Diving Deep into Event Semantics

Zhengzhong Liu (Hector)

Thesis Committee:

Teruko Mitamura (Chair)
Eduard Hovy
Taylor Berg-Kirkpatrick
Vincent Ng (University of Texas at Dallas)

Carnegie Mellon University

October 4, 2018

Carnegie Mellon University

Outline

1. Overview of Event Semantics
   - What are Events?
   - Event Semantic Problems
2. Supervised Approaches for Event Semantics
3. Indirect Supervision for Event Semantics (EMNLP'18)
4. Proposed Work: Denotation, Reference, Facets and More

Carnegie Mellon University

What are Events?

Events are our conceptual understanding of happenings, process and state changes in the world.

Let's start with some events

- Mother died today. — Albert Camus, The Stranger (1942)
- They shoot the white girl first. — Toni Morrison, Paradise (1998)

After reading these event descriptions, and one may wonder: what would happen next?

Events are about states

The textual mentions of events (Event Mentions) describe processes and state changes.

- X died: The state of X changed from "Alive" to "Dead".
- X shoot Y: The state of Y changed to "Being Shot"; The state of X changed to "Have Triggered a Shot".

Events have durations

For example, Vendler (1957) classify events into 4 classes based on their aspectual features:

1. **States**: static, no endpoints, e.g. “know”
2. **Activities**: dynamic, no endpoints, e.g. “run”
3. **Accomplishments**: with endpoint, incremental, e.g. “paint a picture”
4. **Achievements**: with endpoint, instantaneous, e.g.“recognize”

Carnegie Mellon University

Outline

1. Overview of Event Semantics
   - What are Events?
   - Event Semantic Problems
2. Supervised Approaches for Event Semantics
   - Event Detection (TAC-KBP 2015 and 2016)
   - Event Coreference (LREC'14, COLING'18)
   - Problems Beyond Full Coreference: Ellipsis (NAACL CORBON 2016)
   - Problems Beyond Full Coreference: Sequencing (COLING'18)
3. Indirect Supervision for Event Semantics (EMNLP'18)
4. Proposed Work: Denotation, Reference, Facets and More
   - Theoretical Analysis
   - Validation Experiments

Outline

Carnegie Mellon University

Why Event Semantics?

1. They describe "what happened", core content of many documents (e.g. news).
2. Textual event mentions are associated with rich structures.
   - Events connect other discourse elements through frame-semantic structure (e.g. participants, location, time)
   - Event mentions have rich inter-connections (e.g. coreference, temporal, script links)

Event Semantic Problems

``It's not a very meaningful indicator currently because corporations are not behaving in a traditional manner,'' says James H Coxon, head of stock investments for Cigna Corp., the Philadelphia-based insurer.
In particular, Mr. Coxon says, businesses are paying out a smaller percentage of their profits and cash flow in the form of dividends than they have historically.
So, while stock prices may look fairly high relative to dividends, they are not excessive relative to the underlying corporate strength.
Rather than increasing dividends, some companies have used cash to buy back some of their shares, notes Steven G. Einhorn, co-chairman of the investment policy committee at Goldman, Sachs & Co.

Carnegie Mellon University

Event Semantic Problems

Detection:
- Identify mention spans

``It's not a very meaningful indicator currently because corporations are not behaving in a traditional manner,'' says James H Coxon, head of stock investments for Cigna Corp., the Philadelphia-based insurer.
In particular, Mr. Coxon says, businesses are paying out a smaller percentage of their profits and cash flow in the form of dividends than they have historically.
So, while stock prices may look fairly high relative to dividends, they are not excessive relative to the underlying corporate strength.
Rather than increasing dividends, some companies have used cash to buy back some of their shares, notes Steven G. Einhorn, co-chairman of the investment policy committee at Goldman, Sachs & Co.

Event Semantic Problems

Detection:
- Identify mention spans

Event Argument:
- SRL for mentions

``It's not a very meaningful indicator currently because corporations are not behaving in a traditional manner,'' says James H Coxon, head of stock investments for Cigna Corp., the Philadelphia-based insurer.
In particular, Mr. Coxon says, businesses are paying out a smaller percentage of their profits and cash flow in the form of dividends than they have historically.
So, while stock prices may look fairly high relative to dividends, they are not excessive relative to the underlying corporate strength.
Rather than increasing dividends, some companies have used cash to buy back some of their shares, notes Steven G. Einhorn, co-chairman of the investment policy committee at Goldman, Sachs & Co.

Carnegie Mellon University

Event Semantic Problems

Detection:
- Identify mention spans

Event Argument:
- SRL for mentions

Event Attributes:
- Type
- Realis
- etc.

``It's not a very meaningful indicator currently because corporations are not behaving in a traditional manner,'' says James H Coxon, head of stock investments for Cigna Corp., the Philadelphia-based insurer.
In particular, Mr. Coxon says, businesses are paying out a smaller percentage of their profits and cash flow in the form of dividends than they have historically.
So, while stock prices may look fairly high relative to dividends, they are not excessive relative to the underlying corporate strength.
Rather than increasing dividends, some companies have used cash to buy back some of their shares, notes Steven G. Einhorn, co-chairman of the investment policy committee at Goldman, Sachs & Co.

Carnegie Mellon University

Event Semantic Problems

Detection:
- Identify mention spans

Event Argument:
- SRL for mentions

Event Attributes:
- Type
- Realis
- etc.

Type: TransferMoney

``It's not a very meaningful indicator currently because corporations are not behaving in a traditional manner,'' says James H Coxon, head of stock investments for Cigna Corp., the Philadelphia-based insurer.
In particular, Mr. Coxon says, businesses are paying out a smaller percentage of their profits and cash flow in the form of dividends than they have historically.
So, while stock prices may look fairly high relative to dividends, they are not excessive relative to the underlying corporate strength.
Rather than increasing dividends, some companies have used cash to buy back some of their shares, notes Steven G. Einhorn, co-chairman of the investment policy committee at Goldman, Sachs & Co.

