Event Coreference Resolution: A Survey of Two Decades of Research

Jing Lu and Vincent Ng
Human Language Technology Research Institute
University of Texas at Dallas

IJCAI-ECAI 2018

Event Coreference Resolution

Determine event mentions that refer to the same real-world event

Nelson Mandela has died at age 95.
The world has lost a great man, former Prime Minister Julia Gillard said.

Event Coreference Resolution

Determine event mentions that refer to the same real-world event

Nelson Mandela has died(ev1) at age 95.
The world has lost(ev2) a great man, former Prime Minister Julia Gillard said(ev3).

a word or phrase that triggers an event mention

	Trigger Word	Event Type	Arguments	Entity Coreference
ev1	died			
ev2	lost			
ev3	said			

Event Coreference Resolution

Determine event mentions that refer to the same real-world event

Nelson Mandela has died(ev1) at age 95.
The world has lost(ev2) a great man, former Prime Minister Julia Gillard said(ev3).

Coreferent event mentions have the same event type

	Trigger Word	Event Type	Arguments	Entity Coreference
ev1	died	Life.Die		
ev2	lost	Life.Die		
ev3	said	Contact. Broadcast		

Event Coreference Resolution

Determine event mentions that refer to the same real-world event

Nelson Mandela has died(ev1) at age 95.
The world has lost(ev2) a great man, former Prime Minister Julia Gillard said(ev3).

Coreferent event mentions have compatible argum

	Trigger Word	Event Type	Arguments	Entity Coreference
ev1	died	Life.Die	Nelson Mandela (Victim)	
ev2	lost	Life.Die	a great man (Victim)	
ev3	said	Contact. Broadcast	Prime Minister Julia Gillard (Person)	

Event Coreference Resolution

Determine event mentions that refer to the same real-world event

Nelson Mandela has died(ev1) at age 95.
The world has lost(ev2) a great man, former Prime Minister Julia Gillard said(ev3).

To determine the compatibility of event arguments

	Trigger Word	Event Type	Arguments	Entity Coreference
ev1	died	Life.Die	Nelson Mandela (Victim)	(Nelson Mandela, a great man)
ev2	lost	Life.Die	a great man (Victim)	
ev3	said	Contact. Broadcast	Prime Minister Julia Gillard (Person)	

Event Coreference Resolution

Determine event mentions that refer to the same real-world event

Nelson Mandela has died(ev1) at age 95.
The world has lost(ev2) a great man, former Prime Minister Julia Gillard said(ev3).

	Trigger Word	Event Type	Arguments	Entity Coreference
ev1	died	Life.Die	Nelson Mandela (Victim)	(Nelson Mandela, a great man)
ev2	lost	Life.Die	a great man (Victim)	
ev3	said	Contact. Broadcast	Prime Minister Julia Gillard (Person)	

