Building a Cross-document Event-Event Relation Corpus

Yu Hong[1,2]**, Tongtao Zhang**[1]**, Tim O'Gorman**[3]**, Sharone Horowit-Hendler**[4]**,**
Heng Ji[1]**, Martha Palmer**[3]
[1]Computer Science Department, Rensselaer Polytechnic Institute
[2]Department of Computer Science, Soochow University
[3]Department of Linguistics, University of Colorado at Boulder
[4]Institute of Reading Development

Abstract

We propose a new task of extracting event-event relations across documents. We present our efforts at designing an annotation schema and building a corpus for this task. Our schema includes five main types of relations: Inheritance, Expansion, Contingency, Comparison and Temporality, along with 21 subtypes. We also lay out the main challenges based on detailed inter-annotator disagreement and error analysis. We hope these resources can serve as a benchmark to encourage research on this new problem.

1 Introduction

The ultimate goal of Information Extraction (IE) is to construct "Information Networks" (Li et al., 2014) from unstructured texts. Most previous IE work focused on constructing entity-centric Information Networks where each node represents an entity and each edge represents a relation. We propose a novel task to construct a new layer of *event-centric Information Networks* across multiple documents, where each node is an event and the edges capture the relations between two events. This task can provide building blocks for many important applications such as event knowledge base population and temporal event tracking (Do et al., 2012). The nodes can be extracted by existing fine-grained event extraction approaches (Ji and Grishman, 2008; Liao and Grishman, 2010; Hong et al., 2011; Li et al., 2013; Li et al., 2014). However, little previous work can be directly exploited to construct the edges.

In this paper we define a comprehensive schema that includes multiple fine-grained event-event relation types. Some types are similar to those in discourse parsing (Soricut and Marcu, 2003). However, event-event relations are fundamentally different from discourse relations: (1) The input consists of structured events instead of unstructured sentences. (2) For cross-document event pairs, there are neither explicit textual clues nor implicit information about the ordering of clauses that might indicate the relation. Following this schema, we annotated a cross-document event-event relation corpus built on top of the Automatic Content Extraction (ACE2005) [1] event annotations. We will define the task (Section 2), describe the annotation schema (Section 3) and present corpus statistics and annotation challenges (Section 4).

2 Task Definition

In an event-event relation schema, events form a crucial foundation because they serve as nodes and are indispensable in event-centric information networks. We follow the definition of events in the ACE guideline [2]:

Event trigger: the main word which most clearly expresses an event occurrence.

Event arguments: the entities, time expressions and values that are involved in an event.

Event mention: a phrase or sentence within which an event is described, including a trigger and a set of arguments.

Event: a set of coreferential event mentions within one document.

We define the event-event relation task as the annotation of all applicable logical relations between two events. For example, as illustrated in Figure 1, the following events are connected by *Condition* and *Temporality* relations:

Event 1: *Media tycoon Barry Diller on Wednesday quit as chief of Vivendi Universal Entertain-*

[1]http://projects.ldc.upenn.edu/ace

[2]https://www.ldc.upenn.edu/sites/www.ldc.upenn.edu/files/english-events-guidelines-v5.4.3.pdf

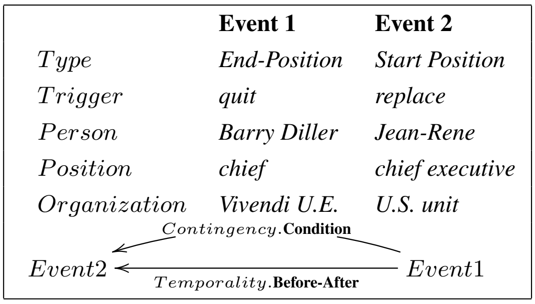


Figure 1: Examples of input and output

ment.

Event 2: *Parent company chairman Jean-Rene Fourtou will replace Diller as chief executive of US unit.*

This example reveals the fact that a successor takes the place only *after the time when (Temporality)* and *under the condition that (Condition)* the predecessor makes room for the successor.