Carnegie Mellon University

Event Semantic Problems

Detection:
- Identify mention spans

Event Argument:
- SRL for mentions

Event Attributes:
- Type
- Realis
- etc.

Realis :
Not happened

``It's not a very meaningful indicator currently because corporations are not behaving in a traditional manner,'' says James H Coxon, head of stock investments for Cigna Corp., the Philadelphia-based insurer.
In particular, Mr. Coxon says, businesses are paying out a smaller percentage of their profits and cash flow in the form of dividends than they have historically.
So, while stock prices may look fairly high relative to dividends, they are not excessive relative to the underlying corporate strength.
Rather than increasing dividends, some companies have used cash to buy back some of their shares, notes Steven G. Einhorn, co-chairman of the investment policy committee at Goldman, Sachs & Co.

Event Semantic Problems

Detection:
- Identify mention spans

Event Argument:
- SRL for mentions

Event Attributes:
- Type
- Realis
- etc.

Event Relations:
- Coreference
- Anaphora
- etc.

``It's not a very meaningful indicator currently because corporations are not behaving in a traditional manner,'' says James H Coxon, head of stock investments for Cigna Corp., the Philadelphia-based insurer.
In particular, Mr. Coxon says, businesses are paying out a smaller percentage of their profits and cash flow in the form of dividends than they have historically.
So, while stock prices may look fairly high relative to dividends, they are not excessive relative to the underlying corporate strength.
Rather than increasing dividends, some companies have used cash to buy back some of their shares, notes Steven G. Einhorn, co-chairman of the investment policy committee at Goldman, Sachs & Co.

Carnegie Mellon University

Event Semantic Problems

Detection:
- Identify mention spans

Event Argument:
- SRL for mentions

Event Attributes:
- Type
- Realis
- etc.

Event Relations:
- Coreference
- Anaphora
- etc.

Coreference?

``It's not a very meaningful indicator currently because corporations are not behaving in a traditional manner,'' says James H Coxon, head of stock investments for Cigna Corp., the Philadelphia-based insurer.
In particular, Mr. Coxon says, businesses are paying out a smaller percentage of their profits and cash flow in the form of dividends than they have historically.
So, while stock prices may look fairly high relative to dividends, they are not excessive relative to the underlying corporate strength.
Rather than increasing dividends, some companies have used cash to buy back some of their shares, notes Steven G. Einhorn, co-chairman of the investment policy committee at Goldman, Sachs & Co.

Event Semantic Problems

Detection:
- Identify mention spans

Event Argument:
- SRL for mentions

Event Attributes:
- Type
- Realis
- etc.

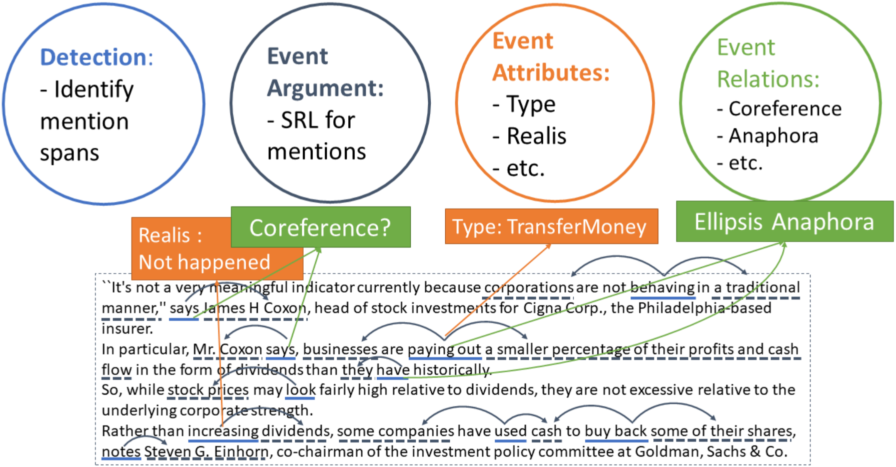
Event Relations:
- Coreference
- Anaphora
- etc.

Ellipsis Anaphora

``It's not a very meaningful indicator currently because corporations are not behaving in a traditional manner,'' says James H Coxon, head of stock investments for Cigna Corp., the Philadelphia-based insurer.
In particular, Mr. Coxon says, businesses are paying out a smaller percentage of their profits and cash flow in the form of dividends than they have historically.
So, while stock prices may look fairly high relative to dividends, they are not excessive relative to the underlying corporate strength.
Rather than increasing dividends, some companies have used cash to buy back some of their shares, notes Steven G. Einhorn, co-chairman of the investment policy committee at Goldman, Sachs & Co.

Carnegie Mellon University

Event Semantic Problems



Carnegie Mellon University

Thesis Goal

1. Build the event semantic components:
   - Our early approaches are based in supervised setting.
   - Recent attempt on indirect supervision is more promising (Proposed Work).
2. Explore the structures between events and other discourse elements:
   - We have identified some obstacles along this line.
   - Our theoretical study shows some possible solutions (Proposed Work).

Outline

Event Detection

Problem Definition

Automatically identity text spans that refer to an event, and assign it with predefined properties.

The predefined properties of event mentions may include:

1. Event Type: an ontology type in the domain, such as "Conflict.Attack".
2. Realis/Epistemic Status: e.g. whether the event actually happened.

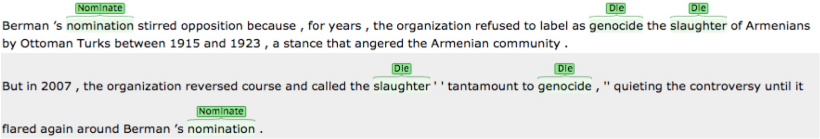


Figure 1: An Example Event Nugget Annotation

Event Detection as Sequence Labeling

1. The problem can be viewed as a sequence labeling task, we model it using a Conditional Random Field (CRF).

Event Detection as Sequence Labeling

1. The problem can be viewed as a sequence labeling task, we model it using a Conditional Random Field (CRF).

Event Detection as Sequence Labeling

1. The problem can be viewed as a sequence labeling task, we model it using a Conditional Random Field (CRF).
2. Challenges comparing to NER.
    - Number of event types are large and some types are only observed a couple times.
    - Some types are very similar in text:
        - Transport-Person vs. Transport-Artifact (e.g. Move, Transport)
        - Transfer-Ownership vs. Transfer-Money (e.g. Give, Buy)
    - Type detection highly depends on other units, such as event arguments.

