0.78	0.84	**0.85±0.01**
ConceptNet 5	-	-	4,757	0.15	0.81	0.92	0.91	0.33	N/A	**0.46±0.02**

ASER (Zhang et al., 2020)

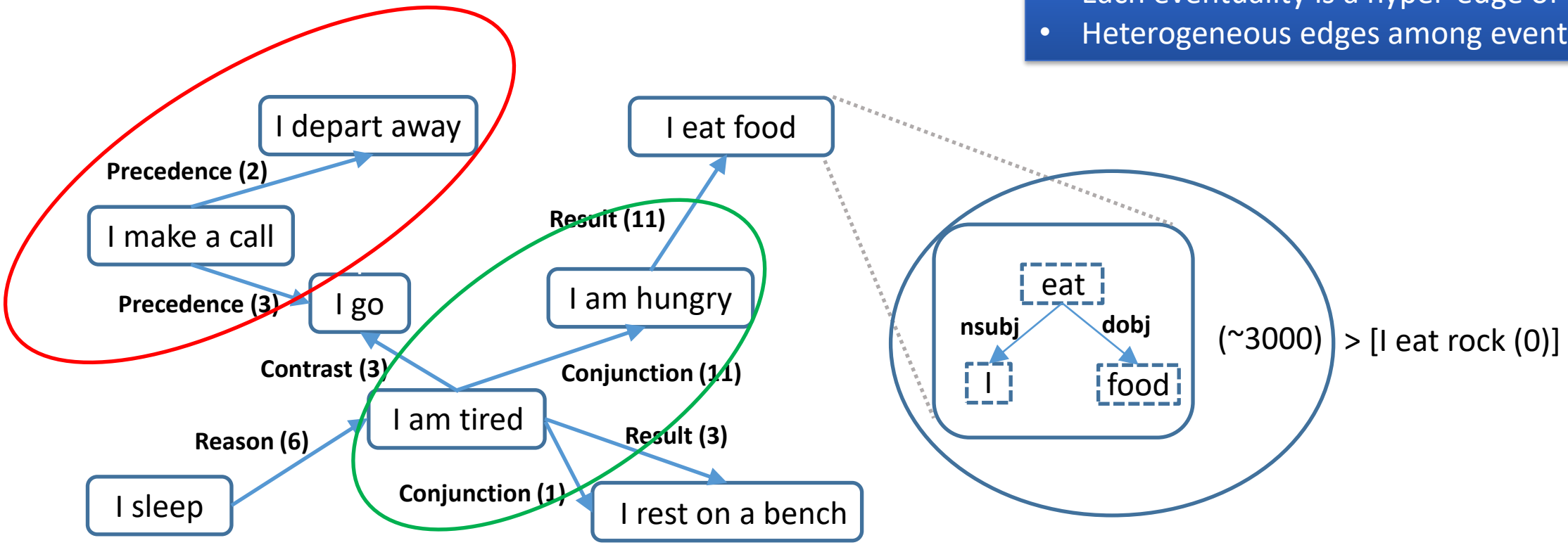
- KG Format
  - A Hybrid graph
  - Node: Eventualities in the format of dependency graphs
  - Edge: All discourse relations
- Resource
  - 11B token textual corpora (i.e., Yelp, NYT, Wikipedia, Reddit, Subtitles, E-books)
- Extraction



Hongming Zhang, Xin Liu, Haojie Pan, Yangqiu Song, and Cane Wing-Ki Leung. ASER: A Large-scale Eventuality Knowledge Graph. WWW 2020.

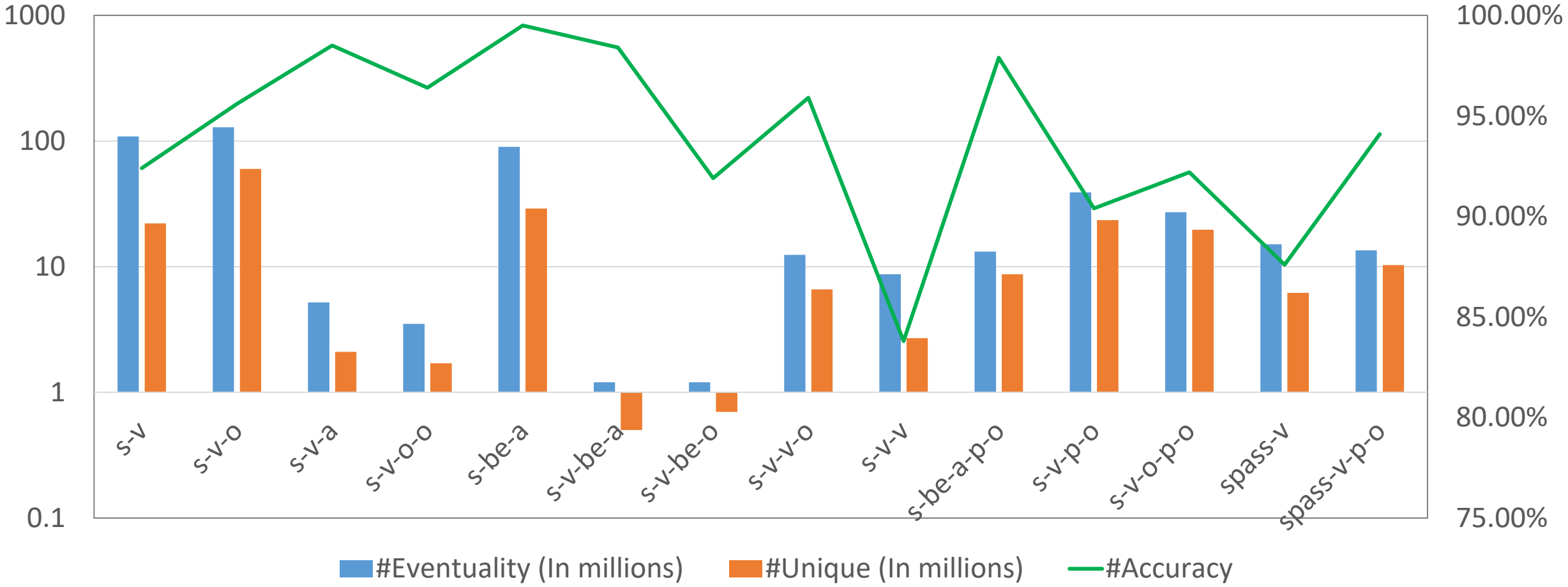
ASER Example

A hybrid graph of
- Each eventuality is a hyper-edge of words
- Heterogeneous edges among eventualities

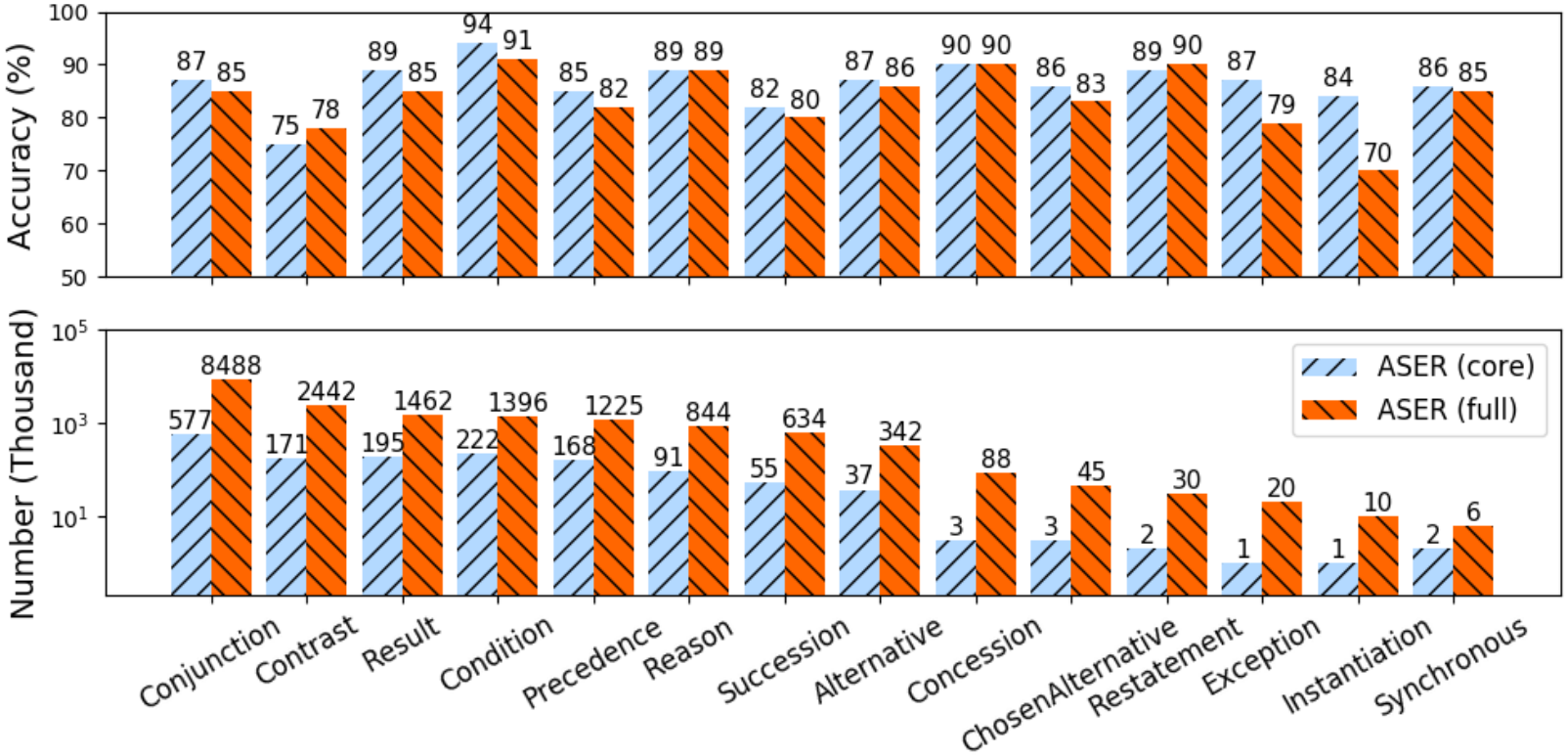


194 million eventualities, 64 million edges

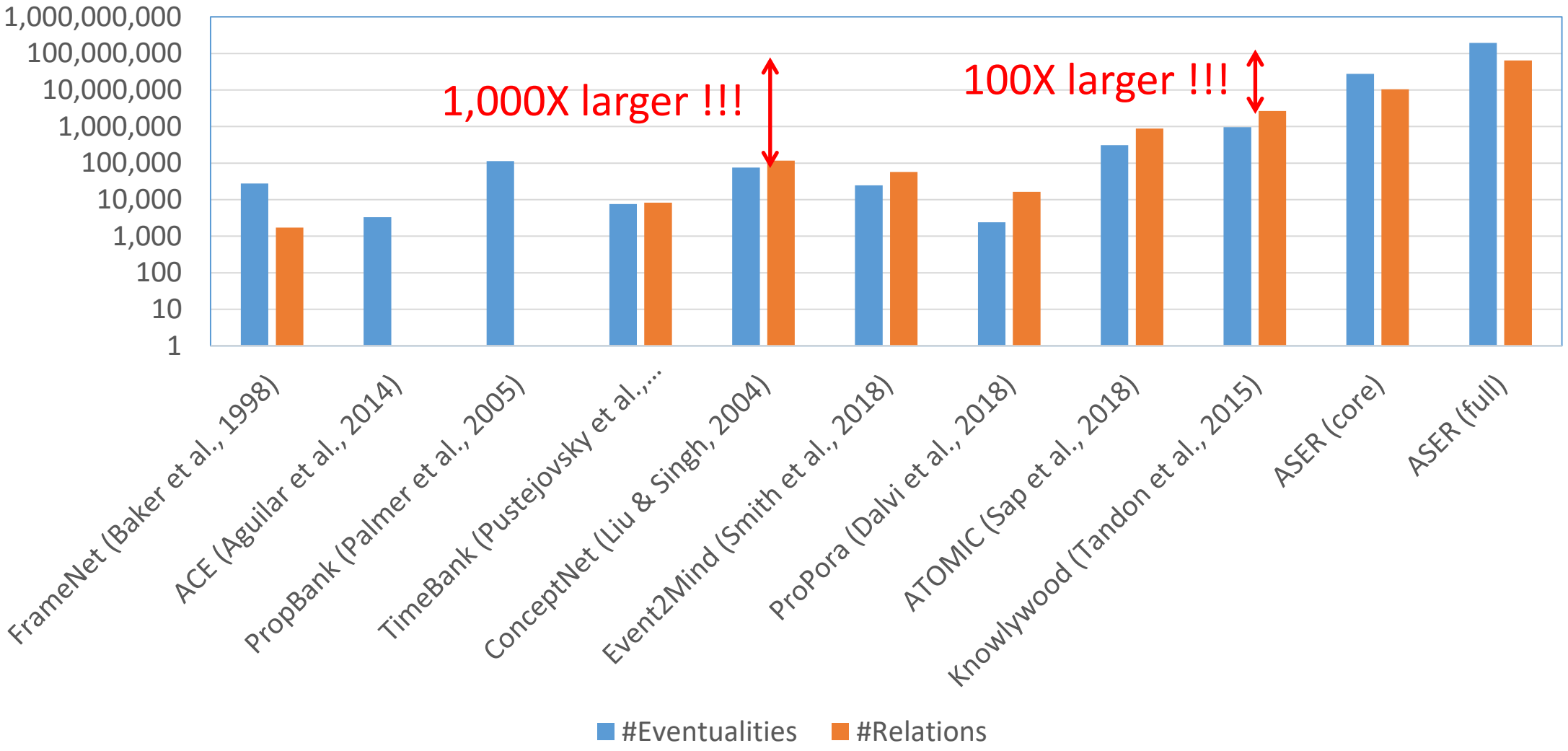
ASER Quantity and Quality (Eventuality)



ASER Quantity and Quality (Edge)



Comparison with Other event KGs



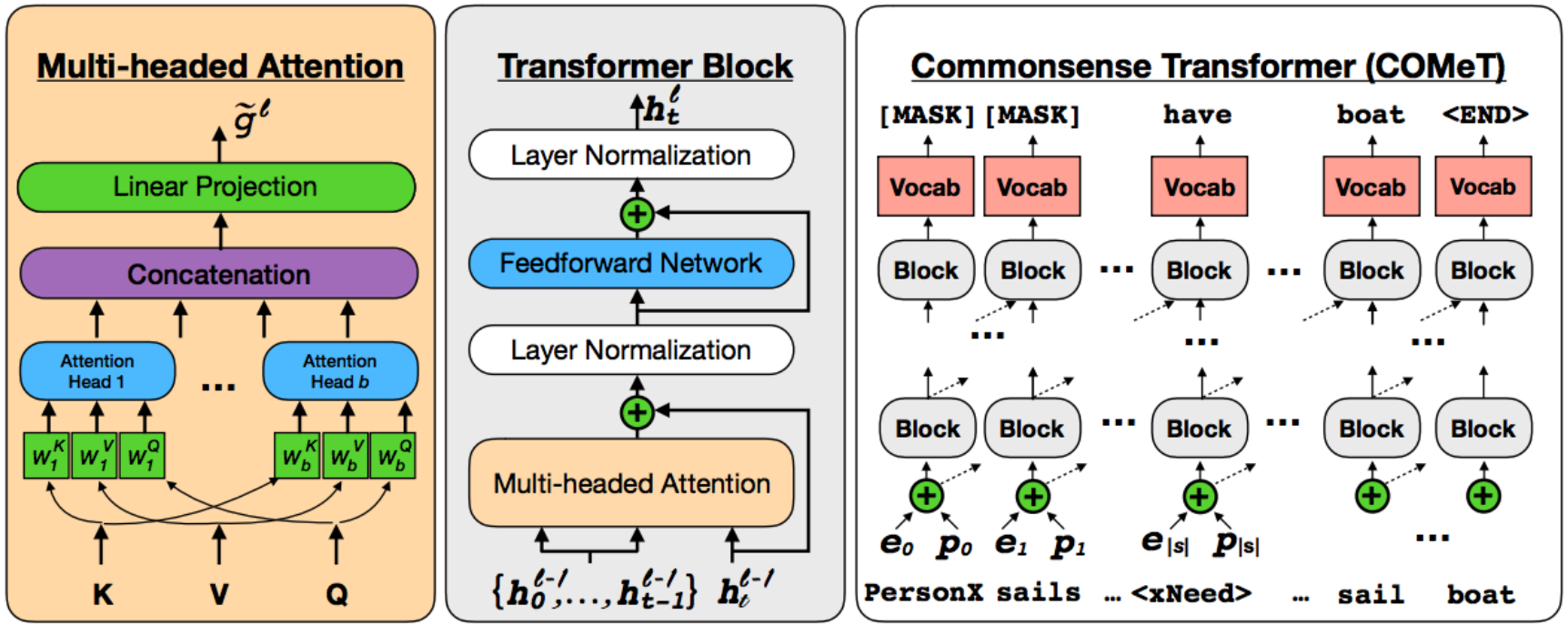
PS: In ConceptNet 5.0, more edges are added, but only the core part, which is inherited from ConceptNet 1.0 (Liu & Singh, 2004), is related to commonsense knowledge.

