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)

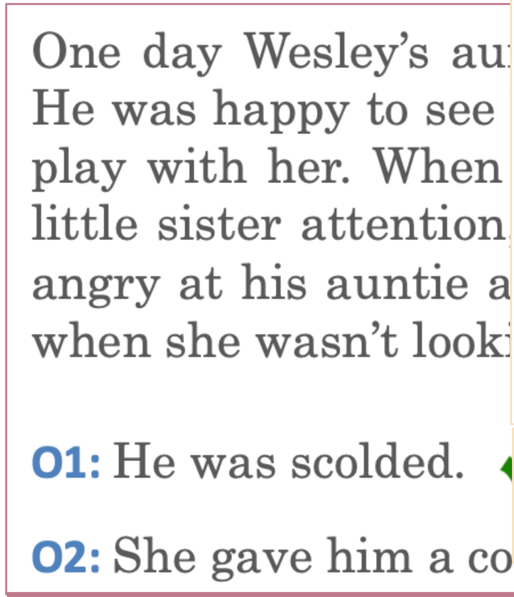
- Given
  - a piece of text (often long enough to contain multiple events)
  - ~~the head phrases of two many events~~
- Extract the relationship(s) between ~~this~~ all event pairs
  - most works focus on one type of relationship, e.g., only predicting coreference relations, or only predicting temporal 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)

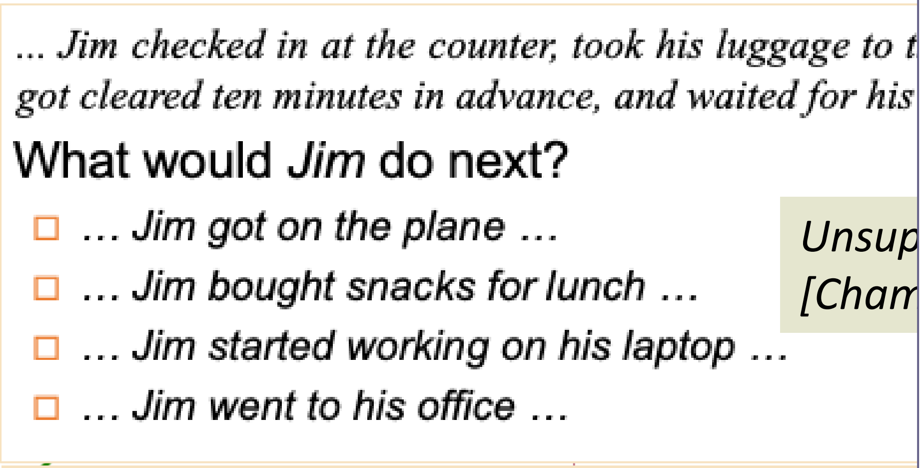
- Given
  - a piece of text (often long enough to contain multiple events)
  - ~~the head phrases of two many events~~
- Extract the events and the relationship(s) between ~~this~~ all event pairs
  - most works focus on one type of relationship, e.g., only predicting coreference relations, or only predicting 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)

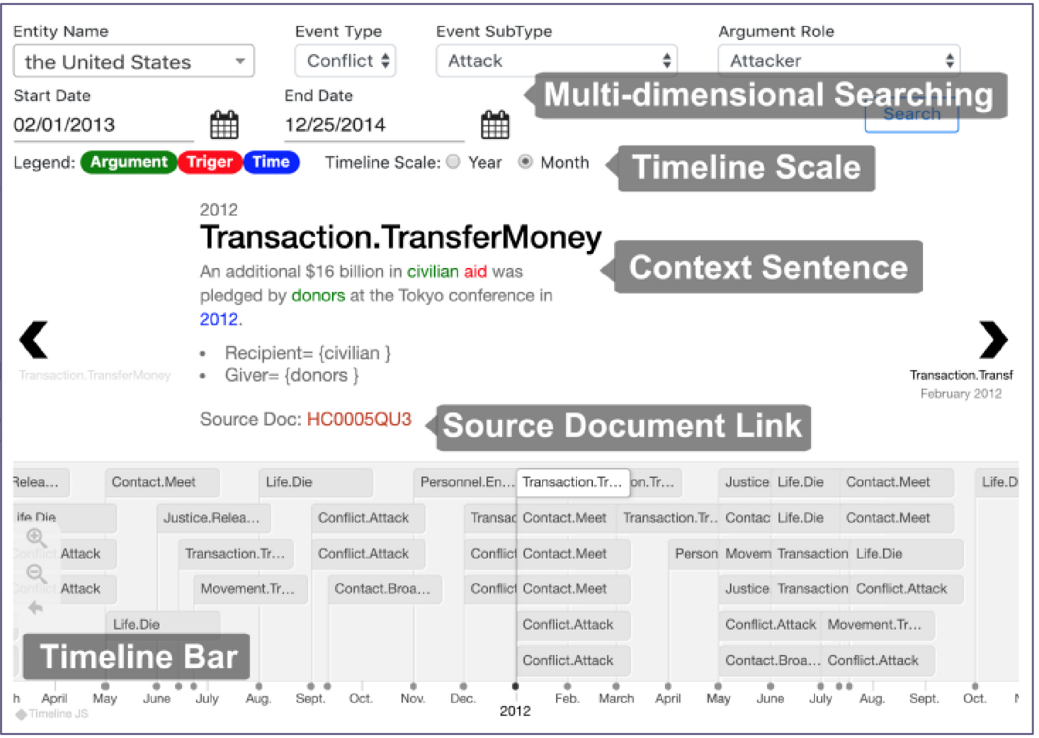
- This part only covers event-event relationships.
- StoryCloze, script learning, schema induction, timeline construction, etc. can also be viewed as tackling relationships among multiple events, but will be covered in later sections of this tutorial.