Feature Design

1. Represent the sense of the trigger.
   - Ontology level representations such as FrameNet.
2. Represent the sense of the surrounding context.
   - Cluster and higher ontology level features (e.g. Brown Cluster, WordNet)
3. Represent the sense of the arguments.
   - Domain Specific Argument Types obtain from various resources (e.g. Money, Weapon)

Event Detection Results Liu et al. (2016a)

	Prec.	Recall	F1
UTD1	47.66	46.35	46.99
LTI-CMU1	61.69	34.94	44.61
wip1	51.76	38.98	44.47
NYU3	41.88	47.21	44.38
SoochowNLP3	49.92	38.81	43.67

Table 1: Top 5 English Nugget Type Detection Systems in TAC-KBP 2016

Event Detection Results (Chinese) Liu et al. (2016a)

		Prec.	Recall	F1
Span	LTI-CMU1	56.46	39.55	46.52
	UTD1	47.23	43.16	45.1
	LTI-CMU3	56.19	35.35	43.4
	UI-CCG1	28.34	39.61	33.04
	RPI-BLENDER1	62.46	18.48	28.52
Type	LTI-CMU1	50.72	35.53	41.79
	UTD1	41.9	38.29	40.01
	LTI-CMU3	49.7	31.26	38.38
	UI-CCG1	24.01	33.55	27.99
	RPI-BLENDER2	59.87	17.5	27.08

Table 2: Top-5 Chinese Nugget Detection System at TAC-KBP 2016.

Multi-tagging Event Nuggets

Annotation phenomenon: one mention span may correspond to multiple event types.

1. The mention "kill" may have types "Conflict.Attack" and "Life.Death"
2. Similarly, 打死 may have types "Conflict.Attack" and "Life.Death".

The Multi-Tagging Problem

How can we explain these multi-tagged events? How do we determine the number of types?

Outline

1. Overview of Event Semantics
2. Supervised Approaches for Event Semantics
   - Event Detection (TAC-KBP 2015 and 2016)
   - Event Coreference (LREC'14, COLING'18)
   - Problems Beyond Full Coreference: Ellipsis (NAACL CORBON 2016)
   - Problems Beyond Full Coreference: Sequencing (COLING'18)
3. Indirect Supervision for Event Semantics (EMNLP'18)
4. Proposed Work: Denotation, Reference, Facets and More

Carnegie Mellon University

Event Coreference

Problem Definition

Event Coreference is a type of relation between event mentions that refer to the same underlying event.

Event Coreference Example

- PLO **picks** new leaders at landmark meeting.
- The parliament of the Palestine Liberation Organization (PLO) **elected** on Thursday six new members to its executive body.

Event Coreference Decoding Structure

1. Pairwise model considers two pair at a time: decisions may conflict.
2. We adopt a tree based model
   - Follow the discourse order.
   - Find one single antecedent for each mention at a time.

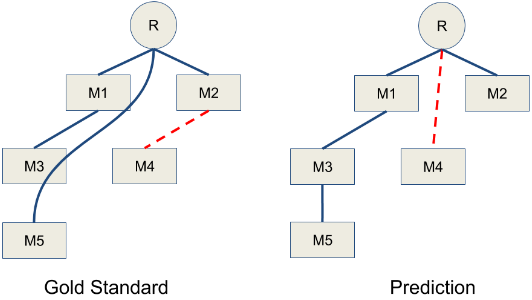


Figure 2: Training with the Latent Antecedent Tree Model

Feature Design

1. Predicate Similarities
   - A range of lexical similarities measures for the predicates.
2. Argument Similarities
   - Compare the argument similarities based on headwords and whole phrase.
3. Discourse Features
   - Such as event and sentence distances.

Event Coreference Results

	B^3	CEAF-E	MUC	BLANC	AVG.
`Singleton`	78.10	68.98	0.00	48.88	52.01
`Matching`	78.40	65.82	**69.83**	76.29	71.94
LCC	82.85	74.66	68.50	**77.61**	75.69
UI-CCG	83.75	75.81	63.78	73.99	74.28
LTI	82.27	75.15	60.93	71.57	72.60
Our System	**85.59**	**79.65**	67.81	77.37	**77.61**

Table 3: Test Results for Event Coreference with the `Singleton` and `Matching` baselines. 4 different popular coreference metrics are used. AVG. is the mean of the 4 values.

Mixed Results from Argument Matching

Event arguments should be important in distinguish different event mentions. However, we've observed mixed results:

1. In Liu et al. (2014), argument information improve our system by 4 absolute scores on Average.
2. In Liu et al. (2018), we find adding the same argument features slightly hurt the performance.

Carnegie Mellon University

The Myth of Arguments

There are some obvious problems:

1. Arguments are often omitted (e.g. The election finished on Saturday.).
2. Argument phrase is often long, increasing the sparsity.

But let's revisit the example

- *PLO* **picks** new leaders at landmark meeting.
- *The parliament of the Palestine Liberation Organization (PLO)* **elected** on Thursday six new members to its executive body.

Why there are some Inconsistencies in the arguments?

Is "PLO" identical to "The parliament of PLO"?

Outline

Carnegie Mellon University

Beyond Full Coreference: Ellipsis Anaphora

There are rich coreference or anaphora phenomenon among events:

Verb Phrase Ellipsis (VPE)

An anaphoric process where a verbal constituent is partially or totally unexpressed, but can be resolved through an antecedent in the context.