Outline

- Understanding Commonsense from the Angle of Events
- Instance-level Event Knowledge Acquisition
  - Human Annotation
  - Automatic Event Knowledge Extraction
  - Language Modeling
- Schema-level Event Knowledge Acquisition
- Conclusion

Knowledge Discovery from Pre-trained LMs

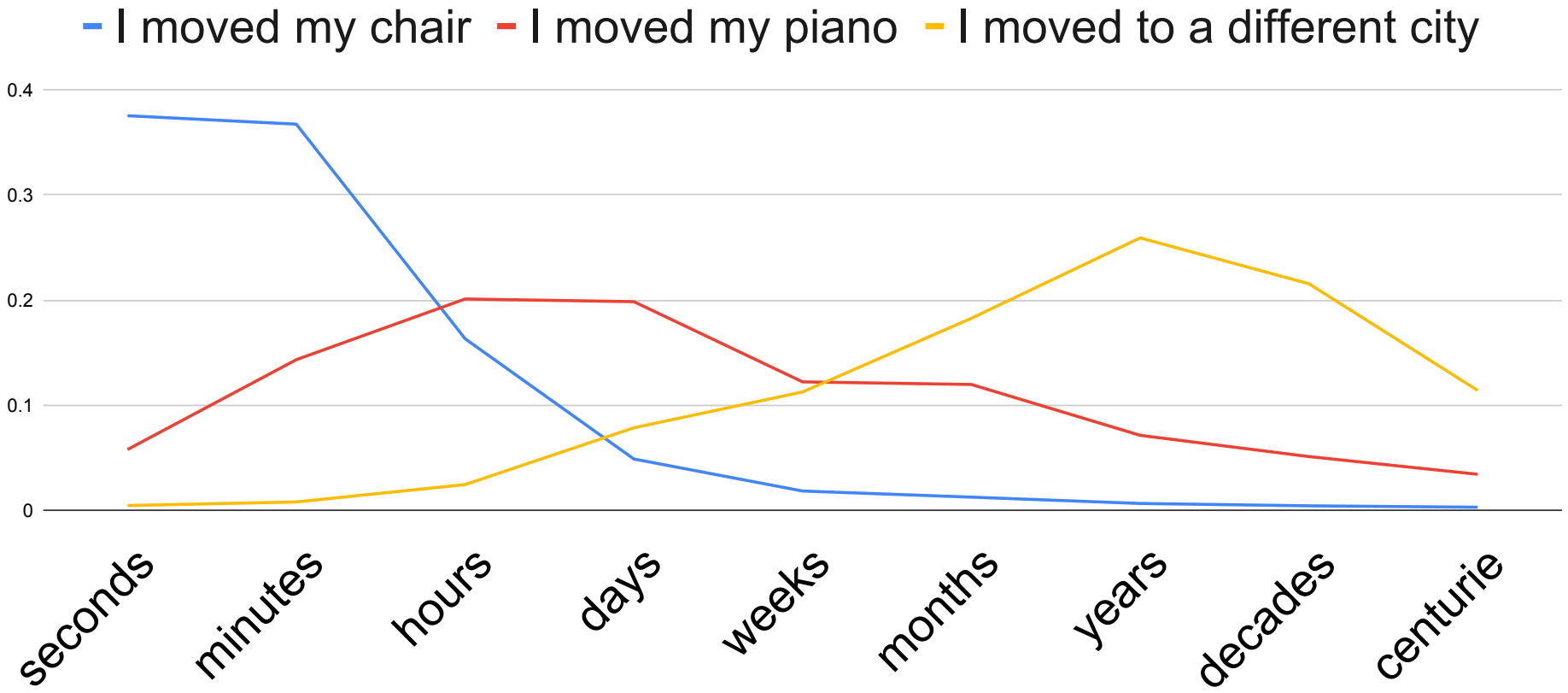
- Language Model
  - Examples: GPT-1/2/3
- COMET (Bosselut et al., 2019):
  - Commonsense Transformers for Automatic Knowledge Graph Construction



Antoine Bosselut, Hannah Rashkin, Maarten Sap, Chaitanya Malaviya, Asli Celikyilmaz, and Yejin Choi. COMET: Commonsense Transformers for Automatic Knowledge Graph Construction. ACL 2019.

Event Temporal Commonsense

- TacoLM (Zhou et al., 2020)
  - a general time-aware language model that distincts temporal properties in fine grained contexts.



Ben Zhou, Qiang Ning, Daniel Khashabi,, and Dan Roth. Temporal Common Sense Acquisition with Minimal Supervision. ACL 2020.

Event Temporal Commonsense

Goal: build a general time-aware LM with minimal supervision

Step 1: Information Extraction

- ☐ Use high-precision patterns to acquire temporal information
  - Unsupervised automatic extraction
- ☐ Overcomes reporting biases with a large amount of natural text

Step 2: Joint Language Model Pre-training

- ☐ Multiple temporal dimensions
  - Duration ~ 1 / Frequency

    "I brush my teeth every morning" → Duration of "brushing teeth" < morning
  - Further generalization to combat reporting biases

Output: TacoLM- a time-aware general BERT

Event Temporal Commonsense

I played basketball for 2 hours.

Original sentence

SRL Parse

Arg-0 Arg-1

I played basketball for 2 hours.

Verb Arg-Tmp

Pattern Matching

for 2 hours: matches Duration pattern

Event Value

I played basketball, Duration, Hours

Dimension

Formatted Output Instance

Information Extraction

I [M] played basketball [SEP] [M] [DUR] [HRS]

- Baseline Model: Pre-trained BERT-base
- Main objective: mask some tokens and recover them
- How we mask:
  - With some probability, mask temporal value while keeping others

I [M] played basketball [SEP] [M] [DUR] **[MASK]**

  - Otherwise, mask a certain portion of E1...En while keeping temporal value unchanged

I [M] **[MASK] [MASK]** [SEP] [M] [DUR] **[HRS]**

  - Max (P(Event|Dim,Val) + P(Val|Event,Dim)); Preserving original LM capability

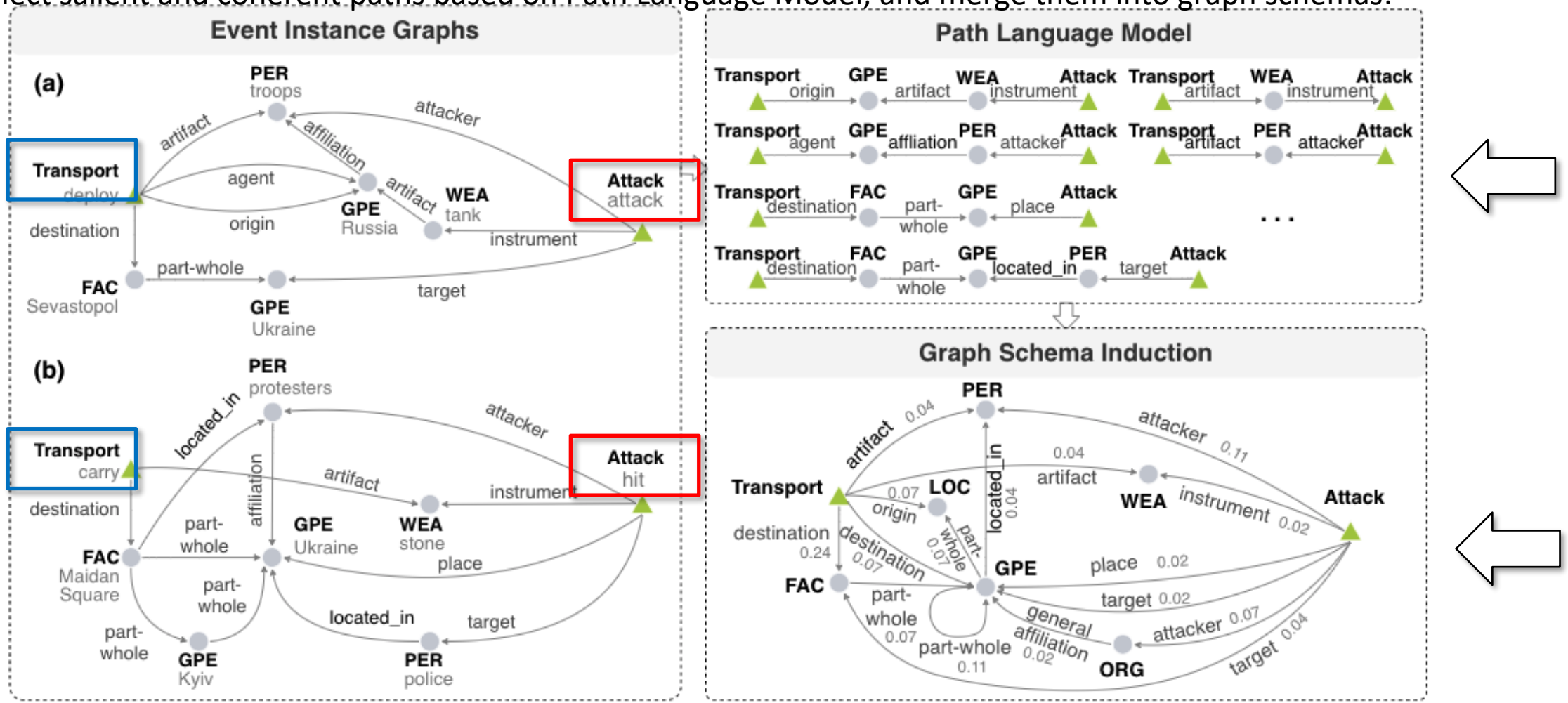
Joint training with language model

Outline

- Understanding Commonsense from the Angle of Events
- Instance-level Event Knowledge Acquisition
  - Human Annotation
  - Automatic Event Knowledge Extraction
  - Language Modeling
- Schema-level Event Knowledge Acquisition
- Conclusion

Event Graph Schema Induction (Li et al., 2020)

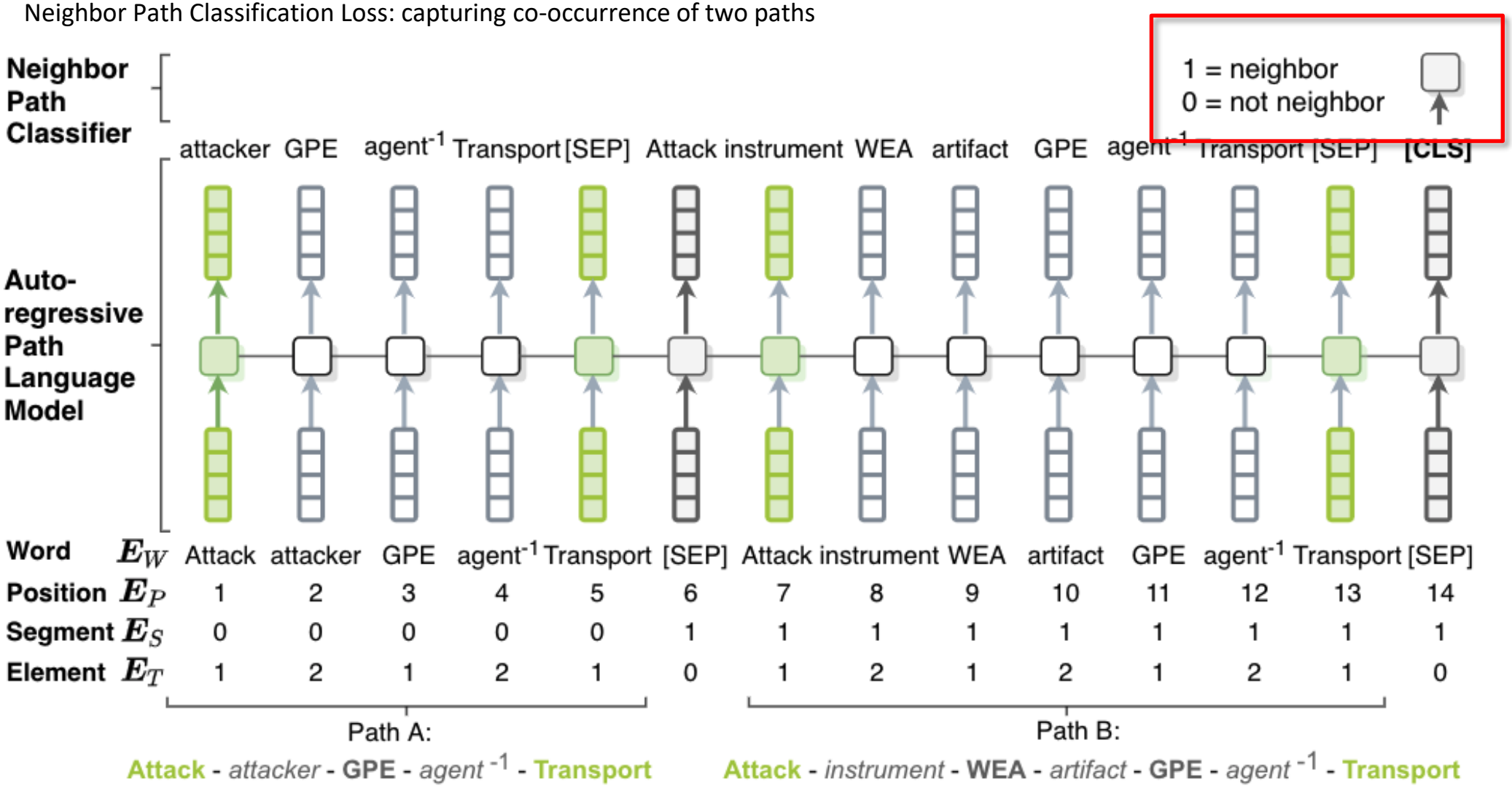
- History repeats itself: Instance graphs (a) and (b) refer to very different event instances, but they both illustrate a same scenario.
- Select salient and coherent paths based on Path Language Model, and merge them into graph schemas.



Manling Li, Qi Zeng, Kyunghyun Cho, Heng Ji, Jonathon May, Nathanael Chambers, Clare Voss. Connecting the Dots: Event Graph Schema Induction with Path Language Modeling. ACL 2020.

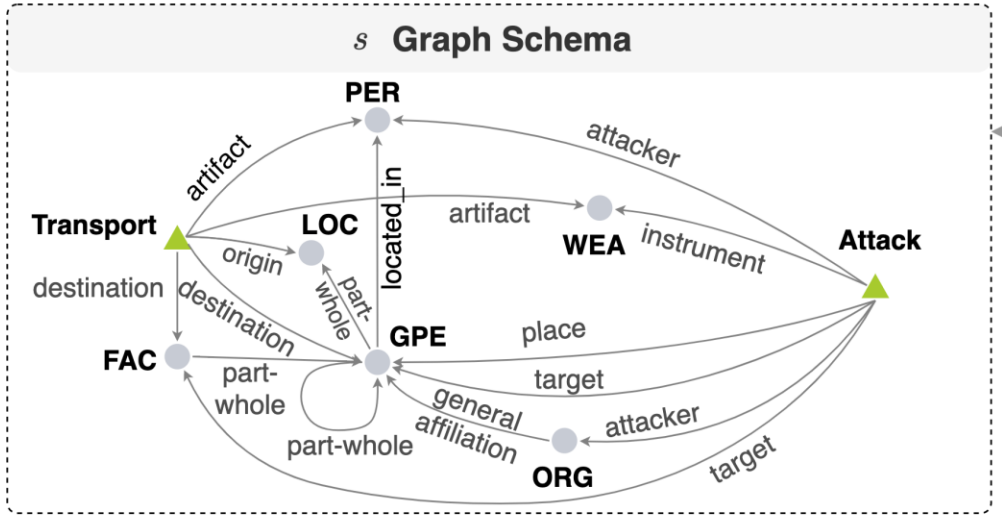
Path Language Model

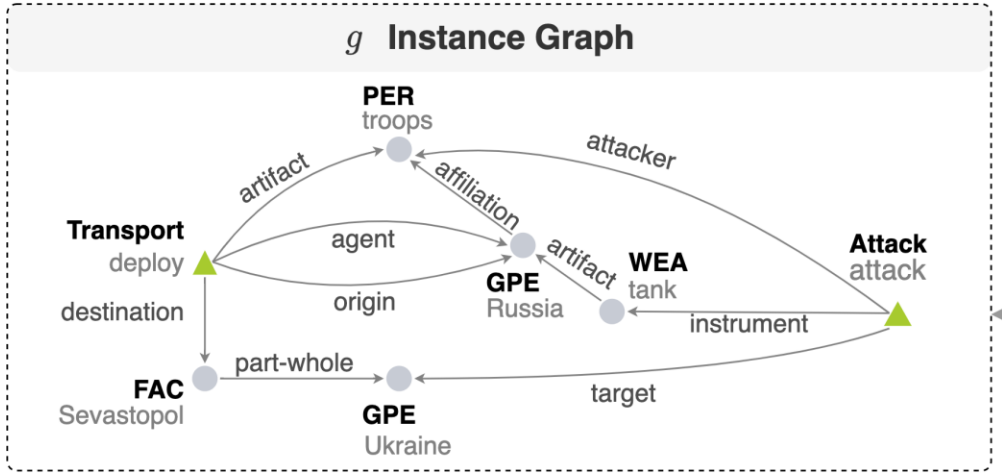
- Path Language Model is trained on two tasks
  - Autoregressive Language Model Loss: capturing the frequency and coherence of a single path
  - Neighbor Path Classification Loss: capturing co-occurrence of two paths



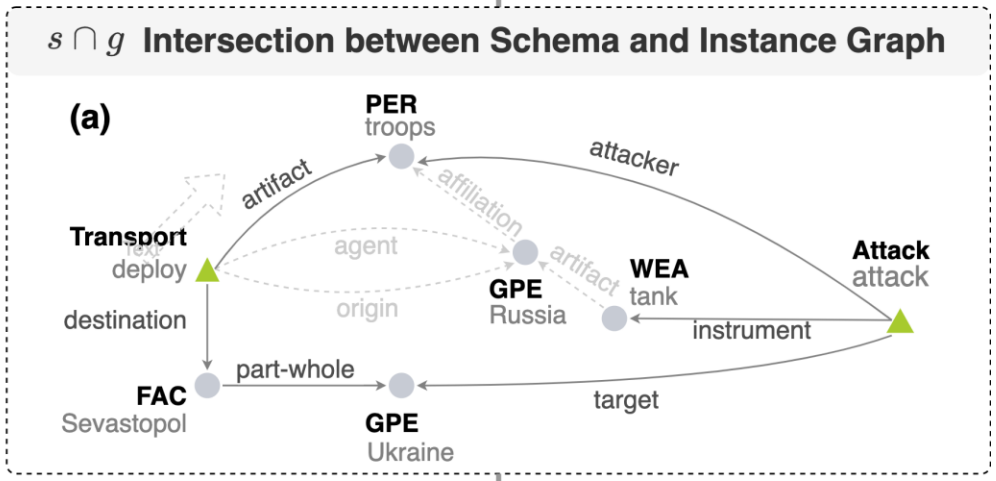
Recover Instance Graph

- A salient schema can serve as a skeleton to recover instance graphs
  - We use each graph schema to match back to each ground-truth instance graph and evaluate their intersection in terms of Precision and Recall.