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



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

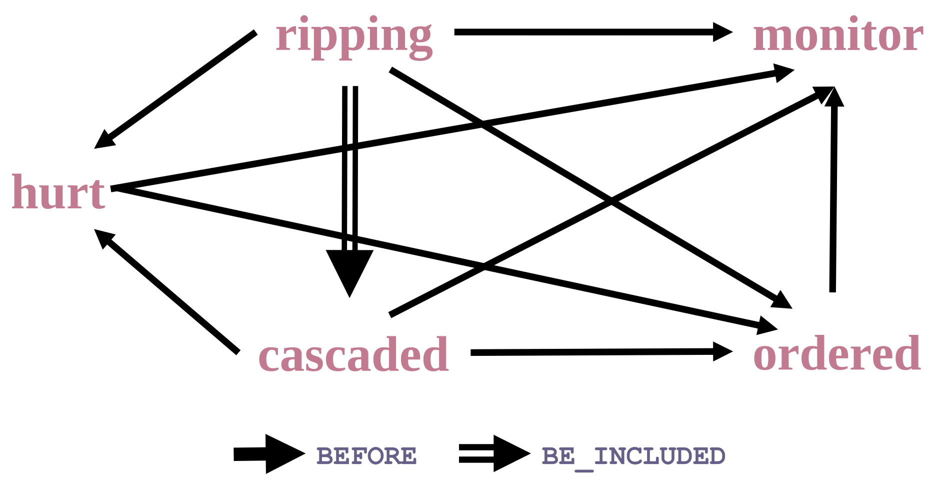


Multilingual Entity, Relation, Event and Human Value Extraction [Li et al., 2019]

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

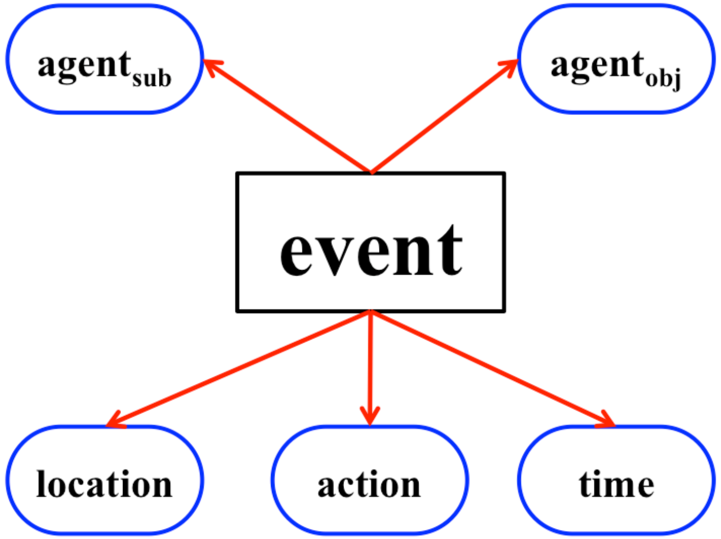
[the pandemic]$_2$ HAPPENS BEFORE [held virtually]$_1$

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 $[\text{walk}]_1$ after $[\text{dinner}]_2$.

He took a $[\text{walk}]_1$ after $[\text{dinner}]_2$ today.

$[\text{walk}]_1$ happens after $[\text{dinner}]_2$ 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?

Multiple events

Event properties and extraction

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: UzZaman SEM'13
P: Glavas LREC'14

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

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

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: Wang EMNLP'20

E, P: Zhou ACL'20

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'20

T: Ning EMNLP'20

*E: Ji ACL'08

E: Naughton PhD'09

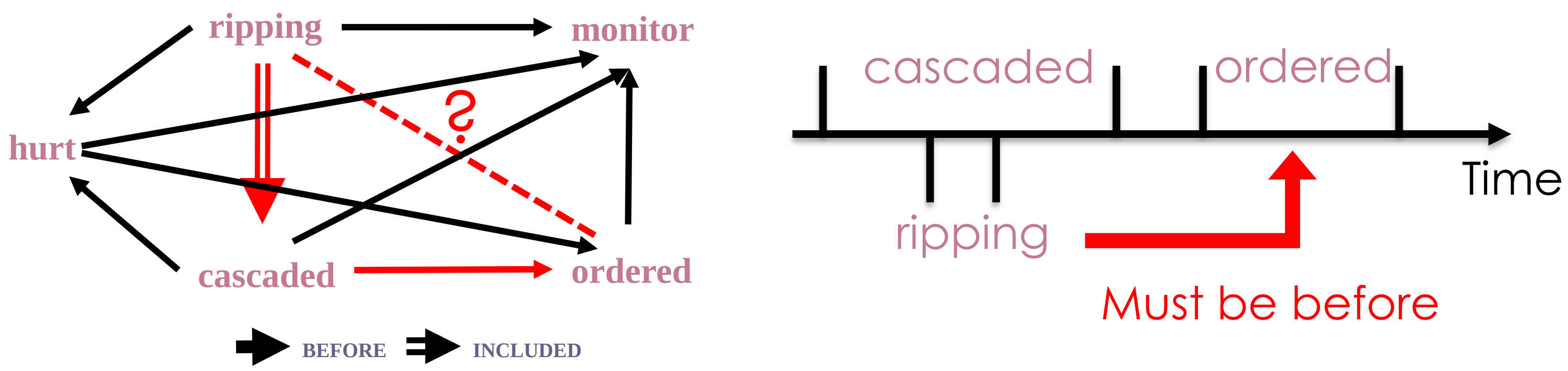
E: Peng EMNLP'16

*E: Liao ACL'10

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

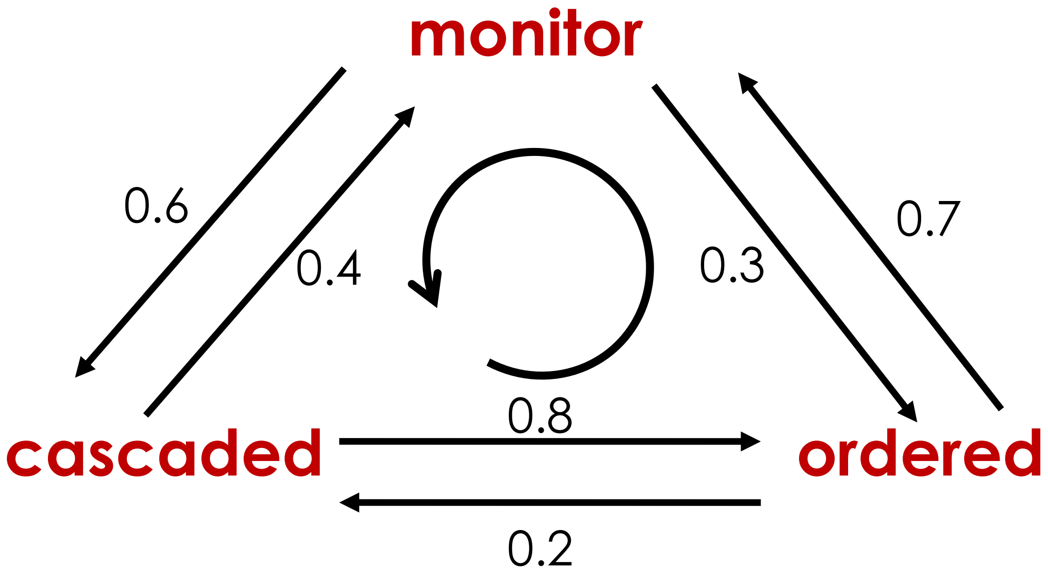
Confidence/softmax

(**cascaded** , **ordered**) ---> model --->

- before: 0.4
- after: 0.2
- includes: 0.1
- included: 0.1
- equal: 0.0
- vague: 0.2

