Meta-Knowledge Annotation of ACE Events

Annotation Guidelines

Paul Thompson, Raheel Nawaz, Sophia Ananiadou

National Centre for Text Mining, School of Computer Science, University of Manchester, UK

Contents

1 Introduction and Background

If a user wishes to search for relevant information located within a collection of electronic documents (e.g., news articles), the usual method is to enter keywords that are typically used within descriptions of the information sought into a search engine. However, such searches typically return a large number of documents, many of which are likely to be irrelevant. This is because keyword searches are usually not sufficiently powerful to allow searches to be well aligned with users' actual information needs. What users are typically looking for are documents that contain specific types of *information chunks,* in which specific concepts are mentioned, and linked to each other in specific ways to describe facts.

Unfortunately, keyword-based searches cannot allow users to isolate only those documents containing information chunks of interest, for a number of reasons. These include the following:

- A given concept may be represented in text in different ways. For example, a particular person may be referred to in different places by their full name, first name, nickname etc., whilst synonyms are likely to exists for many concepts, which may be used interchangeably in the text (e.g. *guns* vs. *artillery*). It can thus be difficult for a search system user to enumerate within their query all possible ways in which a concept of interest could be mentioned in text.
- Even if the user is at least partially successful in retrieving all documents that mention different concepts of interest, it is impossible using a keyword query to ensure that concepts are related to each other in the required ways (i.e., in order to ensure that they are mentioned together to convey a specific type of information of interest). For example, if a search is carried out for documents that contain multiple concepts of interest, then the mentions of these concepts may appear at different places within the retrieved documents, such that the concepts are not related to each other at all.

As an example, let us consider that a user is interested in information about the war in Iraq, and in particular about the types of weapons used by the Iraqis during their attacks. In order to search for this information, the user would enter some terms into a search engine, for example *Iraqi* and *attack.* There are a number of problems here.

Firstly, a number of different verbs could potentially be used to describe attacks, often depending on the type of weapon used. For example, consider sentences (S1) and (S2), which describe attacks using different types of weapons, and which would both be relevant to the user's query:

(S1) *The Iraqis are firing artillery*
(S2) *Iraqi warplanes dropped bombs over the town*

In order to retrieve the highest possible number of potentially relevant documents, the user would need to enumerate several different verbs that could describe attacks as part of their query. The problem is that, especially given that language usage can be highly creative in news articles, it would be virtually impossible to enumerate all potentially relevant verbs as part of the query. It is also the case that information about attacks could be centred around nouns rather than verbs (e.g., *onslaught*).

Even if great efforts are made to think of many variant ways in which attacks could be described, it is not possible to ensure that such words will be linked to *Iraqis* in the text in the way required. With keyword search, it is not possible to guarantee that the entered search terms will appear in the

same sentence. Even if they do, they may not be linked together correctly (i.e., where the Iraqis are the instigators of the attack). . Consider sentence (S3), which would be retrieved if the search terms *fire* and *Iraqi* were submitted to a search engine. However, in this case, it is the U.S. who are carrying out the attack, rather than the Iraqis.

(S3) *The U.S. troops fired artillery at the Iraqis*.

Text mining systems help to cut down on the amount of time that users have to spend sifting through irrelevant documents. Such systems can automatically locate, extract, structure, categorise and link relevant pieces of knowledge from text and store them in a structured database format.

This process is called *event extraction*, since the structured knowledge representations are called *events*. In simple terms, event extraction systems look for knowledge chunks and assign one of a fixed set of semantic labels, according to the type of information being expressed. For example, all of the examples in (S 1-3) above may correspond to the type ATTACK. The different *participants* in the event (e.g., the people, places and other things that contribute towards the description of the event) are identified and also assigned labels (called *semantic roles*), which determine the type of contribution made towards the event description. For example, in (S1), the phrase *The Iraqis* would be assigned the AGENT role, since this phrase corresponds to the instigator of the attack. Likewise, in (S3), the phrase *The U.S. troops* corresponds to the AGENT. In a similar way, an INSTRUMENT role is assigned to *artillery* in (S1) and (S3), and *bombs* in (S2).

After a document collection has been analysed and all events have been extracted, the user can then formulate structured queries over these event representations, instead of using simple keyword-based search. This provides scope for the user to retrieve much more focussed sets of search results, which match more closely to their information needs.

For example, according to the automatic semantic categorisation of events, finding events that mention attacks would simply require that events assigned the type ATTACK are queried, without the need to worry about the exact ways in which attacks may be described in the text. Furthermore, it would be possible to specify that the AGENT of the ATTACK event should be a phrase containing the word *Iraqis.* This would ensure that documents containing sentence (S1) and (S2) would be retrieved, but (S3) would not, because here the phrase containing *Iraqis* does not correspond to the AGENT. Using event-based queries, it is not necessary to specify values for all participant types. For example, by leaving the INSTRUMENT role unspecified, it would be possible to discover the range of different weapons used in the attacks. However, it would also be possible to place a restriction on the value of this role, e.g., phrases containing *artillery,* to further restrict the results returned by the query.

1.1 Event Interpretation

As mentioned above, event extraction systems usually concentrate on identifying events and their participants (e.g., the instigator of an attack, the target, the weapon used, etc.), what they often lack is a means to distinguish adequately between different interpretations of events. Looking back at (S1), (S2) and (S3) can all be understood as introducing new information into the discourse, i.e., the way in which the information is presented shows that the attack in question has not previously been mentioned in the article. In contrast, in sentence (S4), the word "attack" can be understood as referring to information that has previously been introduced into the discourse. In this case, the

use of the definite article *the* prior to the word *attack,* provides evidence that a specific, previously introduced attack is being referenced.

(S4) *The Palestinian leadership said the <u>attack</u> would serve as "a pretext for Israel's government and occupation army to step up its deadly campaign which caused the deaths of 77 Palestinians in February".*
(SOURCE: nw/AFP_ENG_20030305.0918, sentence 26)

NOTE ABOUT EXAMPLE SENTENCES: Where a SOURCE is indicated, this is an actual sentence from a document within the ACE corpus (according to the sentence numbers displayed in *brat*). The corpus is split into a number of different directories, according to the exact source of the data (e.g., the *nw* directory contains newswire reports, while *bn* contains transcripts on broadcast news, etc.)

Whilst the attack in (S4) in certainly factual, i.e., it definitely took place, the sentence (S5) is somewhat different, in that some of the information can be understood to be *potentially* non-factual.

(S5) *The Israeli government said its forces were returning fire, **suggesting** that Miller **could** have been <u>hit</u> by Palestinian gunmen*
(SOURCE: bn/CNN_ENG_20030506_160524.18, sentence 22)

From (S5), we can understand that the event concerning Miller being hit by Palestinian gunmen only **may not be true**. This is denoted both by the presence of the word *suggesting* and the word *could,* which taken together ensure that the reader understands that this is a very tentative analysis. Further examples of non-factual events are shown in (S6-S9).

(S6) *He was in the army and he's going to **probably** be <u>going</u> to Afghanistan.*
(SOURCE: cts/fsh_29783, sentence 118).