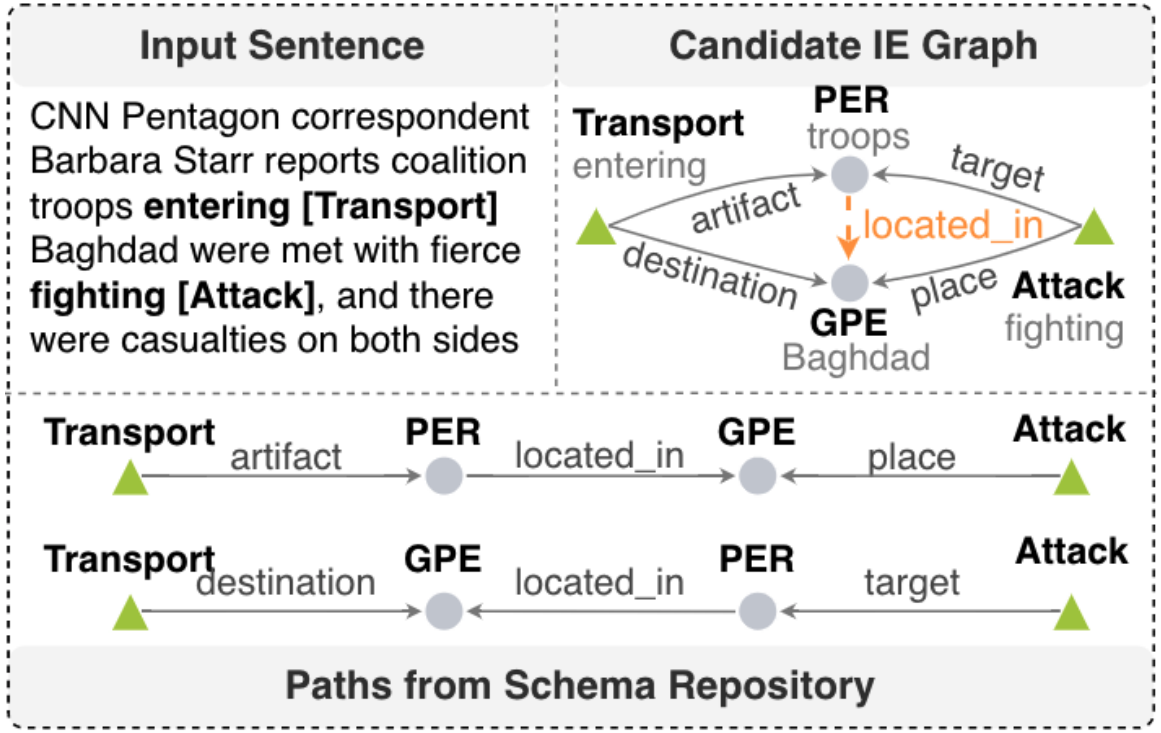
$$\text{Precision} = \frac{\sum_{s \in \mathcal{S}} \sum_{g \in \mathcal{G}} |g \cap s|_{\mathbb{S}}}{\sum_{s \in \mathcal{S}} |s|_{\mathbb{S}}}$$



$$\text{Recall} = \frac{\sum_{s \in \mathcal{S}} \sum_{g \in \mathcal{G}} |g \cap s|_{\mathbb{I}}}{\sum_{g \in \mathcal{G}} |g|_{\mathbb{I}}}$$

Schema-Guided Information Extraction

- Use the state-of-the-art IE system OneIE (Lin et al, 2020) to decode converts each input document into an IE graph
- Each path in the graph schema is encoded as a single global feature for scoring candidate IE graphs
- OneIE promotes candidate IE graphs containing paths matching schema graphs



Dataset	Entity	Event Trigger Identification	Event Trigger Classification	Event Argument Identification	Event Argument Classification	Relation
Baseline	90.3	75.8	72.7	57.8	55.5	44.7
+PathLM	90.2	**76.0**	**73.4**	**59.0**	**56.6**	**60.9**

Outline

- Understanding Commonsense from the Angle of Events
- Instance-level Event Knowledge Acquisition
  - Human Annotation
  - Automatic Event Knowledge Extraction
  - Language Modeling
- Schema-level Event Knowledge Acquisition
- Conclusion

Key takeaways

- There is a transferability from event knowledge to commonsense knowledge
- Compared with commonsense, acquiring event knowledge is cheaper and more scalable.
- All existing acquisition systems have advantages and limitations.

	Quality	Scale	Relation Coverage	Explainability	Robustness	Downstream Task
Human Annotation	**High**	Small	Middle	**High**	High	Difficult
Automatic Event Knowledge Extraction	Middle	**Large**	**High**	**High**	Middle	Difficult
Language Model	Middle	**Large**	**High**	Low	Low	**Easy**

Thanks ☺

Key References

- WSC: Hector Levesque, Ernest Davis, and Leora Morgenstern. The winograd schema challenge. KRR 2012.
- *ConceptNet:* Hugo Liu and Push Singh, ConceptNet - a practical commonsense reasoning tool-kit, BTTJ, 2004
- ATOMIC: Maarten Sap, Ronan LeBras, Emily Allaway, Chandra Bhagavatula, Nicholas Lourie, Hannah Rashkin, Brendan Roof, Noah A. Smith, Yejin Choi, ATOMIC: An Atlas of Machine Commonsense for If-Then Reasoning. AAAI 2019
- COMET: Antoine Bosselut, Hannah Rashkin, Maarten Sap, Chaitanya Malaviya, Asli Celikyilmaz, and Yejin Choi. COMET: commonsense transformers for automatic knowledge graph construction. ACL 2019.
- LAMA: Fabio Petroni, Tim Rocktaschel, Se- ¨ bastian Riedel, Patrick Lewis, Anton Bakhtin, Yuxiang Wu, and Alexander H. Miller. Language models as knowledge bases? EMNLP 2019.
- ASER: Hongming Zhang, Xin Liu, Haojie Pan, Yangqiu Song, and Cane Wing-Ki Leung. ASER: A Large-scale Eventuality Knowledge Graph. WWW 2020.
- TransOMCS: Hongming Zhang, Daniel Khashabi, Yangqiu Song, and Dan Roth. TransOMCS: From Linguistic Graphs to Commonsense Knowledge. International Joint Conference on Artificial Intelligence (IJCAI). 2020.
- KnowlyWood: Niket Tandon, Gerard de Melo, Abir De, and Gerhard Weikum. 2015. Knowlywood: Mining Activity Knowledge From Hollywood Narratives. CIKM 2015.
- Manling Li, Qi Zeng, Kyunghyun Cho, Heng Ji, Jonathon May, Nathanael Chambers, Clare Voss. Connecting the Dots: Event Graph Schema Induction with Path Language Modeling. ACL 2020.
- TacoLM: Ben Zhou, Qiang Ning, Daniel Khashabi, Dan Roth. Temporal Common Sense Acquisition with Minimal Supervision. ACL 2020

Applications, Remaining Challenges, Future Directions and Resources

Event-Centric Natural Language Understanding (Part V)

Heng Ji and Dan Roth
Department of Computer Science
University of Illinois at Urbana-Champaign

Feb 2020

AAAI Tutorials

Recent Advances in Transferable Representation Learning

Where are We and What's Next

- Applications
  - Writing History Book
  - News Understanding and Recommendation
  - Disaster Relief
  - Intelligence Analysis
  - Accelerating Scientific Discovery
- Remaining Challenges and Future Directions
  - External Knowledge Acquisition, Reasoning and Incorporation
  - Document-Level Event Extraction
  - Multimedia Information Extraction and Verification
- Grand Vision
- Resources

Application 1: Writing a History Book

- Problems on Human Written History Books: Highly Biased (History is written by the victors)
- Translated Japanese comments about Nanjing Massacre:
- As my grandma would say,
  This is why we need to
  EDUCATE people

1.日本的“南京大屠杀论争”

日本著名的“南京大屠杀论争”，是以南京大屠杀这一事件是否存在、其规模之大小为焦点的论争。随着中日关系的变化，论争不断受到政治的不同影响。

首先来看一下日本网民的看法。和中国一样，在日本的网络贴吧里，也有很多关于南京大屠杀的讨论。2013年2月，在日本某名为“大舰巨炮主义”的贴吧里，“[中国ＢＢＳ]結局、日本人は南京大虐殺をどう考えているのか? 2013年02月10日 20:32”（military38.com/archives...）这一帖子以“中国人对日本人不反思历史进行批判”这一事实为引子，引出众多网友对南京大屠杀的跟帖议论。

跟帖的人中，有一部分完全否认大屠杀，更大多数人承认在南京发生过大规模杀人事件的。但大部分人虽然承认，却否认30万死者的说法。

来自琦玉县的チーター：大屠杀的数字肯定是捏造啊，30万人从物理角度看都是不可能的。

来自大阪的シンガプーラ：我问过实际去过南京的人的孩子（超过70岁了），据他爸爸说，当时日本军队并不拥有能够杀死几十万中国人的军队人数与武器，30万的数字必然是捏造的。

来自福冈的パンパスネコ：我们要盯着过去到什么时候？我出生的时候战争已经结束了，所以老实说我对大屠杀完全不知道。如果总是拿上一代的责任说事，历史这个东西就会变得一团糟了。

- It’s international human right to know what happened in the history
- We aim to assist historians and librarians at writing more complete and authentic history books

Information Pollution

Conspiracy theories blaming Bill Gates for the coronavirus pandemic are exploding online

Isobel Asher Hamilton Apr 17, 2020, 6:55 AM

- **As the coronavirus pandemic has spread around the world, with millions infected and thousands dead, billionaire Microsoft co-founder and philanthropist Bill Gates has pledged a quarter billion dollars to combat the disease through his foundation.**

- **Gates has been an advocate for pandemic preparedness for years, and his Bill and Melinda Gates Foundation is contributing financing to several coronavirus vaccine initiatives. He famously gave a 2015 TED talk warning of the potential devastation caused by — and urged readiness for — a worldwide pandemic.**

- Why would anyone ever believe these ridiculous rumors?
- Because humans are very good at connecting dots
- And perhaps too good →

What will such a History Book look like?

- Organize chapters by major events clusters and order them on a timeline
- Each chapter looks like a Wikipedia page
  - The description is organized by multimedia timeline
    with detailed source and evidence information, links to original news articles
  - Detailed participants (arguments) and their roles, and their connections and relations
  - Infobox shows event-event relations: temporal, causal and hierarchical
- Never-ending updating over time; put up to the wild for human editing and curation

Protests against the government [edit]

4 December 2011 [edit]

On 4 November 2011, during the annual Russian March event, representatives of "The Russians" movement declared a protest action p
unapproved and took place on 4 December at 21:00 in Moscow. The statement of non-recognition of electoral results spread widely. Cit
said to have occurred during the elections. Alexander Belov declared the beginning of the "Putin, go away!" campaign.[37] The protest a
Alexander Belov, Dmitry Dyomushkin, George Borovikov were arrested along with dozens of other nationalists. The head of the banned
observer. Mass detentions of other public organizations occurred in Moscow. According to police some 258 persons have been detaine

5–7 December 2011 [edit]

On 5 December, around 5,000 opponents of the government began protesting in Moscow, denouncing Vladimir Putin and his governme
Putin step down, whilst some demanded revolution.[15][39] Alexey Navalny, a top blogger and anti-corruption activist who branded Putir
postings on his LiveJournal blog and Twitter account. Navalny's agitation was denounced by United Russia as "typical dirty self-promoti
account.[40][41]

Many pro-government supporters, including the pro-Putin youth group Nashi, were mobilized on 6 December at the site of the planned
Nashi on Manezhnaya Square[43] and an 8,000-strong rally of the Young Guard on Revolution Square.[44] About 500 pro-United Russi
expected anti-government protests. It emerged that 300 protesters had been arrested in Moscow the night before, along with 120 in St.
slogans against Putin,[15] whilst anti-government protesters at Revolution Square clashed with riot police and interior ministry troops. T
Square and dozens of arrests were reported, including Boris Nemtsov, an opposition leader and former deputy prime minister,[48] and A
charged. At least one Russian journalist claimed he was beaten by police officers who stamped on him and hit his legs with batons.[50]
place. After three and a half hours, the Moscow protest came to an end.[51]

Attempts to stage a large protest in Moscow on 7 December fizzled out due to a large police presence in the city.[13]

10 December 2011 [edit]

Rally in Bolotnaya Square in Moscow on 10 December 2011

Via a Facebook group "Суббота на Болотной площади" (Saturday at Bolotnaya Square),[52]
December.[53][54] Prior to the demonstration newspapers commented that tens of thousands o
Moscow,[55][56] and, similarly, over 5,000 in St. Petersburg.[57] A permit had originally been iss
Square. By 8 December, more than 30,000[52] had accepted the Facebook invitation to attend.
demonstration was authorized by the Moscow government for the demonstration which took pla
made by Putin that police and security forces would be deployed to deal with anyone participati
was peaceful and without attempts by the state to prevent or disrupt it.[59][60] Rapper Noize M
specially from Paris for the occasion.[61] Guerrilla theater by FEMEN and the circulation of a ph
protests.[62][63]

Example Chapter: 2019 Hong Kong Protests

The first big protests

On **9 June**, an estimated one million people marched to the government headquarters to show they were against the proposed bill.

Critics say the plan will erode the city's judicial independence

Protesters storm parliament

On **1 July**, the anniversary of Hong Kong's handover from the UK to China, the Legislative Council (LegCo) building was stormed by protesters who sprayed graffiti on the walls, displayed the colonial-era flag and defaced Hong Kong's regional emblem.

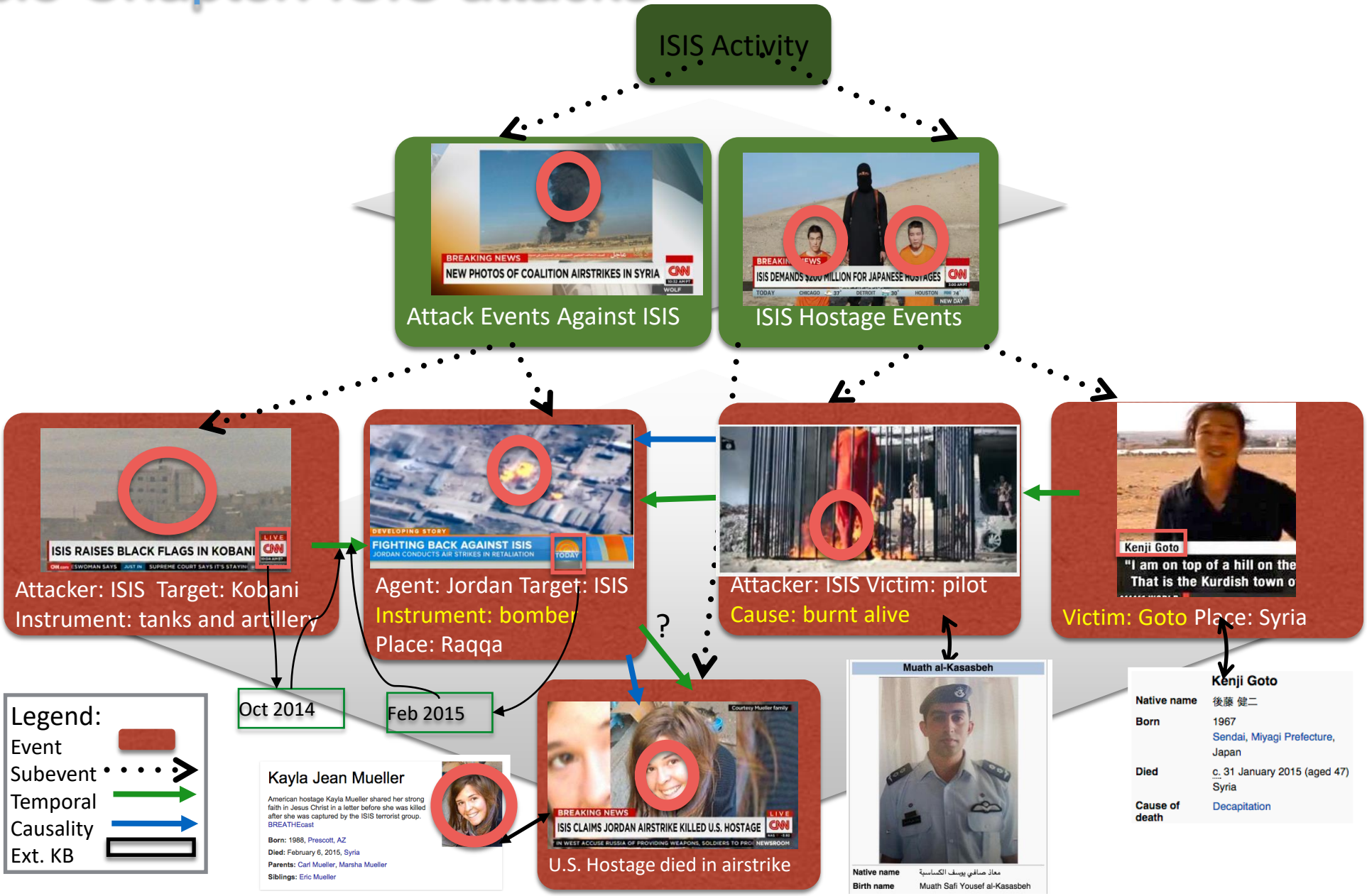
Protesters describe their ransacking of Hong Kong's Legislative Council one month on

One week later, on **7 July**, tens of thousands marched in Kowloon - an area popular with mainland tourists - in a bid to explain their concerns. Until this point the protests had received little if any coverage in state-run mainland media.

