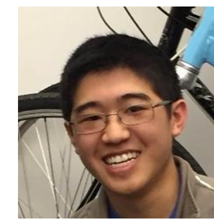
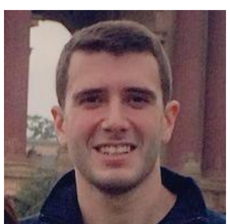
Multi-Sentence Argument Linking



Seth Ebner*, Patrick Xia*,
Ryan Culkin, Kyle Rawlins,
Benjamin Van Durme



Johns Hopkins University

Center for Language and Speech Processing

Motivation

target

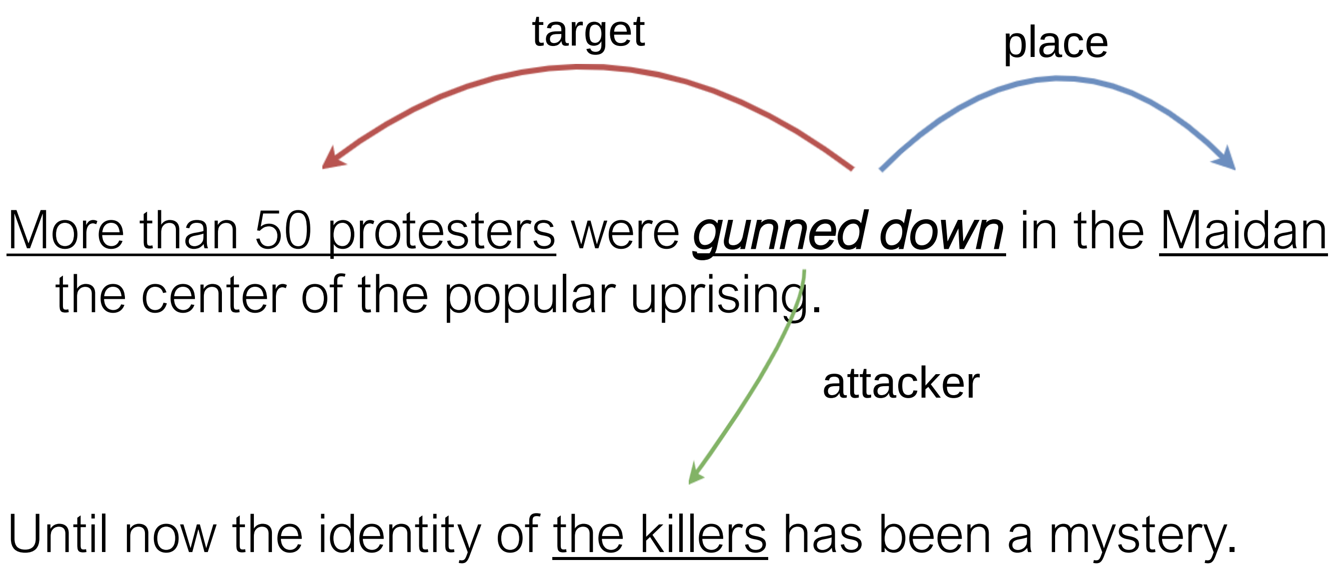
place

<u>More than 50 protesters</u> were ***<u>gunned down</u>*** in the <u>Maidan</u>, the center of the popular uprising.

Firearm-Attack

Attacker:	∅
Target:	More than 50 protesters
Instrument:	∅
Place:	Maidan

Traditional SRL:
in-sentence



Firearm-Attack

Attacker:	the killers
Target:	More than 50 protesters
Instrument:	∅
Place:	Maidan

Argument Linking: cross-sentence

Cross-Sentence Links are Common

Cross-Sentence Links are Common

Beyond NomBank (Gerber and Chai, 2012)

71% relative increase in role coverage on NomBank

Cross-Sentence Links are Common

Beyond NomBank (Gerber and Chai, 2012)

71% relative increase in role coverage on NomBank

DARPA AIDA Phase 1

38% of events have an argument outside the same sentence as the trigger

Cross-Sentence Links are Common

Beyond NomBank (Gerber and Chai, 2012)

71% relative increase in role coverage on NomBank

DARPA AIDA Phase 1

38% of events have an argument outside the same sentence as the trigger

90% of arguments can be found within 2 sentences of the trigger

Motivation

Data: **R**oles **A**cross **M**ultiple **S**entences

Model

Results

Roles Across Multiple Sentences

Roles **A**cross **M**ultiple **S**entences

Based on DARPA AIDA event ontology

Roles Across Multiple Sentences

Based on DARPA AIDA event ontology

9K+ examples from 4K news articles

Roles Across Multiple Sentences

Based on DARPA AIDA event ontology

9K+ examples from 4K news articles

Annotators mark spans in 5-sentence context window

Roles Across Multiple Sentences

Based on DARPA AIDA event ontology

9K+ examples from 4K news articles

Annotators mark spans in 5-sentence context window

Protocol and analysis in the paper!

RAMS Example

When Russian aircraft bombed a remote garrison in southeastern Syria last month, alarm bells sounded at the Pentagon and the Ministry of Defense in London.

instrument

place

attacker

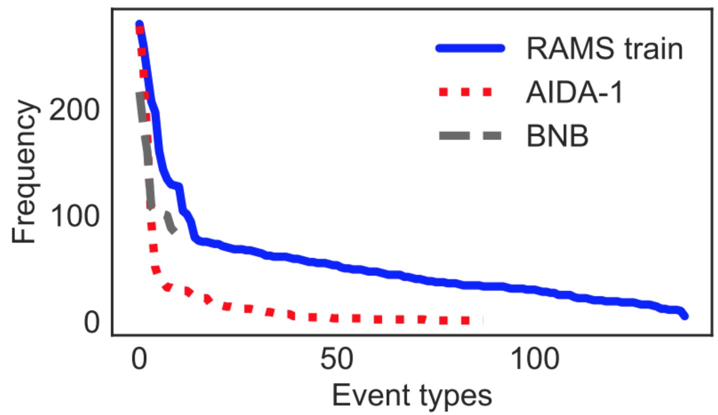
Conflict/Attack/
AirstrikeMissileStrike

The Russians weren't ***bombarding*** a run-of-the-mill rebel outpost, according to U.S. officials.

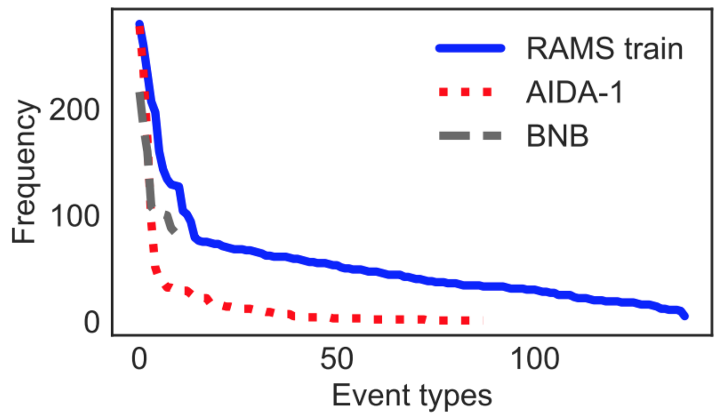
target

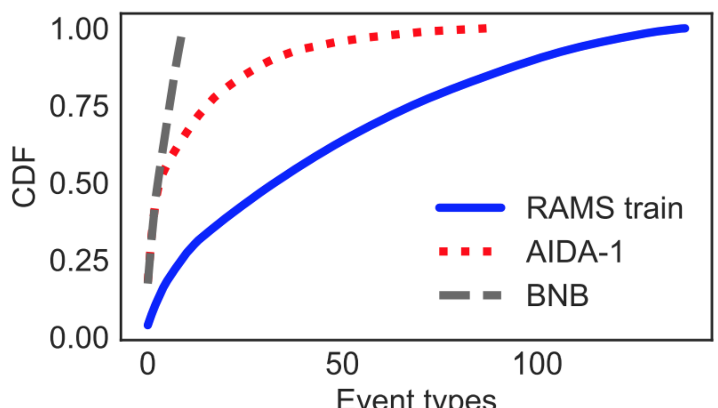
Broad + Diverse Coverage of Ontology

Broad + Diverse Coverage of Ontology



Broad + Diverse Coverage of Ontology





Motivation

Data: Roles Across Multiple Sentences 

Model

Results

Approach

Approach

Roles are evoked by event triggers, forming *implicit arguments* (implicit discourse referents)

Approach

Roles are evoked by event triggers, forming *implicit arguments* (implicit discourse referents)

the <u>place</u> of a *particular* attack event

Approach

Roles are evoked by event triggers, forming *implicit arguments* (implicit discourse referents)

the place of a *particular* attack event

Implicit arguments linked to explicit mentions in text

Model

Model

1) Learn span representations for each trigger span and candidate argument span

Model

1) Learn span representations for each trigger span and candidate argument span

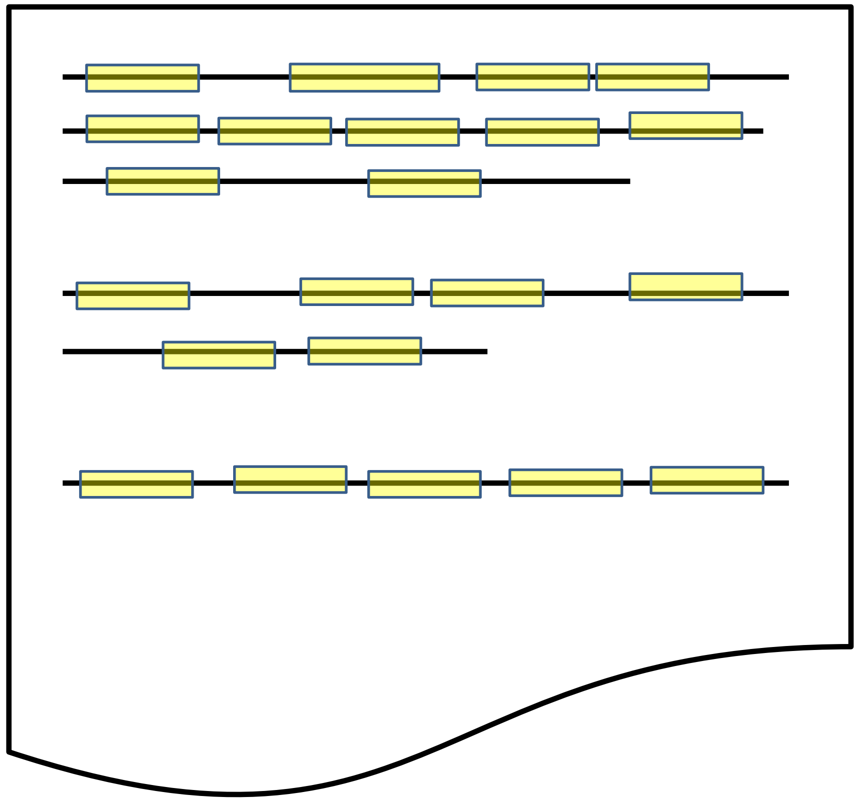
2) For each trigger, prune to top-*k* candidate arguments