(S7) ***I don't know if*** *they've given him <u>the death penalty</u> yet*
(SOURCE: cts/fsh_29302, sentence 174)

(S8) *So, in regard to that, we are **paying great attention** to the Iraqis' **ability** to <u>defend</u> on the ground*
(SOURCE: bc/CNN_IP_20030402.1600.00-2, sentence 44)

In (S6), the word *probably* denotes that, whilst it is very likely that man in question will be travelling to Afghanistan, there is still some uncertainty/speculation surrounding it. Meanwhile, in (S7), the phrase *I don't know if,* denotes that the writer lacks sufficient evidence to say whether the event concerning the death penalty is true. In (S8), analysis is ongoing to determine whether the Iraqis can defend on the ground. Thus at least at the time of writing, it cannot be said definitively that they *can* defend the ground.

A further type of speculated event is one in which the writer expresses an attitude or subjective opinion about a possible future event, as in (S9).

(S9) ***Hopefully****, they find him guilty.*
(SOURCE: cts/fsh_29272, sentence 151).

In (S9), as in the other examples above, it is assumed that the information about the event is coming directly from the author, as there is no evidence to the contrary. Also in (S9), the positive subjectivity towards the event (dented by the word *Hopefully*) is understood to be that of the author. In contrast, consider (S10) and (S11).

(S10) *The Iraqis **promise** an unconventional attack on U.S. troops tonight.*
(SOURCE: bc/CNN_IP_20030404.1600.00-2, sentence 80)

(S11) *Word comes from the grand Ayatollah Sistani that he's **willing** to <u>meet</u> the American commander.*

(SOURCE: bc/CNN_IP_20030403.1600.00-3, sentence 56)

In (S10), there is an event describing a potential attack by the Iraqis on U.S. troops. The wording of the sentence identifies the Iraqis explicitly as the "information source" of the potential event. Additionally, the use of the word *promise* shows that the Iraqis have positive subjectivity towards this event, i.e., they want it to happen. Similarly, in (S11), the information source of the potential event denoted by *meet* is *the grand Ayatollah Sistani,* and the word *willing* shows that he has positive subjectivity towards these meetings, i.e., he is happy for them to take place.

In (S12), negative subjectivity is expressed towards the event denoted by *attacks*. It should also be noted that, in contrast to (S9) - (S11), this is event is not a speculated future event, but rather a definite event that is on-going.

(S12) *"We **condemn** all <u>attacks</u> against civilians including today's attack in Haifa," said information minister Yasser Abed Rabbo*

(SOURCE: nw/AFP_ENG_20030305.0918, sentence 25)

1.2 Annotation of interpretative information

The examples above show, that various different types of interpretative information about an event can be specified in a sentence, often using different types of words and phrases.

We refer to such information *meta-knowledge*; common types of meta-knowledge that can be readily identified in the context of events include the following:

- Does an event introduce new factual information into the discourse? Alternatively, does it represent factual information that has previously been introduced, or is the information uncertain/speculated?
- Is there any attitude/subjectivity expressed towards the event
- What is the source of the information reported in the event?

It should be noted that several of these levels or *dimensions* of meta-knowledge can be simultaneously relevant to describe the interpretation of an event. For example, a particular third-

party source may assert new factual information. Optionally, they may express an opinion towards this factual information. Alternatively, they may present information as uncertain, which may also be accompanied by information relating to subjectivity.

By facilitating the recognition of relevant information about multiple meta-knowledge dimensions at the level of events, both subtle and significant differences in event interpretation can be discovered, by considering different combinations of dimensions.

Our goal is to be able to develop event extraction systems that can automatically distinguish between events with different meta-knowledge interpretations, which will allow meta-knowledge information to be stored in the database as part of the extracted information. This will in turn allow more sophisticated event extraction systems to be developed, which allow meta-knowledge aspects to be specified as part of the search criteria, and to be presented within the results, allowing the interpretation of returned to events to be quickly determined.

Event extraction systems are typically developed using machine learning techniques. A set of documents that is representative of the collection over which search is to be carried out is marked up (or *annotated*) to identify the events of interest that occur within them. Machine learning algorithms use this annotated evidence to learn features of the text that denote the presence of the various parts of an event. The resulting *trained model* can then be used to recognise events in previously un-annotated text. By additionally annotating aspects of meta-knowledge about events, as well as information about the event type and its participants, it will be possible to train models that can recognise not only information about events and their participants, but also assign interpretative information to these events.

The annotation task described in this document is to annotate various types of meta-knowledge information, on top of a corpus in which the basic event annotation has already ben carried out. The annotation task thus involves reading each document in the corpus and, and for each event that has been pre-identified, determining the correct interpretation for each meta-knowledge dimension, and marking-up (annotating) any words and that have been used to determine this interpretation The annotation will be carried out using a user-friendly annotation environment called "brat". The details of the annotation scheme are described later in this document.

1.3 Background to the Task –Searching for Relevant Information

In this section, we provide some more detailed information about events and event based searching. As mentioned above, when searching textual documents for specific types of information, keyword-based searching can present a number of problems, e.g.:

- There a numerous ways of expressing a particular type of information, through different verbs, nouns etc. For example, events describing deaths of people could be described using the words *kill, die, death* etc. Formulating queries that cover all possibilities is complex and time consuming
- Individual search terms will not necessarily be related to each other in the way required by the user. For example, if the user wants to find out who has been killed by the Israelis, it would usually be the case that the word *Israeli* will occur within the grammatical subject of the verb that describes the killing event. However, using an ordinary search engine, there is no way of specifying that such a relationship should hold.

1.3.1 Structured events

Text mining technology can help greatly in searching for information, both to giving extra power to the searching mechanism, thus reducing the number of separate searches that have to be carried out, as well as increasing the relevance of the results that are returned by the search.

Instead of viewing as sequences of words, text mining systems try to *understand* what is being said in the text. This is usually done by identifying and categorising important entities in the text (people, places, weapons, etc.) and then trying to determine how they relate to each other to convey specific types of knowledge. As mentioned above, the end goal is usually to extract structured events from the text. In order to make the concept of an event more concrete, let us reconsider sentence (S5) from above.

(S5) *The Israeli government said its forces were returning fire, suggesting that Miller could have been hit by Palestinian gunmen.*
(SOURCE: bn/CNN_ENG_20030506_160524.18, sentence 22)

Let us consider the event about Miller being hit by Palestinian gunmen. This is one example of an event that is already annotated in the ACE 2005 corpus (which is the collection of texts to which meta-knowledge annotation is to be added). The event is represented in ACE as follows:

ANCHOR: *hit*
TYPE: *Conflict*
SUBTYPE: *Attack*
ATTACKER: *Palestinian gunmen* (TYPE: *PER* SUBTYPE: *Group*)
TARGET: *Miller* (TYPE: *PER* SUBTYPE: *Individual*)
MODALITY: *Other*
POLARITY: *Positive*
GENERICITY: *Specific*
TENSE: *Past*

In ACE 2005, fairly detailed information assigned for each event.