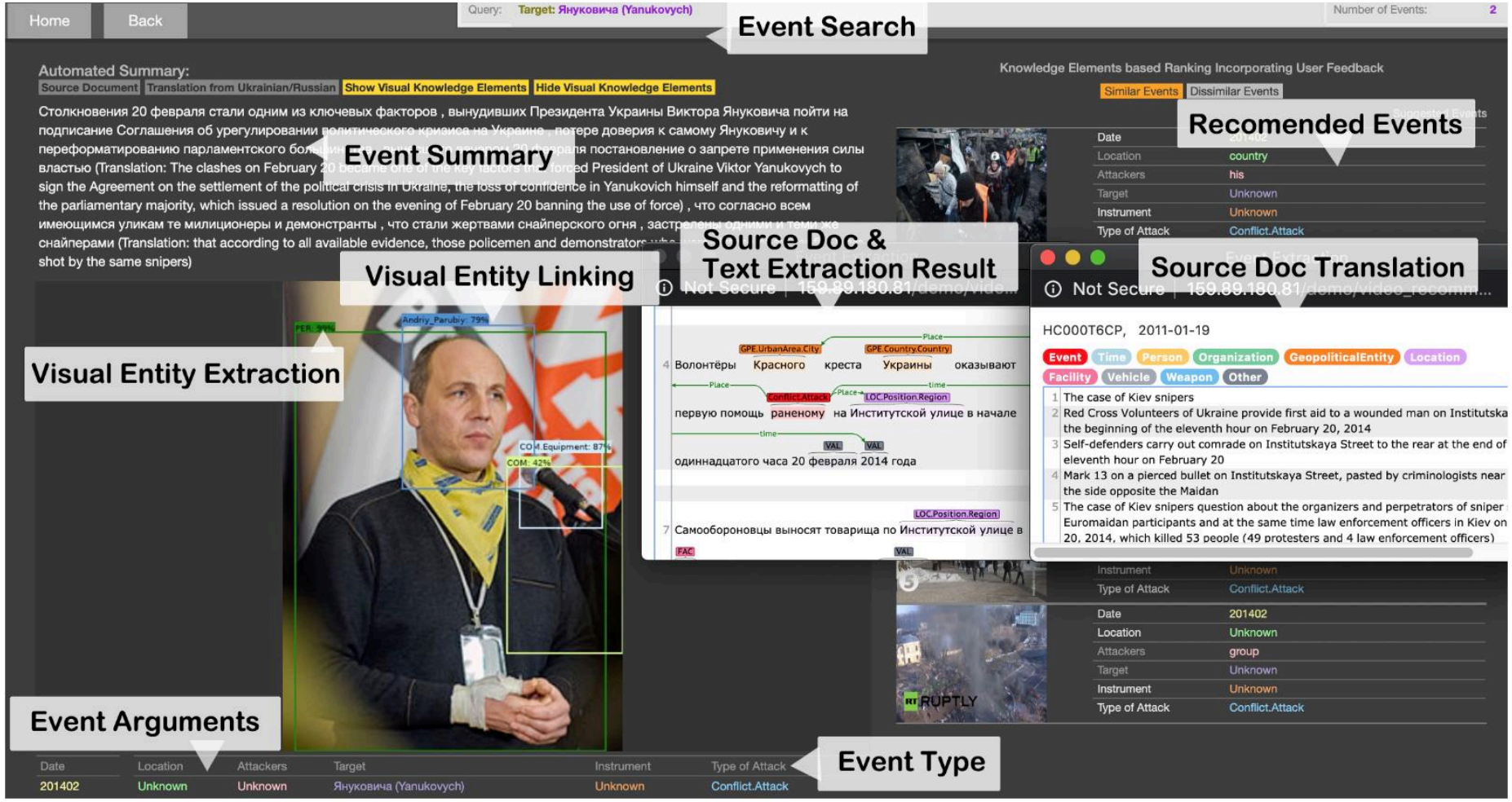
On **9 July**, Carrie Lam reiterated that the extradition bill was "dead" urging protesters to stop their actions. She still refrained from fully withdrawing the bill.

On **21 July**, protesters defaced China's Liaison Office in Hong Kong.

Example Chapter: ISIS attacks



Application 2: News Understanding and Recommendation (Li et al., ACL2020Demo)



(Li et al., ACL2020 Best Demo Paper Award)

GitHub: https://github.com/GAIA-IE/gaia

DockerHub: https://hub.docker.com/orgs/blendernlp/repositories

Demo: http://159.89.180.81/demo/video_recommendation/index_attack_dark.html

Event-centric Question Answering

- Identifying event orders and predicting future events

Heavy snow is causing disruption to transport across the UK, with heavy rainfall bringing flooding to the south-west of England. Rescuers searching for a woman trapped in a landslide at her home in Looe, Cornwall, said they had found a body.

Q1: What events have already finished?
A: searching trapped landslide said found
Q2: What events have begun but has not finished?
A: snow causing disruption rainfall bringing flooding
Q3: What will happen in the future?
A: No answers.
warm-up

Q4: What happened before a woman was trapped?
A: landslide
Q5: What had started before a woman was trapped?
A: snow rainfall landslide
Q6: What happened while a woman was trapped?
A: searching
Q7: What happened after a woman was trapped?
A: searching said found
User-provided

Ning, et al. TORQUE: A Reading Comprehension Dataset of Temporal Ordering Questions. EMNLP, 2020

- 3.2k news snippets with 21k human-generated questions querying temporal relationships

Q: Who will drop Japan as a trading partner in August 2019?

earlier than **timestamp** (2019-08-01)

(1/1/19) Apart from the fact of being one another's closest neighbors, the people of South Korea and Japan have a remarkable amount in common. Economically, they are among one another's biggest trading partners. And yet, time and again, relations between *Seoul and Tokyo* are marked, not by mutual support and co-operation but by *anger, reproach and exasperation.*

A) South Korea [0.41] B) Syria [0.28]
C) South Africa [0.15] D) Portugal [0.16]

Q: Will primary schools in Europe admit non-vaccinated children around September 2019?

earlier than **timestamp** (2019-09-01)

(3/8/18) Public officials and health experts had given several warnings: Do *not allow a student in school* if they had not been vaccinated against measles.

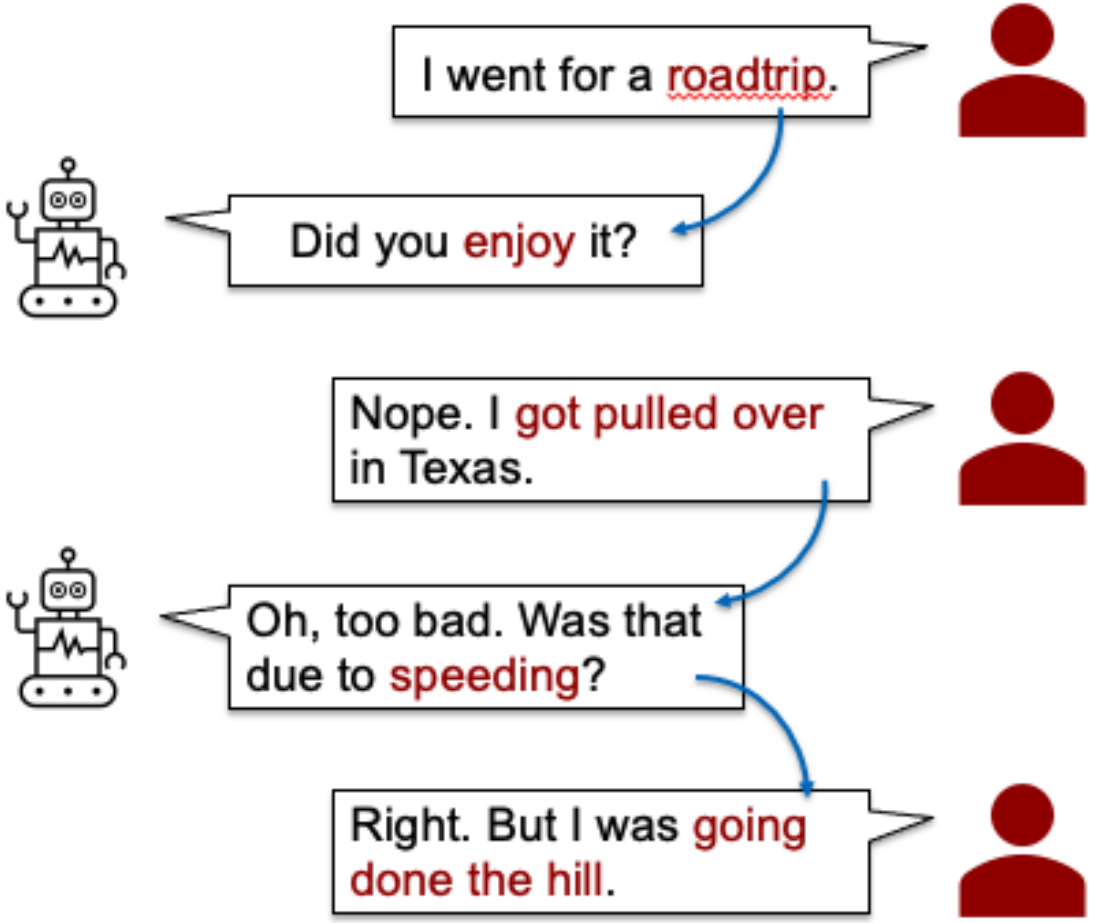
(6/27/19) Fines for parents refusing measles jab. Parents will be fined up to € 2,500 if they don't vaccinate their children against measles under draft legislation in Germany which also *threatens* exclusion from crèches, nurseries and *schools*.

Yes [0.38] / **No [0.62]**

ForecastQA: A Question Answering Challenge for Event Forecasting

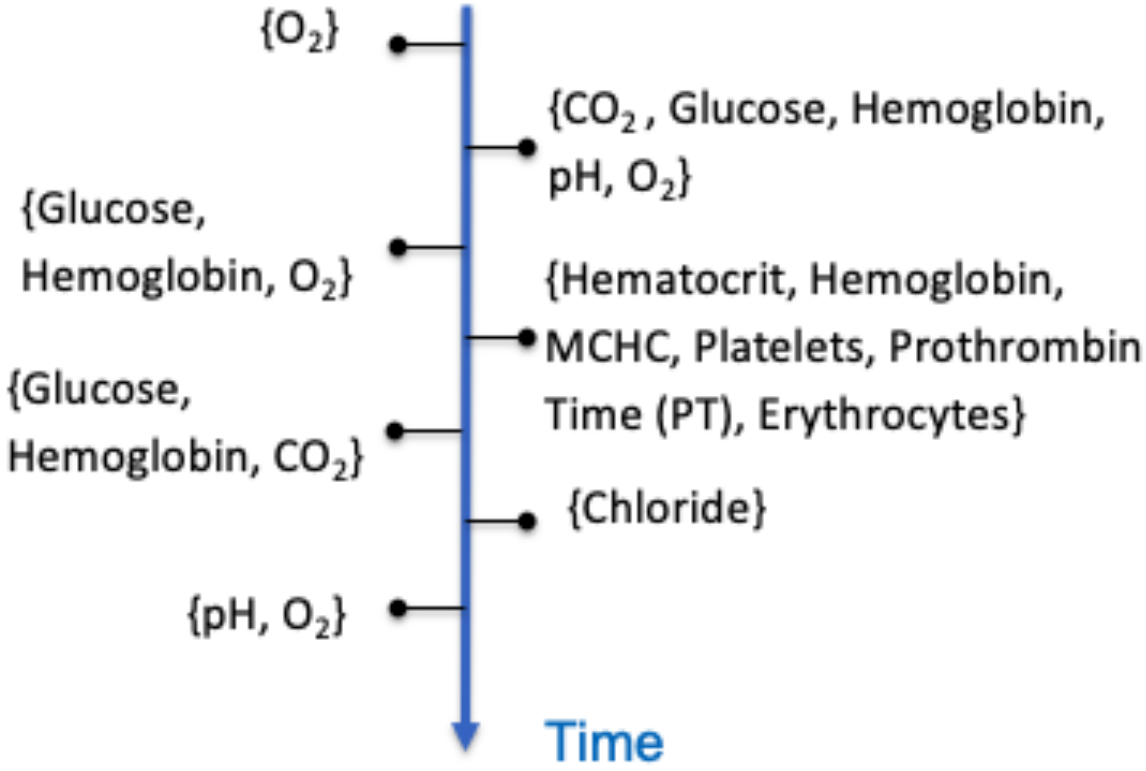
Event-centric Dialogue Systems

- Chatbots



Can event processes improve the consistency of utterance generation/retrieval?

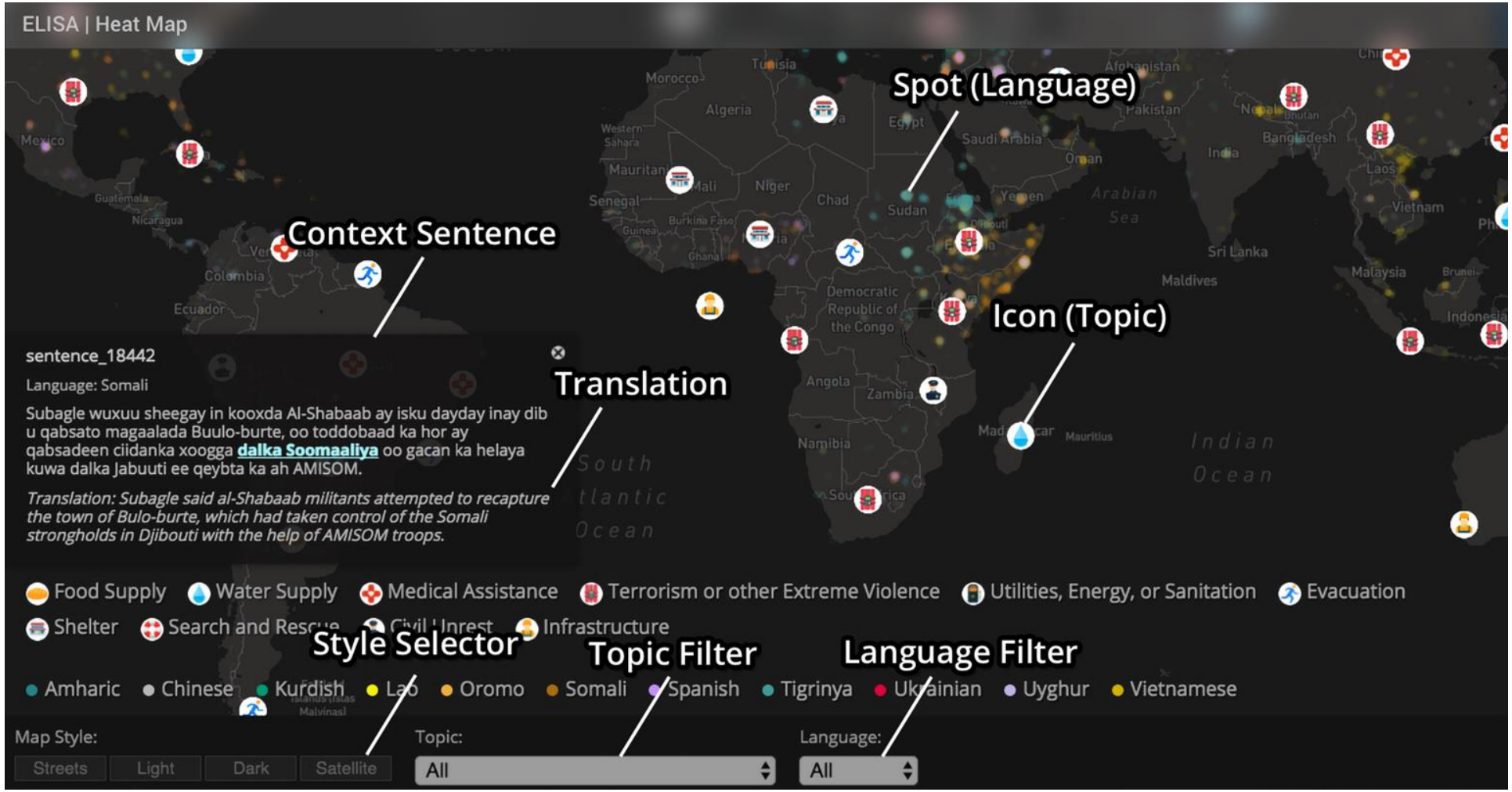
- Clinical event processes



Diagnostic prediction, phenotype prediction, …

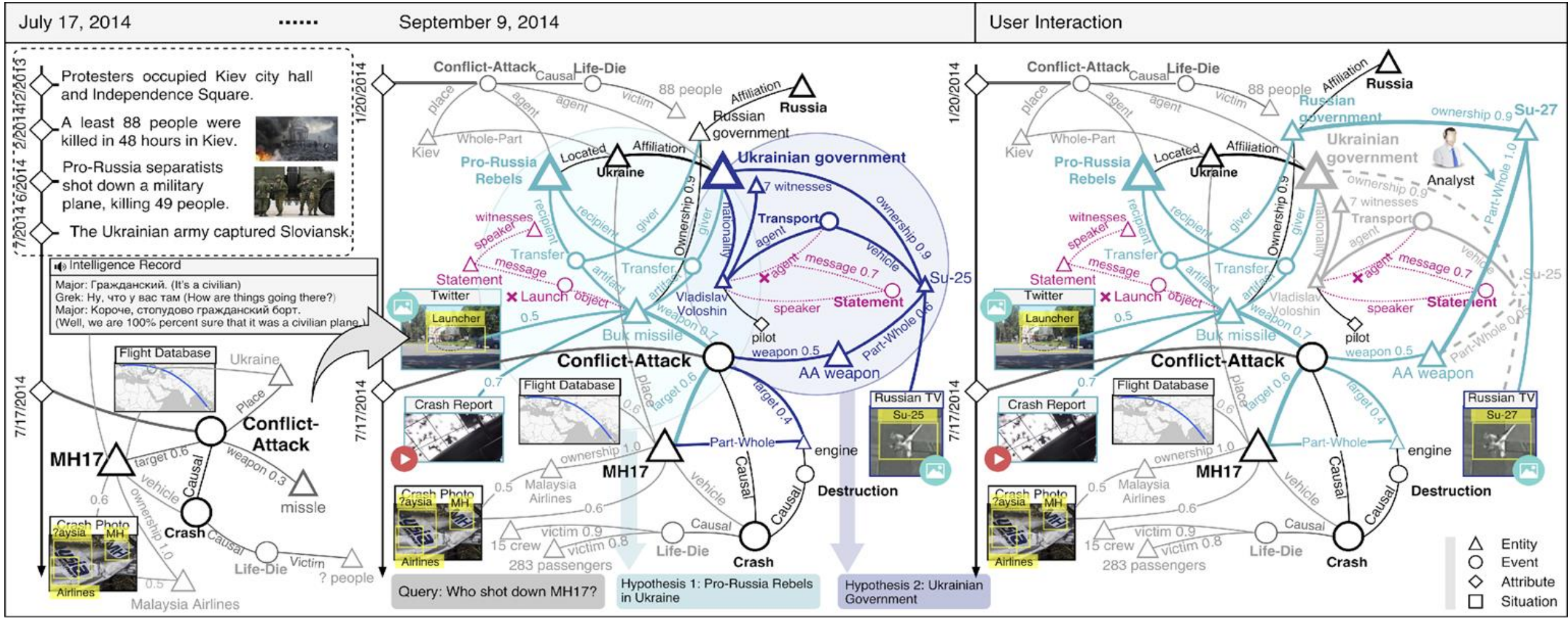
- Transfer learning can important (naturally lack of data)
- Structured prediction can be important (dependency of phenotypes, disease labels)