Verb Phrase Ellipsis Example

- His wife also [antecedent *works for the paper*], as **did** his father.
- In particular, Mr. Coxon says, businesses are [antecedent *paying out a smaller percentage of their profits and cash flow in the form of dividends*] than they **have** historically.

VPE Model

The steps to solve a VPE problem can be divided into two steps:

1. Detect the VPE targets, such as **did**.
   - Classify whether a light verb is elided.
2. Identify the antecedent phrase, which can be further considered as:
   - Antecedent Head Selection: rank the previous verbs as its antecedent.
   - Antecedent Boundary Detection: rank the list of valid syntactic constituents.

VPE Feature Design

1. Target Detection: syntactic roles and context words are informative.
2. Antecedent Selection:
   - Syntactic features to find the governing relationships.
   - Context similarity features based on matching.
   - Distance features.
3. Antecedent Boundary Detection:
   - The context similarity between the antecedent and target.
   - Immediate context are important (e.g. also vs. as)

VPE Results: Target Detection

	WSJ			BNC		
	Prec	Rec	F1	Prec	Rec	F1
Oracle	100.00	93.28	96.52	100.00	92.65	96.18
POS Baseline	42.62	43.7	43.15	35.47	35.29	35.38
Nielsen (2005)	—	—	—	72.50	72.86	72.68
Ours	**80.22**	**61.34**	**69.52**	**80.90**	**70.59**	**75.39**

Table 4: Results for Target Detection on 2 datasets

VPE Results: End-to-End

	WSJ			BNC		
	Prec	Rec	F1	Prec	Rec	F1
Oracle	95.06	88.67	91.76	85.79	79.49	82.52
Rule Baseline	19.04	19.52	19.27	12.81	12.75	12.78
Ours (Pipeline)	52.68	40.28	45.65	**43.03**	**37.54**	**40.10**
Ours (Joint)	**54.82**	**41.92**	**47.51**	41.86	36.52	39.01

Table 5: End-to-End results for VPE. The rule baseline uses the POS for target detection, previous verb for antecedent selection, and largest span for boundary detection. The joint model jointly model the last two steps.

Outline

1. Overview of Event Semantics
2. Supervised Approaches for Event Semantics
   - Event Detection (TAC-KBP 2015 and 2016)
   - Event Coreference (LREC'14, COLING'18)
   - Problems Beyond Full Coreference: Ellipsis (NAACL CORBON 2016)
   - Problems Beyond Full Coreference: Sequencing (COLING'18)
3. Indirect Supervision for Event Semantics (EMNLP'18)
4. Proposed Work: Denotation, Reference, Facets and More

Carnegie Mellon University

Beyond Full Coreference: Event Sequencing

We tend to infer what will happen next when we see an event, this cognitive process is described by Schank and Abelson (1977).

Schank's Script

Human organize events as episodes, prototypical set of events. These structured are called scripts, which allow individuals to make inferences on missing information.

The Event Sequencing task studies the problem of connecting these prototypical event mentions in a text document.

- The list of verbs *order, eat, pay, leave* may trigger the restaurant script.
- The sequencing task aims at identity this set and order them.

Sequencing Model

The structure of sequencing is different from coreference:

- The sequencing links are "directed".
- They can have multiple parents.

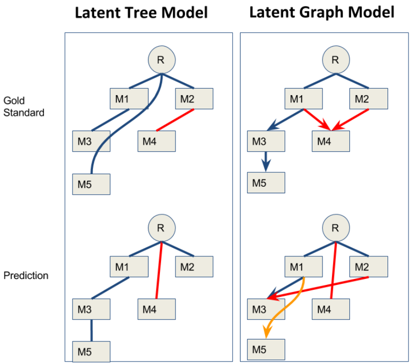


Figure 3: Latent Tree Model (left): tree structure formed by undirected links. Latent Graph Model (right): a DAG form by directed links.

Feature Design

1. Event Compatibility
   - Lexical Compatibility: whether two mentions are compatible lexically (e.g. attack and kill).
   - Discourse Compatibility: whether they are related in terms of discourse (e.g. parse tree, distance).
2. Event Ordering
   - Temporal clues.
   - Discourse ordering.

Carnegie Mellon University

What Connects the Sequence?

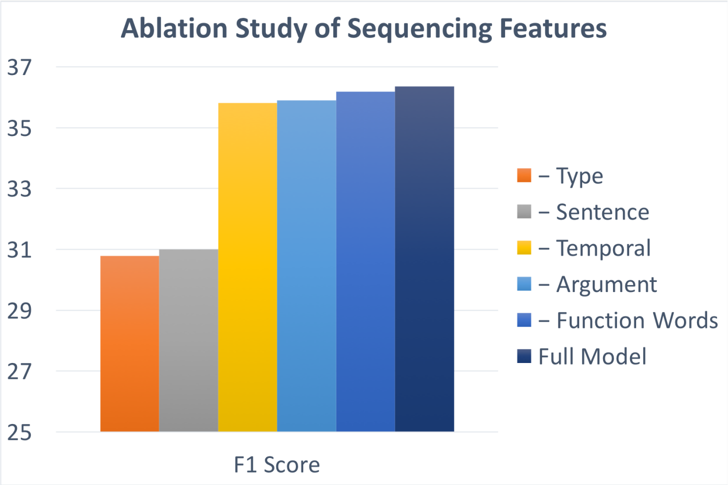


Figure 4: Ablation Study for Event Sequencing

What Connects the Sequence?

Prior script literature suggests that argument is key in connecting the script. But their effect is minimum here.

What happpend?

- Why are event arguments not effective in sequencing?
- What are the alternatives?

Seeking Indirect Supervision

1. All the previous approaches are based some annotated datasets.
2. Event semantics are complex and general, which are difficult to be manually labelled.
    - There are more event types than the domain specific ones.
    - Annotation on event sequencing is slow, and depends on predefined domain constraints.
3. We attempt to seek for other sources of supervisions.
    - We demonstrate our approach with an event salience task.