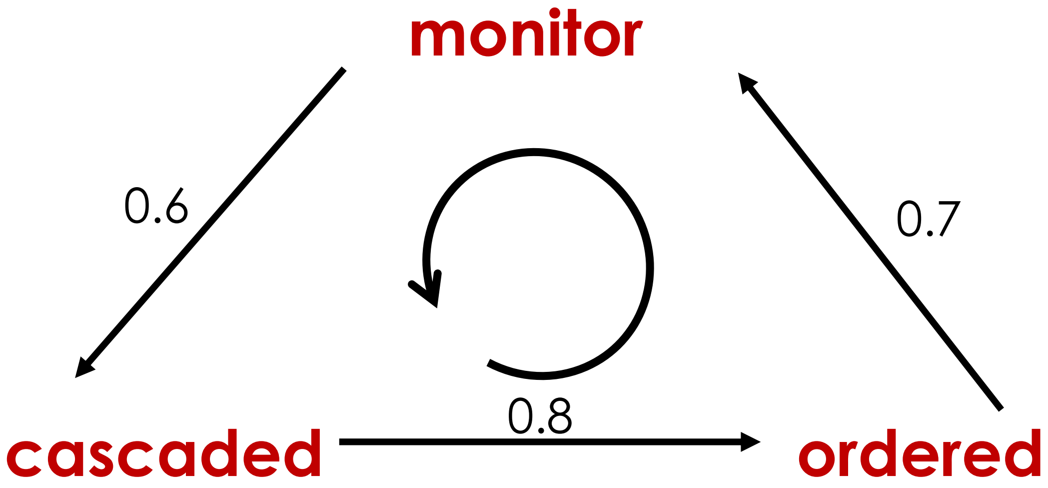
Based on these confidence scores, we need to solve for the final temporal graph.

Global Inference (A Toy Example)



Time cannot be a loop!

Global Inference (A Toy Example)

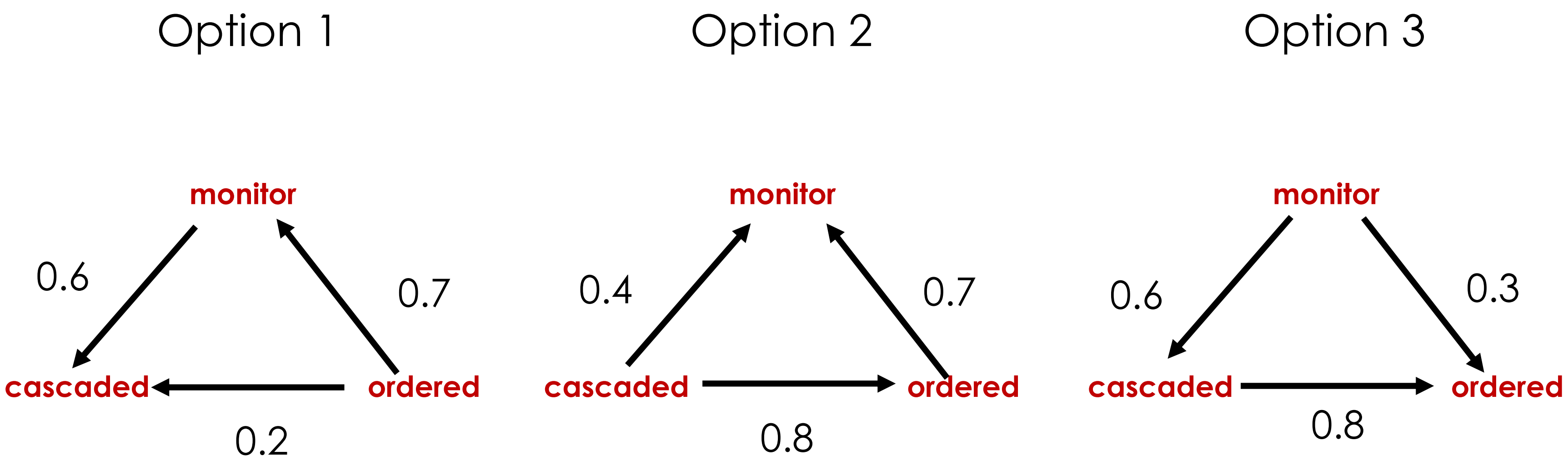


Local inference:

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)