Applications 3: Disaster Relief (Zhang et al., NAACL18Demo)

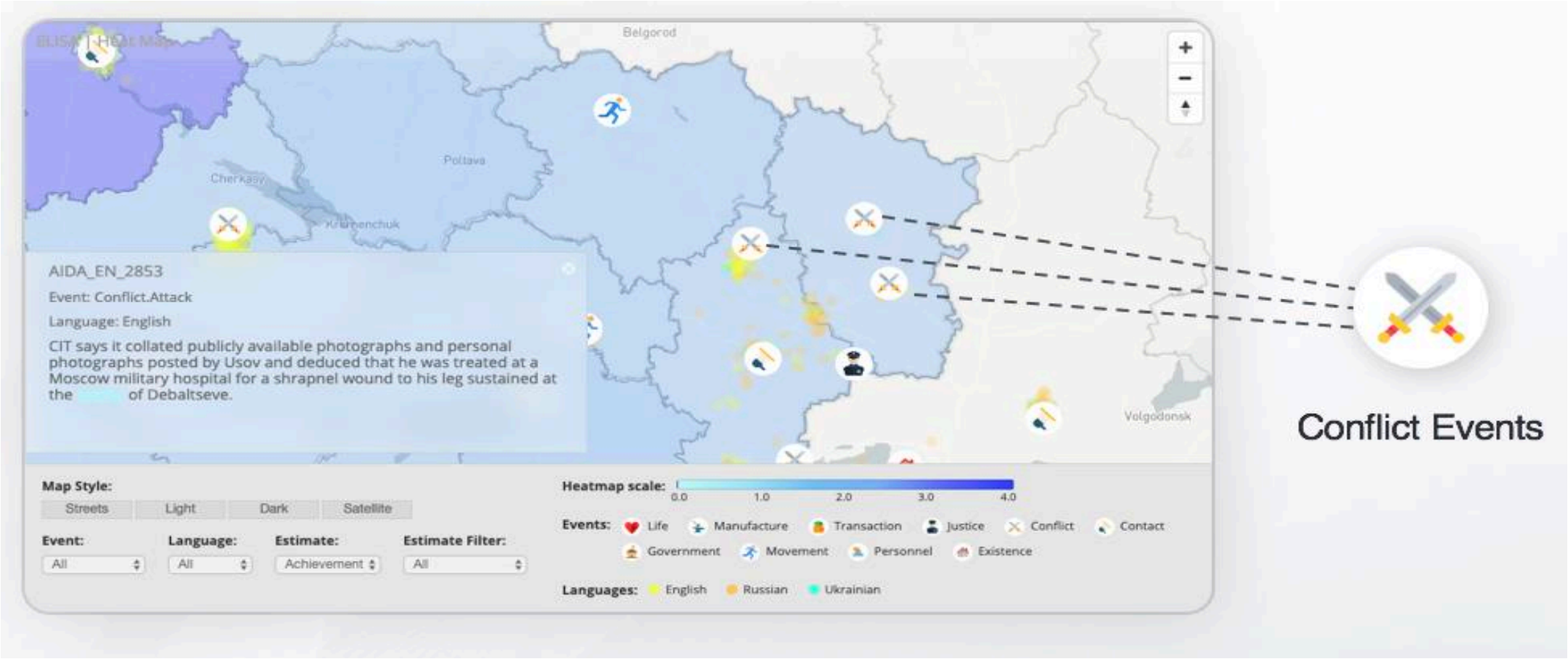


- Re-trainable Systems: http://159.89.180.81:3300/elisa_ie/api
- Demos: http://159.89.180.81:3300/elisa_ie
- Heat map: http://159.89.180.81:8080/

Applications 4: Intelligence Analysis



Applications 4: Intelligence Analysis



http://162.243.120.148:8080/

- Achievement, Benevolence, Conformity, Hedonism, Power, Security, Self-direction, Simulation, Tradition and Universalism (Schwartz, 2012)

Application 5: Accelerating Scientific Discovery

DATA

NeurIPS'19, WWW'20, KDD'20, ACL'20

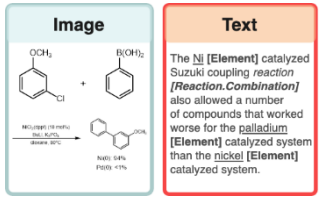
SEMANTICS

NAACL'19, EMNLP'19, ACL'19, ACL'20

KNOWLEDGE BASE

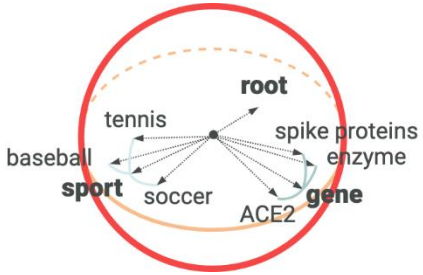
AAAI'19, ACL'20

Graph neural networks
Joint entity/relation/event extraction and ontology construction



Scientific Literature

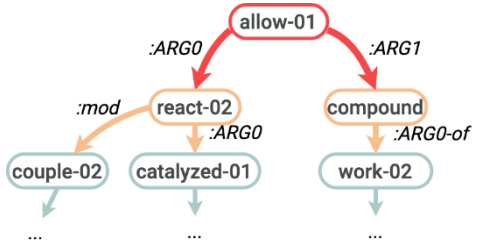
Chemical Ontology & Existing Databases



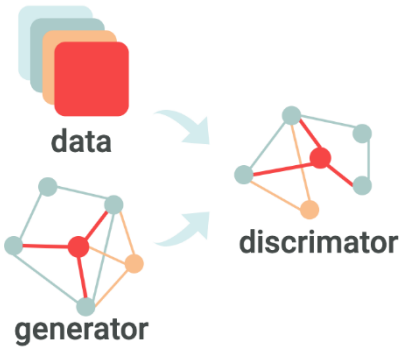
Hierarchical Spherical Embedding

KB concept of Ni	
Name: Nikel	Element category: transition metal
Chemical Symbol: Ni	Appearance: lustrous, metallic, and silver with a gold tinge
Atomic Number: 28	
Type: chemical element	

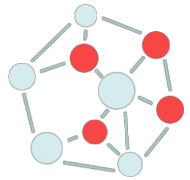
Ontology Enriched Text Embedding



Cross-media Structured Semantic Representation



Generative Adversarial Networks for Data Augmentation and Distant Supervision

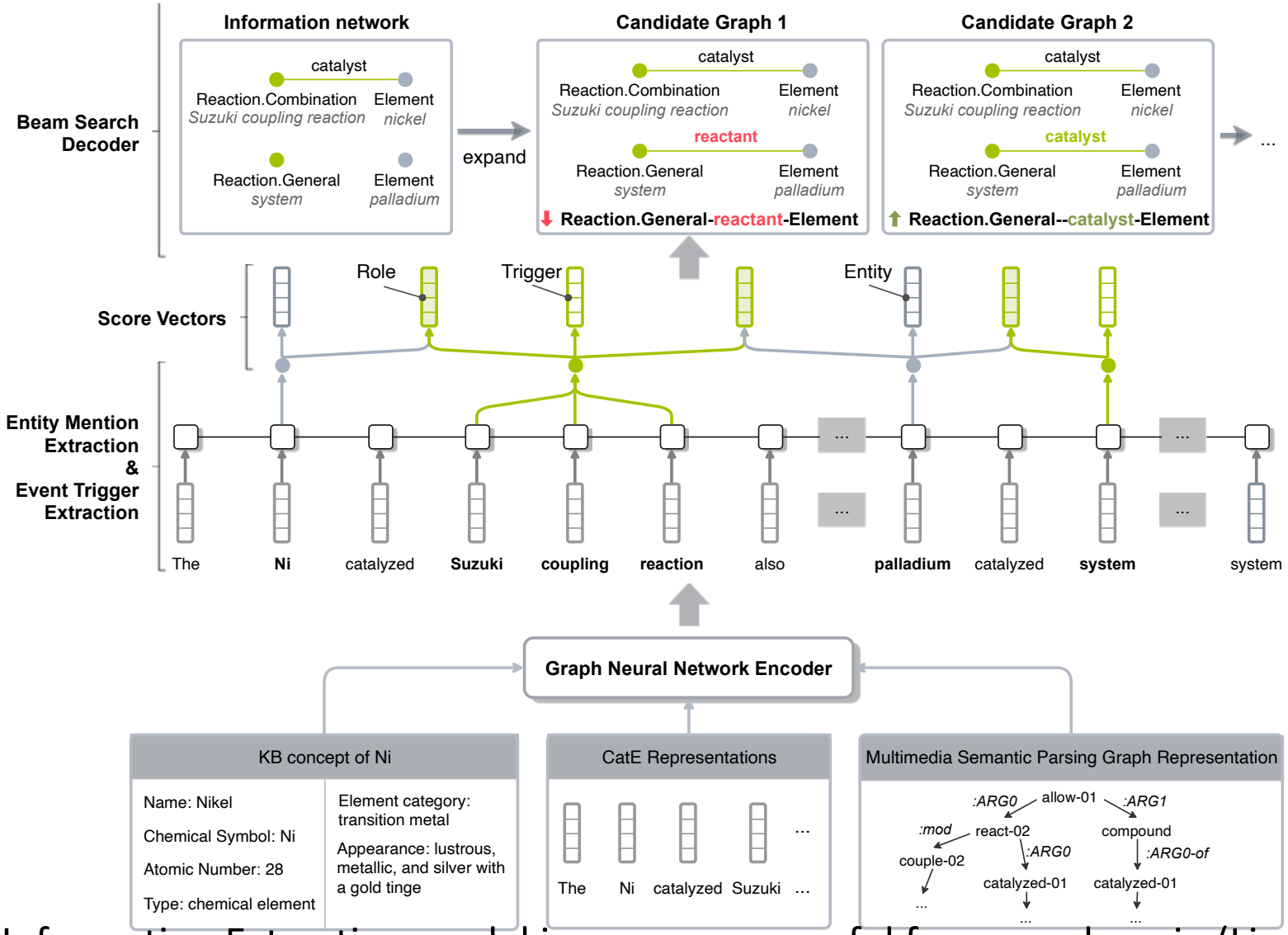


Multimedia Knowledge Base

Multimedia Search and Summarization

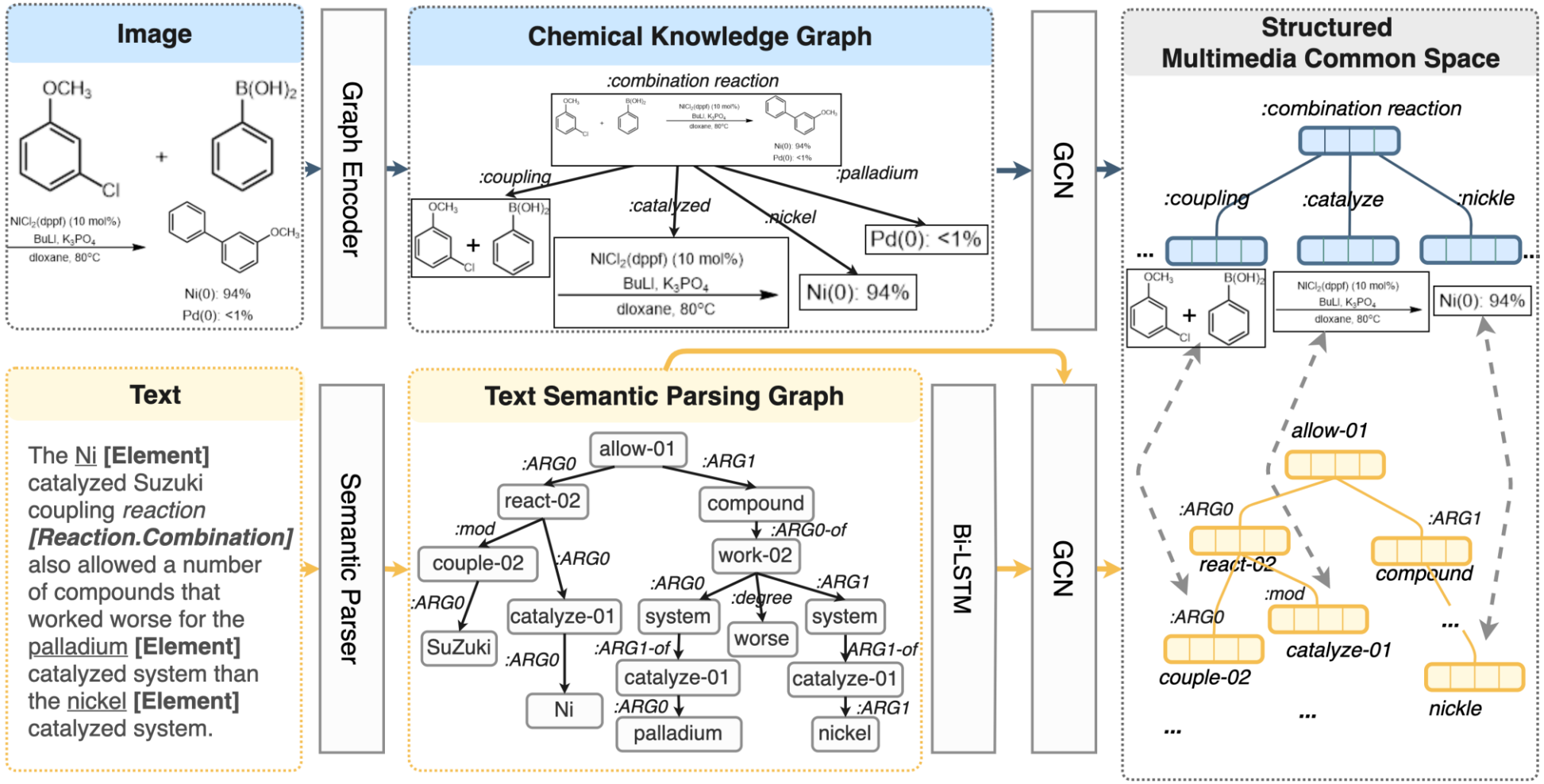
- Extending to other scientific domains including Molecular Synthesis and Agriculture under two new NSF institutes

Fine-grained COVID-19 Event Extraction



- Joint neural Information Extraction model is proven successful for news domain (Lin et al., ACL2020)

Fine-grained COVID-19 Event Extraction



- Multimedia Common Semantic Space Construction is proven successful for news domain (Li et al., ACL2020)

Multimedia Knowledge Graph Construction for Drug Repurposing

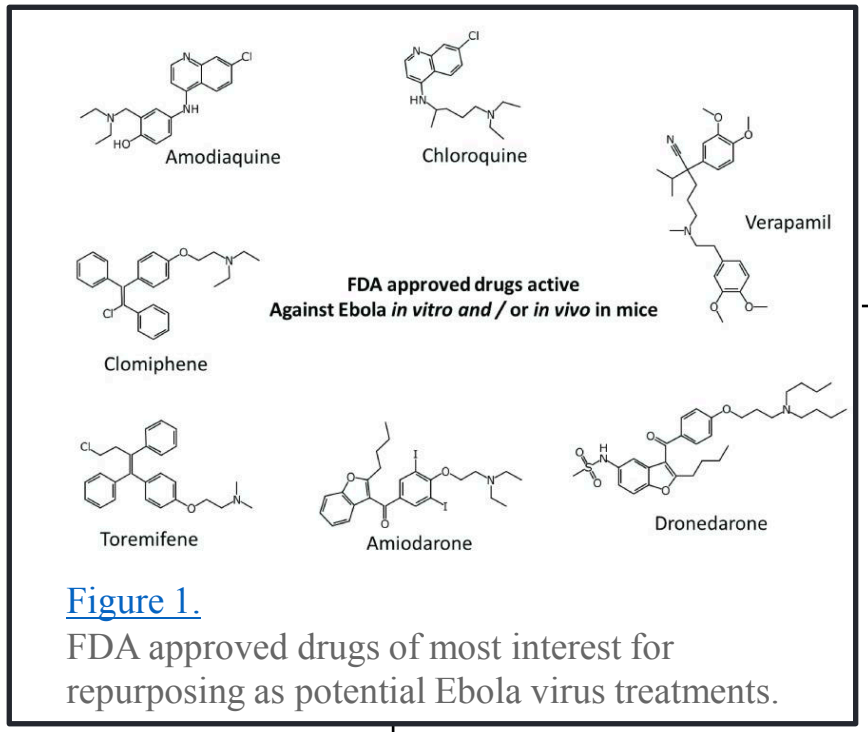
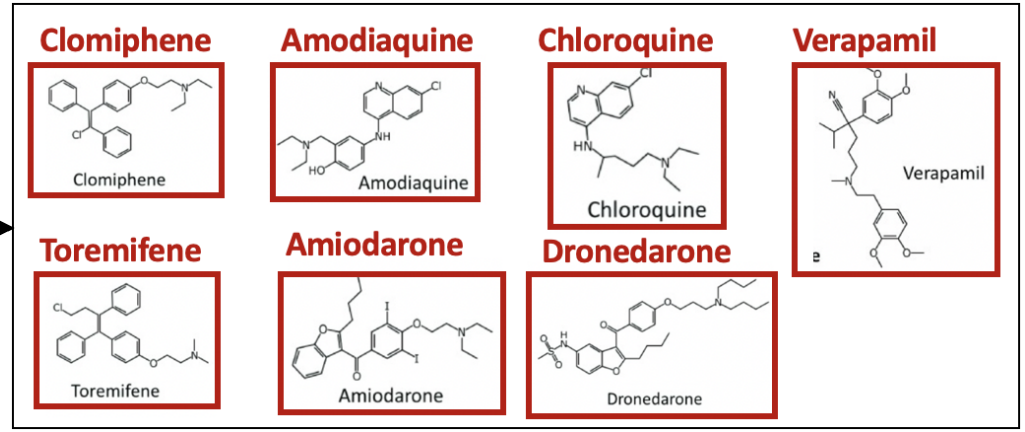
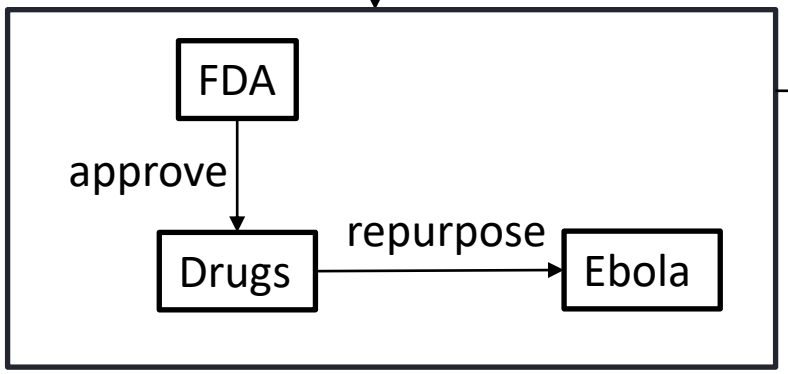


Figure 1.
FDA approved drugs of most interest for repurposing as potential Ebola virus treatments.