Event Salience

Event Salience Task

The task of **event salience detection** is to find events that are most relevant to the main content of documents.

Extracting the most prominent information from text is widely studied. The salience of events has been used before, but haven't been studied thoroughly.

Creating an Event Dataset at Scale

1. Frame Based Event Mention Selection
   - A list of frame and sub-frames related to process and events are used to filter the verbs in the corpus.
2. Annotating Event Salience with the Summarization Test.
   - Hypothesis: Salient events are the ones a person would choose to summarize the document.
   - We use Annotated NYT: a dataset with human-written abstracts.
3. On a small scale validation annotation, the Cohen's Kappa between two annotators and the system are 0.49 and 0.42 respectively.

Federal prosecutors **urged** a **trial** **judge** today to **deny** defense **requests** to **delay** the **trial** of Zacarias Moussaoui and suggested that Mr. Moussaoui, the only person **charged** in the Sept. 11 **attacks**, was to **blame** for many of the **delays** so far. The **attacks** ''were volleys in a declared war against the United States and were more than just acts of terror,'' the prosecutors said in a filing to the Federal District Court in Alexandria, Va. ''Thus, the victims' and the nation's interest in a fair and speedy **trial** is beyond dispute.''.Last week, court-appointed defense lawyers asked that the starting date of the **trial**, now set for Sept. 30, be **delayed** by at least two months to allow them to wade through volumes of evidence that prosecutors have presented to them, including more than 1,300 computer discs.

Figure 5: Examples annotations. Underlying words are annotated event triggers; the red bold ones are annotated as salient.

Event Salience Model

Frequency	Frequency of event lemma in document.
Sentence Location	Location of the first sentence of the event.
Event Voting	Aver. cosine with other document events.
Entity Voting	Aver. cosine with other document entities.
Local Entity Voting	Aver. cosine with entities in the sentence.

Table 6: Baseline Salience Features.

Event Salience Model: Capturing Element Relations

1. Multiple types of relations exist in a discourse:
   - Frame Argument Relations (e.g. purchase vs. product).
   - Script Relations (e.g. purchase vs. ship).
2. Capturing relations beyond similarity voting:
   - We use K different soft gaussian kernels applied to the voting. (K=10 in our experiments.)
   - Produce K additional features: one for each kernel.

Carnegie Mellon University

Event Salience Results (Precision)

Method	P@01		P@05		P@10	
Location	0.3555	–	0.3077	–	0.2505	–
PageRank	0.3628	–	0.3438	–	0.3007	–
Frequency	0.4542	–	0.4024	–	0.3445	–
LeToR	$0.4753^{\dagger}$	+4.64%	$0.4099^{\dagger}$	+1.87%	$0.3517^{\dagger}$	+2.10%
KCE (-EF)	0.4420	−2.69%	0.4038	+0.34%	$0.3464^{\dagger}$	+0.54%
KCE (-E)	$0.4861^{\dagger\ddagger}$	+7.01%	$0.4227^{\dagger\ddagger}$	+5.04%	$0.3603^{\dagger\ddagger}$	+4.58%
KCE	$0.5049^{\dagger\ddagger}$	+11.14%	$0.4277^{\dagger\ddagger}$	+6.29%	$0.3638^{\dagger\ddagger}$	+5.61%

Table 7: Event Salience P@X performance. (-E) and (-F) marks removing Features and Entity information from the full model. The relative performance differences are computed against Frequency. † and ‡ mark the statistically significant improvements over Frequency†, LeToR‡ respectively.

Event Salience Results (Recall)

Method	R@01		R@05		R@10	
Location	0.0807	–	0.2671	–	0.3792	–
PageRank	0.0758	–	0.2760	–	0.4163	–
Frequency	0.0792	–	0.2846	–	0.4270	–
LeToR	0.0836†	+5.61%	0.2980†	+4.70%	0.4454†	+4.31%
KCE (-EF)	0.0714	−9.77%	0.2812	−1.18%	0.4321†	+1.20%
KCE (-E)	0.0925†‡	+16.78%	0.3172†‡	+11.46%	0.4672†‡	+9.41%
KCE	0.0946†‡	+19.44%	0.3215†‡	+12.96%	0.4719†‡	+10.51%

Table 8: Event Salience R@X performance. (-E) and (-F) marks removing Features and Entity information from the full model. The relative performance differences are computed against Frequency. † and ‡ mark the statistically significant improvements over Frequency†, LeToR‡ respectively.

Carnegie Mellon University

Capturing Important Relations

1. Cosine similarities for the following word pairs are all around 0.4 to 0.5, in pre-trained embedding space.
    - business, increase
    - arrest, charge
    - attack, walk
    - hotel, travel
    - charge, murder
2. Kernel model assign high weights to:
    - Script pairs: arrest, charge; charge, murder
    - Frame argument pairs: business, increase; hotel, travel

Carnegie Mellon University

Outline

Revisiting the Questions

1. Why some event mentions can have multiple types?
2. Why the arguments in a coreferent mention pair not matching?
3. Why the argument cannot be used as the connectives between the events?
4. What can be used to connect the events to scripts/sequences?

Carnegie Mellon University

Decomposable Event Semantics

Discourse Elements are Decomposable

Events can be decomposed to smaller semantic units.

1. "The attacker kills the victim."
    - We can analyze the multiple types using the core arguments.
    - The "Conflict.Attack" is more closer to the "Attacker".
    - The "Life.Death" type is closer to the "Victim".
2. Why? Because each core argument is a candidate for state change.
    - And events are descriptors for state changes.
3. Evidence in Chinese events.
    - 坠毁 (crash) can be decomposed as 坠 (fell) and 毁 (destroy)

Partial Denotation Reference

1. Discourse Elements in a document (Events and Entities) often have rich denotation
   - The entity mention "President Trump" refers to a person, which include his political role, his social role, his family role and more.
2. A textual mention normally only focuses on part of the whole denotation space.
3. In the phrase "President Trump's announcement", the reading of this mention is about the political role.