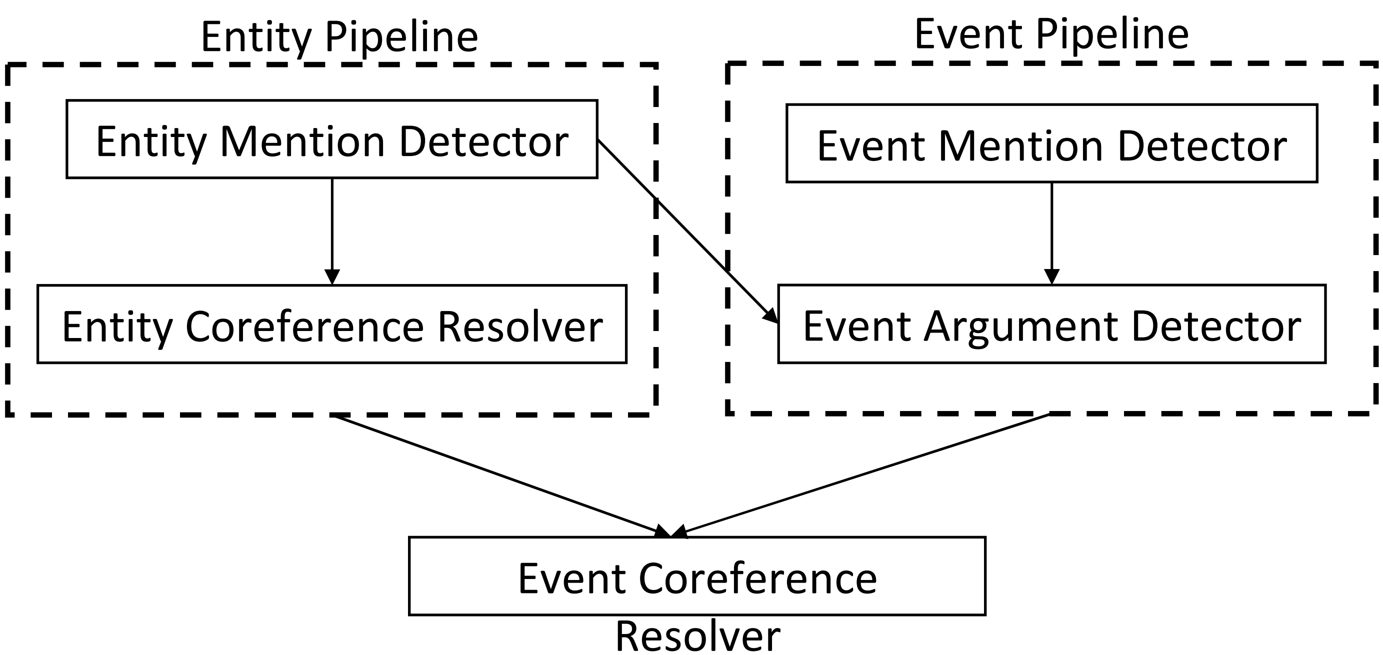
Event Coreference Resolution

More challenging than entity coreference resolution

Event Coreference Resolution

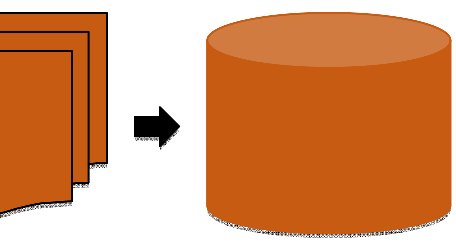
More challenging than entity coreference resolution

– Using the noisy outputs of a larger set of upstream components



Event Coreference Resolution

- Core task in information extraction from text
  - Consolidate the textual information about an event
  - Crucial for high-level NLP applications
    - E.g,. template filling, automated population of knowledge bases, topic detection and trackin question answering, summarization, contradiction detection



ents

Knowledge Base

A separatist group claimed responsibility for an ***explosion (ev1)*** late on Monday which wounded six people in an Istanbul supermarket. Istanbul governor told news agency the ***explosion (ev2)*** in the Bahcelievler district injured six people.

Bombing Template
Perpetrator: A separatist group
Target: six people
Time: Monday
Location: an Istanbul supermarket, Bahcelievler district