0.6+0.2+0.7=1.5

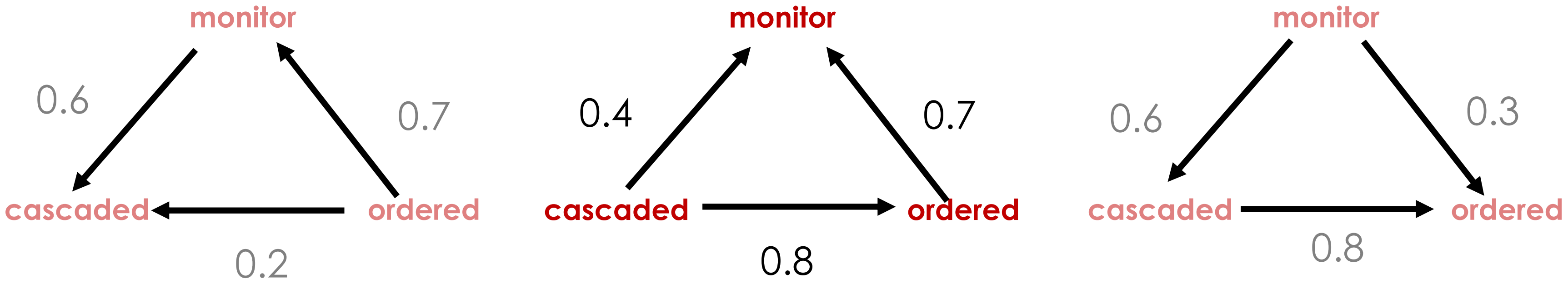
0.4+0.8+0.7=1.9

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

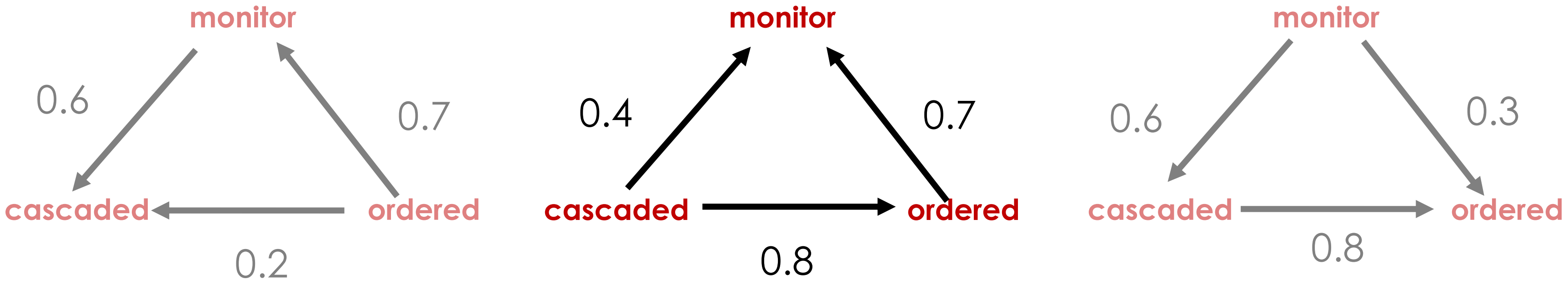


0.6+0.2+0.7=1.5　　0.4+0.8+0.7=1.9　　0.6+0.3+0.8=1.7

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

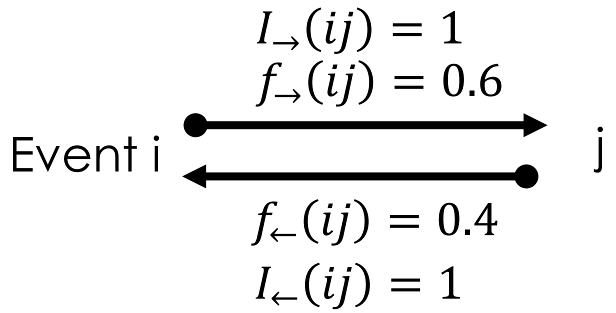
0.4+0.8+0.7=1.9

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)

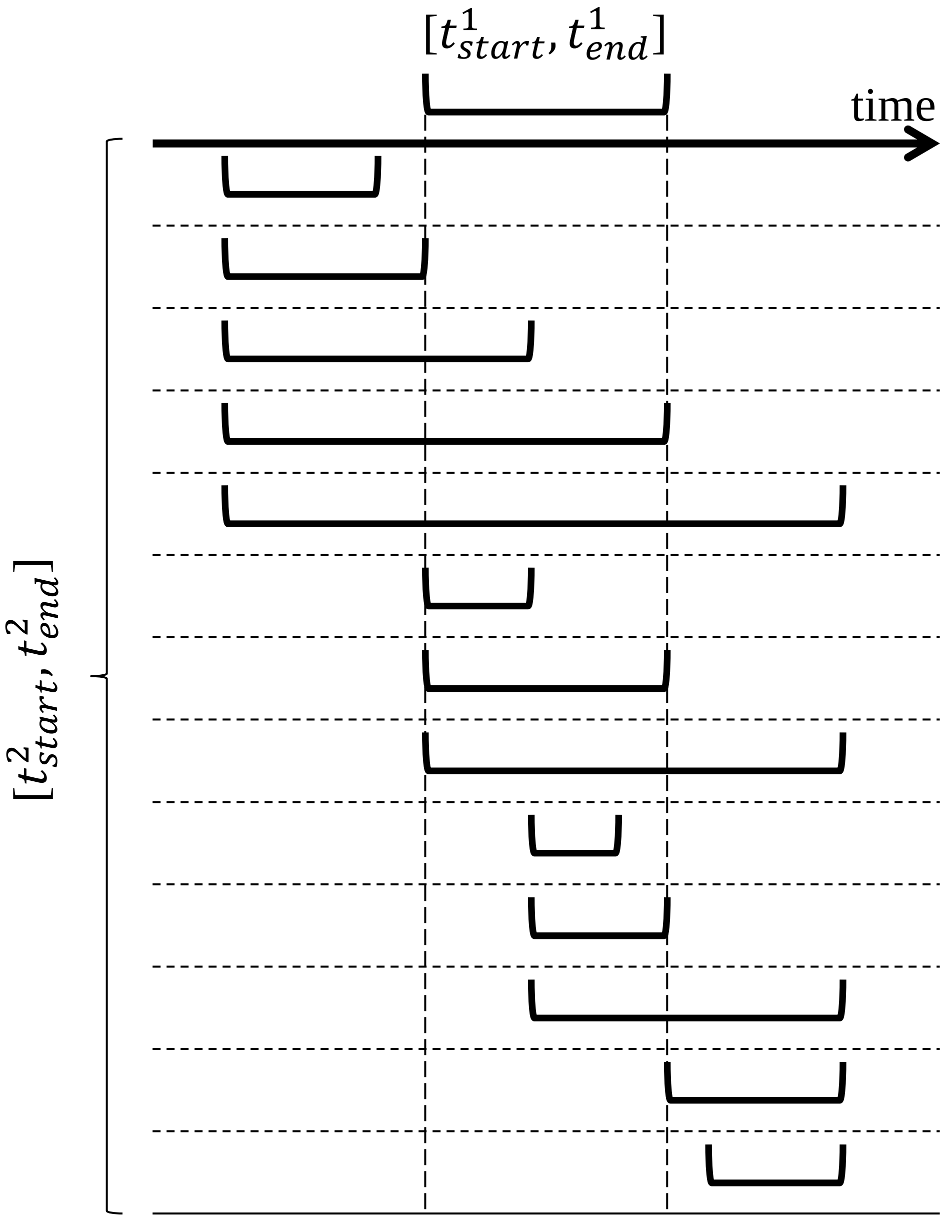
We're maximizing the score of an entire graph **while enforcing transitivity constraints**.

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 → r_1

Relation between Event2 and Event3 → r_2

Relation between Event1 and Event3 → $\mathrm{Trans}(r_1, r_2)$

No.	r_1	r_2	$\mathrm{Trans}(r_1, r_2)$
1	r	r	r
2	r	**s**	r
3	r_1	r_2	$\overline{\mathrm{Trans}(\bar{r}_2, \bar{r}_1)}$
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. $\forall i, j, k$

