NAACL HLT 2013

The 1st Workshop on EVENTS:
Definition, Detection, Coreference, and Representation

Proceedings of the Conference

14 June 2013
Westin Peachtree Plaza Hotel
Atlanta, Georgia

Introduction

The definition and detection of events has its roots in philosophy and linguistics, with seminal works by Davidson (1969; 1985), Quine (1985), and Parsons (1990). Though events have long been a subject of study, the NLP community has yet to achieve a consensus on the treatment of events, in spite of its critical importance to several areas in natural language processing, such as topic detection and tracking (Allan et al., 1998), information extraction (Humphreys et al., 1997), question answering (Narayanan and Harabagiu, 2004), textual entailment (Haghighi et al., 2005), and contradiction detection (De Marneffe et al., 2008). Most attempts to provide annotation of event coreference have been limited to specific scenarios or domains, as in LDC's ACE and Machine Reading event annotation (Humphreys et al., 1997; Bagga and Baldwin, 1999; He, 2007). The recent OnotoNotes annotations include more general event mentions and coreference, but mainly identify coreferences between verbs and nominalizations (Pradhan, 2007). Events are also a central element of the Time-ML temporal relation annotation, with an overlap-ping but slightly different approach (Pustejovsky, et al., 2010).

Truly comprehensive event detection must encompass the detection of events and their subevents, and take into account bridging references (Poesio and Artstein, 2005; 2008). The requisite event representation is clearly related to the information available in lexical resources such as PropBank, VerbNet, and FrameNet, but goes well beyond anything they currently contain. Bejan and Harabagiu (2010) have recently offered broader event coreference annotation for evaluation purposes, which was revised and extended by Lee et al., (2012). The organizers are themselves involved in event coreference projects for deep natural language understanding and medical informatics.

The time is ripe to bring together people interested in a serious discussion about the nature, definition, recognition, and representation of events and their parts and aspects. As a community we have to develop appropriate guidelines, resources, and processes for dealing with events and inter-event coreference.

In this workshop we structure the discussion around three themes:

- Foundations: What are Events? Definition and recognition.

- Coreference: When are two events the same? What kinds of identity are there?

- Representation: How best to represent events and event groups?

This is a genuine "working" workshop. Leading up the workshop, the organizers, with the assistance of the program committee, organized a shared annotation task on event mention and coreference annotation. Data was made available for participants to annotate, and the resulting annotations were analyzed for agreements and disagreements. During the workshop, the principal differences emerging from the different annotation schemes will be highlighted and discussed, with the intention of reaching a consensus on the handling of events and their coreference in future work in the NLP community. We have invited James Pustejovsky, the leader of the Time-ML effort, whose contribution as the keynote address are much appreciated.

We hope that this workshop will be the beginning of a concerted effort to come to grips with the

challenging topic of events.

Reference

1. James Allan, Jaime Carbonell, George Doddington, Jonathan Yamron, and Yiming Yang. 1998. Topic Detection and Tracking Pilot Study: Final Report. *Proceedings of the Broadcast News Understanding and Transcription Workshop*, pp 194–218.

2. Amit Bagga and Breck Baldwin. 1998. Algorithms for scoring coreference chains. *Proceedings of the LREC 1998 Workshop on Linguistic Coreference*, pp 563–566.

3. Amit Bagga and Breck Baldwin. 1999. Cross-document event coreference: Annotations, experiments, and observations. *Proceedings of the ACL 1999 Work-shop on Coreference and its Applications*, pp 1–8.

4. Cosmin Adrian Bejan and Sanda Harabagiu. 2010. Unsupervised Event Coreference Resolution with Rich Linguistic Features. *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics (ACL)*, Uppsala, Sweden.

5. Zheng Chen and Heng Ji. 2009. Graph-based event co-reference resolution. *Proceedings of the ACL-IJCNLP 2009 Workshop on Graph-based Methods for Natural Language Processing*, pp 54–57.

6. Donald Davidson, 1969. The Individuation of Events. In N. Rescher et al., eds., *Essays in Honor of Carl G. Hempel*, Dordrecht: Reidel. Reprinted in D. Davidson, ed., *Essays on Actions and Events*, 2001, Oxford: Clarendon Press.

7. Donald Davidson, 1985. Reply to Quine on Events, pp 172–176. In E. LePore and B. McLaughlin, eds., *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, Oxford: Blackwell.

8. Marie-Catherine de Marneffe, Anna N. Rafferty, and Christopher D. Manning. 2008. Finding Contradictions in Text. *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (ACL-HLT)*, pp 1039–1047.

9. Aria Haghighi, Andrew Ng, and Christopher Manning. 2005. Robust Textual Inference via Graph Matching. *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing (HLT-EMNLP)*, pp 387–394.

10. Tian He. 2007. *Coreference Resolution on Entities and Events for Hospital Discharge Summaries*. Thesis, Massachusetts Institute of Technology.

11. Kevin Humphreys, Robert Gaizauskas, and Saliha Azzam. 1997. Event coreference for information extraction. *Proceedings of the Workshop On Operational Factors In Practical Robust Anaphora Resolution For Unrestricted Texts*, pp 75–81.

12. Srini Narayanan and Sanda Harabagiu. 2004. Question Answering Based on Semantic Structures. *Proceedings of the 20th International Conference on Computational Linguistics (COLING)*, pp 693–701.

13. Massimo Poesio and Ron Artstein. 2005. Annotating (anaphoric) ambiguity. *Corpus Linguistics*, Birmingham, UK.

14. Massimo Poesio and Ron Artstein. 2008. Anaphoric annotation in the ARRAU corpus. *Proceedings of the LREC 2008 conference*, Marrakech, Morocco.

15. Sameer Pradhan, Eduard Hovy, Mitch Marcus, Martha Palmer, Lance Ramshaw and Ralph Weischedel, 2007. OntoNotes: A Unified Relational Semantic Representation, *Proceedings of the First IEEE International Conference on Semantic Computing*, Irvine, CA. Best Paper Award.

16. James Pustejovsky, Kiyong Lee, Harry Bunt, and Laurent Romary. 2010. ISO-TimeML: An International Standard for Semantic Annotation. *Proceedings of the LREC 2010 conference*. Malta.

17. William V. O. Quine, 1985. Events and Reification, pp 162–171. In E. LePore and B. P. McLaughlin, eds.,*Actions and Events: Perspectives on the philosophy of Donald Davidson*, Oxford: Blackwell. Reprinted

Organizers:

Eduard Hovy, CMU
Teruko Mitamura, CMU
Martha Palmer, University of Colorado Boulder

Program Committee:

Marjorie Freeman, BBN
Alan Goldschen, Mitre
Kira Griffit, LDC
Heng Ji, CUNY
Boyan Onyshkevych, DOD
Stephanie Strassel, LDC
Ben Van Durme, JHU-COE
Mihai Surdeanu, University of Arizona
Marta Recasens, then at Stanford University

Invited Speaker:

James Pustejovsky, Brandeis University

Table of Contents

Conference Program

Friday, June 14, 2013 (continued)

4:00–5:30   Working Session III: How best to represent events? What aspects to annotate?

5:30–6:00   General Discussion

6:00   Close