Partial Denotation and the Identity Problem

Ignoring this decomposition of DEs risks inaccurate inference and unintended interpretations. Let's analyze "PLO" vs. "The parliament of the Palestine Liberation Organization (PLO)" again.

Inexact Entity Identity in Event Coreference

- PLO **picks** new leaders at landmark meeting.
- The parliament of the Palestine Liberation Organization (PLO) **elected** on Thursday six new members to its executive body.

Partial Denotation and the Identity Problem

Example of Inexact Argument Identity in Script

The police **warned** the crowd to disperse, but the protesters **refused** to leave Taksim Square and chanted anti-government slogans. *Dozens of police officers* then **moved toward** the crowd and began **spraying** the protesters with water cannons. ...said as she held red carnations in her hand. "I wanted to hand these to them, but instead they **pushed** me away with their shields and said our right to protest was over."

- For example, **warned** and **spraying** should be in the same script.
- Their sharable argument should be the agents.
- But, are they the same? *The Police* vs. *Dozens of police officers*

The Facet-Based Representation

We propose to describe the discourse elements in smaller units:

Facets

The facets of an discourse element are semantic units corresponding to the possible interpretations given the textual description in a discourse model.

1. Recasens et al. (2011) solve the inexact coreference problem by assigning scalar values to the coreference degree. But it is difficult to be consistent on the value.
2. The facet-based representation can provide a more practical model.
3. Facets can also explain the coreference link.

Carnegie Mellon University

Active Facets

The concept of active facets can help us analyze problems.

Active Facets

The active facets of a mention of a DE are the facets that are considered to be relevant to the language user given the linguistic and pragmatic context.

The active facets are the set of denotations for the purpose of this context.

Hypothesis: Active Facets and State Changes

1. Active facets can be derived from the context.
2. Coreference should be done on the active facets.

Are the plants in "France" and "Catalonia" the same one?

The *plant* colonized the South of France, from where *it* entered Catalonia in the 80s, spreading quickly.

Analysis

1. Active Facets: the verb "colonize" make us focus on the organizational perspective.
2. Comparing: Both entity (*plant* and *it*) have the same organization.
3. Irrelevant Facets: The actual facility location. (A facility cannot "colonize")

Active Facets and State Changes

The Relation between State and Facet

We can consider events as functions that change state. So what are the function arguments?

1. Hypothesis: The arguments for the event function are the active facets.
2. Modeling state change on the facet level may allow more generalizability.
   - Since facets can be shared across entities.

Active Facets and State Changes

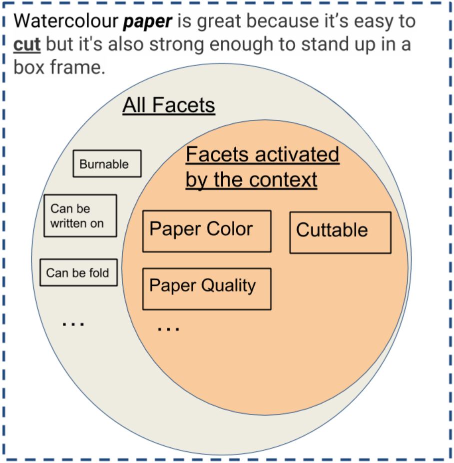


Figure 6: Active facets for the entity mention “paper”.

Active Facets and State Changes

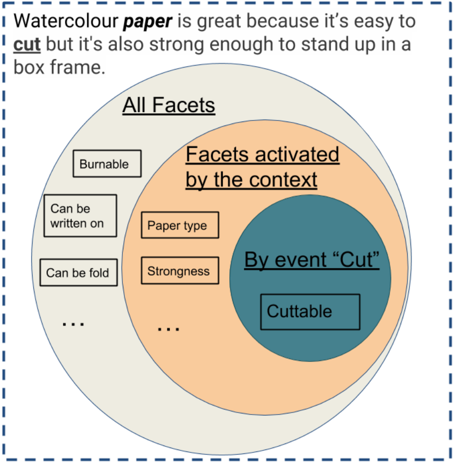


Figure 7: Active facets for the entity mention "paper"

Proposed Answer to Our Questions

Why some event mentions can have multiple types?

- Each event mention have multiple facets, each corresponds to a type. Some of them are irrelevant to the context/application.

Proposed Answer to Our Questions

Why the arguments in a coreferent mention pair not matching?

- Their active facets are matching, the whole argument pairs might not.
- In theory, some phenomenon cannot be easily explained by facets, but can be explained well by quasi-identity theory Recasens et al. (2011).

Proposed Answer to Our Questions

Why the argument cannot be used as the connectives between the events?

- They can, but we should not use identical argument as the connectives, but identical active facets.

Proposed Answer to Our Questions

What can be used to connect the events to scripts/sequences?

- The argument facets.
- The states.

Outline

Validation Experiments: Static

1. Event Coreference (Ongoing)
   - Use facet mapping to improve event coreference.

Validation Experiments: Static

1. Event Coreference (Ongoing)
    - Use facet mapping to improve event coreference.
2. Implicit Semantic Role Labeling (Ongoing)
    - Find semantic arguments beyond the sentence level.
    - Use facet matching to select the argument.

Validation Experiments: Static

1. Event Coreference (Ongoing)
   - Use facet mapping to improve event coreference.
2. Implicit Semantic Role Labeling (Ongoing)
   - Find semantic arguments beyond the sentence level.
   - Use facet matching to select the argument.
3. Quasi-Entity Coreference (the NIDENT corpus)
   - Utilize event coreference and script to predict quasi entity coreference.