Plan for the talk

- Corpora
- Models
- Features
- Evaluations and State-of-the-Art
- Challenges

Plan for the talk

Corpora

Models

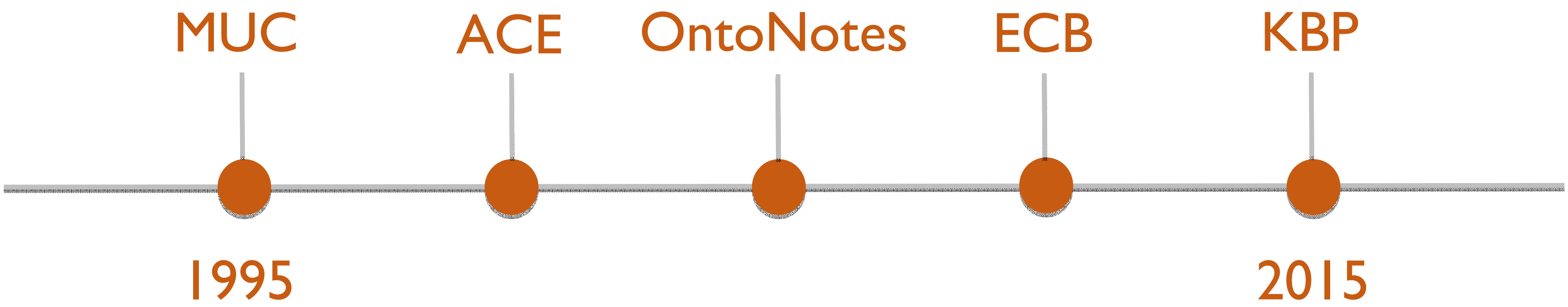
Features

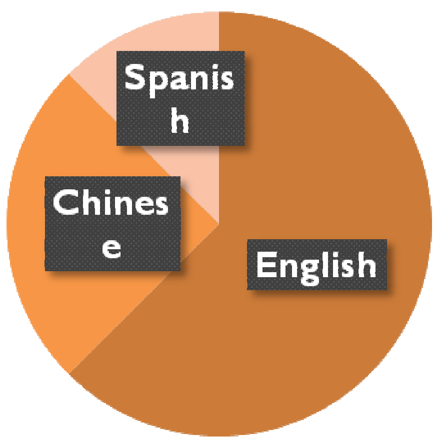
Evaluations and State-of-the-Art

Challenges

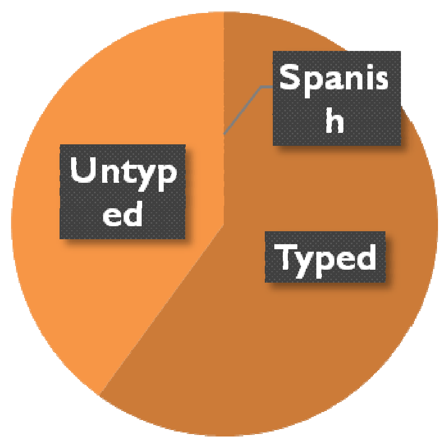
Corpora

5 publicly available corpora

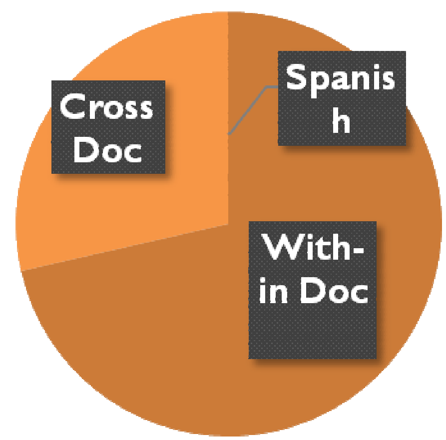




Language



Event Annotation



Coreference Chain Type

Plan for the talk

Corpora

Models

Features

Evaluations and State-of-the-Art

Challenges

Mention-Pair Models

- A two-step resolution framework
  - Step1: use a binary classifier to determine whether two event mentions are coreferent.
    - Pairwise classification decisions could violate transitivity

✓

Microsoft bought Nokia after the acquisition of Skype. It spent $7.2 billion

✓ ✗

Mention-Pair Models

A two-step resolution framework

– Step1: use a binary classifier to determine whether two event mentions are coreferent.

  • Pairwise classification decisions could violate transitivity

Microsoft bought Nokia after the acquisition of Skype. It spent $7.2 billion

– Step2: use clustering to coordinate the pairwise decisions and construct a partition.

  • Agglomerative clustering, graph partitioning

Weaknesses of the Mention-Pair Models

Two-step approach suffers from error propagation

– Errors made by a mention-pair model can propagate to the clustering step

Improvement: Generative Model

– Yang et al. (2015)

– Idea: combine distance-based and distribution-based methods to perform guide
Bayesian clustering

- A clustering model
- A Bayesian model: encode the knowledge provided by a mention-pair model as prior
- Employ rich features in the modeling process

Weaknesses of the Mention-Pair Models

- Can't determine which candidate antecedent is the best
  - only determine how good a candidate antecedent is relative to the event mention, but not how good it is relative to other candidate antecedents

Weaknesses of the Mention-Pair Models

Can't determine which candidate antecedent is the best

– only determine how good a candidate antecedent is relative to the event mention, but not how good it is relative to other candidate antecedents

Solution: Mention-Ranking Models

– allowing candidate antecedents of a mention to be ranked simultaneously so th its correct antecedent has the highest rank.