This consists of:

- **ANCHOR** – each event is anchored to a particular word or phase in the text around which the event is centred (usually a noun or verb). It this case, the anchor is the verb *hit*.
- **TYPE** – One of 8 predefined top level types (e.g., *Conflict, Justice, Life).*
- **SUBTYPE** – One of 33 subtypes that are more specific than the general types (e.g. *Arrest-Jail* is a subtype of *Justice, Injure* is a subtype of *Life*, etc). Only events that fall into one of the pre-defined categories in the text are annotated in the corpus.
- **Participants** – Each participant in the event (an event may have zero or more participants) is characterised according to a set of *semantic roles* that are specific to the event subtype. The sematic roles characterise the contribution that the participant makes to the description of the event. For example, the *CONFLICT_ATTACK* subtype has the possible roles *ATTACKER, TARGET, PLACE, TIME* and *INSTRUMENT* as potential semantic roles. Each role can be assigned to zero or more participants, depending on the information provided in the text. In (S5), a single *ATTACKER*, i.e. *Palestinian gunmen* and a single

TARGET, i.e., *Miller,* are specified. No other information is provided and so the other potential roles are not used

- **Entity types** – The event participants are assigned named entity types. In a similar way to events, each entity type is assigned both a top-level type and a subtype. In the example above, both of the event participants, i.e., *Palestinian gunmen* and *Miller* have been assigned the top-level type *PER*, meaning that they represent persons. The subtypes are *Group* and *Individual,* respectively, to distinguish between groups of people and individuals.

The remaining attributes of the event (MODALITY, POLARITY, GENERICITY and TENSE), which are already specified for each event in the ACE corpus, can be understood to provide certain types of meta-knowledge about the event. However, the annotation work described in this document aims to enrich the existing meta-knowledge to allow a more fine-grained distinction between events. These existing attributes relating to meta-knowledge may be defined more precisely as follows:

- **MODALITY** – Determines whether the event represents a "real" occurrence. There are two possible values: *Asserted* if the author or speaker makes reference to it as though it were a real occurrence, and *Other* otherwise. In the event above, *Other* is assigned, since there is some speculation about the event.
- **POLARITY** - has the value *Positive* unless there is an explicit indication that the event did not take place, in which case *Negative* is assigned.
- **GENERICITY** - has either the *Specific* value, if the event can be understood as describing a singular occurrence at a particular place and time, or a finite set of such occurrences (as the case in the example above), or *Generic* otherwise.
- **TENSE** - specifies the tense of the event with respect to the author, and can be *Past, Present, Future* or *Unspecified* (where the tense cannot be determined from the context)

1.3.2 Event-based searching

The example event structure helps to make it clearer how carrying searching over these extracted event structures, rather than over plain text, can help to lead to more precise and focussed searches. Event structure abstracts from the exact wording in the text, meaning that searches over events can specify the following:

- Event types instead of words used to describe the event. This can vary in terms of the level of specificity. For example, depending on the required breadth of the search, it would be possible to specify either a top-level event type (e.g., *Conflict)* or a more specific event type (e.g., *Attack).* There is no need to worry about the exact textual event anchors of the event, e.g., *hit, fire, bomb*, etc. The event extraction will learn how to recognise and characterise events of different types automatically.

- Restrictions on the event participants in terms of:

  o Semantic roles assigned to participants (e.g., ATTACKER, INSTRUMENT etc.)

  o Values of particular roles, which could be specified as either:

    ▪ Keywords when searching for specific values (e.g., *tank* as the value of the *INSTRUMENT* role)

- Named entity types for a more general search. As with the event types, the two-level system for categorising named entities means that the level of specificity of the search is flexible. For example, it would be possible to search for *CONFLICT_ATTACK* events where the *TARGET* is: a) Any type of phrase they refers to a person or group of people (using the named entity type *PER*). b) A reference to an individual person. In this case, the more specific named entity type *PER_Individual* may be specified. On the other hand, it may be desired to search for events where the instrument of the attack is any type of weapon. In this case, the value of the *INSTRUMENT* slot would be specified as *WEA* (for weapon), rather than a more specific weapon type, like *WEA_shooting*.
  o Restrictions on meta-knowledge values. For example, a user may only be interested in events that represent real occurrences that happened in the past. Therefore, restrictions could also be placed on one or more of the meta-knowledge attributes, e.g., MODALITY=*Asserted* and TENSE=*Past*

1.4 Need for additional meta-knowledge annotation

Considering the meta-knowledge related attributes that already exist in the ACE corpus, it can be appreciated that they cannot adequately distinguish between all of the types of meta-knowledge that were introduced in Section 1.1. For example, in the existing ACE annotation, events assigned *MODALITY=Asserted* may either be introducing new information into the discourse or they may be reintroducing previously asserted information. Likewise, although most events assigned *MODALITY=Other* will correspond to speculations, this is not exclusively the case. Additionally, the existing meta-knowledge attributes do not allow different types of subjectivity to be encoded, nor do they allow different knowledge sources to be identified.

It is for these reasons that we are embarking upon the current annotation effort. All events in the ACE corpus will be manually enriched with additional meta-knowledge information. The scheme for encoding the additional meta-knowledge, together with examples, will be outlined in the remainder of this document.

2 Annotation Task

The task consists of assigning information about a number of meta-knowledge dimensions to each event in the ACE corpus. This corpus consists of a mixture of newswire, transcriptions of broadcast news and broadcast conversations, blogs and Usenet news groups and discussion forums.

The meta-knowledge attributes to be considered consist of a mixture of the existing attributes in the ACE corpus, plus some additional ones, i.e., *Subjectivity* and *Source.* The values of the existing attributes need to be reviewed for two main reasons:
- For the MODALITY attribute, the range of possible values has changed from the original annotation, and hence new values may have to be assigned for certain events.
- We have changed and tightened some of the definitions of attribute values that were provided in the original ACE annotation guidelines. In light of these changed, some assigned values may have to be updated.

The annotation task consists of three main steps, which are further clarified in the subsections below describing the individual dimensions

1) For each event, an appropriate value (from a fixed set) is determined for each dimension, based on evidence from the context in which the event occurs (i.e.,, the sentence in which the event is described). Most often, the evidence comes in the form of a particular word or phrase that is present in the same sentence as the event (such as *promise, suggest* etc.). Evidence that does not occur within the same sentence as the event should not be used.

2) If the evidence for the assignment of a value is a particular word or phrase in the same sentence as the event, then this word or phrase is explicitly annotated as the "clue" for that particular dimension and event.

3) For the *Source* dimension, if a named source of knowledge for the event can be identified, this is also annotated. See section 2.4.2 for more information.