3 Event-Event Relation Schema

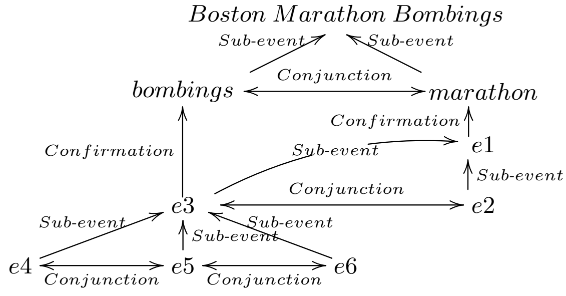
Our event-event relation schema includes 5 main ***Types*** – *Inheritance*, *Expansion*, *Contingency*, *Comparison* and *Temporality* – along with 21 ***Subtypes*** as shown in Table 1. Table 1 also demonstrates *Roles*. Events involved in a relation play certain roles. For example, an *Attack* event and an *Injure* event in a *Contingency.Causality* will play *Cause* and *Result* roles respectively. In the following we will present a detailed definition of each subtype.

3.1 Inheritance and Expansion

Inheritance relations include both traditional *Coreference* relations as well as *Subevent* that marks aggregation-to-component relations. *Reemergence* connects recurrent events while *Variation* summarizes the prototype of an event.

Expansion relations include *Confirmation*, which encodes a concept-to-instance or "subset" relation, and *Conjunction* and *Disjunction*, which relate two subevents within a larger event, and mark two subevents as playing similar (*Conjunction*) or dissimilar (*Disjunction*) roles within the larger event. This kind of relation is useful, since a larger event is often not explicitly mentioned.

The combination of these two kinds of relations allow one to build hierarchical representations of parts of an event network, as shown in Figure 2.



$e1$: $117th$ *annual Boston Marathon*;
$e2$: *winner crossed the finish line*;
$e3$: *explosion near bystanders*;
$e4$: 1_{st} *explosion*;
$e5$: 2_{nd} *explosion*;
$e6$: *the second bomb was placed at the finish line*

Figure 2: A hierarchical event network

3.2 Contingency and Comparison

A ***Contingency*** relation indicates either an event leading to the emergence (*Causality*) or serving as a triggering condition (*Conditional*) of another event.

Comparison relations indicate deeper logical contrasts between relations. *Opposition* indicates a relation in which two events are mutually contradictory, and unlikely to be both true. This has some similarity to *Contrast.Opposition* in the Penn Discourse Treebank ((Miltsakaki et al., 2004)) or specific annotations of opposition ((Feltracco et al., 2015; Takabatake et al., 2015)). *Negation* indicates that while two events could both be true, one shows that the other is no longer true. *Competition* shows that two events are contrasting versions of the same underlying "event" (e.g., *retreat* versus *escape in disorder*).

3.3 Temporality

Last but not least, we also define subtypes of *Temporality*, which represents the temporal order of events. *Temporality* has been an active research topic for a long time. We arrange all categories and normalize the subtype names from the previous work to constitute our *Temporality* schema. Figure 3 illustrates the temporal relation subtypes.

In this work, we elaborate the subtypes *Temporality* in comparison with conventional work by introducing *Meet*, *Start* and *Finish*, which emphasizes the existence of time intervals among events.

The correct subtype of the *Temporality* relation has a great impact on the decision of whether the Start-Position and End-Potion events have a *Comparison.Opposite* or a *Contingency.Condition* re-

Types	*Inheritance*	*Expansion*	*Contingency*	*Contingency*	*Temporality*	
Subtypes	**Reemergence**	**Confirmation**	**Comparison**	**Condition**	**Before-After**	**Start**
Roles	Reference	Generalization	Superior	Condition	Before	Included-Start
	Resurgence	Instantiation	Inferior	Emergence	After	Includes
Subtypes	**Sub-event**	**Conjunction**	**Concession**	**Causality**	**Vagueness**	**Overlap**
Roles	Constituent	Homology-1	Not-Achieved	Cause	Vague-1	Partially-Before
	Synthesis	Homology-2	Achieved	Result	Vague-2	Partially-After
Subtypes	**Variation**	**Disjunction**	**Negation**		**Meet**	**Equality**
Roles	Variance	Heterology-1	Negator		Before	Contemporary-1
	Semina	Heterology-2	Initiator		After	Contemporary-2
Subtypes	**Coreference**		**Opposite**		**During**	**Finish**
Roles	Occurrence		Opponent-1		Includes	Includes
	Paraphrasing		Opponent-2		Included-In	Included-Finish

Table 1: Fine-grained event-event relations and roles.