$$\sum_r I_r(ij) = 1,$$

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

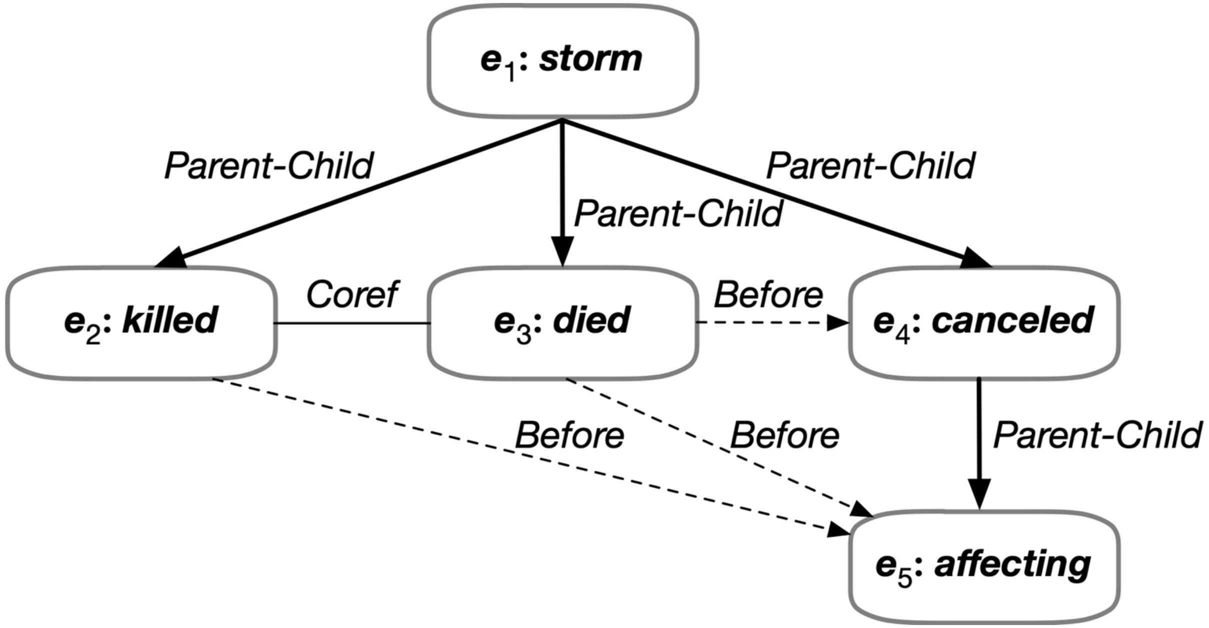
$$J_{causes}(ij) \leq I_{before}(ij)$$

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	–

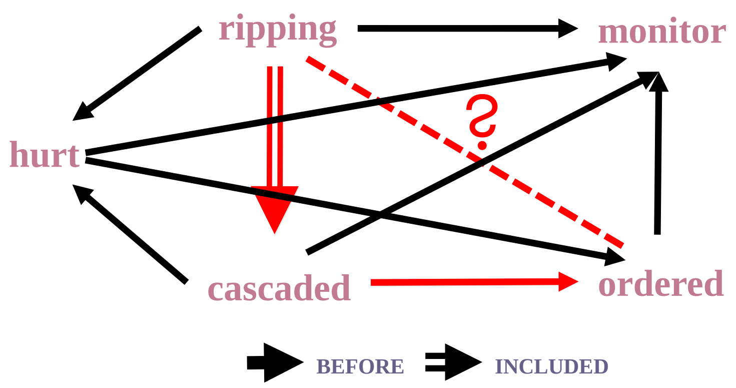
Relation between Event1 and Event2

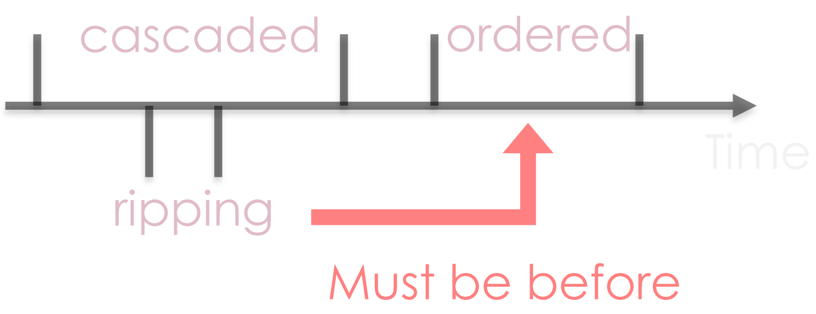
Constrained Learning for Event-Event Relation Extraction. Wang et al., EMNLP2020.

But how do we train the model?

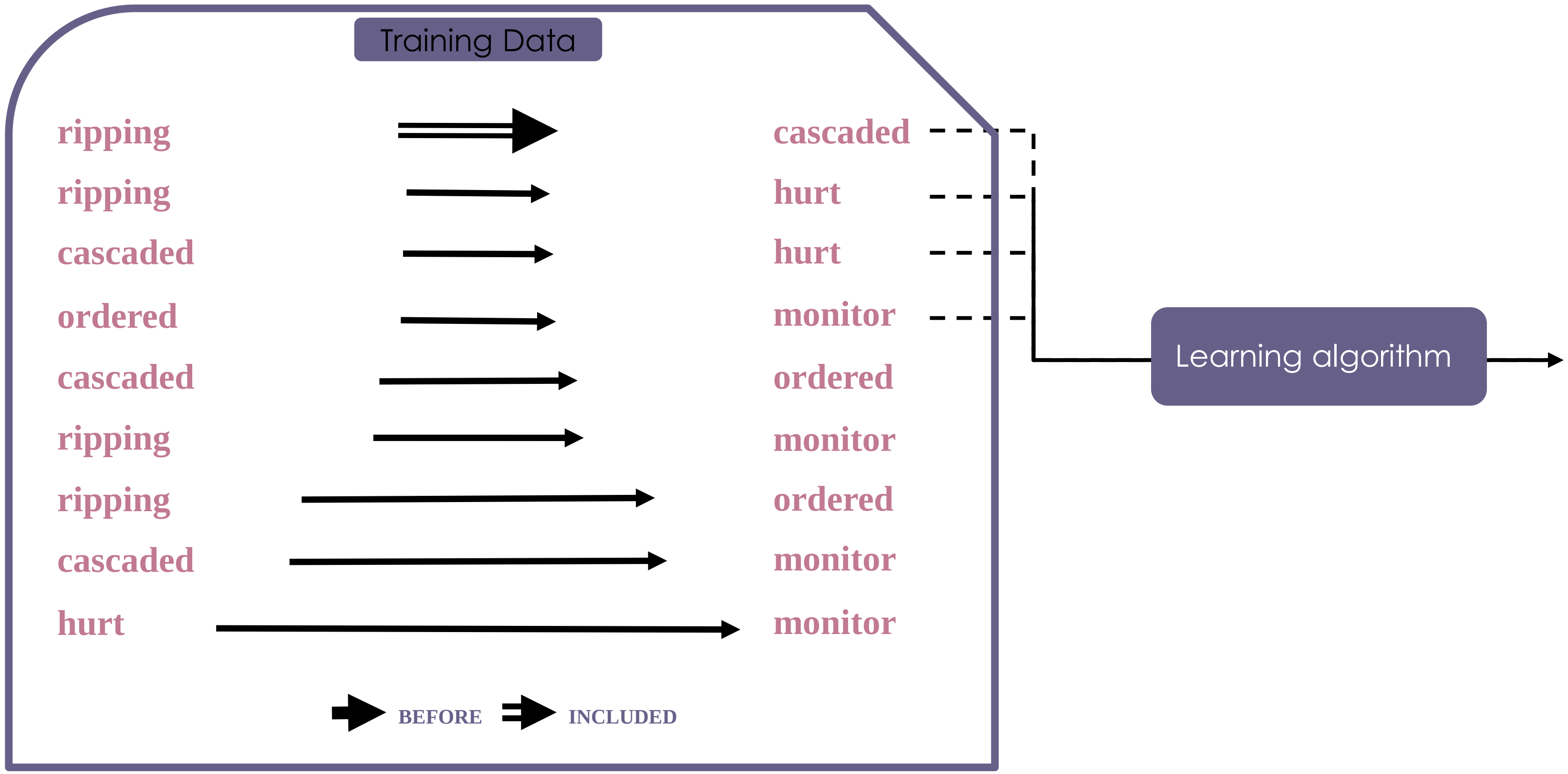
Due to transitivity, temporal relations are not independent