Figure 1 illustrates the scheme and the possible values for each dimension. A distinction is made between:

- Attributes that are present within the existing ACE annotation
- Those that will be added during the current annotation effort, or else whose values have changed from the original ACE annotation.
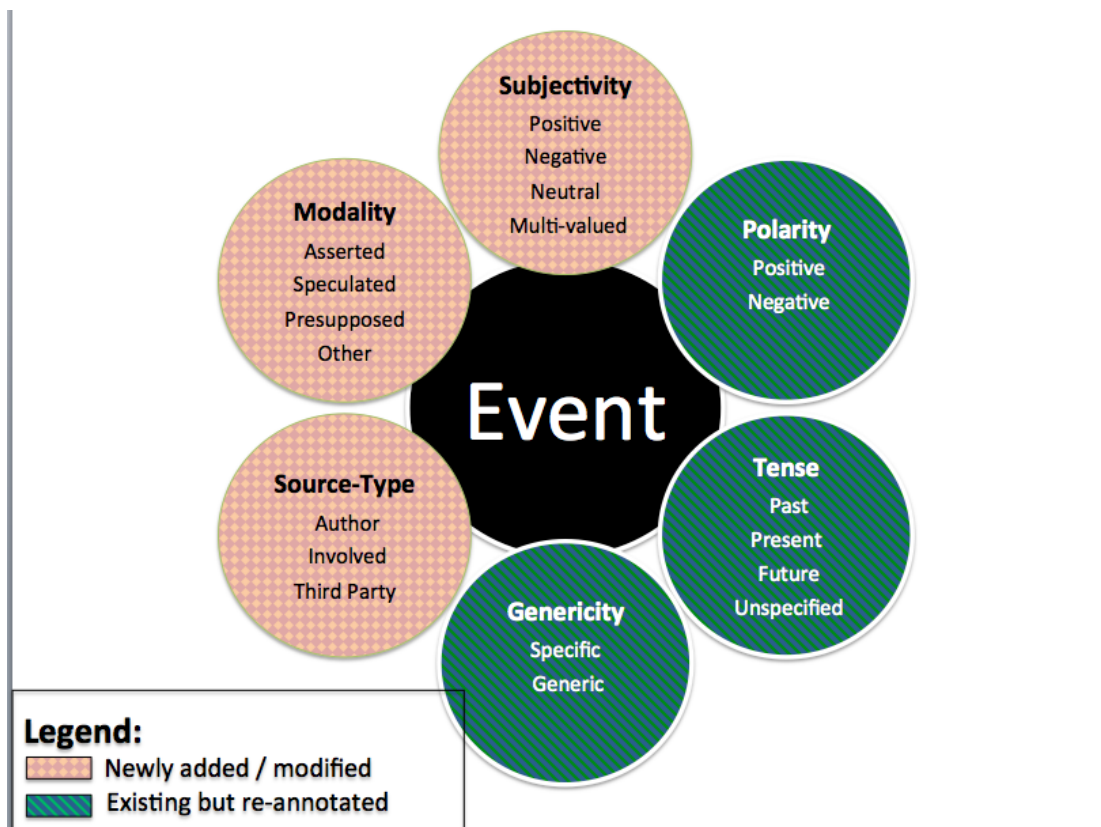

Figure 1: Meta-knowledge annotation dimensions for the ACE corpus

The purpose of the annotation, then, is to discover the different ways in which each value of each dimension can manifest itself as evidence in the text. By annotating the complete ACE corpus with this information, we can train a system to learn patterns based on these annotations. The trained

system will then be able to predict the values of the annotation dimensions for previously unseen events.

In the following sections, we provide detailed information regarding the 6 individual meta-knowledge dimensions. A brief description of each dimension is followed by a subsection describing each possible value of the dimension. We provide

an enumeration of its possible values, together with some examples. In all of the examples, the word(s) on which the event is centered (i.e. the anchor word/phrase) are shown using _underlined italics_, whilst the explicit "clue" words which provide evidence for the assignment of a particular value to a dimension are shown using **bold face.**

Please note that where examples come from broadcast news, capitalisation of words is not used.

2.1 Modality

This dimension attempts to provide a general characterisation of the type of information that is conveyed by the event. The four possible values are covered in detail in the following subsections.

2.1.1 Asserted

This value is assigned to factual or "concrete" events. By "concrete", we mean that the event describes something that actually happens, happened or is happening (i.e., it could be either specific event, or something that happened habitually). NOTE: Events that will happen in the future are always assigned the _Speculated_ modality value (see section 2.1.3).

For this value to be assigned, the event should be phrased as though it is providing new information to the discourse (as long as the information is not speculated), regardless of whether the event constitutes the main assertion of the sentence. Asserted events should be contrasted with "presupposed" events, which are also concrete events, but which are presented as providing background information, or information about events that are already known to have occurred.

Typical evidence

- Event anchor is a verb in the past or present tense (S13 – S16), or an anchor belonging to another part of speech that occurs in a present or past context

- There are no specific clue words in the context of the event to suggest that it should be interpreted in a non-factual way.

- Most typically, the event will have at least one participant.

- The use of an indefinite article will normally denote an asserted event (S18 – S20)

- Events that are reported as being the latest in the series of ongoing events should be annotated as "Asserted" (S21). Note, however, that in most other cases, events preceded by a definite article will be annotated as MODALITY = "Presupposed"

Example sentences

> (S13) After a day of fighting, the Iraqis were _driven back_ five kilometers, about three miles down the main road west towards Mosul from the Kurdish city of Erbil
> (SOURCE: bc/CNN_IP_20030404.1600.00-2, sentence 18)

(S14) The Iraqis are _firing_ artillery like that in response to the Americans are calling in air _strikes_
(SOURCE: bc/CNN_IP_20030404.1600.00-2, sentence 23)

(S15) last week williamson, a mother of four, was found _stabbed_ to death at a condominium in greenbelt, Maryland
(SOURCE: bn/CNN_ENG_20030528_172957.18, sentence 10)

(S16) it's happened, of course, several times during the war in iraq, journalists caught in the line of fire and _killed_
(SOURCE: bn/CNN_ENG_20030506_160524.18, sentence 9)

(S17) When an Al- Jazeera reporter was _killed_ today, the network charged it had been targeted by U.S. force.
(SOURCE: bc/CNN_IP_20030408.1600.04, sentence 18)

(S18) Without claiming responsibility, Hamas and another hardline group, Islamic Jihad, said the blast was revenge for a string of bloody Israeli army _raids_ into the Gaza Strip in recent weeks.
(SOURCE: nw/AFP_ENG_20030305.0918, sentence 23)

(S19) HAIFA, Israel, March 5 (AFP) - Fifteen people were killed and more than 30 wounded Wednesday as a _suicide_ bomber blew himself up on a student bus in the northern town of Haifa, the first bombing in Israel in exactly two months.
(SOURCE: nw/AFP_ENG_20030305.0918, sentence 9)

(S20) In another development, Israeli forensics experts identified a body that washed up on a Tel Aviv beach last week as that of a Briton, Omar Khan Sharif, 27, the suspected accomplice of a British _suicide_ bomber who blew himself up outside a Tel Aviv pub April 30.
(SOURCE: nw/APW_ENG_20030519.0367, sentence 52)

(S21) HAIFA, Israel, March 5 (AFP) - Fifteen people were killed and more than 30 wounded Wednesday as a suicide bomber blew himself up on a student bus in the northern town of Haifa, the first _bombing_ in Israel in exactly two months.
(SOURCE: nw/AFP_ENG_20030305.0918, sentence 9)

2.1.2 Presupposed

This value is assigned to events that are presented as:
- Providing "known" information
- Providing background or secondary information to *asserted* events

Typical Evidence

- Use of nominal event trigger expressions

- Use of definite article (S21, S22, S23) or possessive (S22) before a nominal trigger, indicating that the occurrence of the event is presupposed, i.e., it is not presenting new information. NOTE: the use of the definite article does not always denote a presupposed event. They may be asserted (see S17) or speculated (see S48a).
- Events whose context or trigger alone allows it to be understood that the event is ongoing and known about (S24-S28) or otherwise corresponds to a well-known event (S29)
- Events accompanied by words similar to "continue" or "stop" (e.g., *halt, suspend,* etc), which provide strong evidence that the event or series of events is known to be ongoing (S30)
- Events that are specified in the context of thanking or blaming somebody for an event that has already happened (S30e)
- Events that denote that denote charges, in the context of an outcome (See S32, where the "murdering" event (the charge) is specified in the context of the outcome ("found guilty")
- Copula constructions where it is clear that the event has been previously introduced (S33)