Model

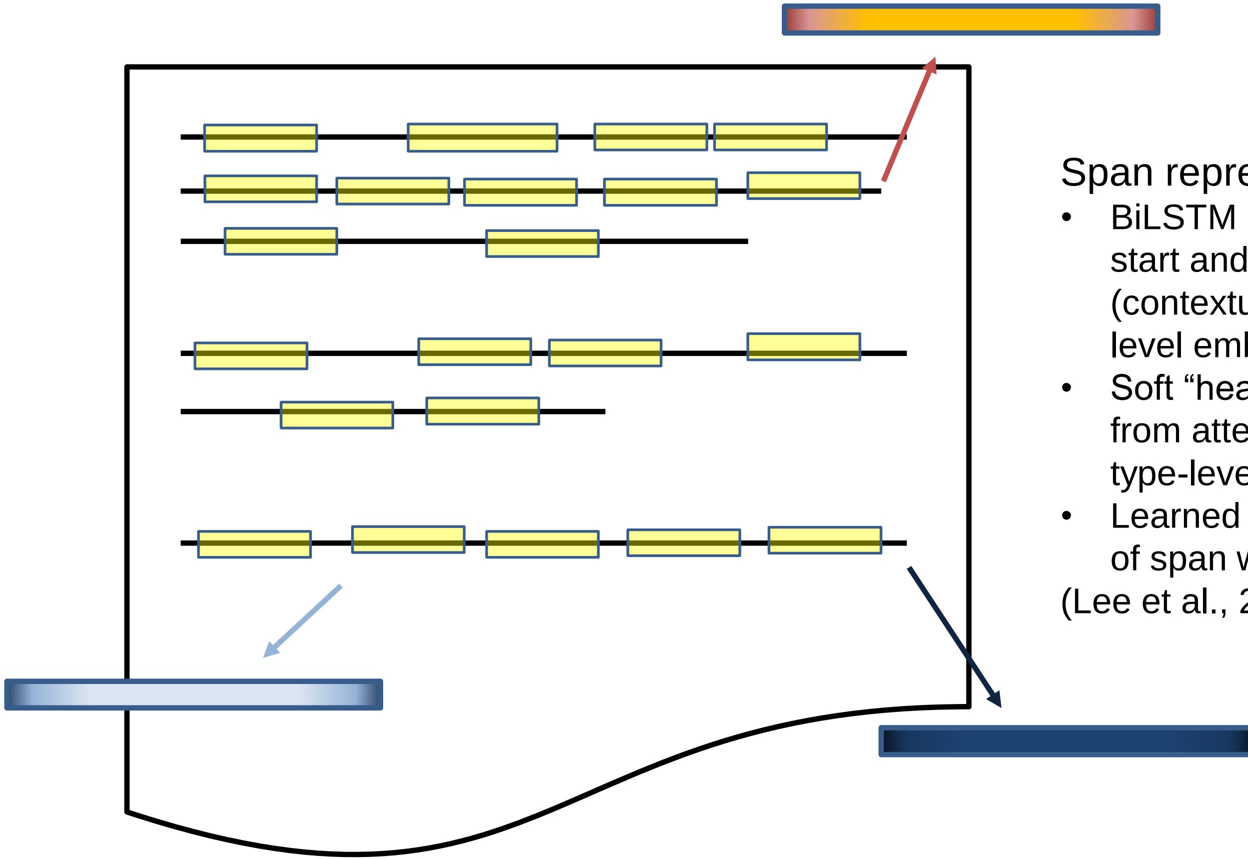
1) Learn span representations for each trigger span and candidate argument span

2) For each trigger, prune to top-*k* candidate arguments

3) Score reps of implicit arguments with reps of explicit arguments using a decomposable scoring function



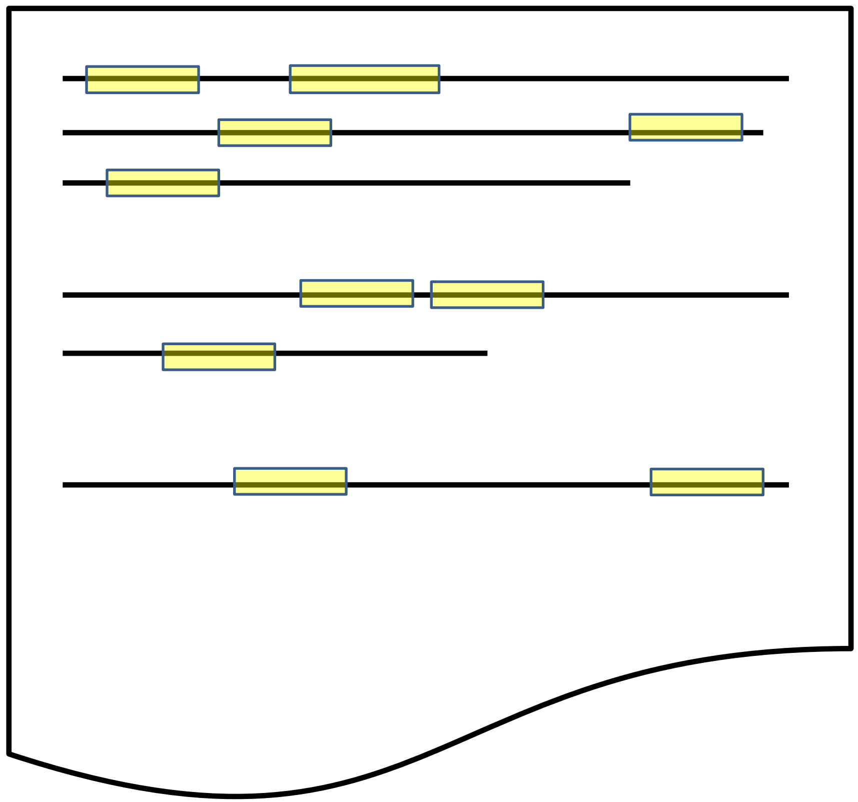
Enumerate candidate argument spans up to a fixed width



Span representation:
- BiLSTM hidden states at start and end of span (contextualized and type-level embeddings)
- Soft “head word” vector from attention mask over type-level embeddings
- Learned representation of span width

(Lee et al., 2017; *i.a.*)

Create argument representations for every span

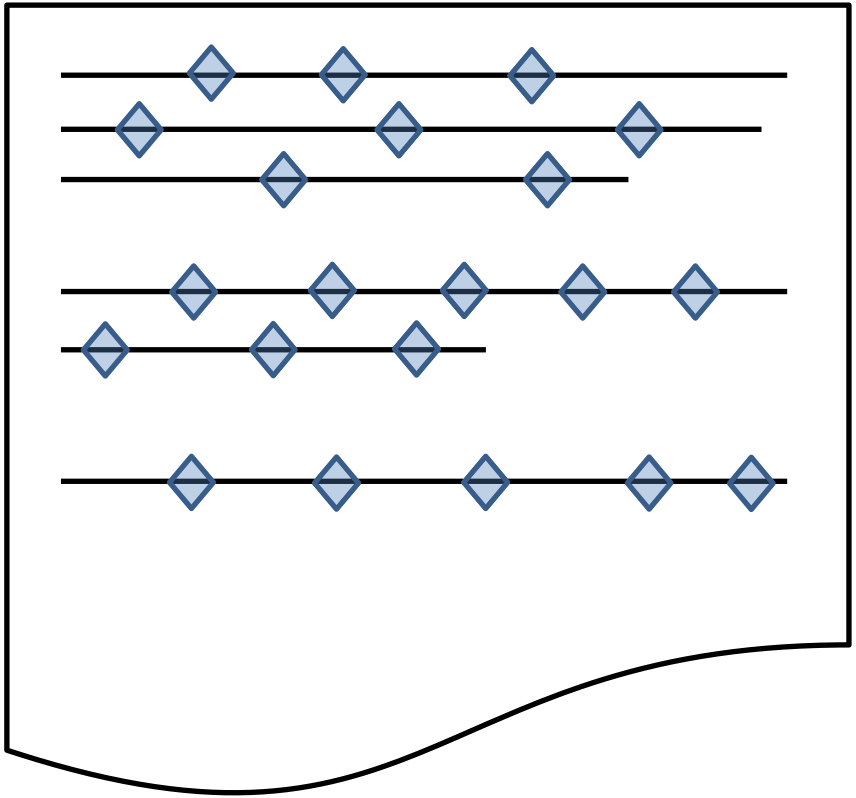


Score each representation, and keep the top $\boldsymbol{\lambda}_{\mathrm{A}} n$

n: document length

$\boldsymbol{\lambda}_{\mathrm{A}}$: hyperparameter

Prune to find good spans

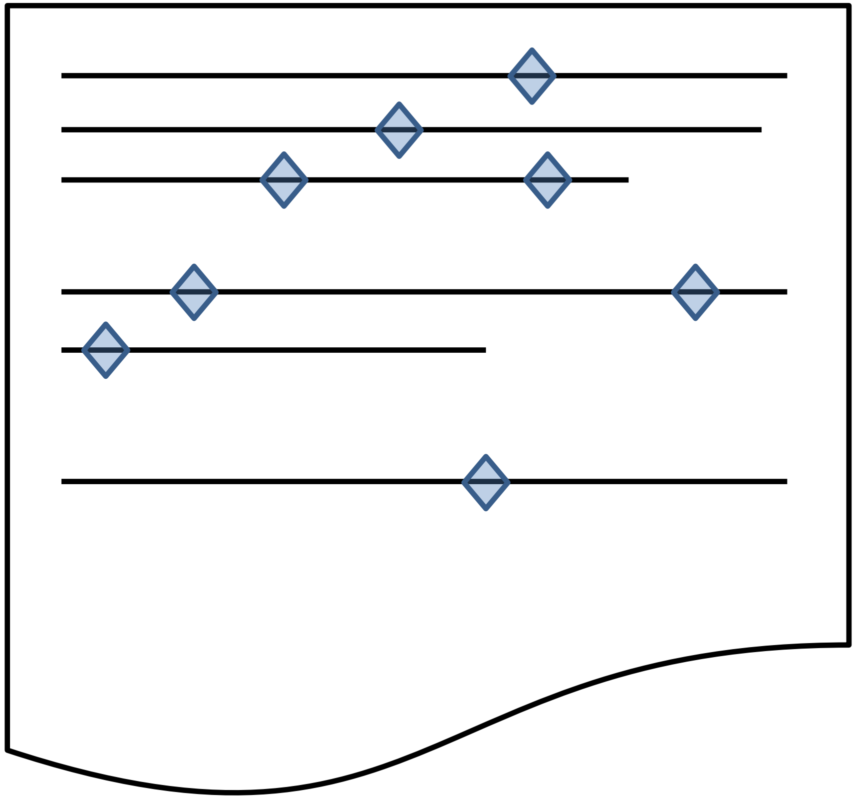


Span representation:

- BiLSTM hidden states at start and end of span (contextualized and type-level embeddings)
- Soft “head word” vector from attention mask over type-level embeddings
- Learned representation of span width

(Lee et al., 2017; *i.a.*)

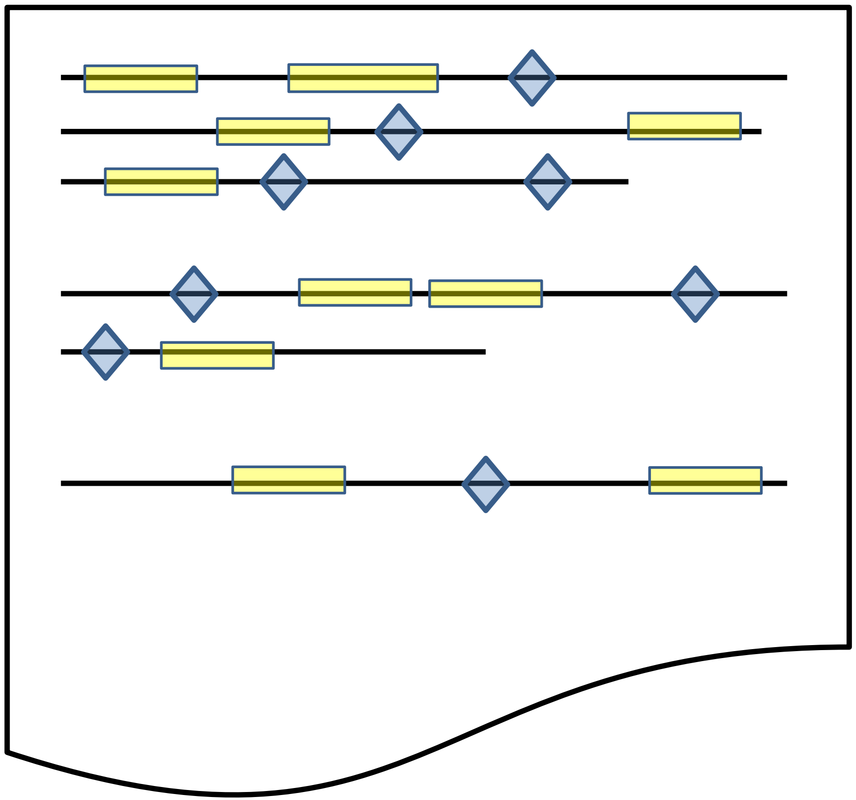
Create event trigger representations for every span