Validation Experiments: State Changes

1. Script/Sequence Induction
    1. Induce event sequence by identify the state changes.

Validation Experiments: State Changes

1. Script/Sequence Induction
   1. Induce event sequence by identify the state changes.
2. State Aware Selectional Preference
   1. Current approaches on verb selection preference do not consider the states of the entity.
   2. We propose to model the state in SRL:
      - Target on domains with frequent state changes. (e.g. cooking recipes)

Timeline

- Event Detection (Done)
  - System for TAC-KBP 2015, 2016 Liu et al. (2015, 2016a)
- Event Coreference (Done)
  - Published at LREC 2014 Liu et al. (2016a, 2014)
- Verb Phrase Ellipsis (Done)
  - Published at NAACL 2016 workshop Liu et al. (2016b)
- Event Sequencing (Done)
  - Published at COLING 2018 Liu et al. (2018)
- Facet Identification:
  - Implicit Argument Modeling (In Progress), Target: NAACL 2019
  - Facet Aware Event Coreference (In Progress), Target: ACL 2019
- State Modeling (Not Started)
  - Application: argument selection or script modeling
  - Target: EMNLP 2019
- Thesis writing: Summer 2019
- Thesis defense: Fall 2019

Thanks You!

Questions?

Carnegie Mellon University

References I

Zhengzhong Liu, Jun Araki, Dheeru Dua, Teruko Mitamura, and Eduard Hovy. 2015. CMU-LTI at KBP 2015 Event Track. In *KBP TAC 2015*. NIST.

Zhengzhong Liu, Jun Araki, Eduard Hovy, and Teruko Mitamura. 2014. Supervised Within-Document Event Coreference using Information Propagation. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, pages 4539–4544, Reykjavik, Iceland. European Language Resources Association (ELRA).

Zhengzhong Liu, Jun Araki, Teruko Mitamura, and Eduard Hovy. 2016a. CMU-LTI at KBP 2016 Event Nugget Track. In *TAC 2016*.

Zhengzhong Liu, Edgar Edgar Gonzàlez, and Dan Gillick. 2016b. Exploring the steps of Verb Phrase Ellipsis. In *NAACL Workshop on Coreference Resolution Beyond OntoNotes (CORBON 2016)*.

Carnegie Mellon University

References II

Zhengzhong Liu, Teruko Mitamura, and Eduard Hovy. 2018. Graph-Based Decoding for Event Sequencing and Coreference Resolution. In *Proceedings of the 27th International Conference on Computational Linguistics*.

Leif Arda Nielsen. 2005. *A corpus-based study of Verb Phrase Ellipsis Identification and Resolution*. Doctor of philosopy, King's College London.

Marta Recasens, Eduard Hovy, and M. Antonia Marti. 2011. Identity, non-identity, and near-identity: Addressing the complexity of coreference. *Lingua*, 121(6):1138–1152.

Roger C Schank and Robert P Abelson. 1977. *Scripts, Plans, Goals and Understanding*. Lawrence Erlbaum Associates.

References III

Stephen Tratz and Eduard Hovy. 2011. A fast, accurate, non-projective, semantically-enriched parser. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 2010, pages 1257–1268.

Zeno Vendler. 1957. Verbs and times. *The Philosophical Review*, 66(2):143–160.

Carnegie Mellon University

Appendix

The followings slides are appendix for features and results.

Event Coreference Features

Head	Headword token and lemma pair.
Type	The pair of event types.
Realis	The pair of realis types.
POS	POS pair of the two mentions.
Exact Match	If the 5-word windows match exactly.
Distance	Sentence distance in between.
Frame	Frame name pair.
Syntactic	Whether a mention is the syntactic ancestor of another.

Table 9: A Simplified List of Coreference Features.

Carnegie Mellon University

VPE Features: Target Detection

1. The POS tag, lemma, and dependency label of the verb, its dependency parent, and the immediately preceding and succeeding words.
2. The POS tags, lemmas and dependency labels of the words in the dependency subtree of the verb, in the 3-word window, and in the same-size window after (as bags of words).
3. Whether the subject of the verb appears to its right (i.e., there is subject-verb inversion).

VPE Features: Antecedent Selection

Type	Feature Description	Purpose
Labels	The POS tag and dependency label of the antecedent head	H
	The POS tag and dependency label of the antecedent's last word	B
	The POS tag and lemma of the antecedent parent	H
	The POS tag, lemma and dependency label of within a 3 word around around the antecedent	B
	The pair of the POS tags of the antecedent head and the target, and of their auxiliary verbs	H
	The pair of the lemmas of the auxiliary verbs of the antecedent head and the target	H

Table 10: Antecedent Features

Carnegie Mellon University

VPE Features: Antecedent Selection

Tree	Whether the antecedent and the target form a comparative construction connecting by *so*, *as* or *than*	H&B
	The dependency labels of the shared lemmas between the parse tree of the antecedent and the target	H
	Label of the dependency between the antecedent and target (if exists)	H
	Whether the antecedent contains any descendant with the same lemma and dependency label as a descendant of the target.	H
	Whether antecedent and target are dependent ancestor of each other	H
	Whether antecedent and target share prepositions in their dependency tree	H

Table 11: Antecedent Features

Carnegie Mellon University

VPE Features: Antecedent Selection

Distance	The distance in sentences between the antecedent and the target (clipped to 2)	H
	The number of verb phrases between the antecedent and the target (clipped to 5)	H
Match	Whether the lemmas of the heads, and words in the the window (=2) before the antecedent and the target match respectively	H
	Whether the lemmas of the ith word before the antecedent and $i-1$th word before the target match respectively (for $i \in \{1, 2, 3\}$, with the 0th word of the target being the target itself)	H&B

Table 12: Antecedent Features

Carnegie Mellon University

VPE Features

Semantic	Whether the subject of the antecedent and the target are coreferent	H
Other	Whether the lemma of the head of the antecedent is *be* and that of the target is *do* (be-do match, used by Hardt and Nielsen)	H
	Whether the antecedent is in quotes and the target is not, or vice versa	H&B