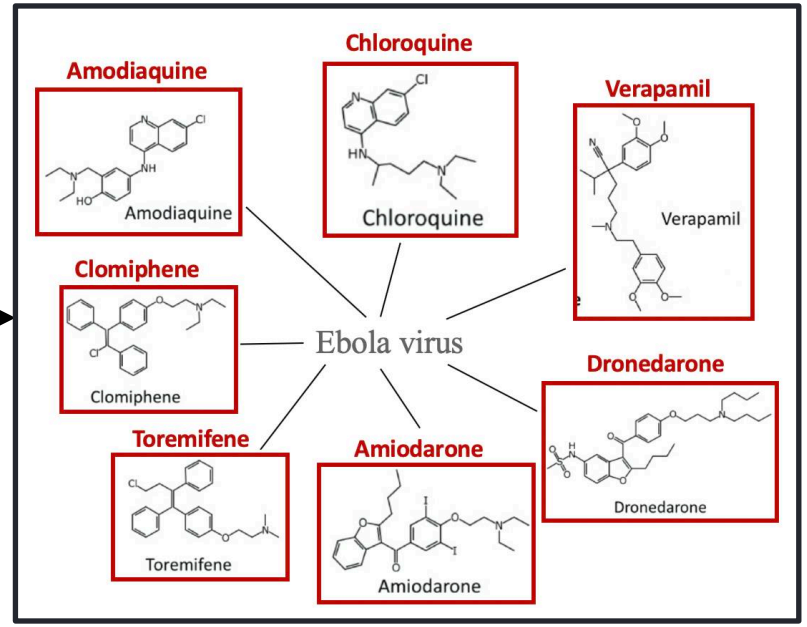
Entity Grounding for Drug Molecular Structure Image



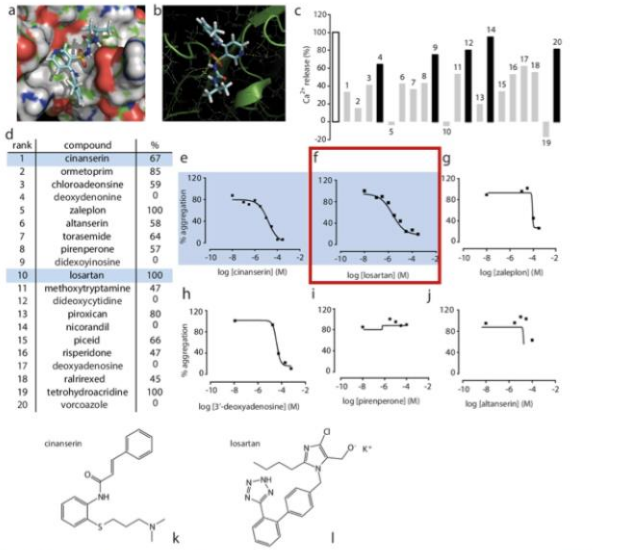
KG from caption text



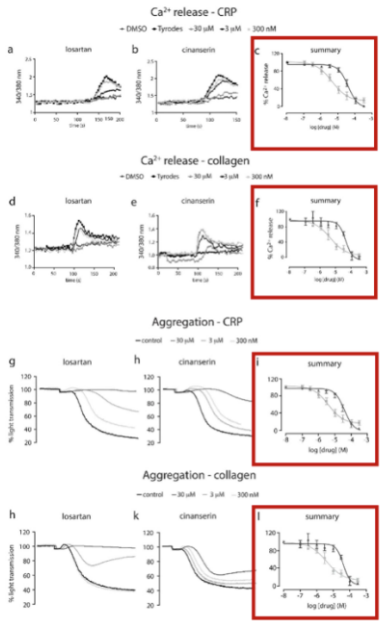
Multimedia Knowledge Graph Expansion



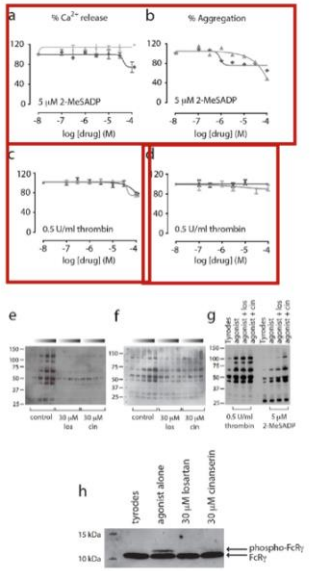
Drug Repurposing Report Generation (Section 6: In vitro Data available)



Src: PMC4074120 Fig. 1. In silico identifies GPVI antagonist. Representative image capture of in silico docking into GPVI using Glide, with space filling model is shown in a, and H-bonding to relevant side chains is detailed in b. The 20 highest ranking compounds were screened for effects on Ca2+ release by the GPVI-specific agonist CRP-XL (10 mg/ml) (c and d, % refers to percent inhibition of Ca2+ release). Maximum Ca2+ release is shown in white, compounds that inhibited Ca2+ release by , 50% or more are in grey, and the remainder in black. Commercially available compounds that inhibited CRP-XL-induced Ca2+ release .50% were further screened by light transmission aggregometry to identify compounds displaying dose-dependent inhibition (e-j). Examples are shown of weak antagonism (g and h) and false positives (i and j). Cinanserin (l) and losartan (k) were taken on for further study.



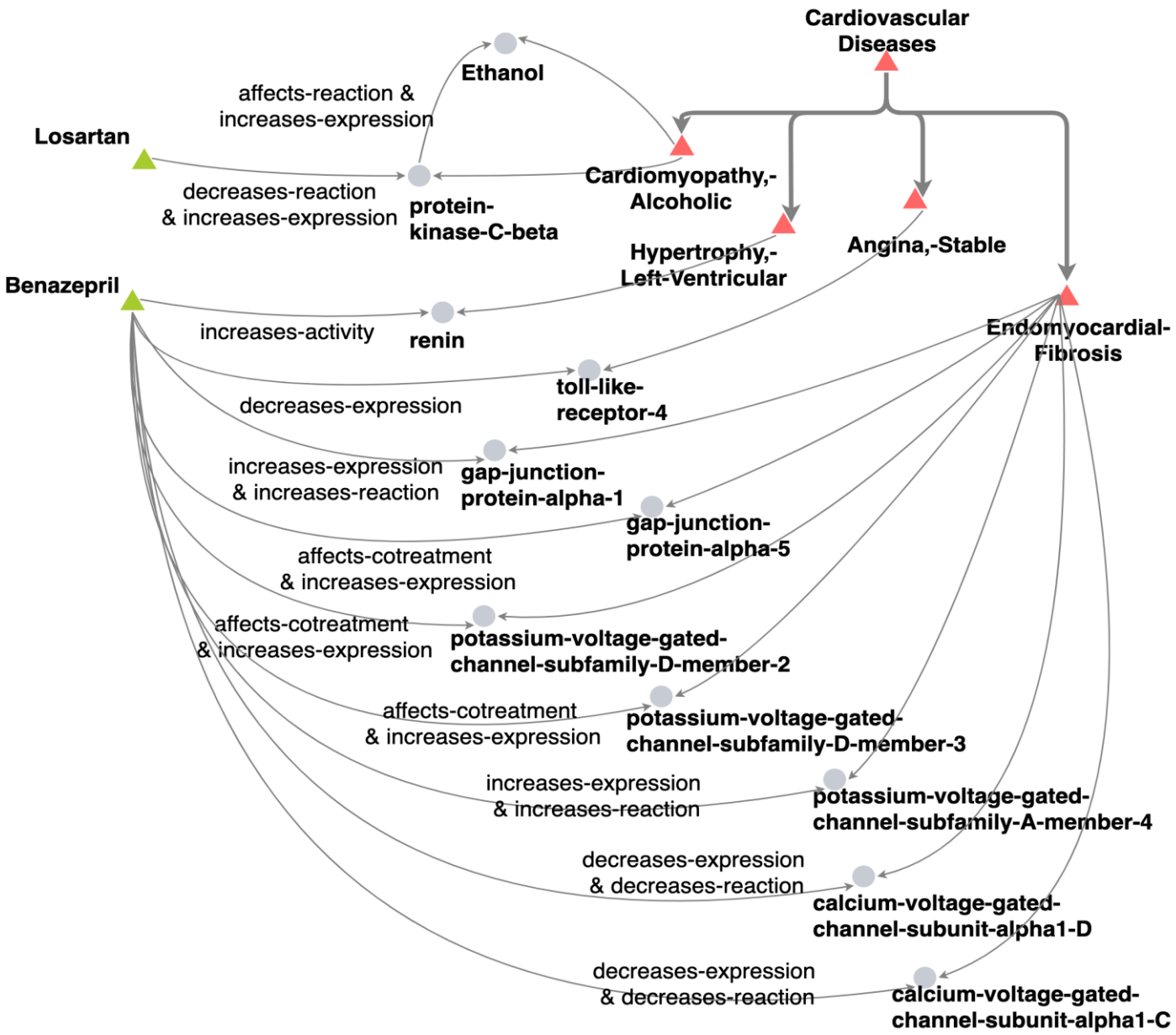
Src: PMC4074120 Fig. 2. Losartan and cinanserin inhibit GPVI-mediated cell activation. Washed human platelets were loaded with fura2-AM and screened for drug-mediated inhibition of Ca2+ release by 1 mg/ml CRP-XL (n = 3, 6 SEM, representative traces and summary, a–c) and 1 mg/ml collagen (n = 3, SEM, representative traces and summary, d–f); losartan (&) and cinanserin (m). To measure aggregation, washed human platelets were incubated with drug for one minute prior to the addition of 1 mg/ml CRP-XL (representative traces and summary shown in g–i) or 1 mg/ml collagen (representative traces and summary shown in j–l).



Src: PMC4074120 Fig. 3. Losartan and cinanserin demonstrate selectivity for GPVI. Ca2+ release and aggregations were carried out with 5 mM of the P2Y12 receptor agonist 2-MeSADP (a, Ca2+ release and b, aggregation), or 0.5 U/ml of the PAR1 and PAR4 receptor agonist thrombin (c, Ca2+ release and d, aggregation). Losartan (X); cinanserin (m), n = 3, 6 SEM. For global tyrosine phosphorylation, washed human platelets were incubated with drug or vehicle alone before addition of 1 mg/ml CRP-XL or collagen. Samples were collected at 10, 30, 60 or 90 seconds (as indicated by the graduated bars with time increasing to the right) in ice cold 26 lysis buffer and separated on 4–12% NuPage gels under reducing conditions. Tyrosine phosphorylation was visualized with 4G10 anti-phosphotyrosine antibody. Losartan and cinanserin reduce CRP-XL- (e) and collagen-(f) induced global tyrosine phosphorylation, but have no effect on thrombin or 2-MeSADP induced global tyrosine phosphorylation (g). Both drugs reduce FcRc phosphorylation (h), (unphosphorylated, lower band; phosphorylated, upper band).

Disease	Cardiovascular disease
PMID, PMCID	Evidence Sentences
22800722 PMC7102827	The in vitro half-maximal inhibitory concentration (IC50) values of food-derived ACE inhibitory peptides are about 1000-fold higher than that of synthetic captopril but they have higher in vivo activities than would be expected from their in vitro activities..... Germinal ACE depends on chloride to a lesser extent compared with the C domain of sACE. Cushman and Cheung reported an optimal in vitro ACE activity of rabbit rung acetone extract in the presence of 300 mM NaCl at pH 8.1-8.3…

Section 9: Has the drug shown evidence of systemic toxicity?



Where are We and What's Next

- Applications
  - Writing History Book
  - News Understanding and Recommendation
  - Disaster Relief
  - Intelligence Analysis
  - Accelerating Scientific Discovery
- Remaining Challenges and Future Directions
  - External Knowledge Acquisition, Reasoning and Incorporation
  - Document-Level Event Extraction
  - Multimedia Information Extraction and Verification
- Grand Vision
- Resources

Event Extraction Challenges: Scene Understanding

- Three Young Boys,ages 2, 5 and 10 survived and are in critical [injure] condition after spending in 18 hours in the cold.

- This was the Italian ship that was taken -- that was captured [transfer-ownership] by Palestinian terrorists back in 1985 [die_time] and some may remember the story of Leon clinghover, he was in a cheal chair and the terrorists shot him and pushed him over the side of the ship into the Mediterranean [die_place] where he [die_victim] obviously, died.

- Then police say the baby's mother pulled out a kitchen knife [die_instrument] opinion on the 911 tape you can hear Williams tape say "go ahead kill me."

- Scenario = ISIS attack → Event = bombing instead of concert

Long-Tail Triggers and Arguments

- Rare Triggers
  - his men **back** to their compound
  - A suicide bomber **detonated** explosives at the entrance to a crowded
  - medical teams **carting** away dozens of wounded victims
  - Today I **was let go** from my job after working there for 4 1/2 years.
  - This morning in Michigan, a second straight night and into this morning, hundreds of people have been **rioting** in Benton harbor.
- Rare Arguments
  - We've seen in the past in Bosnia for example, you held elections and all of the old ethnic thugs [election_person] get into power because they have organization and they have money and they stop the process of genuine building of democracy.
  - He called the case and I'm quoting now, a judge's worst nightmare, but he noted that Maryland Parole Boards and the facility where Michael Serious was held, the institution made what the judge called the final decision on whether to release serious [release-parole_person].
  - Last week Williamson, a mother of four, was found stabbed to death at a condominium [die_place] in Greenbelt, Maryland.
  - A source tell US Enron is considering suing its own investment bankers for giving it bad financial advice [Sue_crime].

Highly Ambiguous Triggers

- Ellison to spend $10.3 billion to **get [org-acquisition]** his company.
- I want to take [transport] this opportunity to stand behind the Mimi and proclaim my solidarity.
- He's left [transport] a lot on the table.
- Stewart has found the road to fortune wherever she has traveled [transport].
- And it's hard to win back that sort of brand equity that she's lost [end-position].
- Still hurts [attack] me to read this.
- We happen to be at a very nice spot by the beach where this is a chance for people to get [transport] away from CNN coverage, everything, and kind of relax.
- He bought the machinery, moved [transport] to a new factory, rehired some of the old workers and started heritage programs.

Event Coreference Resolution Challenges

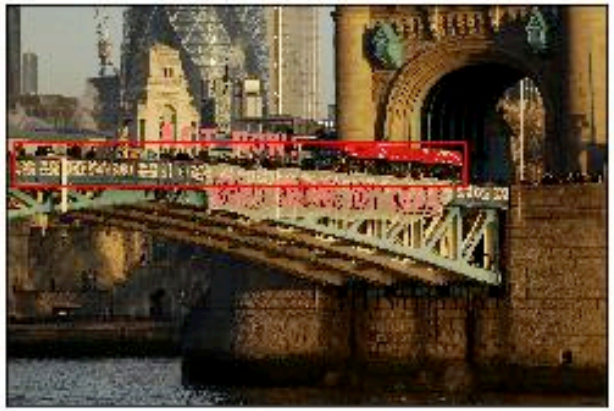
- Super-event and subevent are often mistakenly identified as coreferential
  - These same imposters were later filmed **shooting** handguns and automatic weapons in the direction of the building just minutes after they had switched sides .
  - Right sector goons started the **fire** by throwing Molotov cocktails through the windows

- HC000030E: a sequence of subevents are often mistakenly considered as coreferential
  - Nor does the author speculate on why the police stood by while people hurled themselves from windows to escape the fire or were savagely **beaten** by right wing extremists on the pavement in front of the building .
  - Its true that Washington supports Neo-Nazi extremists who **burned down** the Odessa Trade Unions House .