New	died New	lost died New
Mandela has died…	The world has lost…	His wife passed away...

Easy-First Models

- The aforementioned models resolve mentions from left to right
- Easy-first models operate in an iterative fashion
  – make easy linking decisions first
  – subsequently exploit these easy decisions (as additional knowledge) to make hard linking decisions.

Easy-First Models

lelson Mandela has died at age 95 at his Houghton residence.
he world has lost a great man.
ater President Jacob Zuma confirmed Mandela, one of the greatest man in South frica, passed away at his home.

died	Nelson Mandela(victim), his Houghton residence (location)
lost	a great man (victim)
passed away	Mandela, one of the greatest man in South Africa (victim), his home (location)

lost

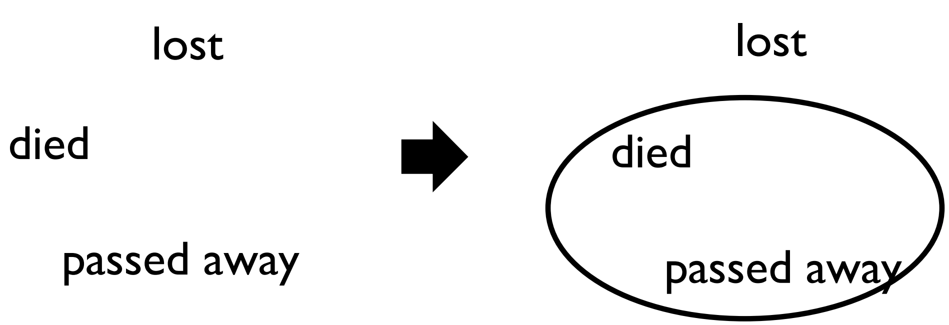
died

passed away

Easy-First Models

elson Mandela has died at age 95 at his Houghton residence.
he world has lost a great man.
ater President Jacob Zuma confirmed Mandela, one of the greatest man in South
frica, passed away at his home.

died passed away	Nelson Mandela/Mandela, one of the greatest man in South Africa (victim), his Houghton residence / his home (location)
lost	a great man (victim)



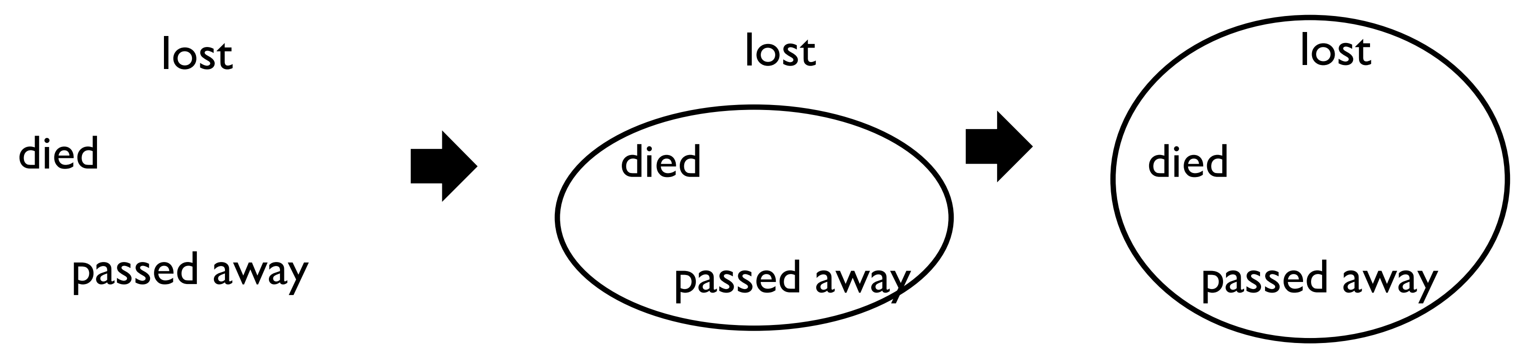
Argument propagation in each iteration

Easy-First Models

lelson Mandela has died at age 95 at his Houghton residence.
he world has lost a great man.
ater President Jacob Zuma confirmed Mandela, one of the greatest man in South frica, passed away at his home.