Table 13: Antecedent Features

Carnegie Mellon University

VPE Results: Target Detection

	WSJ			BNC		
	Prec	Rec	F1	Prec	Rec	F1
Ora^{T}	100.00	93.28	96.52	100.00	92.65	96.18
Log^{T}	80.22	61.34	69.52	80.90	70.59	75.39
Pos^{T}	42.62	43.7	43.15	35.47	35.29	35.38
Log^{T+H}	23.36	26.89	25.00	12.52	38.24	18.86
Rank^{T+H}	0.00	0.00	0.00	15.79	5.88	8.57
Log^{T+H+B}	25.61	17.65	20.90	21.50	32.35	25.83
Rank^{T+H+B}	0.00	0.00	0.00	16.67	11.27	13.45
Nielsen^{T}	—	—	—	72.50	72.86	72.68

Table 14: Results for Target Detection on 2 datasets

VPE Results: End-to-End

	WSJ			BNC		
	Prec	Rec	F1	Prec	Rec	F1
Ora^T+Ora^H+Ora^B	95.06	88.67	91.76	85.79	79.49	82.52
Log^T+Rank^H+Rank^B	52.68	40.28	45.65	43.03	37.54	40.10
Log^T+Rank^H+Log^B	52.82	40.40	45.78	40.21	35.08	37.47
Log^T+Log^H+Rank^B	49.45	37.82	42.86	33.12	28.90	30.86
Log^T+Log^H+Log^B	49.41	37.79	42.83	31.32	27.33	29.19
Pos^T+Prev^H+Max^B	19.04	19.52	19.27	12.81	12.75	12.78
Log^T+Rank^{H+B}	54.82	41.92	47.51	41.86	36.52	39.01
Log^T+Log^{H+B}	38.85	29.71	33.67	26.11	22.78	24.33

Table 15: End-to-End results for VPE (Ora represents the Oracle, Log is the Logistic Regression Model, Rank is the ranking model. The superscripts indicate Target Detection, Head Selection and Boundary Detection respectively.)

Carnegie Mellon University

Sequencing Features: Surface Compatibility

These features capture whether two mentions are script compatible based on the surface information:

- Mention headword pair.
- Event type pair.
- Whether two event mentions appear in the same cluster in Chambers' event schema database.
- Whether the two event mentions share arguments, and the semantic frame name of the shared argument.

Sequencing Features: Discourse Compatibility

These features capture whether two event mentions are related given the discourse context:

- Dependency path between the two mentions.
- Function words (words other than Noun, Verb, Adjective and Adverb) in between the two mentions.
- The types of other event mentions between the two mentions.
- The sentence distance of two event mentions.
- Whether there are temporal expressions in the sentences of the two mentions, extracted from the AGM-TMP slot using a PropBank parser Tratz and Hovy (2011)

Sequencing Features: Event Ordering

This feature set tries to capture the ordering of events.

- Discourse ordering.
- Ordering from a temporal system.

The final feature set is the dot product of these 3 feature sets.

Sequencing Results

The Event Sequencing task is more difficult than what we expect. We compare our system with a oracle informed baseline.

	Prec.	Recall	F-Score
Oracle Cluster+Temporal	**46.21**	8.72	14.68
Our Model	18.28	**16.91**	**17.57**

Table 16: Test results for event sequencing. The Oracle Cluster+Temporal system is running CAEVO on the Oracle Clusters.

Event Salience Model

We compute multi-level interactions between events and entities via kernels: en_i is the ith entity and ev_i is the ith event in document d. $\mathbb{V}$ and $\mathbb{E}$ are the set of events and entities respectively.

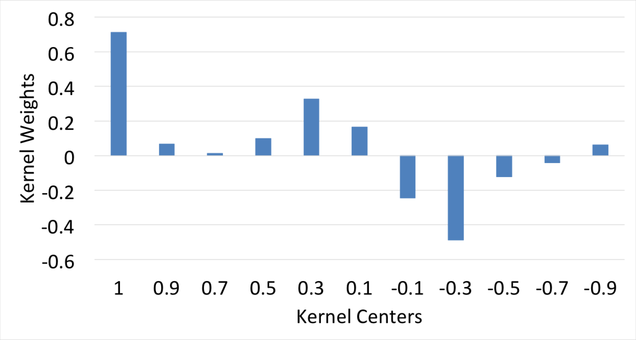
$$\Phi_K(ev_i, \mathbb{V}) = \{\phi_1(\overrightarrow{ev_i}, \mathbb{V}), \ldots, \phi_k(\overrightarrow{ev_i}, \mathbb{V}), \ldots, \phi_K(\overrightarrow{ev_i}, \mathbb{V})\}, \tag{1}$$

$$\phi_k(\overrightarrow{ev_i}, \mathbb{V}) = \sum_{ev_j \in \mathbb{V}} \exp\left(-\frac{\left(\cos(\overrightarrow{ev_i}, \overrightarrow{ev_j}) - \mu_k\right)^2}{2\sigma_k^2}\right). \tag{2}$$

$$\Phi_K(ev_i, \mathbb{E}) = \{\phi_1(\overrightarrow{ev_i}, \mathbb{E}), \ldots, \phi_k(\overrightarrow{ev_i}, \mathbb{E}), \ldots, \phi_K(\overrightarrow{ev_i}, \mathbb{E})\}, \tag{3}$$

$$\phi_k(\overrightarrow{ev_i}, \mathbb{E}) = \sum_{en_j \in \mathbb{E}} \exp\left(-\frac{\left(\cos(\overrightarrow{ev_i}, \overrightarrow{en_j}) - \mu_k\right)^2}{2\sigma_k^2}\right) \tag{4}$$

Learned Kernel Weights



Carnegie Mellon University

Learned Event Entity Relations

		Word2Vec	Kernel
attack	kill	0.69	0.3
arrest	charge	0.53	0.3
USA (**E**)	war	0.46	0.3
911 attack (**E**)	attack	0.72	0.3
attack	trade	0.42	0.9
hotel (**E**)	travel	0.49	0.9
charge	murder	0.49	0.7
business(**E**)	increase	0.43	0.7
attack	walk	0.44	-0.3
people (**E**)	work	0.40	-0.3

Table 17: Examples of pairs of Events/Entities in the kernels. Items marked with (**E**) are entities.