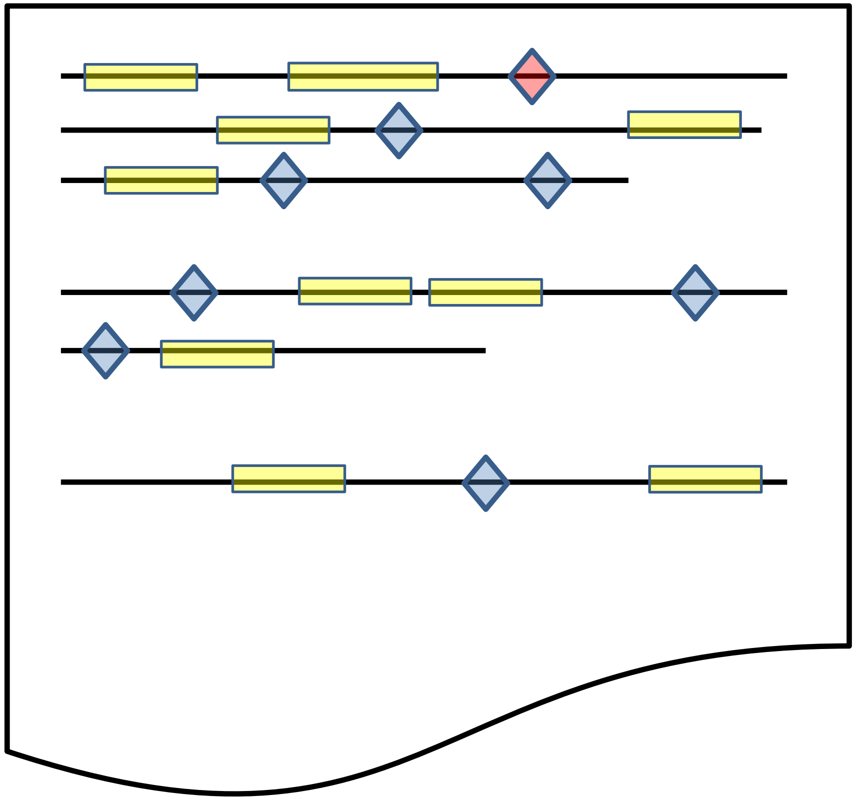


Score each representation,
and keep the top $\lambda_E n$

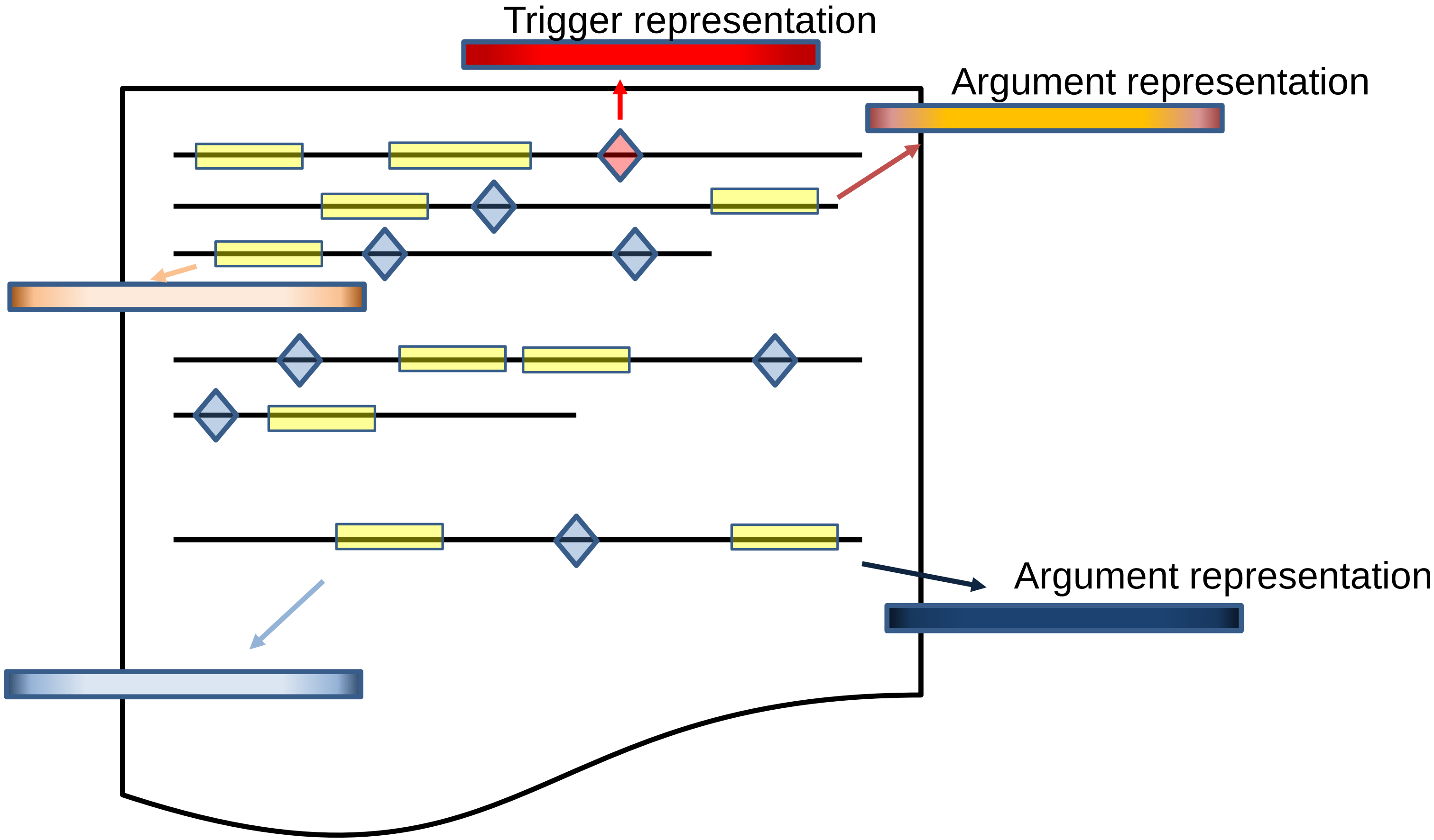
n: document length
λ_E: hyperparameter

Prune to find good spans



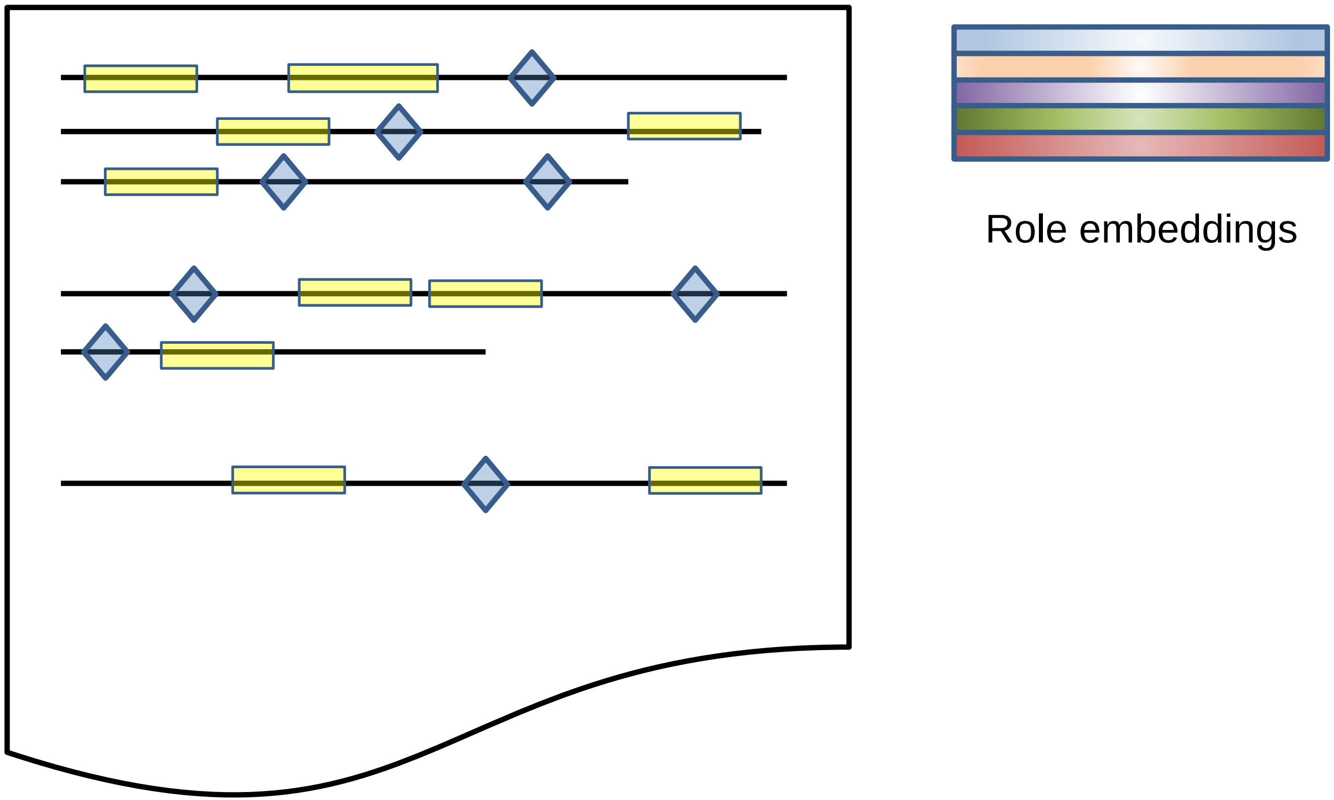


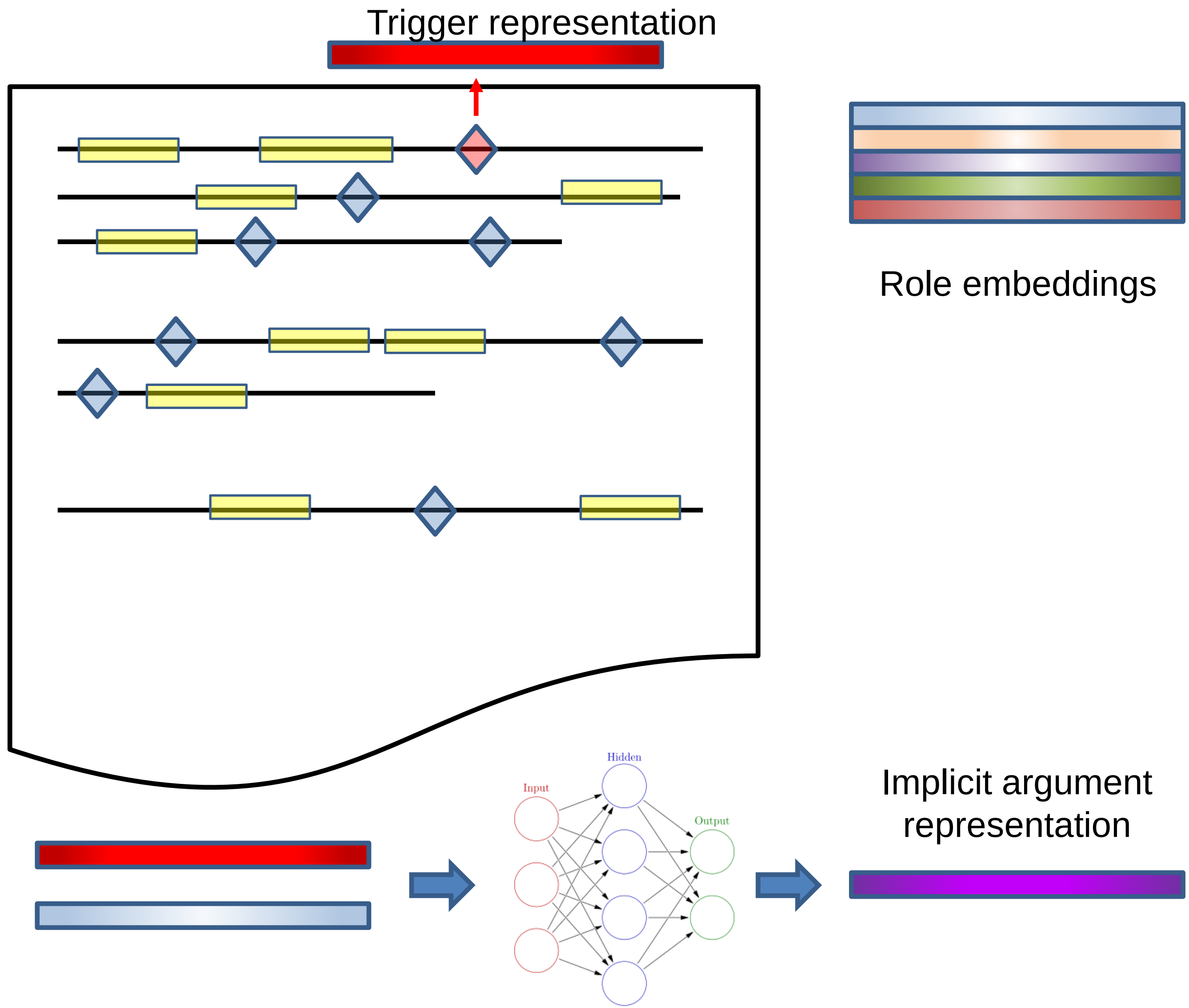
Further prune to find good arguments for a particular event