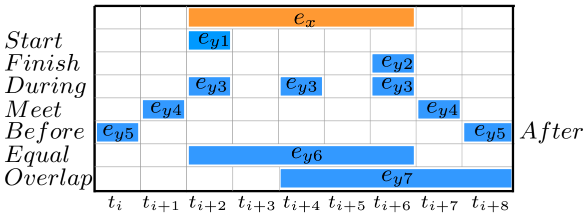


Figure 3: Various temporal relation subtypes between event e_x and event e_{y*} (t is a time interval).

lation, and vice versa.

4 Corpus Annotation

Annotating event-event relations requires an annotator to gain a global view of the overall scenario or topic (e.g., MH17) before exhaustively annotating each event pair. In addition, our relation types are more fine-grained than previous work such as the Richer Event Descriptions (RED) (Ikuta et al., 2014). There are no existing annotation tools to meet these needs, so we developed a new annotation tool to visualize trigger words, arguments and contexts for each event pair to ensure that annotators fully understand documents, background and storyline.

We created an event-event relation corpus based on gold standard events in ACE2005 newswire documents and some additional news documents about *Malaysian Airline 17* (denoted as *MH17*). Table 2 shows the detailed data statistics. Three annotators (*A1* & *A2*: graduate students, *A3*: a linguist) annotate the documents independently. The annotation results from *A1* and *A2* are assessed by *A3* and disagreement among the three annotators is carefully evaluated and *A3* determined the final results.

#Topics	**#Documents**	**#Events**	**#Pairs**
Vivendi	3	22	231
Anwar	3	39	741
SARS	2	9	36
MH17	50	196	19,698
Single	67	597	4,904
Total	125	863	25,610

Table 2: Corpus statistics. *Single* denotes topics that only contain one document, and cross-document annotation is not available in those topics. **Pair** indicates the number of pairs consisting of two events from the same topic.

Table 3 and 4 indicate that this is a very challenging task for annotators. We can see that the major challenge for annotators is the determination of the existence of relations. *Causality* and *Condition* stand as the most challenging types, which require annotators to figure out the storyline of documents and exploit background knowledge. For example, the following two events are from the same document but there are no explicit connectives to indicate the conditional relation between them:

Event 1: *Edward Snowden claimed he was trained as a secret agent.*

Event 2: *The certification would also have given him some of the skills he needed to escape scrutiny.*

A1 and A2 also tend to mistakenly label *Sub-*

Annotators	**κ Value**
A1 and A2	0.1558
A2 and A3	0.1987
A1 and A3	0.1628

Table 3: Stats of Cohen's kappa coefficient.

Transition of Correction	**# of Occurrences**
unrelated → *Condition*	142
unrelated → *Causality*	72
Coreference → *unrelated*	55
Conjunction → *unrelated*	48
Causality → *unrelated*	44

Table 4: Top 5 Error corrections.

event as *Coreference*. Such mistakes happen when the arguments from one event appear as more specific and detailed entities (e.g., *an attack in Baghdad* vs. *an attack in Iraq*). However, when the event network becomes larger and more complicated, errors can be propagated across types, e.g., incorrectly labeled *Sub-event* pairs will also trigger *Conjunction* errors.

Moreover, we have attempted to align the inventory here with other ongoing efforts to annotate within-document event-event relations. Table 5 shows a mapping between a subset of the relations proposed here and those used in the Richer Event Descriptions (RED) (Ikuta et al., 2014). Other similar resources – such as Penn Discourse Treebank (Miltsakaki et al., 2004) – could also be used.

5 Related Work

The proposed schema covers event-event relation types that have been widely studied: (Styler IV et al., 2014; Bethard, 2013; Allen, 1983; Miller et al., 2013; Pustejovsky and Stubbs, 2011; Pustejovsky et al., 2005; UzZaman et al., 2013) also focused on the relation types which are related to *Temporality*. Methods about extracting *Coreference* relation have also been discussed and proposed in (Chen and Ji, 2009; Chen et al., 2009; Bejan and Harabagiu, 2010; Lee et al., 2012; Zhang et al., 2015). (Do et al., 2011; Riaz and Girju, 2013; Mirza and Tonelli, 2014) work on *Causality* relation.

