

TAC KBP Event Detection and Coreference Tasks for English

Version 1.0, April 27, 2015

Track Coordinators:

Carnegie Mellon University

Teruko Mitamura teruko@cs.cmu.edu

Eduard Hovy ehovy@andrew.cmu.edu

OVERVIEW

The Event Detection and Coreference tasks at NIST TAC KBP aims to identify the explicit mentioning of Events in text. Every instance of a mention of the relevant Event types must be identified.

The Event Detection task is required to detect the Event Types and Subtypes taken from **the Rich ERE Annotation Guidelines: Events** (current version is v2.5.1). Also, the task is to identify three REALIS {ACTUAL, GENERIC, OTHER}, which are described in the Rich ERE guidelines. The data sources are provided by LDC. About 150 or more annotated corpora will be provided prior to the evaluation as a training set. For the formal evaluation, about 200 corpora will be given to the participants. We plan to include newswire articles and discussion forums.

The Event Detection Task requires participants to identify all relevant Events Mention instances within each sentence. If the same Event is mentioned in several places in the document, the participants will list them all.

The Event Coreference Task requires participants to identify the coreference links of Events Mention instances in a document.

The eventual benefit of the Event Detection and Coreference will be to detect subevent structures for the future task.

GOALS

There are three subtasks in this evaluation.

1. **Event Nugget Detection**
2. **Event Nugget Detection and Coreference**
3. **Event Nugget Coreference**

EN Task 1: Event Nugget Detection: This task aims to identify the explicit mentioning of Events in text for English. Participating systems must identify all relevant Event Mention instances within each sentence. Every instance of a mention of the relevant Event types/subtypes taken from the Rich ERE Annotation Guidelines must be identified. In addition, systems must identify three REALIS values (ACTUAL, GENERIC, OTHER), which are also described in the Rich ERE guidelines and TAC KBP Event Detection Annotation Guidelines v1.7.

EN Task 2: Event Nugget Detection and Coreference: In addition to the Event Nugget Detection task described in the EN Task 1, this task also aims to identify Full Event Coreference links at the same time. Full Event Coreference is identified when two or more Event Nuggets refer to the same event. This notion is described as Event Hoppers in the Rich ERE Annotation Guidelines. The Full Event Coreference links do not include subevents relations.

EN Task 3: Event Nugget Coreference: This task is to identify Full Event Coreference links, given the annotated Event Nuggets in the text.

TASKS

EN Task 1: Event Nugget Detection

The input of this task will be unannotated documents. The output will be Event Nugget Identification, Event Types and Subtypes, and REALIS information.

1. Event Type and Subtypes (listed below)
2. REALIS Value (one of: ACTUAL, GENERIC, OTHER)
3. Event Nugget Identification (offset and character length)

Event Types and Subtypes:

For purposes of this evaluation, an event must fall under one of the event types and subtypes below. For more details, see the **Rich ERE Annotation Guidelines: Events**.

Type	Subtype	Type	Subtype
Business	Start Org	Movement	Transport.Person
Business	End Org	Movement	Transport.Artifact
Business	Declare Bankruptcy	Personnel	Start Position
Business	Merge Org	Personnel	End Position
Conflict	Attack	Personnel	Nominate
Conflict	Demonstrate	Personnel	Elect
Contact	Meet	Justice	Arrest-Jail
Contact	Correspondence	Justice	Release-Parole
Contact	Broadcast	Justice	Trial-Hearing

Contact	Contact	Justice	Sentence
Manufacture	Artifact	Justice	Fine
Life	Be Born	Justice	Charge-Indict
Life	Marry	Justice	Sue
Life	Divorce	Justice	Extradite
Life	Injure	Justice	Acquit
Life	Die	Justice	Convict
Transaction	Transfer Ownership	Justice	Appeal
Transaction	Transfer Money	Justice	Execute
Transaction	Transaction	Justice	Pardon

REALIS Identification

Event mentions will refer to **ACTUAL** (events that actually occurred); **GENERIC** (events that are not specific events with a (known or unknown) time and/or place); or **OTHER** (which includes failed events, future events, and conditional statements, and all other non-generic variations). For more detail explanation, see the **Rich ERE Annotation Guidelines** and TAC KBP Event Detection Annotation Guidelines v1.7.

Event Nugget Identification

A system will choose to identify Event Nuggets in the text. The definition of the Event Nugget generally follows the **Rich ERE Annotation Guidelines**. Each Nugget is the actual string of words that indicate the mentioned event, and must correspond to the Event type and subtype above. When a sentence mentions more than one event type both must be mentioned, e.g., in the example sentence “he **shot** the soldier **dead**,” both [conflict.ATTACK] and [life.DIE] are events. For more details, see the **Rich ERE Annotation Guidelines**.