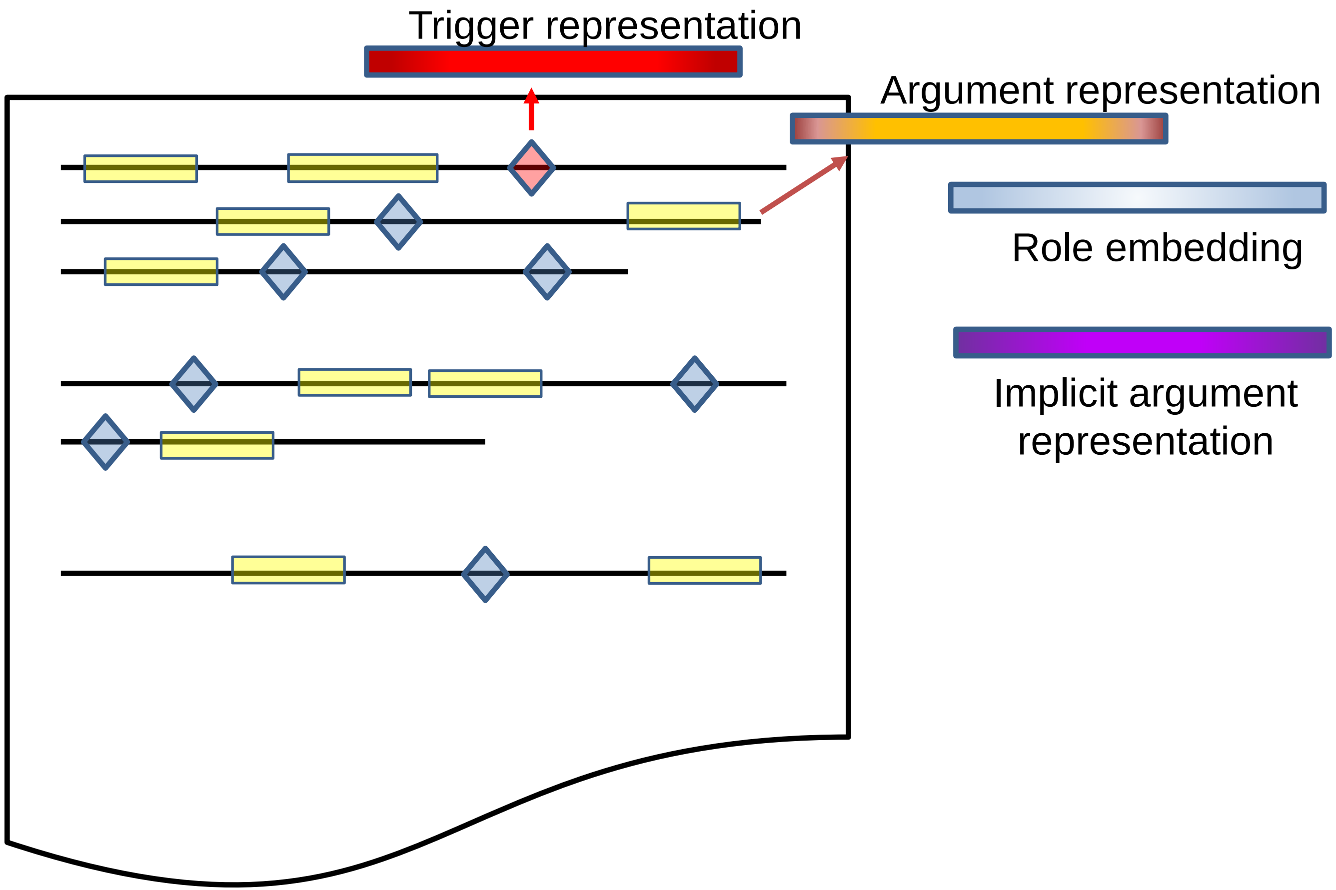


$$s_c(e,a) = \mathbf{e}^\top \mathbf{W}_c \mathbf{a} + s_A(a) + s_E(e) + \phi_c(e,a)$$

Further prune to find good arguments for a particular event


Role embeddings


Trigger representation
Role embeddings
Input
Hidden
Output
Implicit argument
representation



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

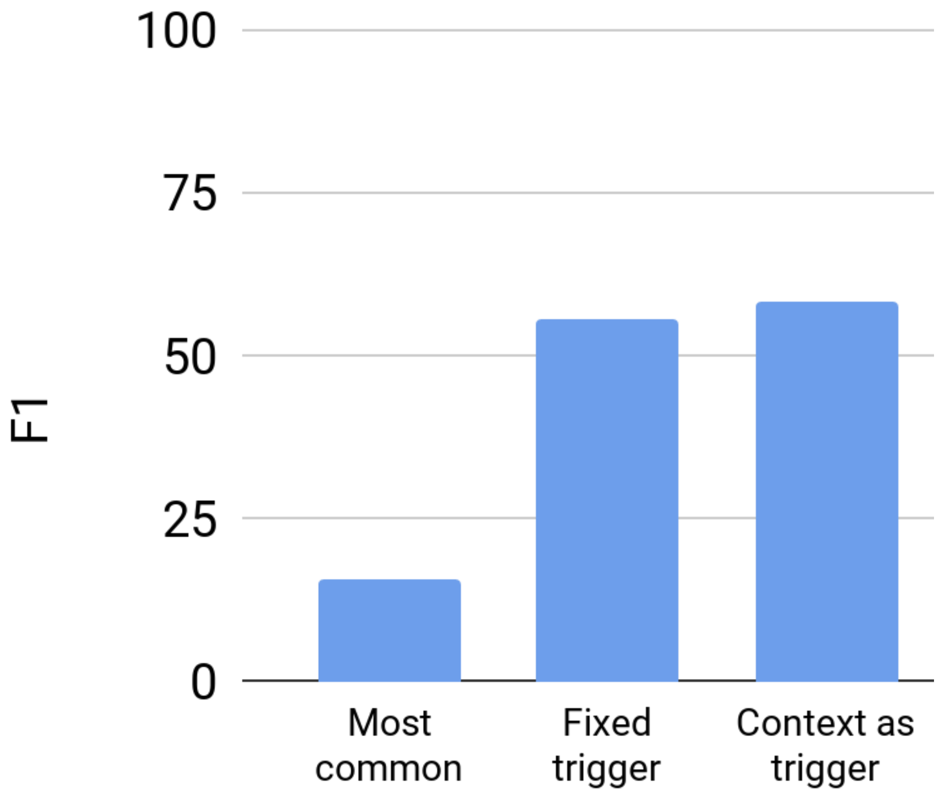
Motivation

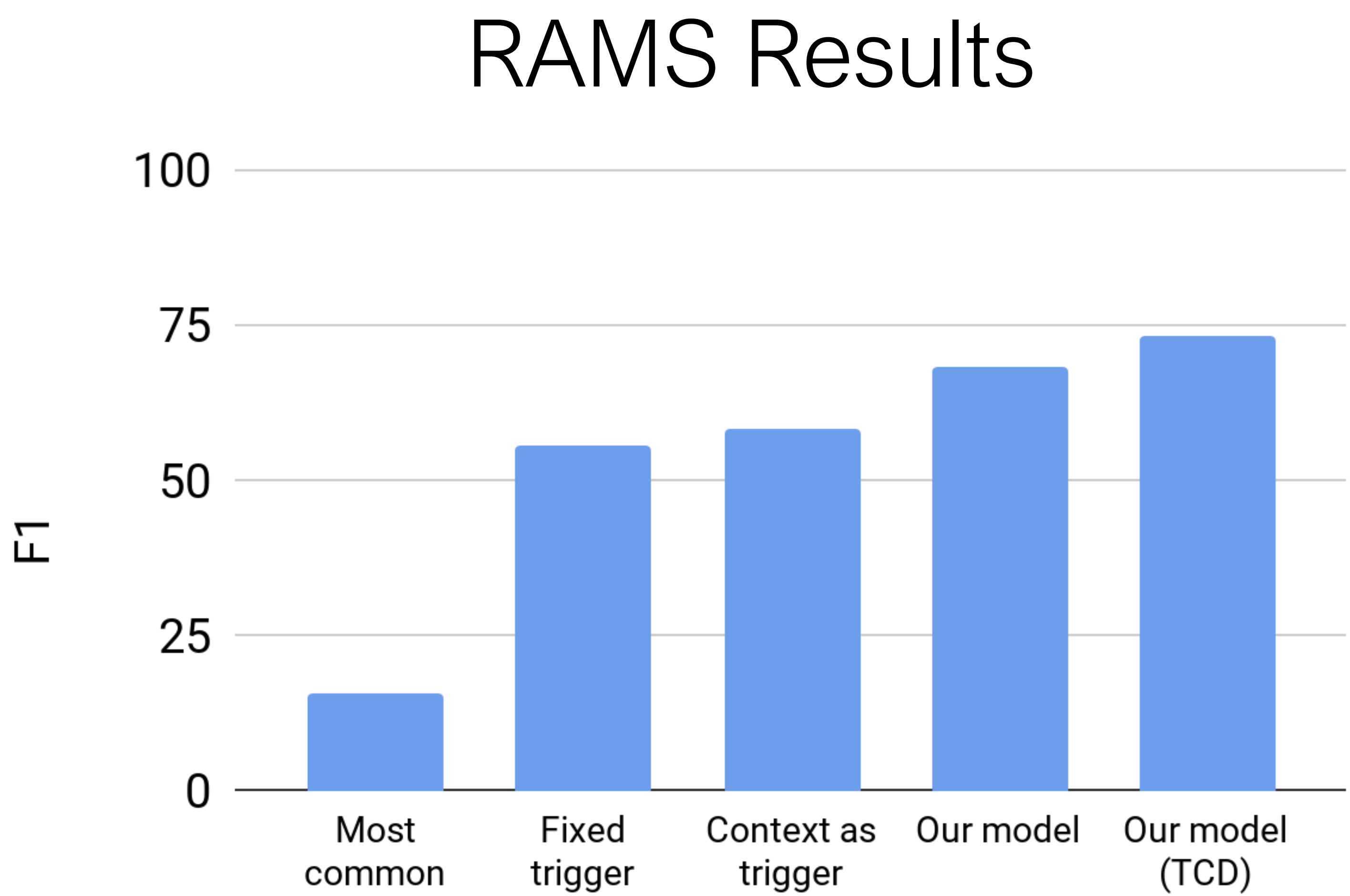
Data: Roles Across Multiple Sentences

Model

Results

RAMS Results





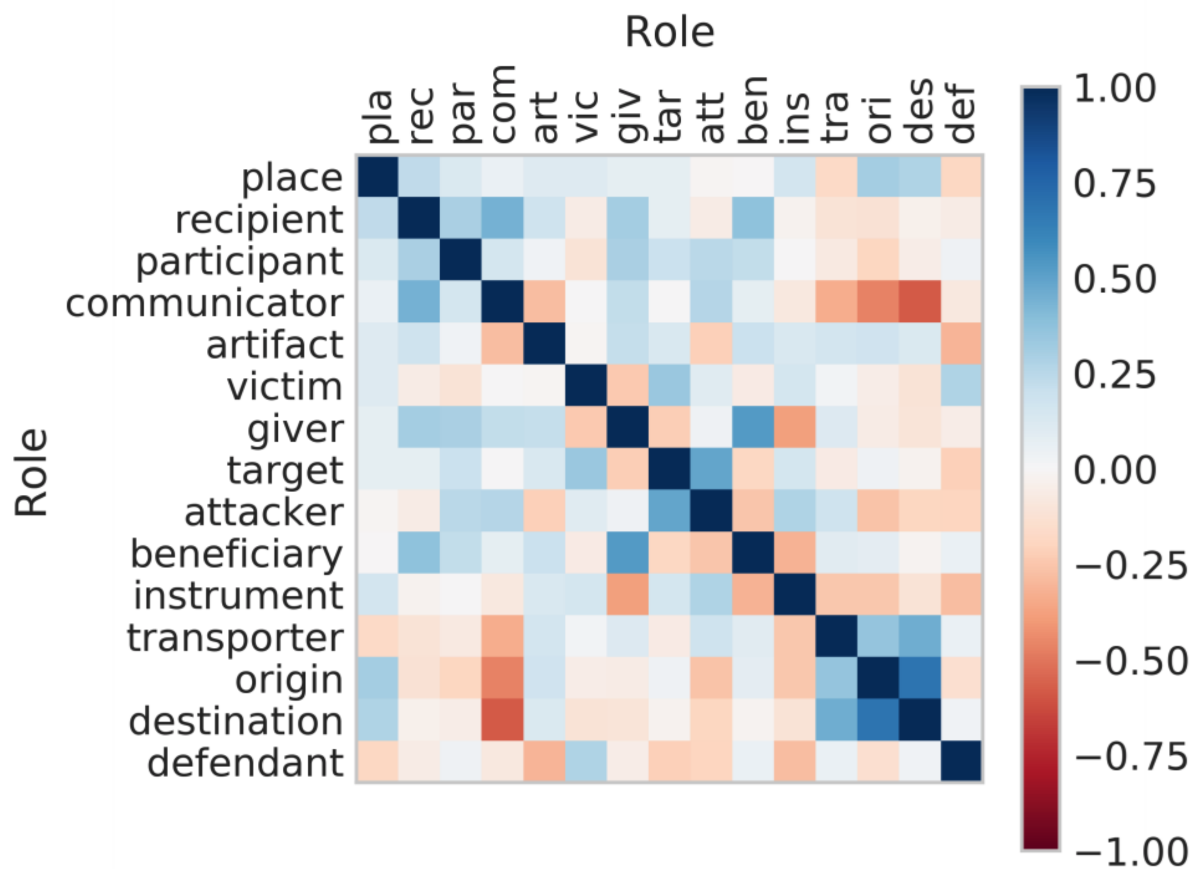
RAMS Results
100
75
50
25
0
F1
Most common
Fixed trigger
Context as trigger
Our model
Our model (TCD)