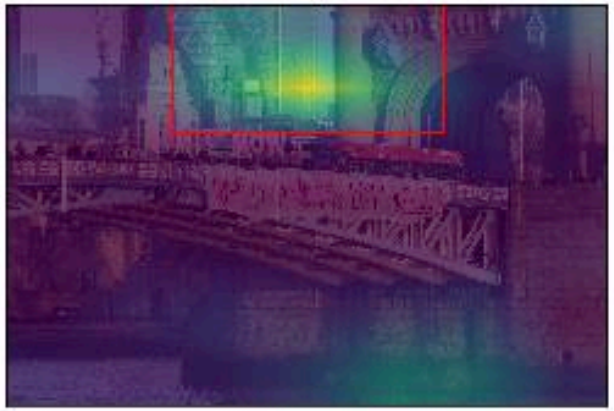
Remaining Challenges for Multimedia Extraction: wrong localization

```
Predicted event: CONFLICT||DEMONSTRATE, ground truth: CONFLICT||DEMONSTRATE
Predicted verb: parading
```

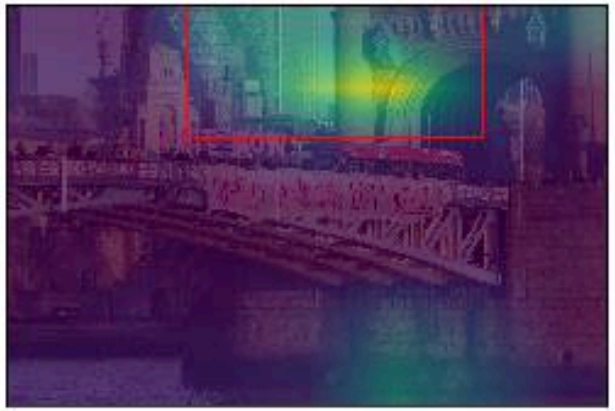
ENTITY: people



PLACE: open air



ENTITY: demonstrator



```
Predicted event: CONFLICT||DEMONSTRATE, ground truth: CONTACT||MEET
Predicted verb: parading
```

ENTITY: people

PLACE: street

ENTITY: troops

Remaining Challenges for Multimedia Extraction: too many instances

```
Predicted event: CONFLICT||DEMONSTRATE, ground truth: CONFLICT||DEMONSTRATE
Predicted verb: parading
```

ENTITY: people

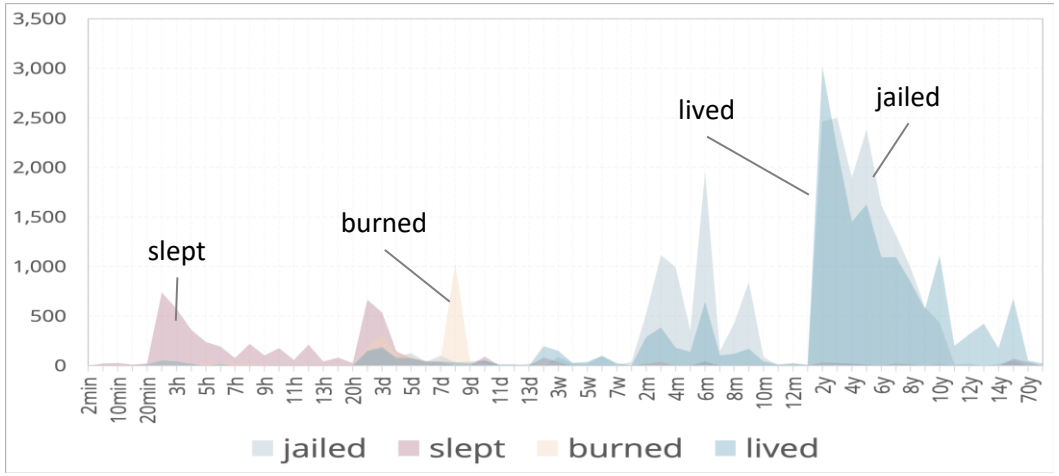
PLACE: street

ENTITY: dissent

Future Direction 1: Knowledge Acquisition and Reasoning

❿ Discover from Google N-grams using lexical patterns

☐ Verb + for + Time / Spent + Time + Verb-ing

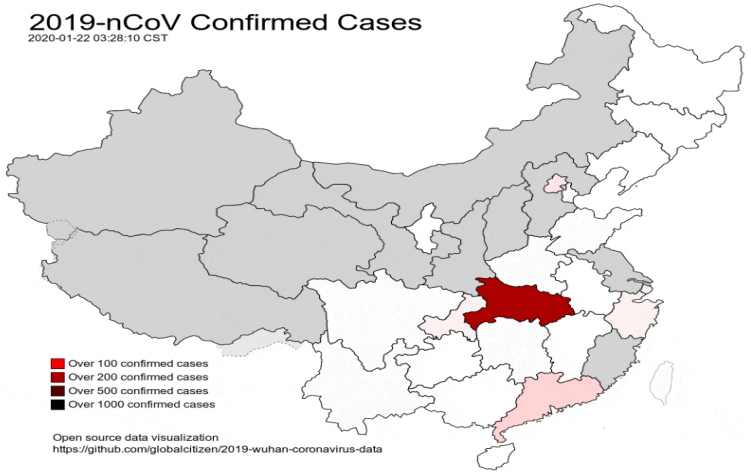


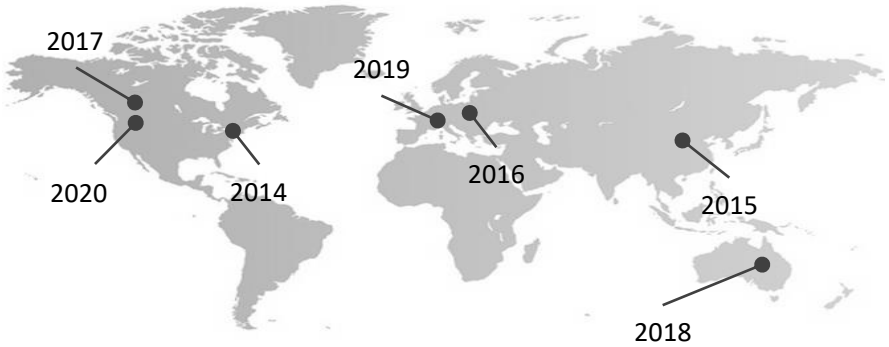
❿ Discover from Wikipedia event infoboxes and categories

Date	1 September 1939 – 2 September 1945 (6 years and 1 day)[a]

Categories: World Wars | World War II | Conflicts in 1939 | Conflicts in 1940 | Conflicts in 1941 | Conflicts in 1942 | Conflicts in 1943 | Conflicts in 1944 | Conflicts in 1945 | Global conflicts | Modern Europe | Modern history | Nuclear warfare

Event Spatial Evolution Pattern

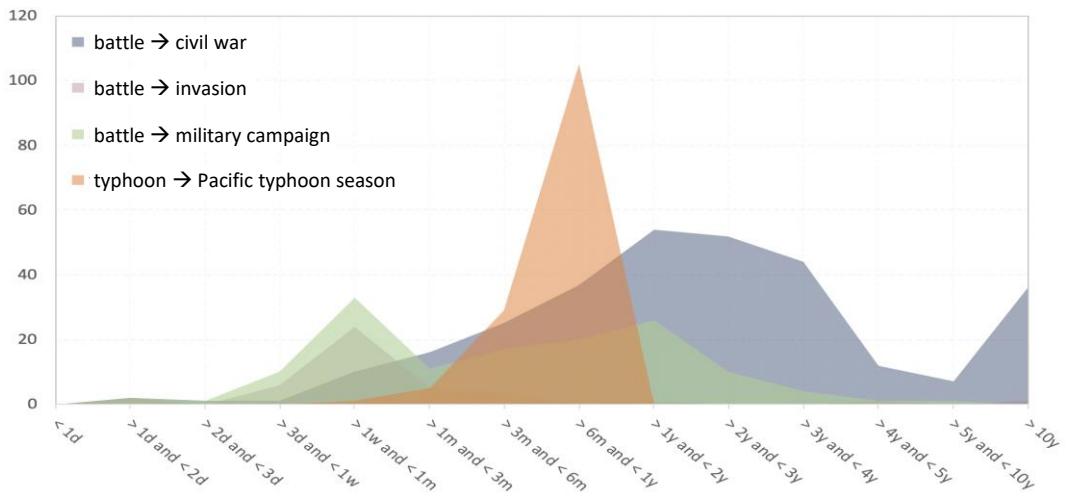




Future Direction 1: Knowledge Acquisition and Reasoning

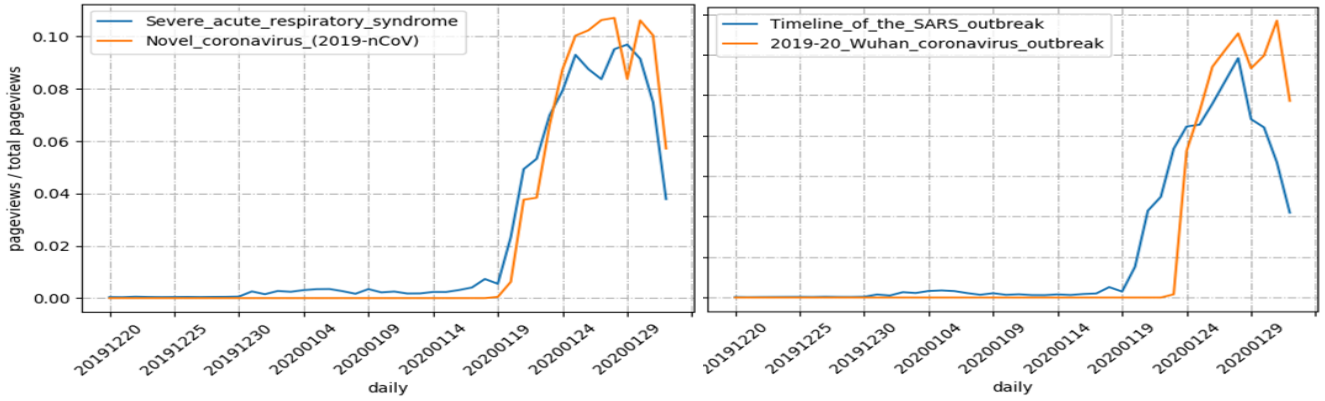
- Exploit Wikipedia/Wikidata
  - layout, category (event type), part of relation, has cause, has immediate cause, has contributing factor, has effect, immediate cause of, start/end time
  - Mining external links in Wikipedia articles and edit history; measuring the similarity of two temporal histograms

Cross-event Time Gap Pattern Discovery



Hierarchical/Causal Composition

Wikipedia Pageview Co-burst



But How to Encode such Symbolic Knowledge?

- We just need to encode them in an elegant way
- The following position slides are from a working group at Dagstuhl Seminar 2019 (Yoav Goldberg, Heng Ji, Dan Roth, Ido Dagan):

Why? Because they are complementary

Representation	Pros	Cons
Symbolic Semantics	Easy to manipulate	Not trivial to define symbols
	Effective for ordering and composition	Hard to integrate incompatible representations across multiple data modalities or languages
	More explainable to users; Easy for users to specify needs / influence results	Has limited coverage; learning not scalable
		Complexity grows very fast with the number of symbols

Why? Because they are complementary

Representation	Pros	Cons
Distributional Semantics	Easy to extend to multi-media and multi-lingual, can serve as a bridge for cross-media cross-lingual common representation	Not explainable
	provide continuous instead of discrete representation: allow soft decisions and soft matching	Hard to control the output and its properties
	more generalizable	hard/impossible to verify/validate the correctness of the result

How to Combine?

1. Use symbolic semantics to represent input structure (e.g., edges in AMR graphs) and distributional semantics to represent nodes (e.g., concept nodes in AMR)
   - ❑ Use in composition based classifiers such as CNN and GCN
   - ❑ Use in composition based methods like Tree/Graph NN.
   - ❑ Match sub-structures using soft-similarity on nodes and hard similarity on edges/structure
2. Combine graph embedding (e.g., knowledge graph embedding, social graph embedding) with text embedding to:
   - ❑ Extend the coverage (Graph embedding provides additional ‘neighbors’.)
   - ❑ Validate
   - ❑ Allow flexible matching

How to Combine?

3. Use symbolic structure to enforce output structure
   - inter-dependency between labels and constraints (e.g., we use Bi-LSTM + CRFs for name tagging, where CRFs layer is used to capture the inter-dependency among labels, I-ORG cannot appear after B-PER, Lin et al., ACL2020)
4. Convert distributional semantic representation to symbolic semantic representation (to show to a user / to edit / to perform symbolic inference later)
   - ground/map the distributions of concepts to taxonomy/ontology/knowledge bases
   - perform hierarchical clustering on event trigger words based on their distributional semantic representations, and then select the centroid entity or predicate lexical dictionary in Propbank/Ontonotes/VerbNet to assign a name to the event type (Huang et al., 2016)
5. "reason with symbolic, compute with distributional"
   - Use local prediction/scoring based on distributional representation + global inference based on symbolic representation (Li et al., EMNLP2020)

Future Direction 2: Document-Level Event Extraction

Seven people **convicted** *last week* in *Vietnam*'s biggest-ever criminal trial, including two former senior government officials, have requested an **appeal** of the verdicts, a court official said Tuesday.

The trial by a *Ho Chi Minh City* court was seen as a litmus test of the communist government's resolve to fight widespread corruption.

The ``godfather'' of organized crime, *Truong Van Cam*, better known as Nam Cam, was convicted of seven crimes, including murder. He was **sentenced** to face a firing squad, and his lawyer has said he also plans to appeal.

Hanh, also a former member of the powerful Communist Party Central Committee, was **convicted** of receiving US$8,500 in bribes from Nam Cam's family to secure the crime boss' early release from labor camp in 1990s. Hanh was **sentenced** to 10 years in jail.

Chien was **convicted** of receiving a stereo set worth 27 million dong (US$1,750) from Nam Cam's family and **sentenced** to six years in jail.

—> Where was Hanh / Chien sentenced?

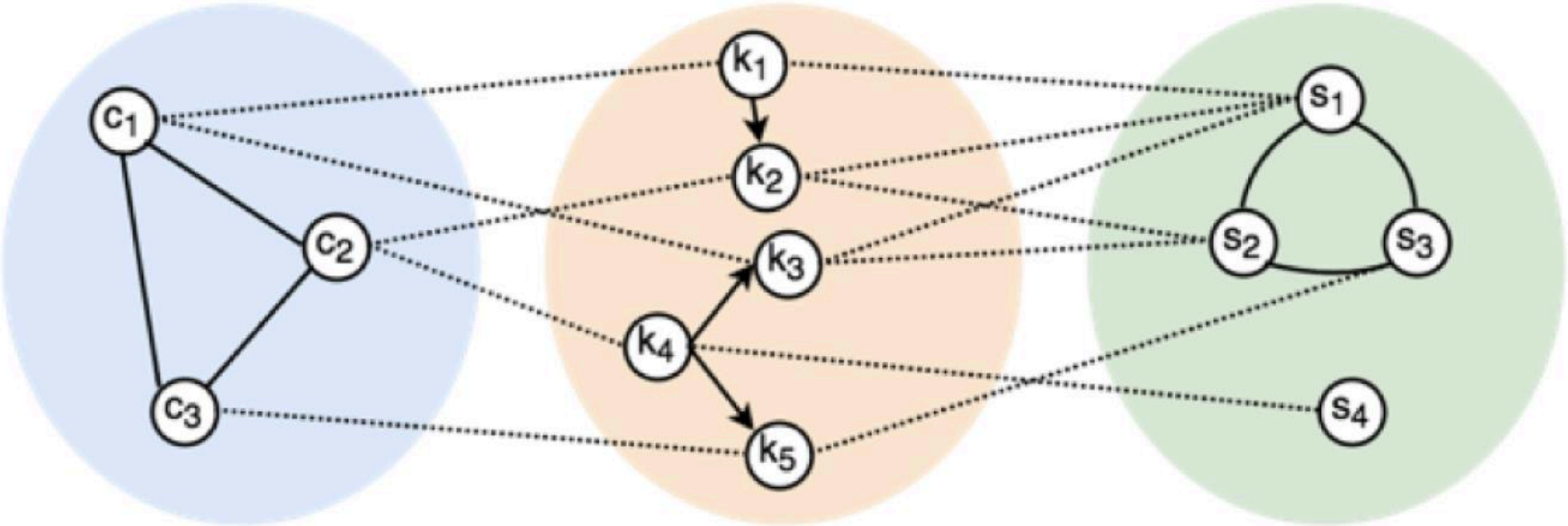
Future Direction 3: Knowledge-element Level Information Consistency Checking

Attacker	Demand	Protester	Time	Location
police	democracy	130K	07-01-2019	Legislative Council Complex

Type	Attacker	Possible Action	Protester	Time	Location
riot	protesters	Sending People's Liberation Army	550K	08-12-2019	Hong Kong International Airport

Future Direction 3: Knowledge-element Level Information Consistency Checking

Context	Knowledge Element	Source (Narrator/User)	
• Time • Location • Quality • Perplexity • Background • Censorship	• Confidence • Provenance • Salience • Novelty • Temporal/Causal/Subevent/ Conflict relations between knowledge elements	• Emotion • Sentiment • Aspect • Confidence • Political Stance • Social Norm • Language	• Perceived Control • Moral value • Power • Expertise • Deception • Data Modality • Human Feedback



Where are We and What's Next

- Applications
  - Writing History Book
  - News Understanding and Recommendation
  - Disaster Relief
  - Intelligence Analysis
  - Accelerating Scientific Discovery
- Remaining Challenges and Future Directions
  - External Knowledge Acquisition, Reasoning and Incorporation
  - Document-Level Event Extraction
  - Multimedia Information Extraction and Verification
- Grand Vision
- Resources

Detailed Event-Centric Challenges

Representing Time

- Representing time of events for the purpose of reasoning about it is still an open problem
- There are other challenges that we still don't have good enough formulations for:
- Semantic clarity
  - What does it mean for a relation to be temporally qualified by some temporal argument T?
    - All time points? At least one? None outside?

      - In 1944 US and Germany were at war;
      - John got married in 1944;
      - Jim was born in 1944.

  - How informative are temporal relations?

      - The great depression started before FDR took office.
      - WWI started before FDR took office.

- Expressiveness
  - Multiple perspectives; Nesting; Negation; Uncertainty

- The lion had a large meal and slept for 24 hours.
  - When did it eat? Exactly 24 hours?
- The lion didn't sleep after having a large meal.
  - [Negated]
- The lion may have had a large meal before sleeping.
  - [Uncertain]
- If the lion has a large meal, it will sleep for 24 hours.
  - [Hypothetical]
- The lion typically sleeps for 24 hours after having large meals.
  - [Repetitive]
- After having a large meal, lions may sleep longer.
  - [Generic]

- Mary graduated from college in 1969 (or 1971)
- John believed in 1995 that Mary attended college in the 1960s
- Mary did not attend college in the 1970s (but at some other time)

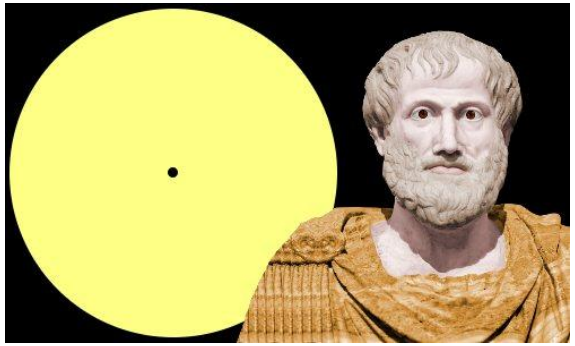
Planning & Reasoning about Events

- Reasoning about Events often requires the ability to consider and plan around implicit events and reason about quantities and time.