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





Local learning



Local learning is not sufficient

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

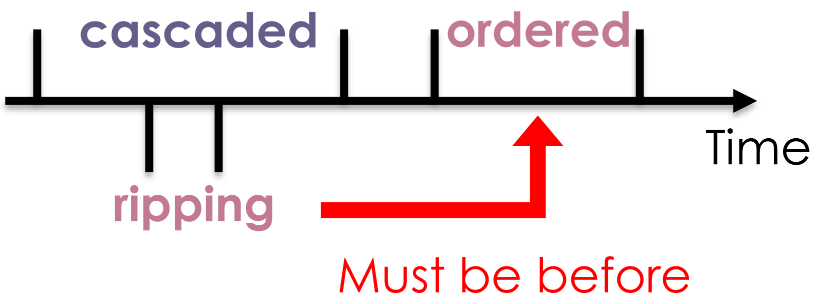
…**ripping** two houses…firefighters **ordered** the evacuation of nearby homes…

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)

$\boxed{\hat{Y} = \text{"solution to ILP"}}$

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

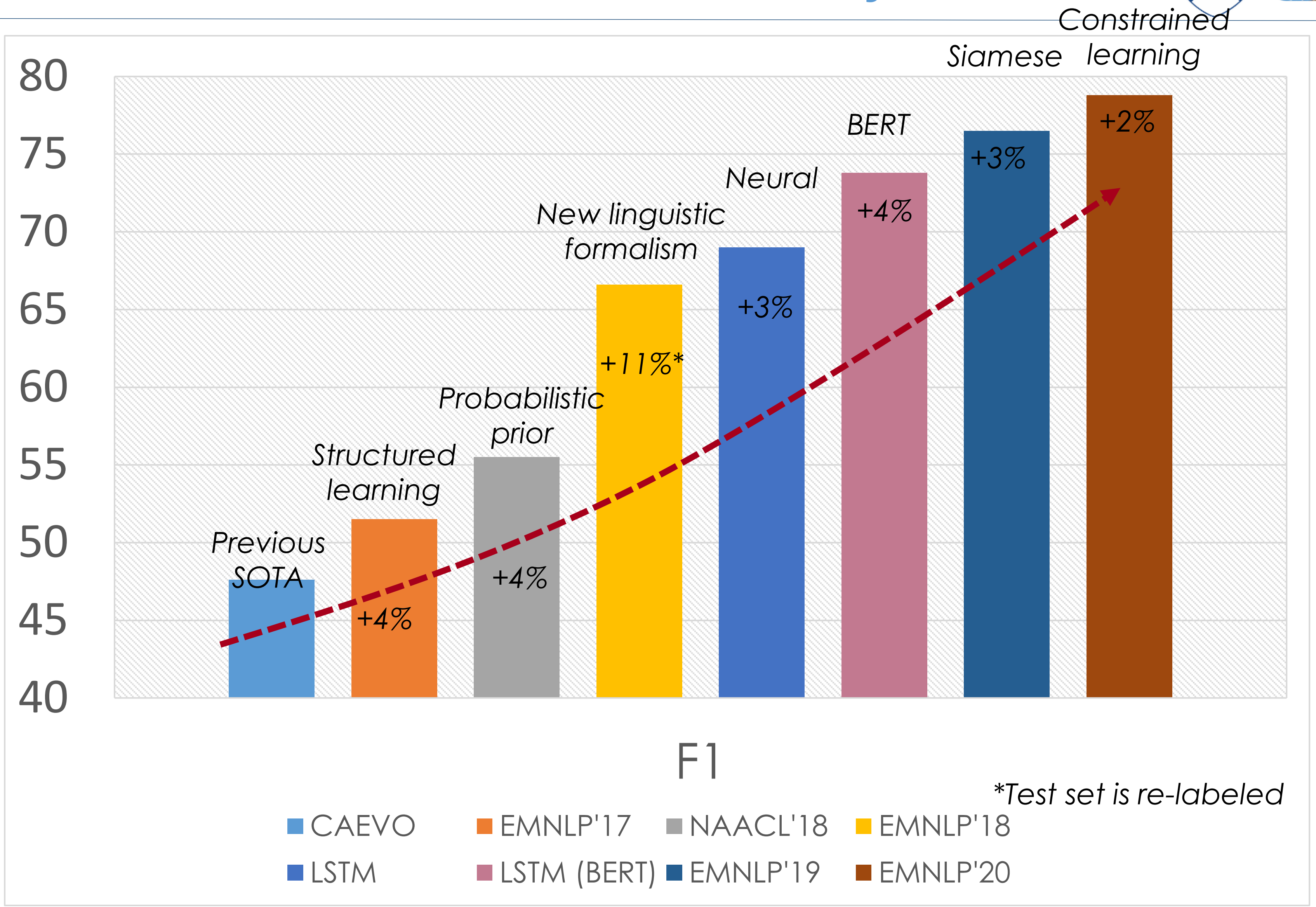
Symmetry constraints
$$L_S = \sum_{e_1, e_2 \in \mathcal{E}, \alpha \in \mathcal{R}_S} |\log \alpha_{(e_1, e_2)} - \log \bar{\alpha}_{(e_2, e_1)}|$$