RAMS Results vs. Distance

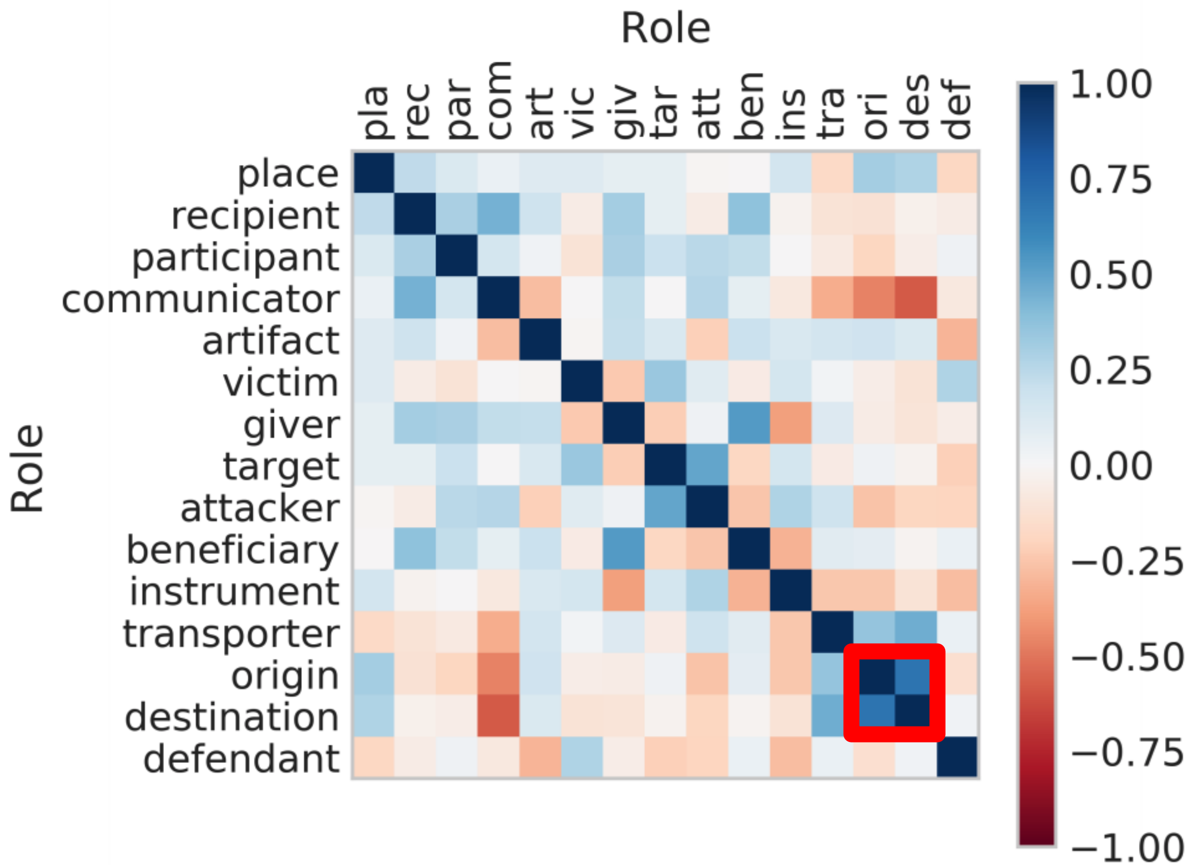
RAMS Results vs. Distance

Performance does not degrade over distance

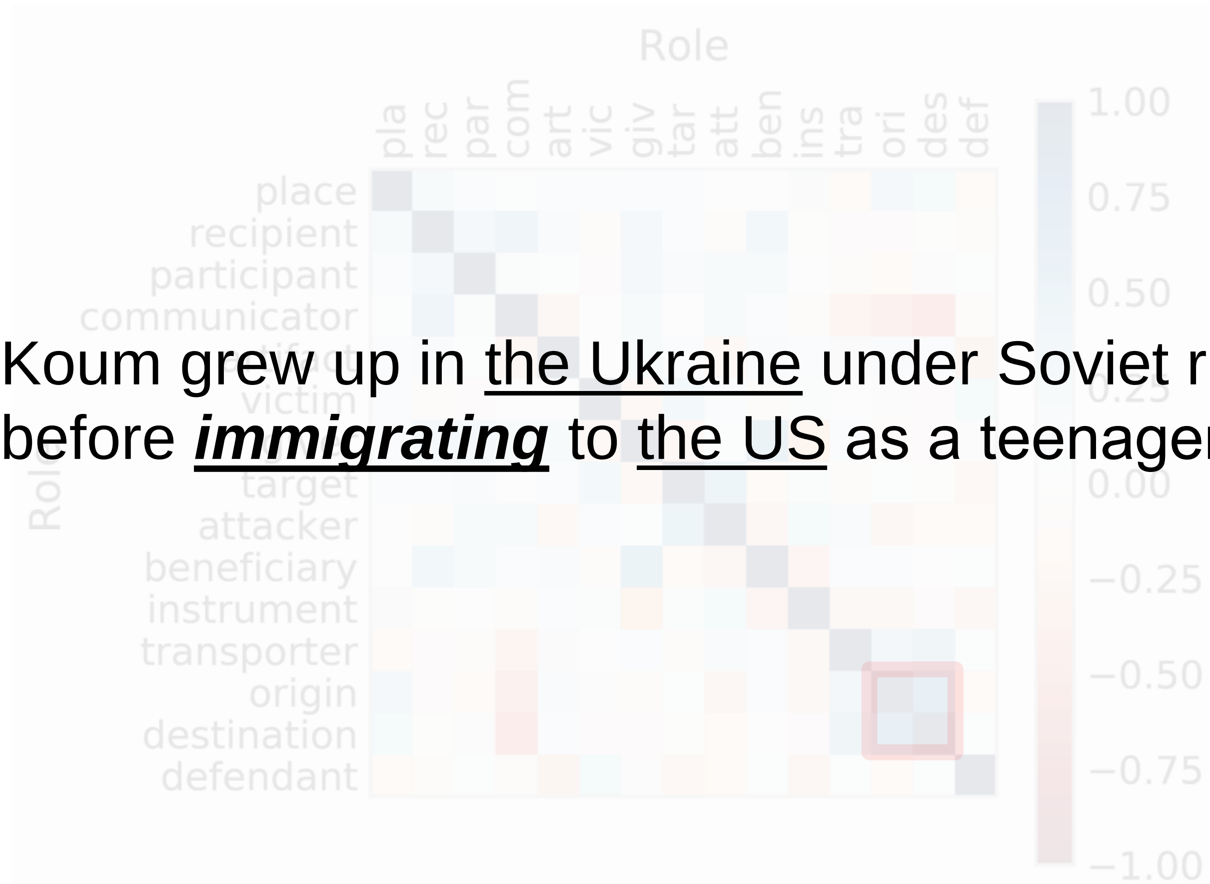
Role Embedding Similarity



Role Embedding Similarity



Role Embedding Similarity



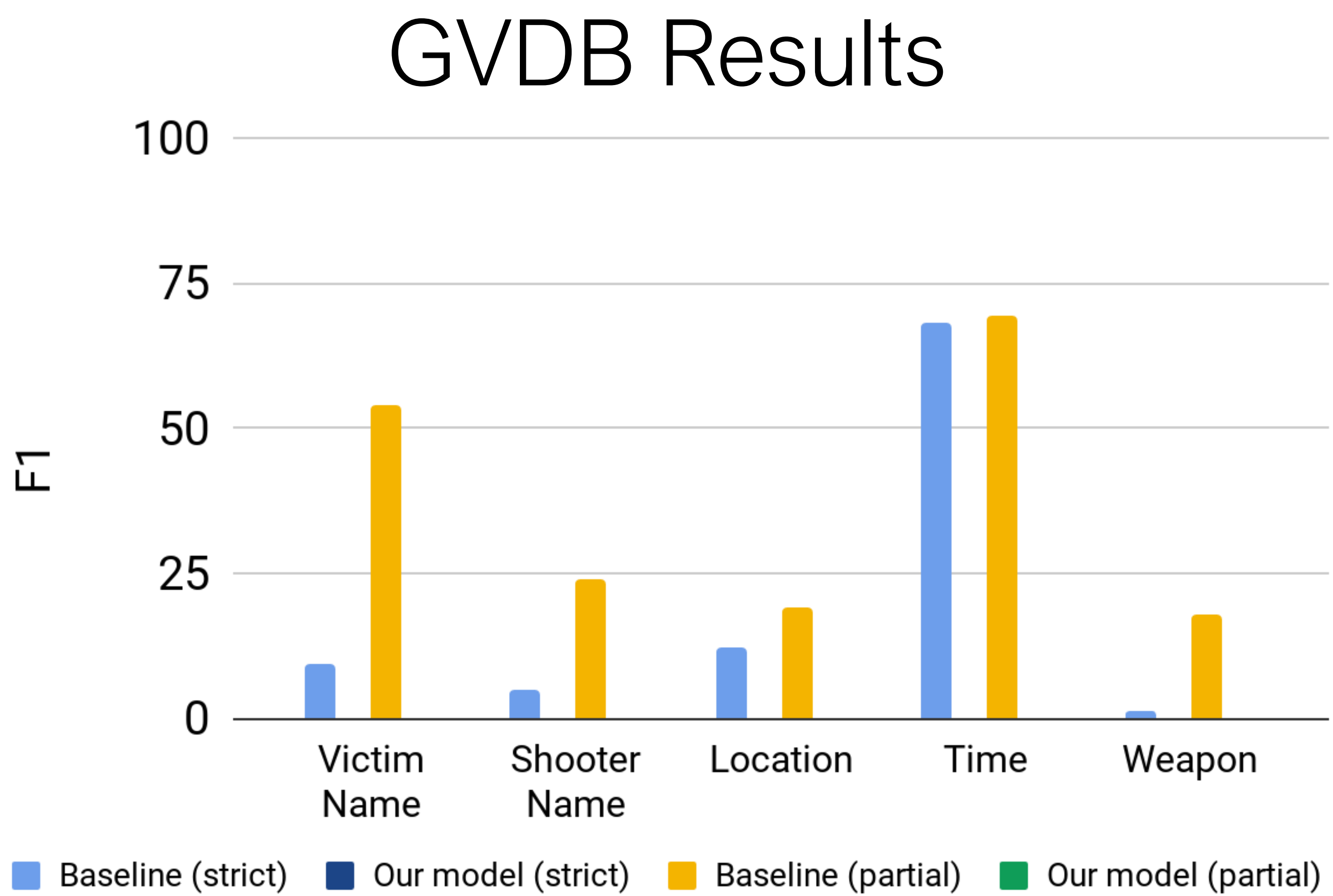
Koum grew up in the Ukraine under Soviet rule before ***immigrating*** to the US as a teenager…