Example sentences

(S21) The Palestinian leadership said the *attack* would serve as "a pretext for Israel's government and occupation army to step up its deadly campaign which caused the deaths of 77 Palestinians in February".
(SOURCE: nw/AFP_ENG_20030305.0918, sentence 26)

(S22) Israel's night-time *raid* in Gaza involving around 40 tanks and armoured vehicles came after Prime Minister Ariel Sharon convened his security cabinet to discuss a response to the bomb *attack*.
(SOURCE: nw/AFP_ENG_20030305.0918, sentence 17)

(S23) Those *deaths* are not yet included in the official *casualty* count.
(SOURCE: bc/CNN_IP_20030404.1600.00-2, sentence 68)

(S24) israel and the united states say arafat has not been tough enough on *terrorism*
(SOURCE: bn/CNN_ENG_20030614_173123.4, sentence 28)

(S25) It is an elementary fact of life during *wartime*: A press pass provides no protection
(SOURCE:bc/CNN_IP_20030408.1600.04, sentence 23)

(S26) Al-Jazeera correspondent Omar Al Issawi will talk to Larry about the death of his colleague and the continuing plight of journalists in a *war* zone.
(SOURCE: bc/CNN_IP_20030408.1600.04, sentence 48)

(S27) We condemn all *attacks* against civilians including today's attack in Haifa," said information minister Yasser Abed Rabbo.
(SOURCE: nw/AFP_ENG_20030305.0918, sentence 25)

(S28) if you want an end to *terrorism*, if you want an end to violence, you have to give people hope and you have to give them a stake in the political process.

(SOURCE: nw/APW_ENG_20030519.0367, sentence 21)

(S29) Famed *World War II* reporter Ernie Pyle was killed by a sniper on an island in the Pacific
(SOURCE: bc/CNN_IP_20030408.1600.04, sentence 24)

(S30) Hamas said Monday it has no intention of **halting** *attacks*, despite Egypt's efforts to have Palestinian militant groups agree to a one-year **suspension** of *shootings* and *bombings*.
(SOURCE: nw/APW_ENG_20030519.0367, sentence 21)

(S31) She also **thanks** the Rockefeller Brothers Fund "for helping to *fund* the project." The book is about "one of the finest and most important governing bodies," she says.
(SOURCE: wl/FLOPPINGACES_20050203.1953.038, sentence 21)

(S32) Peterson Trial Scott Peterson has been found guilty of *murdering* his wife Laci and their unborn son, and he now faces the death penalty.
(SOURCE: cts/fsh_29302)

(S33) It was a *war* that Japan assumed would be ended in "3 to 6 months"; at the time of Pearl Harbor, it had been going on for 5 years, at the cost of over half a million Japanese lives
(SOURCE: un/soc.history.war.world-war-ii_20050127.2403, sentence 55)

2.1.3 Speculated

This value is assigned to events where there is an explicit indication that the truth value is unknown, is questioned, is hypothetical, or that the event has not yet taken place.

Typical Evidence

- The event is based on some kind of assessment carried by the author/reporter or a third party. The assessment may be an inference, analysis, opinion, interpretation, speculation or other types of cognitive analysis. These will normally be indicated by a an explicit verb form or a nominalisation, e.g., *believe, think, suggest, indicate, appear, seem, assume* (S34 – S36)).

- Use of modal auxiliaries like *may, might* and *could*, as well as adverbs/adjectives like *probably/probable, likely, unlikely, perhaps, maybe* and *allegedly,* which can all indicate uncertainty on the part of the author (S37- S39)

- Constructions that explicitly indicate that the truth value of an event is unknown. This may be a question (S40), a conditional clause (S41), choice (S42) or explicit phrase, such as *It is not known whether* (S43) or a verb such as *investigated* (S44)

- Frequency indicators such as *often, frequently, normally.* These denote that there is speculation about whether the event occurs all of the time.

- Events that are were being planned or prepared, and where it is not certain that the event has taken place (S45 – S47)

- Future events (S48 – S49)
    - Sometimes subjectivity (See section 2.3 for more details) is expressed towards these potential future events (S50 – S53) specific subject

- Events that can be considered as hypothetical (S54 – S55).

(S34) But the Pentagon *believes* some Iraqi Republican Guard units have <u>withdrawn</u> to the capital for what could be their last stand.
(SOURCE: bc/CNN_IP_20030403.1600.00-1, sentence 15)

(S35) We don't *think* that the <u>fighting</u> is over yet.
(SOURCE: bc/CNN_IP_20030403.1600.00-1, sentence 39)

(S36) And *the way seems to be open* for coalition forces to <u>close</u> in on the city's southern limits
(SOURCE: bc/CNN_IP_20030403.1600.00-1, sentence 30)

(S37) And I think she fell and *maybe* she <u>hit</u> her head or something and it <u>killed</u> her, so he panicked
(SOURCE: cts/fsh_29272, sentence 77)

(S38) Well, when they do finally enter Baghdad, U.S. and coalition troops *could* face urban <u>combat</u> with the Republican Guard
(SOURCE: bc/CNN_IP_20030403.1600.00-3, sentence 105)

(S39) Israeli media said the <u>retaliation</u> was *unlikely* to be on a scale that would risk disrupting US preparations for war on Iraq.
(SOURCE: nw/AFP_ENG_20030305.0918, sentence 18)

(S40) *Are you* going to <u>demonstrate</u> against them in the streets?
(SOURCE: bc/CNN_CF_20030304.1900.04, sentence 34)

(S41) the Iraqi leader calling on people to be resolute, to be strong, and saying that *if* they put up a good <u>fight</u>, then the coalition would eventually back down.
(SOURCE: bc/CNN_IP_20030405.1600.00-2, sentence 32)

(S42) But by and large, the Republican Guard has *either* been <u>destroyed</u> or has withdrawn.
(SOURCE: bc/CNN_IP_20030403.1600.00-1, sentence 29)

(S43) *I don't know if* they've given him the <u>death</u> penalty yet
(SOURCE: cts/fsh_29302, sentence 174)

(S44) French authorities are *investigating* <u>transfers</u> of $15 million from Swiss banks to Paris accounts in Suha's name at the Arab Bank and at BNP Paribas Bank, a French bank, TIME reports. Senior Palestinian security officials tell TIME that Arafat also shipped money to the gunmen of the Aqsa Martyrs Brigades.