Similar event-event relation schema such as

This work		**RED**
Inheritance	*Subevent*	***Contains-subevent***
	Coreference	***Identity***
Continency	*Cause*	***Cause***
	Condition	***Precondition***
Comparison	*Opposite*	N/A
	Concession	N/A
Expansion	*Confirmation*	***Set/Member***
Temporality	*Before, After*	***Before***
	During	***Contains***
	Overlap	***Overlap***
	Equality	***Simultaneous***
	Start	***Begins-on***
	Finish	***Ends-on***

Table 5: Mappings to RED (Ikuta et al., 2014)

RED (Ikuta et al., 2014) is in general more coarse-grained and has fewer types and subtypes.

Event-event relations differ from textual entailment (Dagan et al., 2013) or discourse relations (Soricut and Marcu, 2003; Miltsakaki et al., 2004; Radev, 2000), which focus on the relatedness between two sentences, by tackling a full document or multiple documents. We adopted some terminology (e.g., *Causality* and *Expansion*) from the taxonomy of discourse relations (Soricut and Marcu, 2003). We focus on a wider scope of cross-document events with richer and more fine-grained structured event representations.

If we consider each event-event relation instance as a frame (e.g., a contingency/causality event-event relation is similar to the frame causation), the architecture of the Event Networks is also similar to FrameNet (Baker and Sato, 2003) and thus the ontological analysis and constraints in (Ovchinnikova et al., 2010) are also applicable to our task.

6 Conclusions and Future Work

Our work will expand the research venue of IE from entity-centric to event-centric. In the future we will further expand the corpus[3], and compare and integrate with other within-document event-event relation schemas such as RED. We also plan to develop a pilot system using these resources.

[3]The annotated corpus is available at `http://nlp.cs.rpi.edu/data/event_relation.zip`

Acknowledgements

This work was supported by the U.S. ARL NS-CTA No. W911NF-09-2-0053, DARPA DEFT No. FA8750-13-2-0041 and NSF IIS-1523198. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation here on.

References

James F Allen. 1983. Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11):832–843.

Collin F. Baker and Hiroaki Sato. 2003. The framenet data and software. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*.

Cosmin Adrian Bejan and Sanda Harabagiu. 2010. Unsupervised event coreference resolution with rich linguistic features. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*.

Steven Bethard. 2013. Cleartk-timeml: A minimalist approach to tempeval 2013. In *Proceedings of Joint Conference on Lexical and Computational Semantics*, volume 2, pages 10–14.

Zheng Chen and Heng Ji. 2009. Graph-based event coreference resolution. In *Proceedings of Workshop on Graph-based Methods for Natural Language Processing*.

Zheng Chen, Heng Ji, and Robert Haralick. 2009. A pairwise event coreference model, feature impact and evaluation for event coreference resolution. In *Proceedings of the Workshop on Events in Emerging Text Types*, pages 17–22. Association for Computational Linguistics.

Ido Dagan, Dan Roth, Mark Sammons, and Fabio Massimo Zanzotto. 2013. Recognizing textual entailment: Models and applications. *Synthesis Lectures on Human Language Technologies*, 6(4):1–220.

Quang Do, Yee Seng Chan, and Dan Roth. 2011. Minimally supervised event causality identification. In *Proceedings of Conference on Empirical Methods in Natural Language Processing*.

Quang Xuan Do, Wei Lu, and Dan Roth. 2012. Joint inference for event timeline construction. In *Proceedings of the Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pages 677–687.

Anna Feltracco, Elisabetta Jezek, and Bernardo Magnini. 2015. Opposition relations among verb frames. In *Proceedings of the Workshop on EVENTS at the Conference of the North American Chapter of the Association for Computational Linguistics on Human Language*.

Yu Hong, Jianfeng Zhang, Bin Ma, Jian-Min Yao, Guodong Zhou, and Qiaoming Zhu. 2011. Using cross-entity inference to improve event extraction. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pages 1127–1136, Portland, OR, USA.