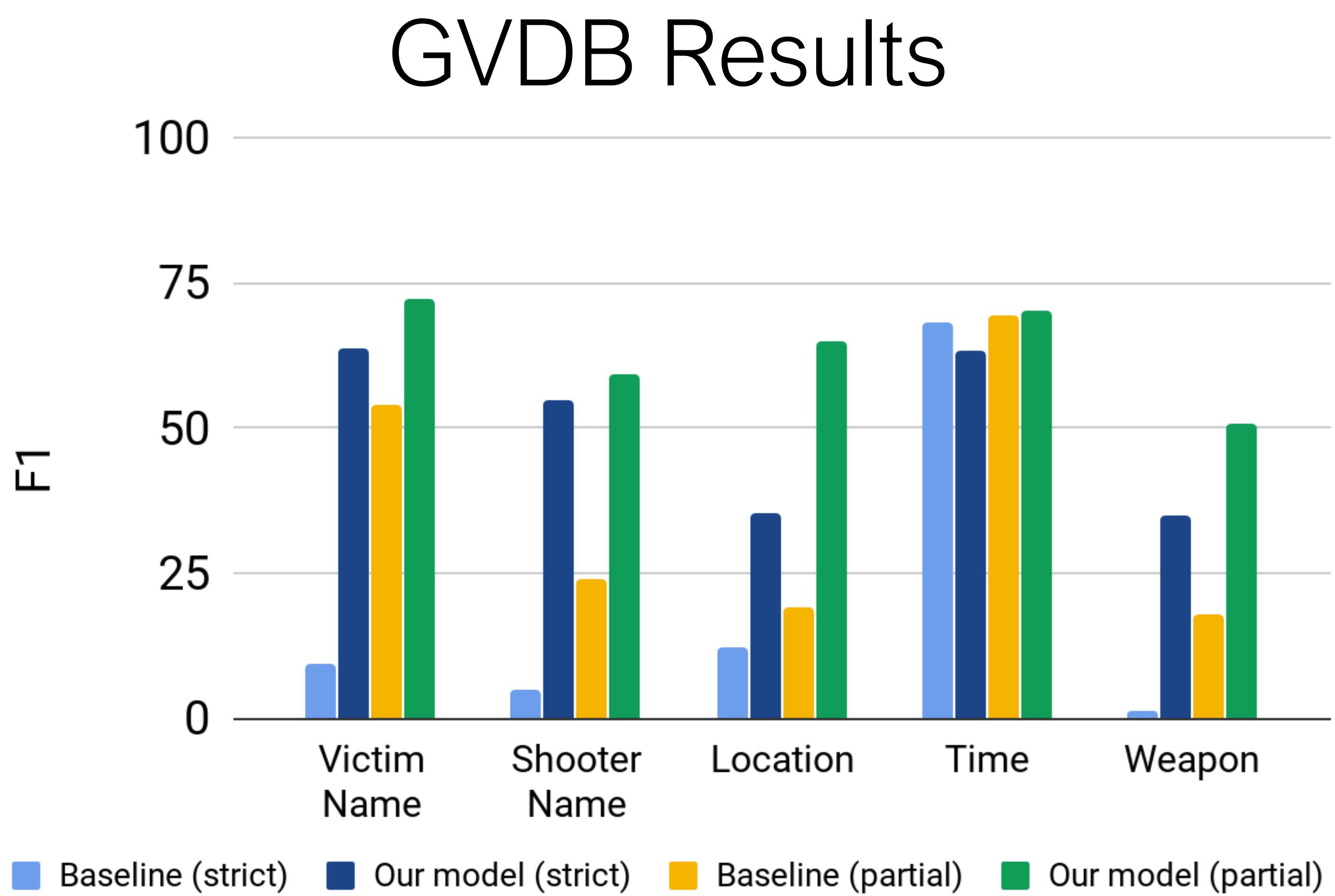
Gun Violence Database

Pavlick et al., 2016

ST. LOUIS (KTVI) – The St. Circuit Attorney's Office charged a 50-year-old Friday for allegedly gunning another man down in a north city neighborhood back in December. According to Schron Jackson, a spokeswoman for the St. Louis Metropolitan Police Department, the shooting happened at 12:10 p.m. on December 2, 2015 in the 3000 block of N. 20th Street in the St. Louis Place neighborhood. The victim, identified as 32-year-old Leonard Arnold, was found lying unconscious outside. Arnold had been shot multiple times and was pronounced dead at the scene, Jackson said. The suspect, Maurice Alexander, was charged with first-degree murder, first-degree assault, and two counts of armed criminal action. …

Victim Name:	Leonard Arnold
Shooter Name:	Maurice Alexander
Location:	3000 block of …
Weapon:	ø
# Shots:	shot multiple times
…	


GVDB Results
100
75
50
25
0
F1
Victim Name
Shooter Name
Location
Time
Weapon
Baseline (strict)
Our model (strict)
Baseline (partial)
Our model (partial)


GVDB Results
100
75
50
25
0
F1
Victim Name
Shooter Name
Location
Time
Weapon
Baseline (strict)
Our model (strict)
Baseline (partial)
Our model (partial)

Summary

Summary

Model finds argument spans anywhere in a document that fill roles for events

Summary

Model finds argument spans anywhere in a document that fill roles for events

RAMS dataset has broader coverage than prior datasets

Summary

Model finds argument spans anywhere in a document that fill roles for events

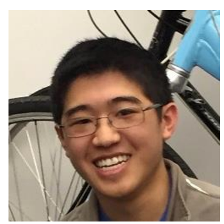
RAMS dataset has broader coverage than prior datasets

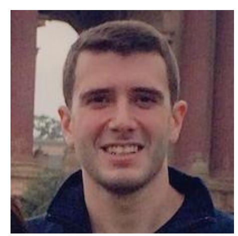
Model works well on other tasks too (more results in the paper!)

Thanks!

Data, code, and models

nlp.jhu.edu/rams

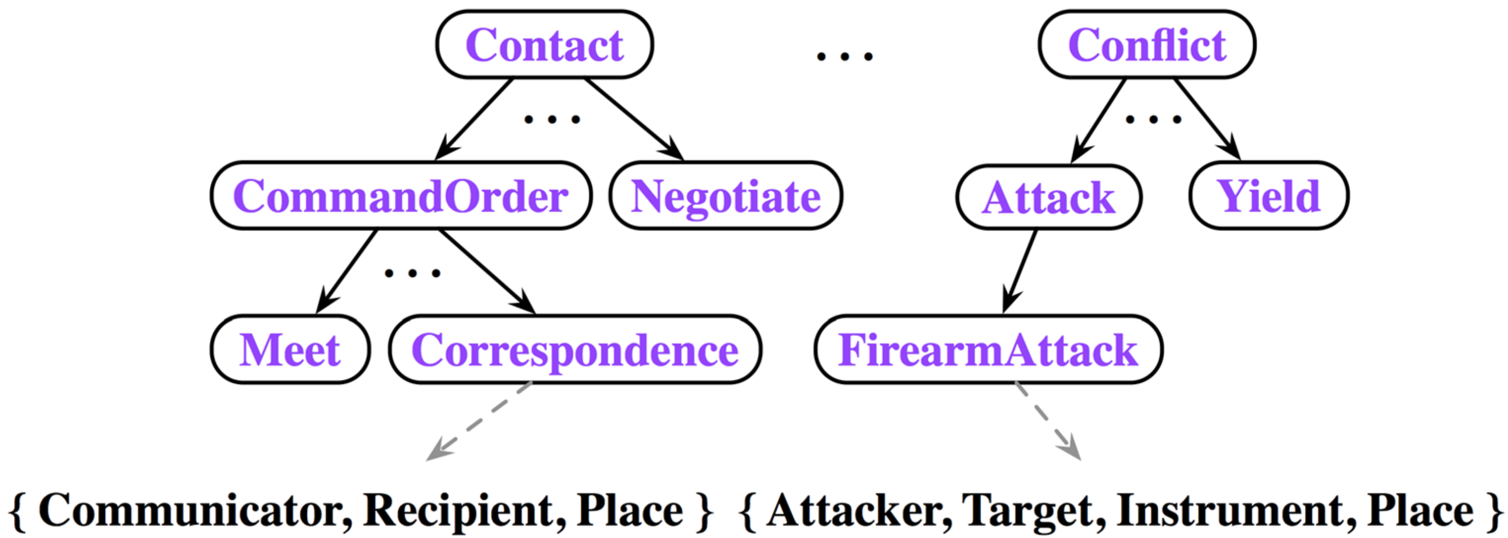




Additional Slides

Note: these slides may be either outdated or even incorrect (based on much earlier versions of the model or buggy graphing code). Use at your own risk.

Motivation


Contact
...
Conflict
...
CommandOrder
Negotiate
...
Attack
Yield
...
Meet
Correspondence
FirearmAttack
{ Communicator, Recipient, Place }
{ Attacker, Target, Instrument, Place }

Roles Across Multiple Sentences (RAMS)

Gloss

the banks lent $ 380 million to Trump for the benefit of Trump at someplace

Context

Meanwhile , another Trump disaster was brewing .

Eastern Air Lines , which had been struggling , put its northeastern air shuttle up for sale .

Trump persuaded the banks to lend him $ 380 million to purchase the route , and in June 1989 the Trump Shuttle began flying .

Trump introduced the airline with his usual style — by insulting the competition .

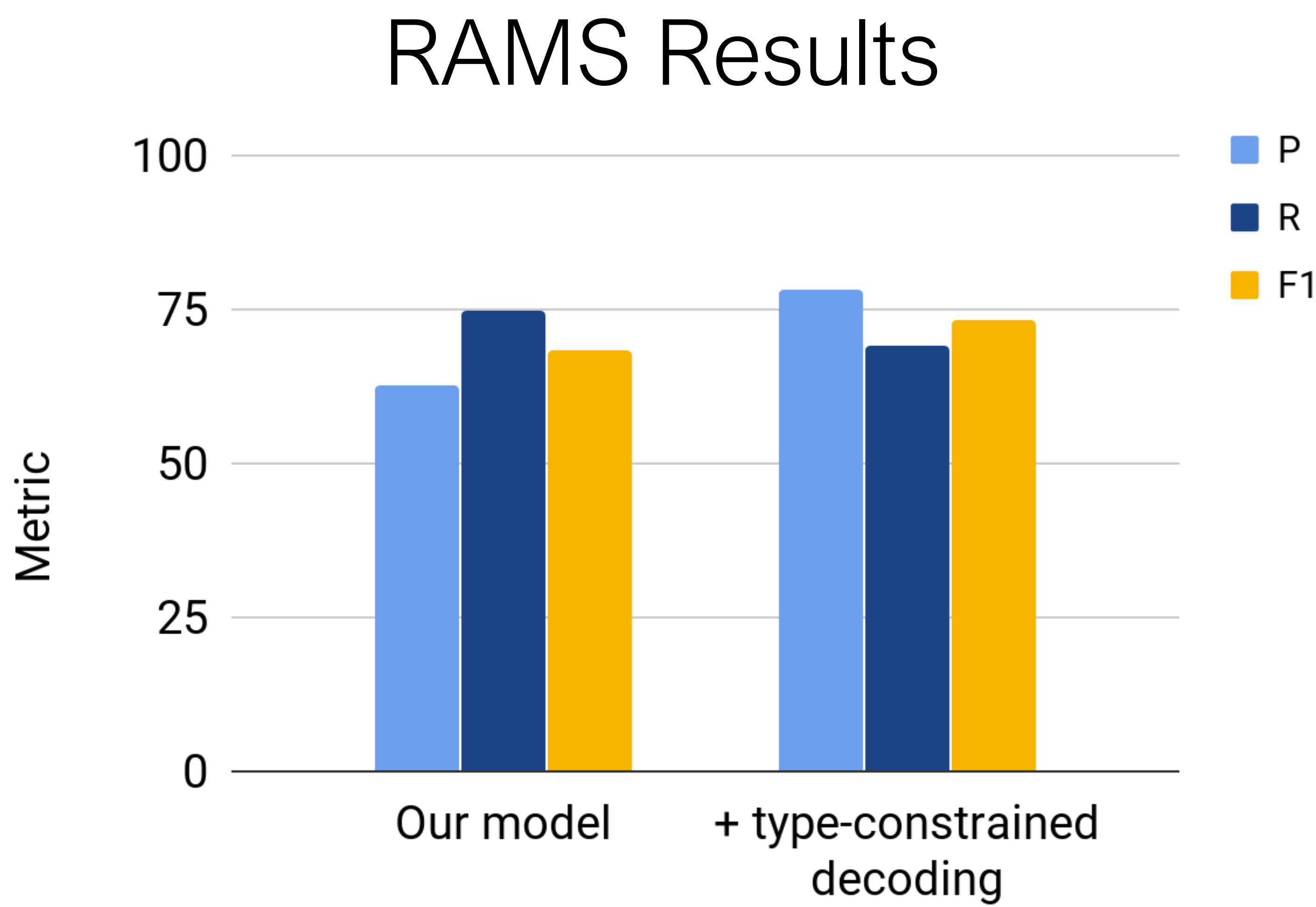
At an elegant event at Logan Airport in Boston , Trump took the stage and suggested that the other airline with a northeastern shuttle , Pan Am , flew unsafe planes .