(S45) He ***was to have*** <u>discussed</u> his objections with U.S. President George W. Bush this week
(SOURCE: nw/APW_ENG_20030519.0367, sentence 41)

(S46) Apart from his ID card, also found in the wreckage were the remains of a letter allegedly outlining his ***plan*** to carry out a suicide <u>attack</u>
(SOURCE: nw/AFP_ENG_20030305.0918, sentence 20

(S47) The forces encountered limited resistance along the way and are now ***poised*** for the final <u>assault</u> not far from the edge of the city
(SOURCE: bc/CNN_IP_20030403.1600.00-1, sentence 21)

(S48) Well, ***when*** they do finally <u>enter</u> Baghdad, ...
(SOURCE: bc/CNN_IP_20030403.1600.00-3, sentence 105)

(S49) That blast turned out to be a rocket-propelled grenade, but there's plenty of artillery and mortar <u>fire</u> to come.
(SOURCE: bc/CNN_IP_20030404.1600.00-2, sentence 25)

(S50) And Iraqi officials ***promise*** to <u>take</u> journalists to this particular hospital, the Yarmuk Hospital.
(SOURCE: bc/CNN_IP_20030405.1600.00-2, sentence 55)

(S51) The spate of bombings underscored how ***difficult*** it will be to carry out the U.S.-backed ``road map'' plan, a three-stage prescription for ending violence immediately and <u>setting up</u> a Palestinian state by 2005
(SOURCE: nw/APW_ENG_20030519.0367, sentence 11)

(S52) Word comes from the grand Ayatollah Sistani that he's ***willing*** to <u>meet</u> the American commander.
(SOURCE: bc/CNN_IP_20030403.1600.00-3, sentence 56)

(S53) And obviously we're going to ***pray*** that there's fewer <u>casualties</u>
(SOURCE: bc/CNN_CF_20030304.1900.04, sentence 30)

(S54) Whoever wants to visit Arafat can <u>visit</u> Arafat but he ***won't be allowed*** to <u>meet</u> senior Israeli officials,'' said a senior Sharon adviser, Raanan Gissin
(SOURCE: nw/APW_ENG_20030519.0367, sentence 29)

(S55) This is very worrying and i believe the israeli army should take special measures to prevent this. the message from the israeli government is that its soldiers are not targeting journalists, but that journalists who <u>travel</u> to places where there could be live fire exchange between israeli forces and palestinian gunmen have a responsibility to take greater precautions.
(SOURCE: bn/CNN_ENG_20030506_160524.18, sentence 26)

2.1.4 Other

This value should be assigned to any events whose modality is not clear, or does not fall into one of the four categories described above. This value should be used as sparingly as possible.

<u>Examples</u>

(S56) It's a life and *death* situation

2.2 Polarity

This dimension aims to capture whether the event describes a positive or negative situation.

2.2.1 Positive

Where there is no indicated negation of the event.

2.2.2 Negative

Where the event has been negated, or is presented in a negative context.

<u>Typical Evidence</u>

Always indicated through an explicit word or phrase in same sentence as event. Typical indicators are as follows:

- The most common means of expressing negation is through the use of the words *not* or *no* (S58-S59)

- Other words can also be used to express the fact that an event did not take place, when occurring in certain contexts. Examples include *cancel*, *stop*, *forbid* (S60-S61)

- It should be noted that some examples may at first look as though they are denoting negative polarity, when actually they are expressing subjectivity towards the event. See section 2.3 for examples of these.

<u>Example Sentences</u>

    (S58) ***No*** *injuries* reported
        (SOURCE: bc/CNN_IP_20030329.1600.00-4, sentence 41)

    (S59) No, I ***don't*** *demonstrate* against anybody during a war
        (SOURCE: bc/CNN_CF_20030304.1900.04, sentence 38)

    (S60) Where did it happen that a president was ***forbidden*** from *moving* out of his office?
        (SOURCE: nw/APW_ENG_20030519.0367, sentence 36)

    (S61) He ***cancelled*** his *trip* after Sunday's bus bombing.

2.3 Subjectivity

Many sentences express some attitude towards the event. An asserted or presupposed event can be either praised or condemned. A hypothetical event can be planned, proposed, wished for, or alternatively feared. These subjective attitudes may include, but are not limited to, the following:

Commands
Requests
Threats
Proposals
Discussions
Desires
Hopes
Fears
Promises
Plans
Intentions

Subjectivity is normally expressed by a specific cue phrase in the context of the event trigger, although the event trigger itself may sometimes denote subjectivity/

NOTE: Subjectivity value other than the default value (i.e., "neutral") should ONLY be assigned in cases where the validity of the subjectivity can be verified. Verifiable cases include the following:

1. Where the subjectivity expressed in the author's own subjectivity (e.g., S56)
2. Where the subjectivity is provided in a quote provided from a named source (S58-S59)
3. Where the wording of the report is such where it is clear that a named source has directly said what follows, even if it is not directly quoted. Examples would include verbs like "think", "support", "oppose", "condemn", etc. (S54, S55, S57, S60, S61)

In cases 2 and 3, the *Source-Type* of the event should be set to "involved" or "Third party" (as appropriate, see section 2.4).

Subjectivity should NOT be annotated where a third party is speculating about the subjectivity of somebody else. As an example, consider the (S62)

(S62) israel and the united states say arafat has not been tough enough on *terrorism*
(SOURCE: bn/CNN_ENG_20030614_173123.4, sentence 28)

Here, the Israel and the United States are implying that Arafat has a positive subjectivity towards terrorism, there is no evidence that he has said this directly. However, what we can conclude directly from the sentence is that Israel and the United States have negative subjectivity towards terrorism, since we know that they said this about Arafat.

2.3.1 Positive

This value should be assigned if the information sources of the event evaluates the event as good for themselves, for social groups with whose interests they identify, or for the wider community, whether or not the event could be considered harmful to others.

Typical Evidence

- Words indicating that the event is anticipated, wanted, hoped for or encouraged. Examples would include verbs like *promise, want, pray, urge, allow*, *support, plan, intend*, etc. and adjectives like *willing* (S63 – S65).

- Adjectives providing an assessment of the relative ease with which an event can be achieved, e.g. *difficult, easy, hard* (S65). In many cases, where such adjectives occur, there is a presupposition that the event is actually "wanted", no matter how difficult it may be.

  (S63) And Iraqi officials ***promise*** to <u>take</u> journalists to this particular hospital, the Yarmuk Hospital.
  (SOURCE: bc/CNN_IP_20030405.1600.00-2, sentence 55)

  (S64) Word comes from the grand Ayatollah Sistani that he's ***willing*** to <u>meet</u> the American commander.
  (SOURCE: bc/CNN_IP_20030403.1600.00-3, sentence 63)

  (S65) And obviously we're going to ***pray*** that there's fewer <u>casualties</u>
  (SOURCE: bc/CNN_CF_20030304.1900.04, sentence 30)

  (S66) The spate of bombings underscored how ***difficult*** it will be to carry out the U.S.-backed ``road map'' plan, a three-stage prescription for ending violence immediately and <u>setting up</u> a Palestinian state by 2005
  (SOURCE: nw/APW_ENG_20030519.0367, sentence 11)

  (S67) Hamas said Monday it has ***no intention of halting*** <u>attacks,</u> despite Egypt's efforts to have Palestinian militant groups agree to a one-year suspension of shootings and bombings.
  (SOURCE: nw/APW_ENG_20030519.0367, sentence 21)

Let us consider (S67) more closely. At first glance, it may seem that "halting" is negating the event whose trigger is "attacks". However "attacks" is a presupposed, ongoing event. The phrase "no intention of halting" is actually expressing the fact that Hamas has positive subjectivity towards these attacks, as they do not want the attacks to stop.

2.3.2 Negative

Negative subjectivity applies when an event is evaluated as bad or harmful from the perspective of

the source.