Rei Ikuta, William F Styler IV, Mariah Hamang, Tim OGorman, and Martha Palmer. 2014. Challenges of adding causation to richer event descriptions. *Proceedings of the Annual Meeting of the Association for Computational Linguistics*.

Heng Ji and Ralph Grishman. 2008. Refining event extraction through cross-document inference. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, Columbus, Ohio, USA.

Heeyoung Lee, Marta Recasens, Angel Chang, Mihai Surdeanu, and Dan Jurafsky. 2012. Joint entity and event coreference resolution across documents. In *Proceedings of Conference on Empirical Methods in Natural Language Processing*.

Qi Li, Heng Ji, and Liang Huang. 2013. Joint event extraction via structured prediction with global features. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pages 73–82, Sofia, Bulgaria.

Qi Li, Heng Ji, Yu Hong, and Sujian Li. 2014. Constructing information networks using one single model. In *Proceedings of the Conference on Empirical Methods on Natural Language Processing*.

Shasha Liao and Ralph Grishman. 2010. Using document level cross-event inference to improve event extraction. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pages 789–797, Uppsala, Sweden.

Timothy Miller, Steven Bethard, Dmitriy Dligach, Sameer Pradhan, Chen Lin, and Guergana Savova. 2013. Discovering temporal narrative containers in clinical text. In *Proceedings of the 2013 Workshop on Biomedical Natural Langua ge Processing*.

Eleni Miltsakaki, Rashmi Prasad, Aravind K Joshi, and Bonnie L Webber. 2004. The penn discourse treebank. In *Proceedings of the International Conference on Language Resources and Evaluation*.

Paramita Mirza and Sara Tonelli. 2014. An analysis of causality between events and its relation to temporal information. In *Proceedings of the International Conference on Computational Linguistics*.

Ekaterina Ovchinnikova, Laure Vieu, Alessandro Oltramari, Stefano Borgo, and Theodore Alexandrov. 2010. Data-driven and ontological analysis of framenet for natural language reasoning. In *Proceedings of the International Conference on Language Resources and Evaluation*.

James Pustejovsky and Amber Stubbs. 2011. Increasing informativeness in temporal annotation. In *Proceedings of the 5th Linguistic Annotation Workshop*.

James Pustejovsky, Robert Knippen, Jessica Littman, and Roser Saurí. 2005. Temporal and event information in natural language text. *Language resources and evaluation*, 39(2-3):123–164.

Dragomir Radev. 2000. A common theory of information fusion from multiple text sources, step one: Cross-document structure. In *Proceedings of Special Interest Group on Discourse and Dialogue*.

Mehwish Riaz and Roxana Girju. 2013. Toward a better understanding of causality between verbal events: Extraction and analysis of the causal power of verb-verb associations. In *Proceedings of Special Interest Group on Discourse and Dialogue*.

Radu Soricut and Daniel Marcu. 2003. Sentence level discourse parsing using syntactic and lexical information. In *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics on Human Language*, pages 149–156. Association for Computational Linguistics.

William F Styler IV, Steven Bethard, Sean Finan, Martha Palmer, Sameer Pradhan, Piet C de Groen, Brad Erickson, Timothy Miller, Chen Lin, Guergana Savova, et al. 2014. Temporal annotation in the clinical domain. *Transactions of the Association for Computational Linguistics*, 2:143–154.

Yu Takabatake, Hajime Morita, Daisuke Kawahara, Sadao Kurohashi, Ryuichiro Higashinaka, and Yoshihiro Matsuo. 2015. Classification and acquisition of contradictory event pairs using crowdsourcing. In *Proceedings of the Workshop on EVENTS at the Conference of the North American Chapter of the Association for Computational Linguistics on Human Language*.

Naushad UzZaman, Hector Llorens, Leon Derczynski, Marc Verhagen, James F. Allen, and James Pustejovsky. 2013. SemEval-2013 Task 1: TempEval-3: Evaluating Time Expressions, Events, and Temporal Relations. In *Proceedings of the International Workshop on Semantic Evaluation*, pages 1–9.

Tongtao Zhang, Hongzhi Li, Heng Ji, and Shih-Fu Chang. 2015. Cross-document event coreference resolution based on cross-media features. In *Proceedings of Conference on Empirical Methods in Natural Language Processing*.