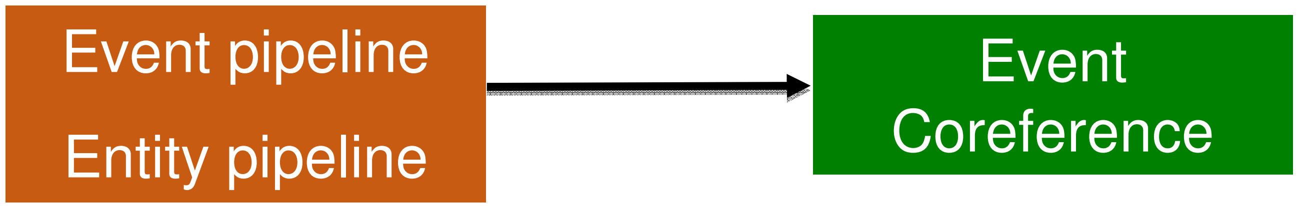
died passed away lost	Nelson Mandela/Mandela, one of the greatest man in South Africa/a great man (victim) his Houghton residence / his home (location)

Could bootstrap entity coreference output (Lee et. al 2012)



Joint Models

- The aforementioned models all adopt a pipeline architecture
  - Error propagation



Joint Models

The aforementioned models all adopt a pipeline architecture

– Error propagation

Event pipeline

Entity pipeline

→

Event Coreference

– Not able to employ the interdependence between event detection and event coreference resolution

Life.Die

Nelson Mandela has died at age 95

↕

The world has lost a great man…

Life.Die ~~Personnel.End-Position~~ ~~Life.Injury~~ ~~NULL~~

Joint Models

- Joint inference over the outputs of different tasks in the IE pipeline
  - Integer Linear Programming (ILP)
    - encode hard constraints, e.g. two triggers that do not have the same event subtype cannot coreferent.
  - Markov Logic Networks (MLNs)
    - encode both soft and hard formulas
- Joint learning over different tasks
  - E.g., jointly learn event extraction and coreference
  - Recast as a structured prediction problem
    - Use the structured perceptron training algorithm (Araki and Mitamura, 2015)
    - Use a structured conditional random field model (Lu and Ng, 2017)

Semi-Supervised Models

Data acquisition bottleneck

– manually annotating data for all the components in the IE pipeline is expensive

Solution

– Use active learning to select informative instances (Chen and Ng, 2016)

– Utilize large amounts of out of domain text data (Peng et. al, 2016)

- represent event structures by five event semantic components
- Convert each event component to its corresponding vector representation

$$\underbrace{[\ \ldots\ldots\ ]}_{\textbf{event}} = \Big[\ \underbrace{[\ \ldots\ldots\ ]}_{\textbf{action}}\ \underbrace{[\ \ldots\ldots\ ]}_{\textbf{agent}_{\textbf{sub}}}\ \underbrace{[\ \ldots\ldots\ ]}_{\textbf{agent}_{\textbf{obj}}}\ \underbrace{[\ \ldots\ldots\ ]}_{\textbf{location}}\ \underbrace{[\ \ldots\ldots\ ]}_{\textbf{time}}\ \underbrace{[\ \ldots\ldots\ ]}_{\textbf{sentence or clause}}\ \Big]$$

Unsupervised Models

- Eliminate a model's reliance on annotated data.
- Probabilistic generative models
  – Use a set of Dirichlet Processes (DPs), in which each DP is associated with eac document, and each mixture component is an event coreference cluster share across documents (Bejan and Harabagiu, 2014)
  – Use a latent variable to represent coreference decisions (Chen and Ng, 2015)

Plan for the talk

- Corpora
- Models
- Features
- Evaluations and State-of-the-Art
- Challenges

Features

Lexical features

- Explicitly or implicitly compare the event triggers of a pair of event mentions
- Commonly used features:
  - Pair features: Trigger pairs. POS pairs of triggers
  - String-matching features
  - Trigger similarity features
- String-matching features have been shown to contribute significantly