<u>Typical Evidence</u>

- Words indicating that the event is threatened, feared, unwanted, unapproved or considered harmful. Examples would include verbs like *threaten, condemn, boycott, oppose* and adjectives like *bad* and *grave* (S68 – S71).
- Sometimes, event triggers themselves may express strong negative subjectivity. Examples include *terrorism, genocide, holocaust, massacre, and ambush* (S72 – S73)

<u>Example sentences</u>

(S68) "We ***condemn*** all <u>*attacks*</u> against civilians including today's attack in Haifa," said information minister Yasser Abed Rabbo
(SOURCE: nw/AFP_ENG_20030305.0918, sentence 25)

(S69) Israeli government spokesman Avi Pazner called the latest deadly blast a "***very grave attack***" and said the Jewish state would take "vigorous action against the terrorist organisations".
(SOURCE: nw/AFP_ENG_20030305.0918, sentence 29)

(S70) In response the spate of bombings, Israel threatened to ***boycott*** foreign envoys who <u>*meet*</u> with Yasser Arafat
(SOURCE: nw/APW_ENG_20030519.0367, sentence 19)

(S71) Israel ***holds*** the Palestinian leader ***responsible*** for the latest <u>*violence*</u>, even though the recent attacks were carried out by Islamic militants.
(SOURCE: nw/APW_ENG_20030519.0367, sentence 20)

(S72) The truth is that, as director of U.N. peacekeeping, Annan personally refused requests to authorize U.N. peacekeepers in Rwanda to seize weapons and prevent <u>*genocide*</u>
(SOURCE: wl/FLOPPINGACES_20050203.1953.038, sentence 36)

(S73) But for five and a half years, he noted, Annan "refused to accept any responsibility for the Rwandan <u>*holocaust*</u> until Mr. Gourevitch and others revealed that less than 5,000 U.N. troops could have stopped the killings if Mr. Annan had not closed his eyes." It is astonishing that a journalist for NBC New would close her eyes to Annan's role.

(SOURCE: wl/FLOPPINGACES_20050203.1953.038, sentence 36)

(S74) Hamas said Monday it has no intention of halting <u>*attacks,*</u> despite Egypt's efforts to have Palestinian militant groups ***agree to a one-year suspension*** of <u>*shootings*</u> and <u>*bombings*</u>
(SOURCE: nw/APW_ENG_20030519.0367, sentence 21)

Let us consider (S74) in more detail. The word *suspension* firstly appears to negate the shootings and bombings. However, with closer consideration, it can be understood that the shootings and bombings are ongoing, and that Egypt has demonstrated that they want them to stop. Therefore, we can annotate that Egypt has negative subjectivity towards the shootings and bombings.

2.3.3 Multi-valued

In some cases, there is both positive and negative subjectivity expressed towards a particular event, by different sources. In these cases, the "multi-valued" value should be assigned, and each of the sources should be annotated and linked to the event. However, it is not necessary to identify which source corresponds to which subjectivity.

Example Sentences

> (S75) hamas and other palestinian militant group ***rejected*** abbas' ***call*** to end their <u>attacks</u> with a deadly raid on an israeli army outpost in gaza
> (SOURCE: bn/CNN_ENG_20030614_173123.4, sentence 44)

> (S76) Hamas vowed to continue its attacks, while the Palestinian Authority ***accused*** Israel of trying to ***disrupt*** top-level Palestinian <u>meetings</u> to discuss reforms of Arafat's administration.
> (SOURCE: nw/AFP_ENG_20030305.0918, sentence 38)

In (S75), Abbas is calling to end the attacks, and so he has negative subjectivity towards them. On the other hand, Hamas and other Palestinian militant groups are rejecting this call by Abbas. In other words, they have positive subjectivity towards the attacks.

In (S76), Israel has negative subjectivity towards the meetings, as they are trying to disrupt the meeting. However, since the Palestinian Authority accused Israel of this disruption, they have positive subjectivity towards the meeting.

2.3.4 Neutral

The default value for subjectivity is neutral, which applies if an author or a disinterested party merely reports an event, or speculates whether it will occur, without any specific evaluation about whether this is good or bad.

2.4 Source

This dimension encodes to the source or origin of the knowledge being expressed by the event. Specifically, we wish to distinguish between events that can be attributed to the speaker/author, to those who were/are closely involved in the event, or to other third parties not closely involved in the event. Furthermore, we annotate and link the named sources to the event if present in the text. All of the above are important to allow potential distinctions to be made in the interpretation of events. For example, people who are closely involved in an event may have more biased views than those who are more distanced from an event. Likewise, an event that can be traced to a

particular information source will often be considered more reliable than if the source is not named (e.g. the event is stated based on hearsay).

The three categories of the *Source* dimension are defined as follows:

2.4.1 Author

Assigned to events that are presented as corresponding information provided by the author, or representing their own information point of view. This is the default value, assigned to events unless there is any evidence for one of the other values.

<u>Typical Evidence</u>

- Any event where there is a lack of explicit evidence/clues about the information source in the context of the event.

2.4.2 Involved

This value indicates that the information expressed by the event is attributed to a specified third party who is someway involved or has close links to the actions described by the event. This may be an individual, group, government, political or terrorist organisation who is clearly involved in the event.

<u>Typical Evidence</u>

- A named source will always be present and should be annotated.
- Always indicated through an explicit word or phrase. These may include the following:
  - o Reporting verbs like *say* and *report* (S77-S78), or expressions like *word from* (S79).
  - o Verbs that are used to denote speculated events, e.g. *believe* (S80)
  - o Verbs that are used to denote positive or negative subjectivity towards the event, e.g. *condemn, praise, promise,* etc. (S81)

<u>Example sentences</u>

(S77) <u>Military officials</u> ***say*** 39 Americans have been <u>*killed*</u> in combat.
(SOURCE: bc/CNN_IP_20030402.1600.00-2, sentence 13)

(S78) <u>The Americans</u> ***say*** they've <u>*destroyed*</u> two divisions of the elite Republican Guard
(SOURCE: bc/CNN_IP_20030403.1600.00-1, sentence 46)

(S79) ***Word comes from*** <u>the grand Ayatollah Sistani</u> that he's willing to <u>*meet*</u> the American commander
(SOURCE: bc/CNN_IP_20030403.1600.00-3, sentence 63)

(S80) But <u>the Pentagon</u> ***believes*** some Iraqi Republican Guard units have <u>*withdrawn*</u> to the capital for what could be their last stand.
(SOURCE: bc/CNN_IP_20030403.1600.00-1, sentence 15)

(S81) <u>EU foreign policy supremo Javier Solana</u> likewise ***slammed*** the <u>*attack*</u>
      (SOURCE: nw/AFP_ENG_20030305.0918, sentence 33)

2.4.3 Third Party

This value indicates that the information expressed by the event can be attributed to a third party source without does not have direct links with participants in the event. Examples include media agencies, police, etc.

<u>Typical Evidence</u>

- There will always be some type of explicit specification that the information comes from a source other than the author. This can be done in several ways.
    - There may be an explicit phrase denoting the source. In this case, typical clue expressions will be the same as those introduced above for *Involved* sources (e.g., verbs used for reporting, denoting speculation or subjectivity). The explicit phrase denoting the source could be:
        - A phrase naming the source (S82-S83), in which case the named source should be annotated and linked to the event as a *named source.*
        - A more vague phrase (e.g., *Some sources report that ....*). In this case, the vague phrase should not be annotated as an *unnamed source*.
    - There may be no explicit phrase corresponding to the source. However, there are other ways of denoting that the information comes from a third party source, including the following:
        - Reporting verbs *say* or *report,* used in the passive voice (S84)
        - Verbs like *hear that* in the active voice (S85).