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

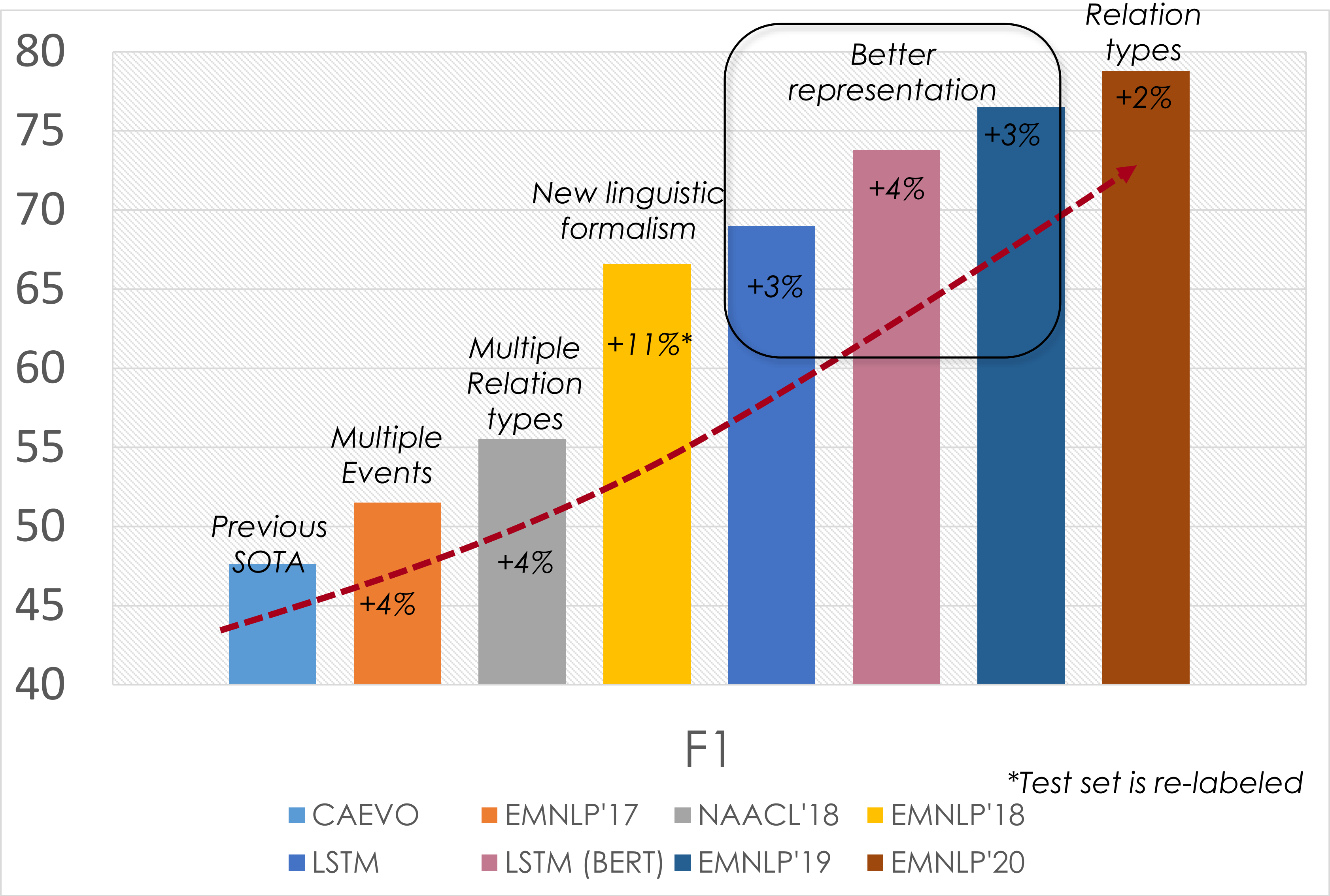
$$L_{t_1} = \log \alpha_{(e_1, e_2)} + \log \beta_{(e_2, e_3)} - \log \gamma_{(e_1, e_3)}$$
$$L_{t_2} = \log \alpha_{(e_1, e_2)} + \log \beta_{(e_2, e_3)} - \log(1 - \delta_{(e_1, e_3)})$$