EN Task 2: Event Nugget Detection and Coreference

The input of this task will be unannotated documents. The output will be Event Nugget Identification, Event Types and Subtypes, REALIS information, described in EN Task 1 plus Event Coreference relations.

EN Task 3: Event Nugget Coreference

The input of this task will be the documents in which Event Nuggets are annotated in the text. The output will be Event Coreference relations for these annotated Event Nuggets.

System Outputs

1. Submission Format (for all tasks)

For each nugget detected, the system must output one line in a text file:
If the system chooses not to provide the confidence scores, the last two fields are empty. Details of evaluation file formats are described in the **Event Nugget Detection and Coreference Scoring** document.

System submissions should follow the guideline strictly in order to ensure correct evaluation. A separated validator script will be provided to help diagnostic submission format. However, participants should try to make sure the formats to be correct. If you have any doubts about the formats, please contact the task coordinators.

Each nugget will be represented with one tab-separated line, using the following formats:

- system-ID: unique ID assigned to each system run
- doc-ID: unique ID assigned to each source document
- mention ID: unique ID of the event nugget assigned by the system
- token ID list: list of IDs for the token(s) of the current mention (according to a token ID table released to participants)
- mention-string: character string of event mention (see **Event Detection Annotation Guidelines**)
- event-type: *type.subtype* from the hierarchy given above
- Realis-value: one of *ACTUAL, GENERIC, OTHER*
- Confidence scores of event span: score between 0 and 1 inclusive (optional)
- Confidence scores of event type: score between 0 and 1 inclusive (optional)
- Confidence scores of Realis-value: score between 0 and 1 inclusive (optional)

Coreference decisions should be attached after listing all nuggets in a document. Note that even the system is participating Event Nugget Coreference task only, the system should also first copy all provided event nuggets, and then append the coreference lines.

Each coreference cluster should also be represented in one tab-separated line, using the following columns:

- Relation name: this should always be @Coreference
- Relation id: This is for bookkeeping purposes, which will not be read by the scorer. The relation id used in the gold standard files will be in form of "R<id>" (e.g. R3)

- Mentions Id List: list of event mentions in this coreference cluster, separated by comma (,). In terms of coreference, the ordering of event mentions does not matter.

Special header and footers are used to mark boundaries of documents.

Corpus

The corpus for this task will consist of 200 documents from two different types of documents: newswire and discussion forum documents. About half will be taken from each genre. The documents are in XML format.

Scoring

For event nugget detection, systems will be scored on the F-1 score of precision and recall over the gold standard. For event nugget coreference, systems will be scored using the evaluation metrics used in CoNLL shared tasks¹.

The scorer reads the output of event mention detection systems and compares them to the gold standard.

Input:

Gold standard annotation for a text, in evaluation file format (tbf)

System output annotation for the same text, in evaluation file format (tbf)

Output:

System score report for event nugget detection and coreference

The detail scoring definition is described in the separate document: **Event Mention Detection Scoring and Event Coreference Scoring**.

Submissions and Schedule

Systems will have at least one week to process the evaluation documents.

Submissions should be fully automatic and no changes should be made to the system once evaluation corpus has been downloaded. Up to **three** alternate system runs for each task may be submitted per-team. Submitted runs should be ranked according to their expected overall score.

The deadline for the system reports for the final Proceedings version is in February, 2016.

¹ <http://conll.cemantix.org/2012/>

End of May, 2015	Release Event Nuggets and Coreference training data
August 31- Sept 11, 2015	EN Task 1: Event Nugget Detection evaluation window
August 31- Sept 11, 2015	EN Task 2: Event Nugget Detection and Coreference evaluation window
Sept 14 - 21, 2015	EN Task 3: Event Nugget Coreference evaluation window
October, 2015	Release of individual evaluated results to participants
October 15, 2015	Deadline for short system descriptions
October 15, 2015	Deadline for workshop presentation proposals
October 21, 2015	Notification of acceptance of presentation proposals
November 1, 2015	Deadline for system reports (workshop notebook version)
November 16-17, 2015	TAC 2015 workshop in Gaithersburg, Maryland, USA

References

DEFT Rich ERE Annotation Guidelines: Events v.2.5.1, Linguistic Data Consortium, February 24, 2015

Event Mention Detection Scoring document, Language Technologies Institute, Carnegie Mellon University

TAC KBP Event Detection Annotation Guidelines v1.7, Language Technologies Institute, Carnegie Mellon University