Event type: Transaction.TransferMoney.BorrowLend

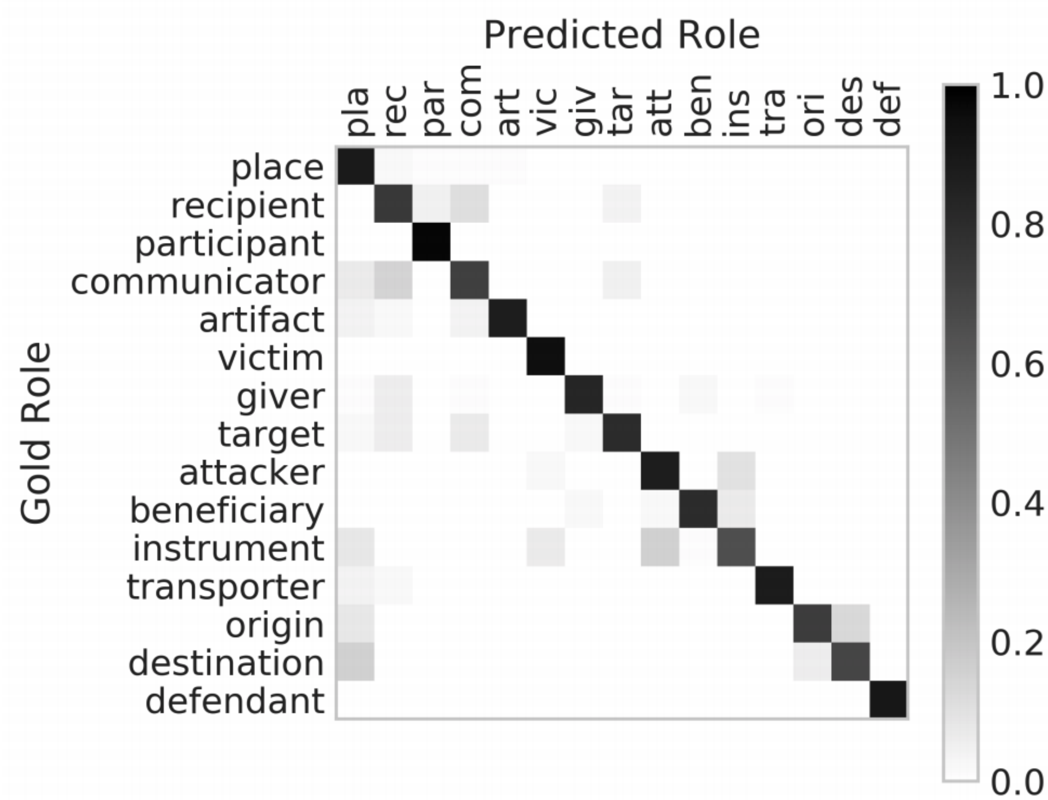
Event definition: The borrowing or lending or money

Arguments

Giver	someone or something	the banks	☐ Argument not present
Recipient	someone or something	Trump	☐ Argument not present
Beneficiary	someone	Trump	☐ Argument not present
Money	some money	$ 380 million	☐ Argument not present
Place	someplace		☑ Argument not present


RAMS Results
P
R
F1
100
75
50
25
0
Metric
Our model
+ type-constrained decoding

Error Analysis



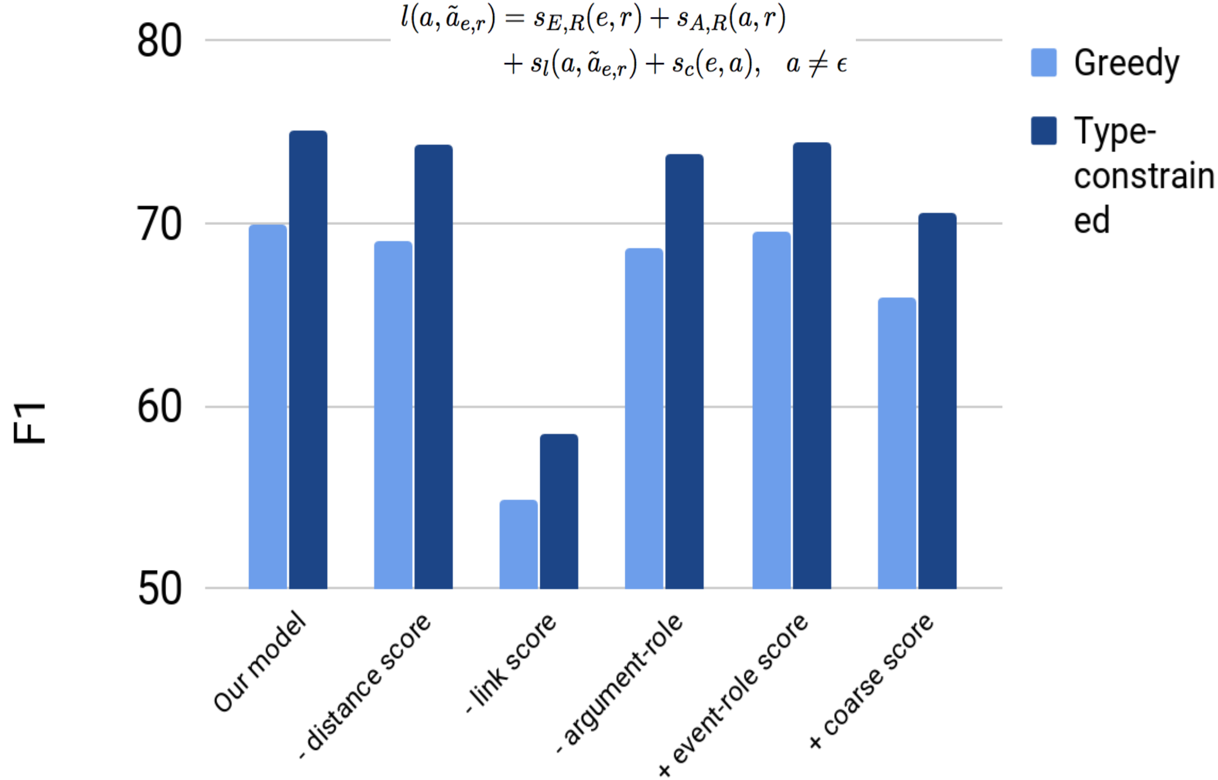
Ablations

Model	Greedy	TCD
Our model	**69.9**	75.1
- distance score	69.0	74.3
- $s_l(a, \tilde{a}_{e,r})$	54.9	58.4
- $s_{A,R}(a, r)$	68.6	73.8
+ $s_{E,R}(e, r)$	69.5	74.4
+ $s_c(e, a)$	65.9	70.6
w/ argmax decoding	**69.9**	75.1
BERT 6–9	69.6	**75.3**
ELMo	68.5	75.2

Table 3: F_1 on RAMS dev data when link score components are separately included/excluded (Equation 2) or other contextualized encoders are used in the best performing model. TCD = type-constrained decoding.

Ablations

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



RAMS Results vs. Distance

Dist.	# Gold	# Predict	P	R	F_1
-2	79	69	81.2	70.9	75.7
-1	164	151	76.8	70.7	73.7
0	1,811	1,688	77.7	72.4	75.0
1	87	83	78.3	74.7	76.5
2	47	39	87.2	72.3	79.1
Total	2,189	2,030	78.0	72.3	75.1

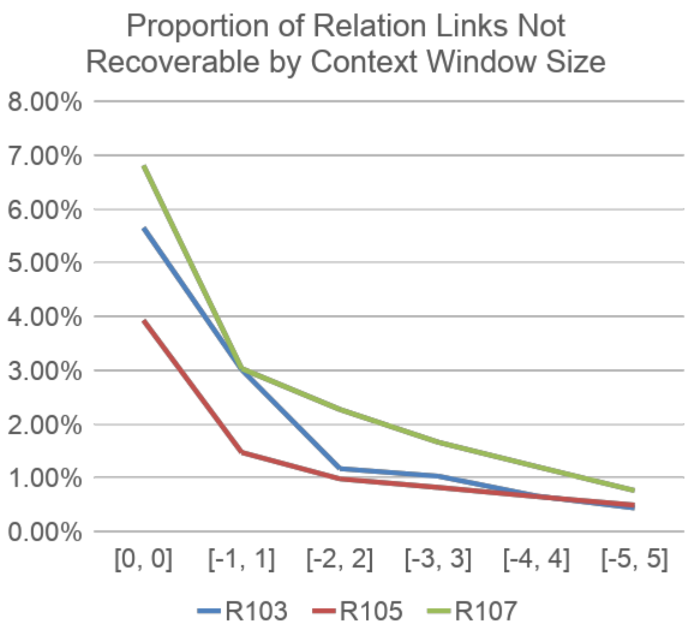
Table 4: Performance breakdown by distance (number of sentences) between argument and event trigger for our model using TCD over the development data.

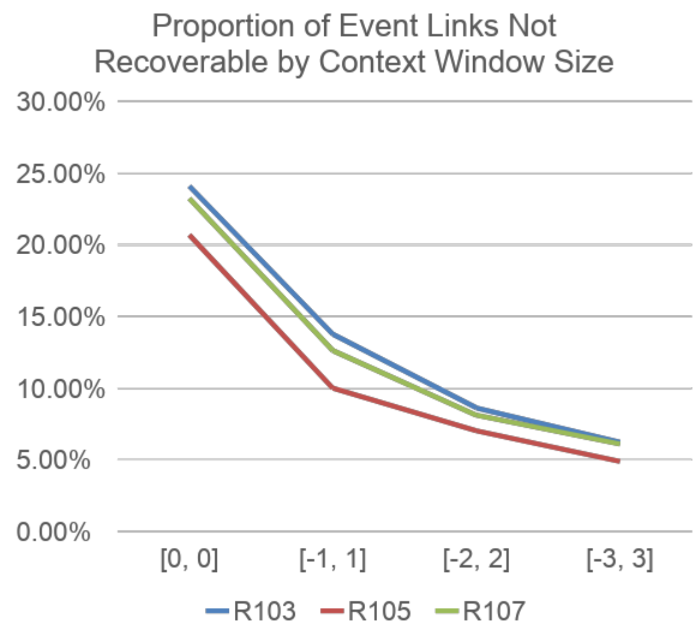
Performance doesn't degrade over distance

RAMS works as pre-training

Strategy	Dev. F_1	P	R	F_1
No pre-training	25.0	36.6	12.9	19.1
No pre-training^T	27.1	53.5	11.0	18.2
RAMS pre-training	34.1	43.9	**16.9**	24.4
RAMS pre-training^T	**34.2**	**62.5**	15.4	**24.8**

Table 8: P(recision), R(ecall), and F_1 on AIDA-1 English development and test data. T designates the use of ontology-aware type-constrained decoding.