Will we make it to dinner before the movie?

Did Aristotle have a laptop?

See [Geva et al. TACL'21]



Mayor Rahm Emanuel now has raised more than $10 million toward his bid for a third term – more than five times the total raised by his 10 challengers combined, campaign finance records show.

Grounding

- There is a need to ground in order to reason about events.
  - Temporally, Spatially
  - Multimodal: Where is this event? (and When did it happen)?
  - Manchester Bomber Believed Muslims Were Mistreated, Sought Revenge
  - The visit to Berlin took place just after the collapse of the wall.

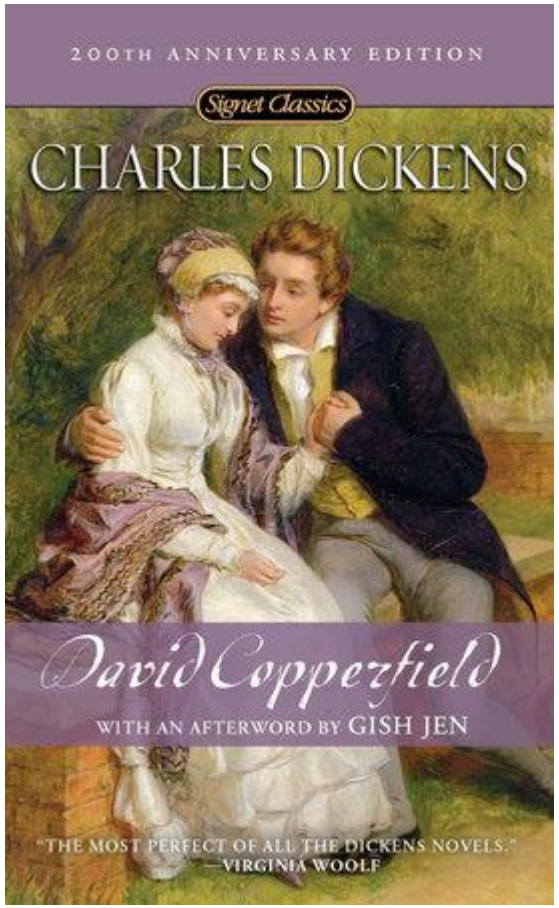
- What is the event?
- When/where did it happen?

- What is the event?
- When/where did it happen?

High Level Event-Centric Challenges

CliffsNotes

- How can we support readers in better ways?
  - The novel features the character David Copperfield, his journey of change and growth from infancy to maturity, as many people enter and leave his life and he passes through the stages of his development. (**Fiction, and you know it**)
  - London and England in the 19-th century; socio-economic state, child exploitation; schools, prisons, emigration to Australia (**True historical facts**)
- Currently, we can't even tell the difference
  - What are the computational tasks we should think about?



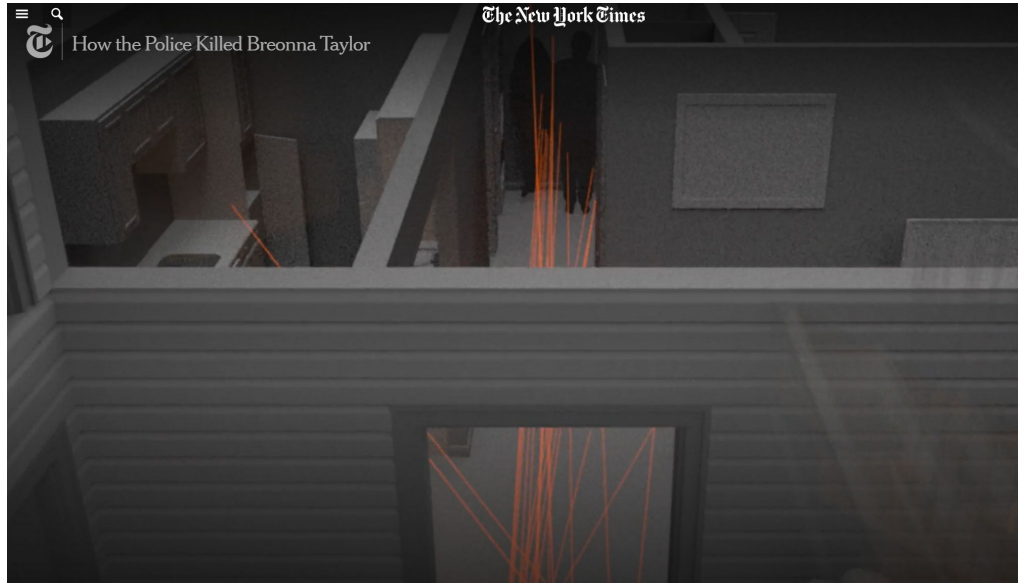
Event Reconstruction

Thank You!

- The NYT put together a reconstruction video of the killing of Breonna Taylor.
  - Reconstruction is done by
    - Analyzing written evidence
    - Transcripts from interviews with witnesses, policemen, neighbors
    - Analysis of physical evidence collected in the scene
  - It took weeks to produce

- Can this be automated to form a reconstruction of a (chain of) events?

  - Ideally, this should be automated as a support tool for analysts
    - Requires multimodal capabilities
    - Understanding and reasoning about quantities & geometry
    - Reasoning about uncertainty
    - Reasoning about (possible) conflicting events/reports
      - Notion of multiple perspective & trustworthiness



Multimodal + 3D reconstruction:
https://www.nytimes.com/video/us/100000007348445/breonna-taylor-death-cops.html

Some Key Challenges

- History Writing:
  - Organize chapters by major events clusters and order them on a timeline
  - Each chapter looks like a Wikipedia page
    - The description is organized by multimedia timeline with detailed source and evidence information, links to original news articles
    - Detailed participants (arguments) and their roles, and their connections and relations
    - Infobox shows event-event relations: temporal, causal and hierarchical
  - Never-ending updating over time; put up to the wild for human editing and curation

- This introduces multiple challenges in Event-Centric NLP, but also in summarization and argumentation

Protests against the government [edit]

4 December 2011 [edit]

On 4 November 2011, during the annual Russian March event, representatives of "The Russians" movement declared a protest action
unapproved and took place on 4 December at 21:00 in Moscow. The statement of non-recognition of electoral results spread widely. Cit
said to have occurred during the elections. Alexander Belov declared the beginning of the "Putin, go away!" campaign.[37] The protest a
Alexander Belov, Dmitry Dyomushkin, George Borovikov were arrested along with dozens of other nationalists. The head of the banned
observer. Mass detentions of other public organizations occurred in Moscow. According to police some 258 persons have been detaine

5–7 December 2011 [edit]

On 5 December, around 5,000 opponents of the government began protesting in Moscow, denouncing Vladimir Putin and his governme
Putin step down, whilst some demanded revolution.[15][39] Alexey Navalny, a top blogger and anti-corruption activist who branded Putir
postings on his LiveJournal blog and Twitter account. Navalny's agitation was denounced by United Russia as "typical dirty self-promoti
account.[40][41]

Many pro-government supporters, including the pro-Putin youth group Nashi, were mobilized on 6 December at the site of the planned
Nashi on Manezhnaya Square[43] and an 8,000-strong rally of the Young Guard on Revolution Square.[44] About 500 pro-United Russi
expected anti-government protests. It emerged that 300 protesters had been arrested in Moscow the night before, along with 120 in St.
slogans against Putin,[15] whilst anti-government protesters at Revolution Square clashed with riot police and interior ministry troops. T
Square and dozens of arrests were reported, including Boris Nemtsov, an opposition leader and former deputy prime minister,[48] and
charged. At least one Russian journalist claimed he was beaten by police officers who stamped on him and hit his legs with batons.[50]
place. After three and a half hours, the Moscow protest came to an end.[51]

Attempts to stage a large protest in Moscow on 7 December fizzled out due to a large police presence in the city.[13]

10 December 2011 [edit]

Rally in Bolotnaya Square in Moscow on 10 December 2011

Via a Facebook group "Суббота на Болотной площади" (Saturday at Bolotnaya Square),[52]
December.[53][54] Prior to the demonstration newspapers commented that tens of thousands
Moscow,[55][56] and, similarly, over 5,000 in St. Petersburg.[57] A permit had originally been iss
Square. By 8 December, more than 30,000[52] had accepted the Facebook invitation to attend.
demonstration was authorized by the Moscow government for the demonstration which took pl
made by Putin that police and security forces would be deployed to deal with anyone participat
was peaceful and without attempts by the state to prevent or disrupt it.[59][60] Rapper Noize M
specially from Paris for the occasion.[61] Guerrilla theater by FEMEN and the circulation of a p
protests.[62][63]

Where are We and What's Next

- Applications
  - Writing History Book
  - News Understanding and Recommendation
  - Disaster Relief
  - Intelligence Analysis
  - Accelerating Scientific Discovery
- Remaining Challenges and Future Directions
  - External Knowledge Acquisition, Reasoning and Incorporation
  - Document-Level Event Extraction
  - Multimedia Information Extraction and Verification
- Grand Vision
- Resources

Software and Resources

- Joint Neural Information Extraction system
  - http://blender.cs.illinois.edu/software/oneie/

- Multimedia Event Extraction system and new benchmark with annotated data set
  - GitHub: https://github.com/GAIA-AIDA/uiuc_ie_pipeline_fine_grained
  - Text IE DockerHub: https://hub.docker.com/orgs/blendernlp/
  - Visual IE repositories: https://hub.docker.com/u/dannapierskitoptal

- COVID-19 knowledge graph construction and drug repurposing generation
  - http://blender.cs.illinois.edu/covid19/

Event Extraction References

[1] Ying Lin, Heng Ji, Fei Huang, Lingfei Wu. 2020. A Joint End-to-End Neural Model for Information Extraction with Global Features. Proc. ACL2020.
[2] Manling Li, Alireza Zareian, Qi Zeng, Spencer Whitehead, Di Lu, Heng Ji, Shih-Fu Chang. 2020. Cross-media Structured Common Space for Multimedia Event Extraction. Proc. ACL2020.
[3] Manling Li, Alireza Zareian, Ying Lin, Xiaoman Pan, Spencer Whitehead, Brian Chen, Bo Wu, Heng Ji, Shih-Fu Chang, Clare Voss, Daniel Napierski and Marjorie Freedman. 2020.GAIA: A Fine-grained Multimedia Knowledge Extraction System. Proc. ACL2020 Demo [Best Demo Paper Award].
[4] Heng Ji and Clare Voss. 2020. Low-resource Event Extraction based on Share-and-Transfer Invited Book Chapter for Events and Stories, Cambridge Press.
[5] Lifu Huang, Heng Ji, Kyunghyun Cho, Ido Dagan, Sebastian Riedel and Clare Voss. 2018. Proc. The 56th Annual Meeting of the Association for Computational Linguistics (ACL2018).
[6] Hongming Zhang, Haoyu Wang, Dan Roth. 2020. Unsupervised Label-aware Event Trigger and Argument Classification. arXiv:2012.15243.
[7] Seth Ebner, Patrick Xia, Ryan Culkin, Kyle Rawlins and Benjamin Van Durme. 2020. Multi-Sentence Argument Linking. Proc. the 58th Annual Meeting of the Association for Computational Linguistics.
[8] Xinya Du and Claire Cardie. 2020. Event Extraction by Answering (Almost) Natural Questions. Proc. the 58th Annual Meeting of the Association for Computational Linguistics.

Event-Event Relation References

1. Algorithms for scoring coreference chains. Bagga & Baldwin, 1998.
2. Discriminative training methods for hidden markov models: Theory and experiments with perceptron algorithms. Collins, 2002.
3. A Linear Programming Formulation for Global Inference in Natural Language Tasks. Roth & Yih, 2004.
4. Inducing temporal graphs. Bramsen et al., 2006.
5. Refining Event Extraction through Cross-document Inference. Ji & Grishman, 2008.
6. Unsupervised Learning of Narrative Event Chains. Chambers & Jurafsky, 2008.
7. Jointly combining implicit constraints improves temporal ordering. Chambers & Jurafsky, 2008.
8. Sentence Level Event Detection and Coreference Resolution. Naughton, 2009.
9. Graph-based event coreference resolution. Chen & Ji, 2009.
10. Using document level cross-event inference to improve event extraction. Liao & Grishman, 2010.
11. Evaluation Metrics For End-to-End Coreference Resolution Systems. Cai & Strube, 2010.
12. Predicting globally-coherent temporal structures from texts via endpoint inference and graph decomposition. Denis & Muller, 2011.
13. Minimally supervised event causality identification. Do et al., 2011.
14. Temporal evaluation. UzZaman & Allen, 2011.
15. Joint inference for event timeline construction. Do et al., 2012.
16. Semantic Relations between Events and their Time, Locations and Participants for Event Coreference Resolution. Cybulska & Vossen, 2013.
17. TEMPEVAL-3: Evaluating time expressions, events, and temporal relations. UzZaman et al., 2013.
18. HiEve: A Corpus for Extracting Event Hierarchies from News Stories. Glavas et al., 2014.
19. CATENA: CAusal and TEmporal relation extraction from NAtural language texts. Mirza & Tonelli, 2016.
20. Event Detection and Co-reference with Minimal Supervision. Peng et al., 2016.
21. Which Coreference Evaluation Metric Do You Trust? A Proposal for a Link-based Entity Aware Metric. Moosavi & Strube, 2016.
22. Story Comprehension for Predic3ng What Happens Next. Chaturvedi et al., 2017.
23. A Structured Learning Approach to Temporal Relation Extraction. Ning et al., 2017.
24. Joint Reasoning for Temporal and Causal Relations. Ning et al., 2018.
25. A Multi-Axis Annotation Scheme for Event Temporal Relations. Ning et al., 2018.
26. Improving Temporal Relation Extraction with a Globally Acquired Statistical Resource. Ning et al., 2018.
27. KnowSemLM: A Knowledge Infused Semantic Language Model. Peng et al., 2019.
28. Multilingual Entity, Relation, Event and Human Value Extraction. Li et al., 2019.

Event Prediction References

[1] Zack, et al. Human brain activity time-locked to perceptual event boundaries. Nature neuroscience, 4(6):651–655.
[2] Chambers and Jurafsky. Unsupervised learning of narrative event chains. ACL, 2008
[3] Radinsky and Horvitz. Mining the Web to Predict Future Events. WSDM, 2013
[4] Berant, et al. Modeling Biological Processes for Reading Comprehension. EMNLP, 2014
[5] Chaturvedi, et al. Story comprehension for predicting what happens next. EMNLP, 2017
[6] Rashkin, et al. Event2Mind: Commonsense Inference on Events, Intents, and Reactions. ACL, 2108
[7] Liu, et al. Automatic event salience identification. EMNLP, 2018
[8] Zhukov et al. Cross-task weakly supervised learning from instructional videos. CVPR, 2019
[9] Zhang, et al. Analogous Process Structure Induction for Sub-event Sequence Prediction. EMNLP, 2020
[10] Chen, et al. “What are you trying to do?” Semantic typing of event processes. CoNLL, 2020
[11] Ning et al. TORQUE: A Reading Comprehension Dataset of Temporal Ordering Questions. EMNLP, 2020
[12] Lyu, et al. Reasoning about Goals, Steps, and Temporal Ordering with WikiHow. EMNLP, 2020
[13] Jindai, et al. Is Killed More Significant than Fled? A Contextual Model for Salient Event Detection. COLING, 2020
[14] Fried, et al. Learning to Segment Actions from Observation and Narration. ACL, 2020
[15] Zhang, et al. Diagnostic Prediction with Sequence-of-sets Representation Learning for Clinical Events. AIME, 2020
[16] Surís, et al. Learning the Predictability of the Future. arXiv:2101.01600, 2021

Event Knowledge Acquisition References

[1] WSC: Hector Levesque, Ernest Davis, and Leora Morgenstern. The winograd schema challenge. KRR 2012.
[2] *ConceptNet:* Hugo Liu and Push Singh, ConceptNet - a practical commonsense reasoning tool-kit, BTTJ, 2004
[3] ATOMIC: Maarten Sap, Ronan LeBras, Emily Allaway, Chandra Bhagavatula, Nicholas Lourie, Hannah Rashkin, Brendan Roof, Noah A. Smith, Yejin Choi, ATOMIC: An Atlas of Machine Commonsense for If-Then Reasoning. AAAI 2019
[4] COMET: Antoine Bosselut, Hannah Rashkin, Maarten Sap, Chaitanya Malaviya, Asli Celikyilmaz, and Yejin Choi. COMET: commonsense transformers for automatic knowledge graph construction. ACL 2019.
[5] LAMA: Fabio Petroni, Tim Rocktaschel, Sebastian Riedel, Patrick Lewis, Anton Bakhtin, Yuxiang Wu, and Alexander H. Miller. Language models as knowledge bases? EMNLP 2019.
[6] ASER: Hongming Zhang, Xin Liu, Haojie Pan, Yangqiu Song, and Cane Wing-Ki Leung. ASER: A Large-scale Eventuality Knowledge Graph. WWW 2020.
[7] TransOMCS: Hongming Zhang, Daniel Khashabi, Yangqiu Song, and Dan Roth. TransOMCS: From Linguistic Graphs to Commonsense Knowledge. International Joint Conference on Artificial Intelligence (IJCAI). 2020.
[8] KnowlyWood: Niket Tandon, Gerard de Melo, Abir De, and Gerhard Weikum. 2015. Knowlywood: Mining Activity Knowledge From Hollywood Narratives. CIKM 2015.
[9] Manling Li, Qi Zeng, Kyunghyun Cho, Heng Ji, Jonathon May, Nathanael Chambers, Clare Voss. Connecting the Dots: Event Graph Schema Induction with Path Language Mod

Acknowledgement

This research is based upon work supported in part by U.S. DARPA KAIROS Program No. FA8750- 19-2-1004, and U.S. DARPA AIDA Program No. FA8750-18-2-0014. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of DARPA, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright annotation therein.