
Eric Davis, Lori Levin, Teruko Mitamura
March 15, 2008

Abstract

This is an annotation manual for event co-reference.

1 Introduction: Goal of Annotation

We annotate related events and group them into the same semantic frame. For example, a frame such as 'war' could have events such as 'bomb,' 'air raid,' 'attack,' 'dethrone...'. The 'bombing,' 'raiding,' and possibly 'attacking' events would all fit into the 'war' frame and be co-referent. In addition to annotating co-referent events in the same semantic frame, we also annotate entity relations between these events (Described in the Entity Relations manual). A small examples is that the agents of the 'bombing' event from above may be related via an 'ally' relation.

The interest of the annotation is clear: we want to be able to know which events are related to other events at the document level. Entity relations within these related events will then provide further information about the main actors, targets, causes, et al of the events. This annotation is a step towards that goal.

2 Annotation Units

Annotation units for events:

- Propositions identified by PropBank
- Nominals identified by NomBank

3 Annotating event co-reference

Our intentions are to annotate the following:
• **happened:** Any event (verb or nominal) that definitely occurred should be annotated.

(1) Britain bombed Iraq last night

In (1) above, the *bombing* event definitely occurred, as evidenced by the past tense of *bombed* and the adverbial phrase *last night*.

• **did not happen:** Any event (verb or nominal) that definitely did not occur should also be annotated.

(2) Hall did not speak about the bombings

In (2) above, the *speaking* event should be annotated despite the negation. Also, assuming that *bombing* from above and *bombing* from this sentence were in the same document, they would be co-referent.

• **planned:** Any event (verb or nominal) that is planned or expected to happen should also be annotated. Example verbs include: *planned, expected* and *agreed to*.

(3) Hall planned to meet with Saddam

In (3) above, the *meeting* event should be annotated because it is planned. Note that *planning* is not annotated because it is not an event.

4 **Tricky cases**

Sometimes the above guidelines on what to annotate do not apply fully. Here are a few cases that should be annotated but do not meet the above guidelines. Note that such events are not necessarily denoted with past tense.

• **pre-supposed definites:** Any event (verb or nominal) that is pre-supposed to have happened should be annotated.

(4) It may well be that theft will become a bigger problem

In (4) above, the *stealing* event has definitely occurred, and the writer is merely speculating that more may happen in the future.

• **habituals in present tense:** Any event (verb or nominal) that occurs numerous times in the present tense should be annotated

(5) Disneyland opens at 8 am
In (5) above, the opening event is a habitual event that occurs every morning at 8. As such, it should be annotated.

5 Arguments

Annotated events fitting into semantic frames may also have arguments. Arguments include agent, patient, theme, cause, location, time, et al. These arguments can be used to create entity relations within the frame of events.

- **agent**: Agent is typically a human or animate subject. It is mainly used as a volitional agent but also includes internally controlled subjects such as forces and machines. Agent has also been collapsed with the experiencer role to include cognition (thinking), perception (seeing), and emotion (feeling).

  (6) Pablo hit the man

  In (6) above, the agent of the hitting event is Pablo.

- **aim**: The aim argument denotes events that are done for a certain purpose or aim. Note that events that the primary event for which the purpose event is done are still annotated as separate events.

  (7) He must inform the officials to settle the matter

  In (7) above, the settling event is the purpose/aim of the informing event. As such, settling fits in the aim argument slot.

- **beneficiary**: The beneficiary argument denotes people or organization who benefit from information, money, et al. This role is only used if it is obvious that the person or organization is truly benefitting. When in doubt, use the recipient role.

  (8) Hostile and even friendly nations ... steal information ... and share it with their own companies

  In (8) above, the companies are benefitting from stolen information that they would otherwise not be privy to. As such, companies fits in the beneficiary argument slot.

- **cause**: The cause argument denotes events that occurred as the result of another event. Note that events that are caused are still annotated as separate events.
(9) The bombings killed 23 and injured 2

In (9) above, the bombing event caused the killing event and injuring event. As such the bombing event would fit in the cause argument slot.

• exception: This role denotes events or persons that are exceptions to the general rule. The key triggering word is typically except or something similar.

(10) It had been submitted in a timely manner, except in special cases

In (10) above, the documents had typically been submitted in a timely manner. However, there are special cases that are exceptions. As such, special cases would fit in the exception role.

• goal: This role denotes what the action is directed towards. However, telling, giving, hearing, et al should be labeled recipient, theme. Also, final destinations of verbs like go are labeled with the goal role.

(11) I ain’t going anywhere

In (11) above, the anywhere entity would not be labeled location but instead is the destination of going. As such, anywhere would be labeled with the goal role.

• instrument: This role denotes objects or forces that change or impact other objects and cause change. Such propositions are typically introduced with 'with.'

(12) They should use wet blanket over the flames

In (12) above, the wet blanket entity is being used to put out the fire. Thus, it fits the instrument role.

• location: The location argument denotes where an event took place.

(13) Maria kissed Tony in Harlem

In (13) above, the kissing event happened in Harlem. Harlem would thus be the argument in the location role.

• recipient: The recipient role denotes a person or organization that received an object or entity. The key point here is that it is not obvious that the received object benefitted the recipient in any way. When in doubt, choose recipient over beneficiary.
(14) I came to give you important information

In (14) above, you will receive information. However, the source of the information is not known to be reliable, so it is not certain that you will benefit from the information. As such, you is the recipient in the above example.

- **source**: The source role denotes where an action took place. However, the vast majority of the time, the theme role is a better fit, so theme is used much more frequently than source. When in doubt, use theme over source.

(15) ... Previews Magazine, detailing upcoming events

In (15) above, *Previews Magazine* is the source of information about *upcoming events*. Therefore, *Previews Magazine* would be labeled with the source role.

- **theme**: A theme denotes participants in a location or undergoing a change of location. This role has been collapsed with the patient role, so theme can be used with participants undergoing a change. Moreover, theme can be used with participants who have been affected. It also denotes events that are arguments of other events. Also, if a source is not clear or obvious, the entity is tagged theme rather than source.

(16) One should adopt whatever makes him look good and avoid what degrades him

In (16) above, the degrading event is an argument of the avoiding event. Since, avoiding requires an argument, degrading fits in the theme role for the avoiding event.

- **time**: The time role denotes when an event took place. This could be a date or a temporal adverb such as recently.

(17) Hall recently returned from a trip to Iraq

In (17) above, recently denotes when the returning event took place. As such it fits in as the time argument for the returning event.

6 Annotation Tool

Annotation Tools
• GATE (Sheffield)
• PropBank

7 Sources

Sources

• Wikipedia
• VerbNet