Proportion of Relation Links Not Recoverable by Context Window Size
8.00%
7.00%
6.00%
5.00%
4.00%
3.00%
2.00%
1.00%
0.00%
[0, 0]
[-1, 1]
[-2, 2]
[-3, 3]
[-4, 4]
[-5, 5]
R103
R105
R107


Proportion of Event Links Not Recoverable by Context Window Size
30.00%
25.00%
20.00%
15.00%
10.00%
5.00%
0.00%
[0, 0]
[-1, 1]
[-2, 2]
[-3, 3]
R103
R105
R107

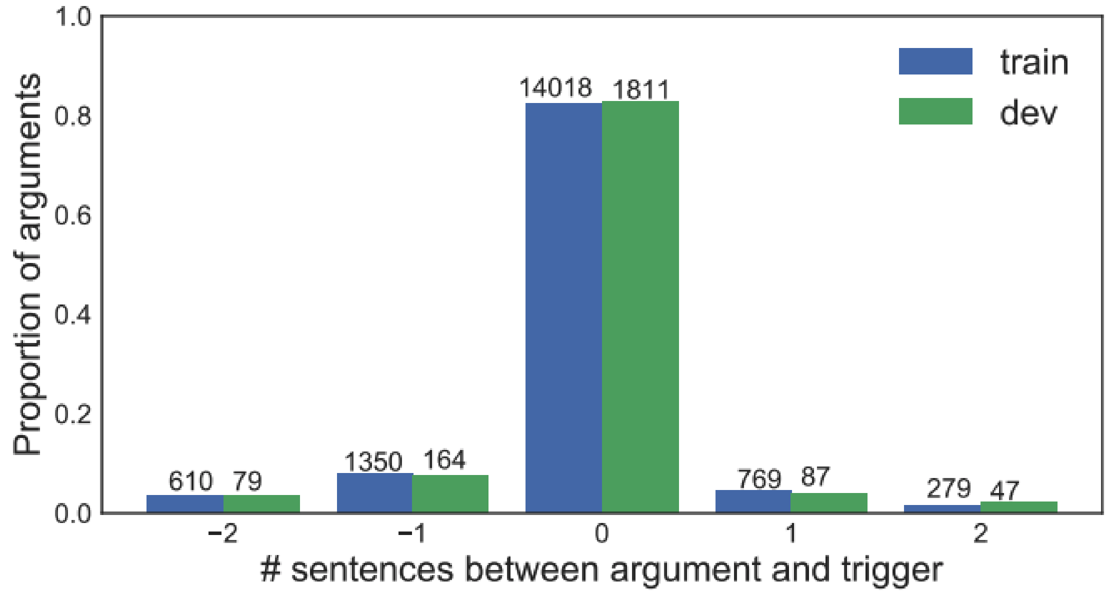


Figure 5: Distances between triggers and arguments in RAMS and proportion of arguments at that distance (counts are shown above each bar). Negative distances indicate that the argument occurs before the trigger.

Threshold	Conjunctive	Disjunctive	Start	End
0	55.3	78.0	59.8	73.5
1	69.9	80.3	74.9	75.3
2	73.9	82.0	78.2	77.8
3	76.4	83.6	80.9	79.1
4	78.8	84.3	82.7	80.4

Table 6: Pairwise span boundary inter-annotator agreement statistics for various span difference thresholds.

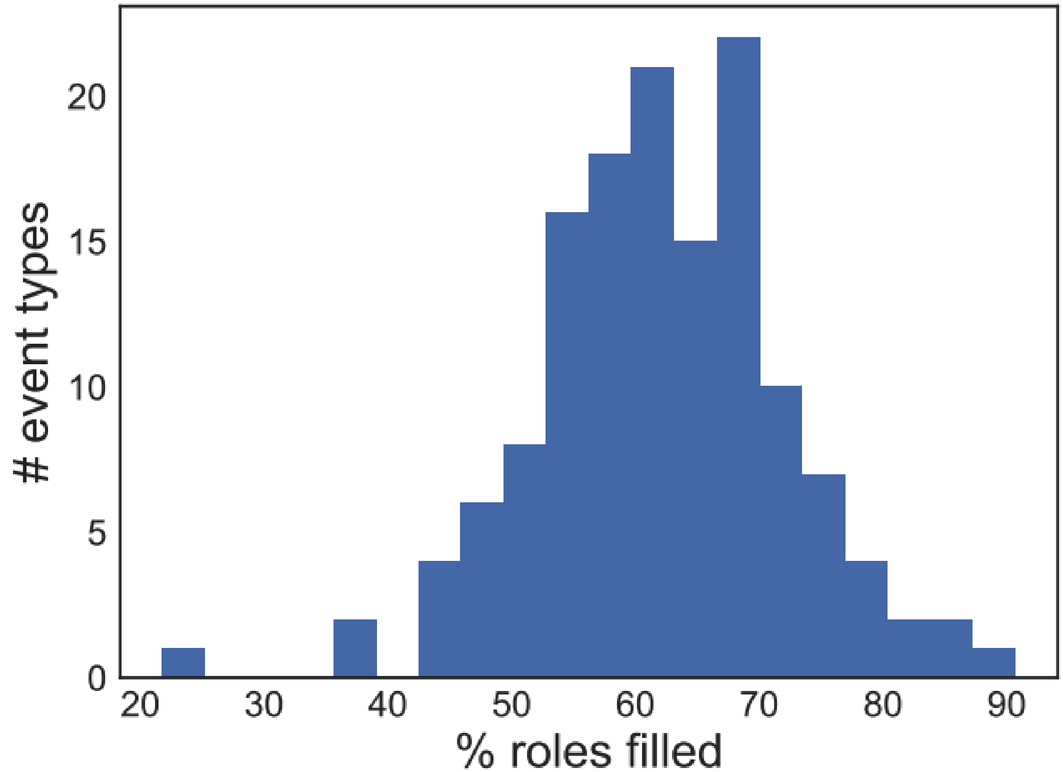


Figure 8: Number of event types for which a given percentage of roles are filled in RAMS train set.

RAMS Results

Model	Dev. F_1	P	R	F_1
Our model	69.9	62.8	**74.9**	68.3
Our model$^{\text{TCD}}$	**75.1**	78.1	69.2	**73.3**
Most common	17.3	15.7	15.7	15.7
Fixed trigger$^{\text{TCD}}$	60.2	**83.7**	41.9	55.8
Context as trigger$^{\text{TCD}}$	62.1	80.5	45.8	58.4

Table 2: P(recision), R(ecall), and F_1 on RAMS development and test data. TCD designates the use of ontology-aware type-constrained decoding.

RAMS Results

Model	Dev. F_1	P	R	F_1
Our model	69.9	62.8	**74.9**	68.3
Our model$^{\text{TCD}}$	**75.1**	78.1	69.2	**73.3**
Most common	17.3	15.7	15.7	15.7
Fixed trigger$^{\text{TCD}}$	60.2	**83.7**	41.9	55.8
Context as trigger$^{\text{TCD}}$	62.1	80.5	45.8	58.4

Table 2: P(recision), R(ecall), and F_1 on RAMS development and test data. TCD designates the use of ontology-aware type-constrained decoding.

RAMS Results

Model	Dev. F_1	P	R	F_1
Our model	69.9	62.8	**74.9**	68.3
Our model^{TCD}	**75.1**	78.1	69.2	**73.3**
Most common	17.3	15.7	15.7	15.7
Fixed trigger^{TCD}	60.2	**83.7**	41.9	55.8
Context as trigger^{TCD}	62.1	80.5	45.8	58.4

Table 2: P(recision), R(ecall), and F_1 on RAMS development and test data. TCD designates the use of ontology-aware type-constrained decoding.

Roles Across Multiple Sentences (RAMS)

Gloss

the banks lent $ 380 million to Trump for the benefit of Trump at someplace

Context

Meanwhile , another Trump disaster was brewing .

Eastern Air Lines , which had been struggling , put its northeastern air shuttle up for sale .

Trump persuaded the banks to lend him $ 380 million to purchase the route , and in June 1989 the Trump Shuttle began flying .

Trump introduced the airline with his usual style — by insulting the competition .

At an elegant event at Logan Airport in Boston , Trump took the stage and suggested that the other airline with a northeastern shuttle , Pan Am , flew unsafe planes .

Event type: Transaction.TransferMoney.BorrowLend

Event definition: The borrowing or lending or money

Arguments

Giver	someone or something	the banks	☐ Argument not present
Recipient	someone or something	Trump	☐ Argument not present
Beneficiary	someone	Trump	☐ Argument not present
Money	some money	$ 380 million	☐ Argument not present
Place	**someplace**		☑ **Argument not present**

Roles Across Multiple Sentences (RAMS)

Gloss

the banks lent $ 380 million to Trump for the benefit of Trump at someplace

Context

Meanwhile , another Trump disaster was brewing .
Eastern Air Lines , which had been struggling , put its northeastern air shuttle up for sale .
Trump persuaded the banks to lend him $ 380 million to purchase the route , and in June 1989 the Trump Shuttle began flying .
Trump introduced the airline with his usual style — by insulting the competition .
At an elegant event at Logan Airport in Boston , Trump took the stage and suggested that the other airline with a northeastern shuttle , Pan Am , flew unsafe planes .

Event type: Transaction.TransferMoney.BorrowLend

Event definition: The borrowing or lending or money

Arguments

Giver	someone or something	the banks	☐ Argument not present
Recipient	someone or something	Trump	☐ Argument not present
Beneficiary	someone	Trump	☐ Argument not present
Money	some money	$ 380 million	☐ Argument not present
Place	someplace		☑ Argument not present

Roles Across Multiple Sentences (RAMS)

Gloss

the banks lent $ 380 million to Trump for the benefit of Trump at someplace

Context

Meanwhile , another Trump disaster was brewing .

Eastern Air Lines , which had been struggling , put its northeastern air shuttle up for sale .

Trump persuaded the banks to lend him $ 380 million to purchase the route , and in June 1989 the Trump Shuttle began flying .

Trump introduced the airline with his usual style — by insulting the competition .

At an elegant event at Logan Airport in Boston , Trump took the stage and suggested that the other airline with a northeastern shuttle , Pan Am , flew unsafe planes .

Event type: Transaction.TransferMoney.BorrowLend

Event definition: The borrowing or lending or money

Arguments

Giver	someone or something	the banks	☐ Argument not present
Recipient	someone or something	Trump	☐ Argument not present
Beneficiary	someone	Trump	☐ Argument not present
Money	some money	$ 380 million	☐ Argument not present
Place	someplace		☑ **Argument not present**

Roles Across Multiple Sentences (RAMS)

Gloss

the banks lent $ 380 million to Trump for the benefit of Trump at someplace

Context

Meanwhile , another Trump disaster was brewing .

Eastern Air Lines , which had been struggling , put its northeastern air shuttle up for sale .

Trump persuaded the banks to lend him $ 380 million to purchase the route , and in June 1989 the Trump Shuttle began flying .

Trump introduced the airline with his usual style — by insulting the competition .

At an elegant event at Logan Airport in Boston , Trump took the stage and suggested that the other airline with a northeastern shuttle , Pan Am , flew unsafe planes .

Event type: Transaction.TransferMoney.BorrowLend

Event definition: The borrowing or lending or money

Arguments

Giver	someone or something	the banks	☐ Argument not present
Recipient	someone or something	Trump	☐ Argument not present
Beneficiary	someone	Trump	☐ Argument not present
Money	some money	$ 380 million	☐ Argument not present
Place	**someplace**		☑ **Argument not present**

Roles Across Multiple Sentences (RAMS)

Gloss

the banks lent $ 380 million to Trump for the benefit of Trump at someplace

Context

Meanwhile , another Trump disaster was brewing .

Eastern Air Lines , which had been struggling , put its northeastern air shuttle up for sale .

Trump persuaded the banks to lend him $ 380 million to purchase the route , and in June 1989 the Trump Shuttle began flying .

Trump introduced the airline with his usual style — by insulting the competition .

At an elegant event at Logan Airport in Boston , Trump took the stage and suggested that the other airline with a northeastern shuttle , Pan Am , flew unsafe planes .

Event type: Transaction.TransferMoney.BorrowLend

Event definition: The borrowing or lending or money

Arguments

Giver	someone or something	the banks	☐ Argument not present
Recipient	someone or something	Trump	☐ Argument not present
Beneficiary	someone	Trump	☐ Argument not present
Money	some money	$ 380 million	☐ Argument not present
Place	someplace		☑ Argument not present

Roles Across Multiple Sentences (RAMS)

Gloss

the banks lent $ 380 million to Trump for the benefit of Trump at someplace

Context

Meanwhile , another Trump disaster was brewing .

Eastern Air Lines , which had been struggling , put its northeastern air shuttle up for sale .

Trump persuaded the banks to lend him $ 380 million to purchase the route , and in June 1989 the Trump Shuttle began flying .

Trump introduced the airline with his usual style — by insulting the competition .

At an elegant event at Logan Airport in Boston , Trump took the stage and suggested that the other airline with a northeastern shuttle , Pan Am , flew unsafe planes .

Event type: Transaction.TransferMoney.BorrowLend

Event definition: The borrowing or lending or money

Arguments

Giver	someone or something	the banks	☐ Argument not present
Recipient	someone or something	Trump	☐ Argument not present
Beneficiary	someone	Trump	☐ Argument not present
Money	some money	$ 380 million	☐ Argument not present
Place	someplace		☑ Argument not present