Features

Argument features

- Event mentions having incompatible arguments are unlikely to be coreferent
- Commonly used features:
  - Number of overlapping and unique arguments
  - Argument similarity: whether entities are coreferent, similarity metrics
- An event argument extractor and an entity coreference resolver are typically needed
- Improving existing argument extractors and entity coreference resolvers can further improve event coreference resolver

Features

Semantic features

- Compute the similarity between two event mentions from lexical semantic resources (e.g., WordNet), Brown clusters, and the word embeddings
- Experiments show that the embedding-based similarity feature has the highest weight among all features, hence suggesting its usefulness
- Event (sub)type match has also been shown to be a strong indicator
  - died (Life.die) vs. lost (Personnel.End-Position)

Features

Discourse features

- Encode the token, event and sentence distance between two event mentions
- Experiment shows discourse features do not contribute to event coreference performance as much as other types of features

Plan for the talk

Corpora

Models

Features

Evaluations and State-of-the-Art

Challenges

Evaluation

End-to-end event coreference resolver

– event coreference performance is significantly affected by event mention (i.e., trigger) detection performance

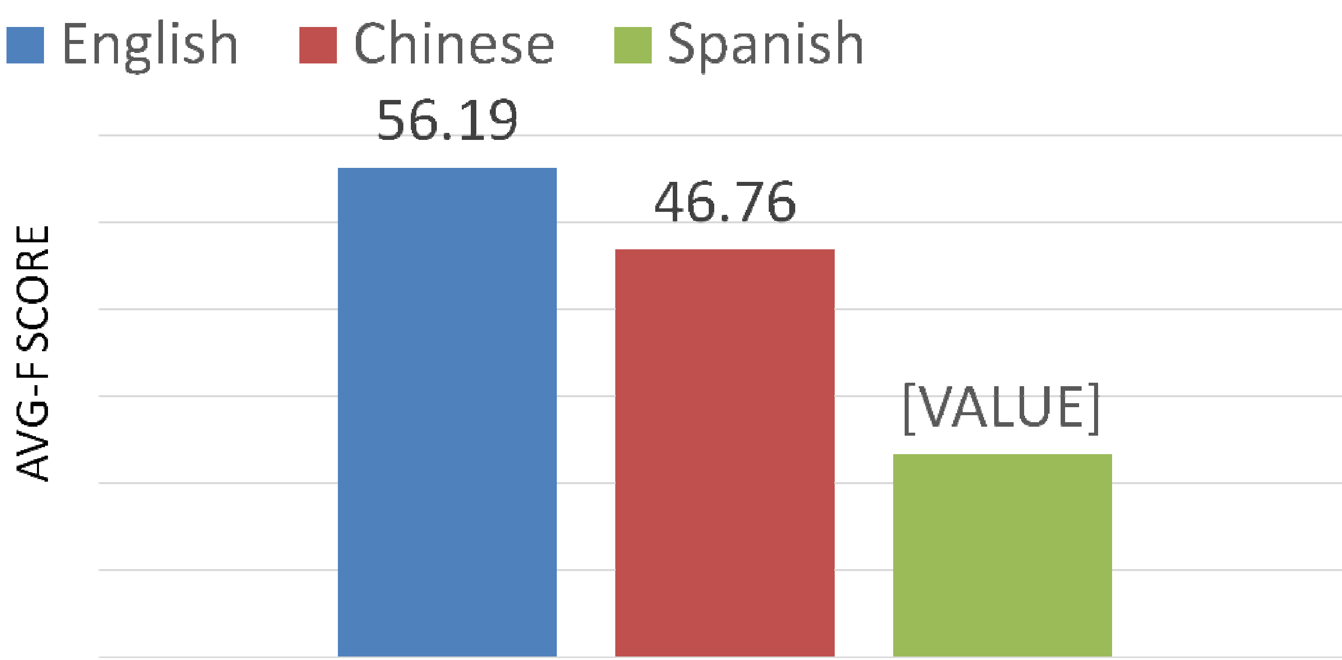
Evaluation Metrics

– Since researchers cannot agree on which evaluation metric is the best to use, multiple metrics are typically used
  - Link-based MUC metric (Vilain et al., 1995)
  - Mention based B^3 metric (Bagga and Baldwin, 1998)
  - Entity based $CEAF_e$ metric (Luo, 2005)
  - Rand index-based BLANC metric (Recasens and Hovy, 2011)
  - CoNLL score and AVG score