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

Temporal relation extraction in recent years

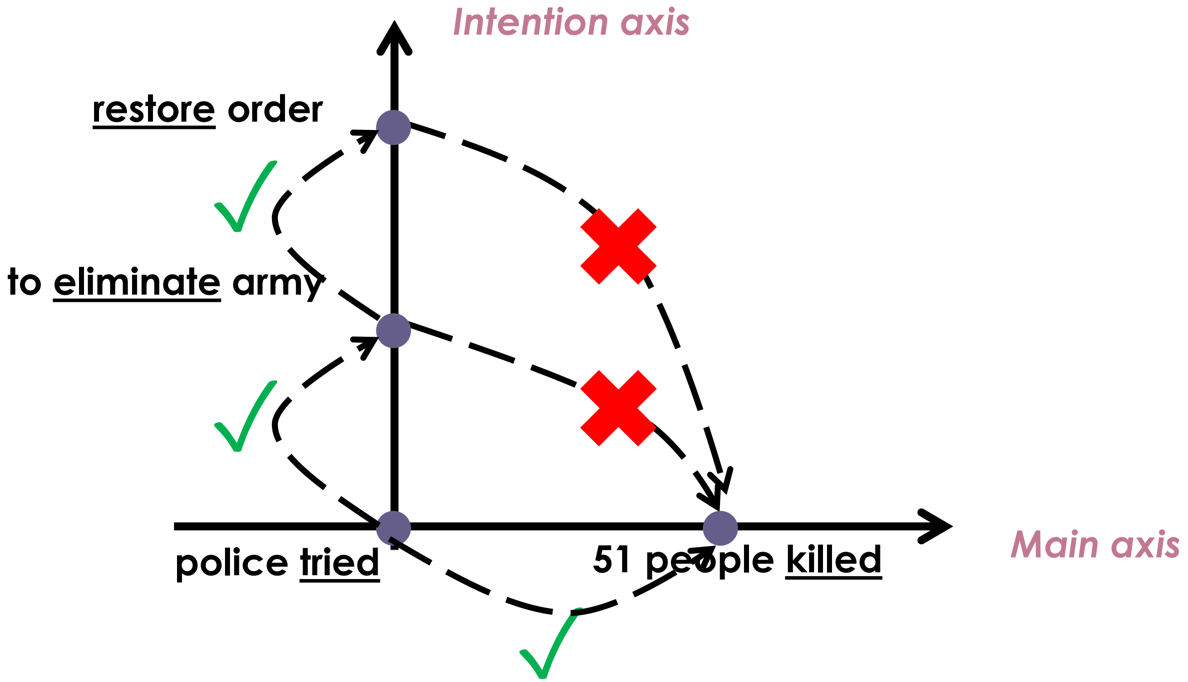


Temporal relation extraction in recent years



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.*

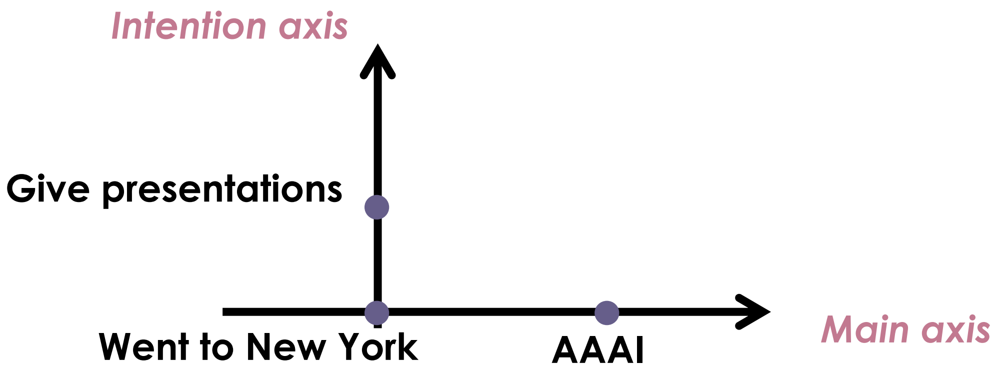


MATRES: A Multi-Axis Annotation Scheme for Event Temporal Relations. Ning et al., ACL2018.

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?



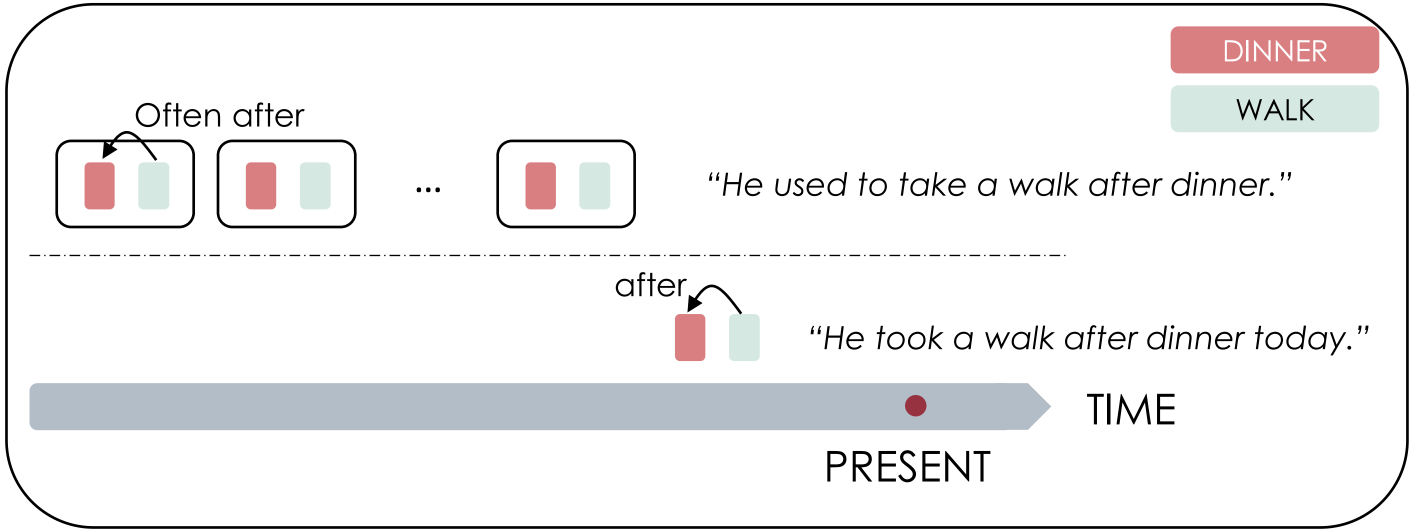
MATRES: A Multi-Axis Annotation Scheme for Event Temporal Relations. Ning et al., ACL2018.

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

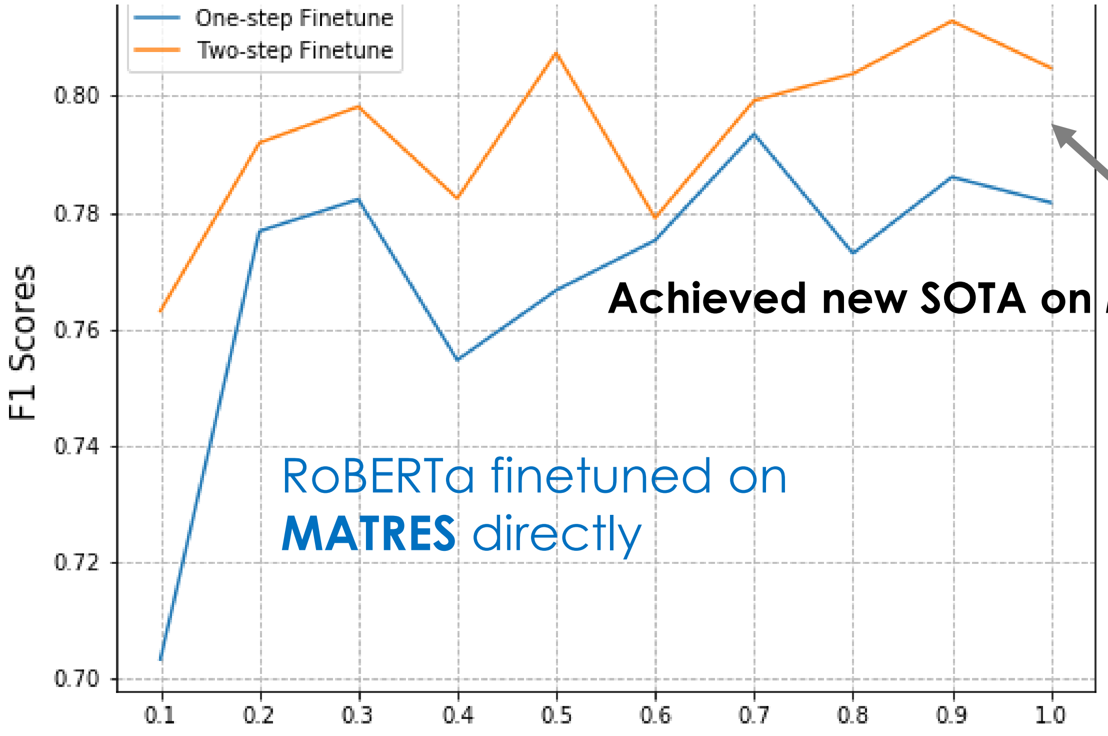
Q4: What happened before a woman was trapped?
A: landslide
Q5: What had started before a woman was trapped?
A: snow rainfall landslide
Q6: What happened while a woman was trapped?
A: searching
Q7: What happened after a woman was trapped?
A: searching said found

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**



Achieved new SOTA on MATRES

RoBERTa finetuned on **MATRES** directly

% MATRES data used for training

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