Results

The best results to date on multi-lingual KBP 2017 dataset

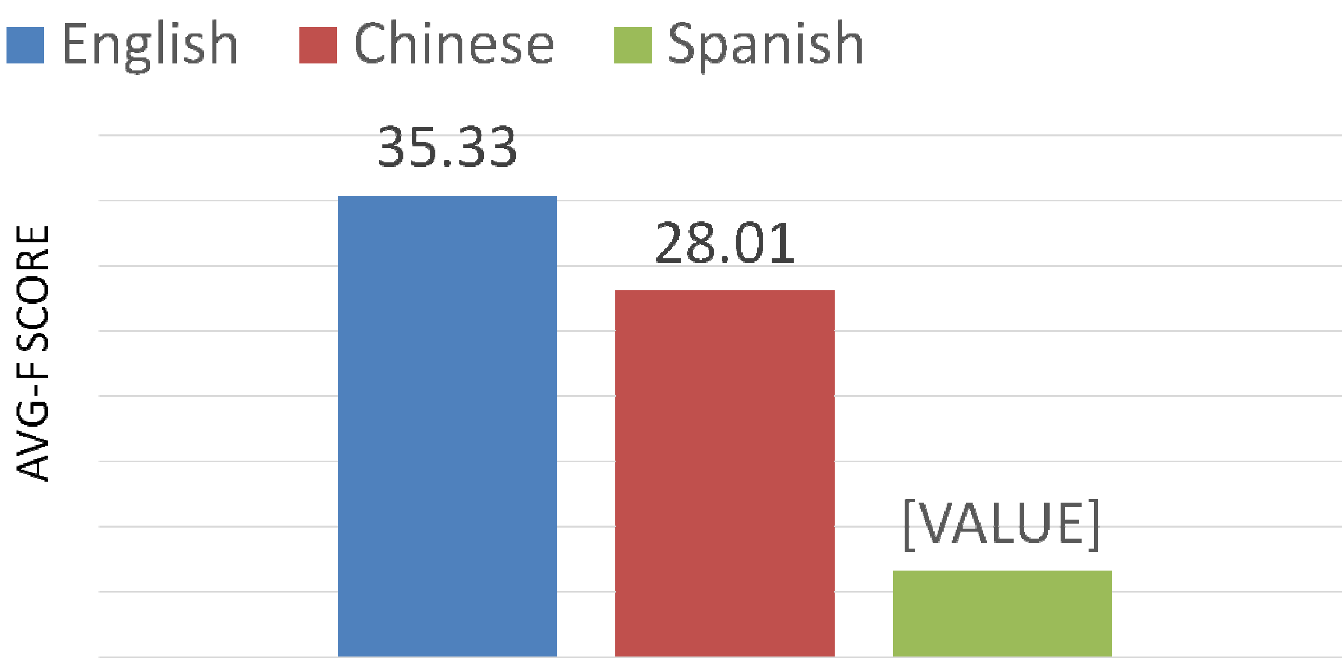
Event Trigger Detection result



Results

The best results to date on multi-lingual KBP 2017 dataset

Event Coreference result



Plan for the talk

Corpora

Models

Features

Evaluations and State-of-the-Art

Challenges

Challenges

- It may be worthwhile to investigate joint models further.
  - Previous work has applied joint inference to the four key tasks in IE
  - Can we jointly learn these four tasks?

Challenges

- It may be worthwhile to investigate joint models further.
  - Previous work has applied joint inference to the four key tasks in IE
  - Can we jointly learn these four tasks?

Challenges

- We may need to employ sophisticated features
  – Given recent successes on employing word vectors for event coreference resolution, can we learn representations from complex features?
    - features that are derived from computed arguments and entity coreference chains.

Challenges

Low-resource languages

– How can we learn models if coreference-annotated corpora may not be availal
?

– How can we obtain semantic knowledge if large lexical knowledge bases do no
exist?