- <u>Example sentences:</u>

(S82) Two 13-year-old children were among those <u>*killed*</u> in the Haifa bus bombing, <u>Israeli public radio</u> ***said***
      (SOURCE: nw/AFP_ENG_20030305.0918, sentence 13)

(S83) The explosion <u>*killed*</u> the attacker and four shoppers, <u>police</u> said
      (SOURCE: nw/ nw/APW_ENG_20030519.0367, sentence 16)

(S84) Of the 27 British troops, in all who have been killed, at least 19 of them ***were said to be*** victims of friendly <u>*fire*</u> or accidents.
      (SOURCE: bc/CNN_IP_20030404.1600.00-2, sentence 70)

(S85) I ***hear*** though, that he's going to be <u>*killed*</u> for a crime in California in like, thirty five years or something
      (SOURCE: cts/fsh_29302, sentence 165)

2.5 Genericity

This dimension is concerned with determining whether the event refers to a specific event or series or events, or whether it refers to a more general situation that's occurs habitually. This is part of the original ACE annotation, but the values will be reviewed.

2.5.1 Specific

A specific event is either an event that describes either a single occurrence of an action, or a finite set of such occurrences, which have a definite start and end time.

Typical Evidence

- No explicit clues, but rather determined by the way in which the event is presented.
- Single events presented as being instantaneous (S86)
- Single events presented having duration, but which have finished or which will are expected to be of finite duration (S87)
- Sets of events which are presented to be of finite duration (S88-S90)

Examples

(S86) Eight people, including a pregnant woman and a 13-year-old child were _killed_ in Monday's Gaza _raid_, provoking US-led international calls for Israeli restraint
(SOURCE: nw/AFP_ENG_20030305.0918, sentence 40)

(S87) The bus was ripped to shreds while _travelling_ between a residential area and Haifa university.
(SOURCE: nw/AFP_ENG_20030305.0918, sentence 15)

(S88) There was no immediate claim of responsibility for Monday's attack at the Shaarei Amakim mall in the working class town of Afula, which has been _targeted_ repeatedly by Palestinian militants because of its proximity to the West Bank.
(SOURCE: nw/APW_ENG_20030519.0367, sentence 12)

(S89) He told AFP that Israeli intelligence had been dealing with at least 40 tip-offs of impending _attacks_ when the Haifa bus was blown up
(SOURCE: nw/AFP_ENG_20030305.0918, sentence 30)

(S90) The spate of _bombings_ underscored how difficult it will be to carry out the U.S.-backed "road map'" plan, a three-stage prescription for ending violence immediately and _setting up_ a Palestinian state by 2005
(SOURCE: nw/APW_ENG_20030519.0367, sentence 11)

2.5.2 Generic

An event is generic if it is something that is presented as happening continuously or habitually, or if there is no clear indication of whether the event is of finite duration.

Typical Evidence

- No explicit clues, but rather determined by the way in which the event is presented.
- Events which are presented as happening continuously or habitually (S91)
- Sets of events that are ongoing, and which are presented as having no clear end. (S92-S93)

Examples

(S90) Palestinian chief negotiator Saeb Erakat said international intervention was needed to break the cycle of *violence*
(SOURCE: nw/AFP_ENG_20030305.0918, sentence 39)

(S91) The bus attack came just two days after Israeli forces staged a bloody raid into a refugee camp in central Gaza targeting a founding member of Hamas, which has spearheaded suicide *bombings* against the Jewish state.
(SOURCE: nw/AFP_ENG_20030305.0918, sentence 37)

(S92) Hamas vowed to continue its *attacks*, while the Palestinian Authority accused Israel of trying to disrupt top-level Palestinian meetings to discuss reforms of Arafat's administration.
(SOURCE: nw/AFP_ENG_20030305.0918, sentence 38)

2.6 Tense

This dimension determines the time when the event took place, with respect to the time of reporting or textual anchor time, i.e., the time when the event is written by the author or spoken by the speaker of the article/broadcast. This is part of the original ACE annotation, but values will be reviewed.

It is important to note that value of this dimension is determined by considering the *textual context* of the event., i.e., it is not necessarily determined by the tense of the event trigger, especially as triggers can belong to parts of speech other than verbs.

2.6.1 Past
This value is used for those events that occur prior to the textual anchor time.

Typical Evidence

- Anchor is a verb in the past tense (S93)

- Anchor is a noun in a past tense context (e.g. the main verb in the clause in which is contained is in the past tense) (S94-S95)

Examples

(S93) Palestinian medical sources said 60-year-old Mohammed al-Biyari was _killed_ in his home near Jabaliya refugee camp by the rocket fire and another _wounded_.
(SOURCE: nw/AFP_ENG_20030305.0918, sentence 11)

(S94) Without claiming responsibility, Hamas and another hardline group, Islamic Jihad, said the _blast_ was revenge for a string of bloody Israeli army raids into the Gaza Strip in recent weeks.
(SOURCE: nw/AFP_ENG_20030305.0918, sentence 23)

(S95) In the January _attack_, two Palestinian _suicide_ bombers _blew_ themselves up in central Tel Aviv, _killing_ 23 other people.
(SOURCE: nw/AFP_ENG_20030305.0918, sentence 28)

2.6.2 Future

This value used for those events that have not yet occurred at the textual anchor time.

NOTE: Future events should NOT be assigned MODILTY=Asserted. Normally, the modality value will be "Speculated", as the source of the information (i.e., the author/speaker) is speculating that the event will happen in the future.

Typical Evidence

- An event that is expected to occur, denoted using either the future tense or a phrase the explicitly refers to the future, as in _to come_ (S83)
- The event is in a speculated or hypothetical context, as denoted by the presence of an appropriate clue word (S84-S85)

Examples

(S96) That blast turned out to be a rocket-propelled grenade, but there's plenty of artillery and mortar _fire **to come**_.
(SOURCE: bc/CNN_IP_20030404.1600.00-2, sentence 25)

(S97) translator: hamas will not drop our weapons, even **if** all leaders are _assassinated_
(SOURCE: bn/CNN_ENG_20030614_173123.4, sentence 53)

(S98) Yeah. Hopefully, you know, if he's guilty **hopefully**, you know, they find him _guilty_ -- and hopefully they have evidence, you know.
(SOURCE: cts/fsh_29272, sentence 151)

2.6.3 Present

This value is used for those events that either occur at the textual anchor time, or are ongoing the textual anchor time.

Typical evidence

- Anchor may be a verb or other part of speech (e.g., a noun). Indeed, the trigger is a noun in all of the examples below (S99 –S101), showing the importance of considering context as well as the trigger itself.
- Context shows that the event or series of events denoted by the anchor started at some point in the past, but does not mention an end (S99)
- Context provides explicit information that the event or series of events is ongoing (S100-S101)

> (S99) Malls and other public places have significantly stepped up security since the outbreak of Israeli-Palestinian *fighting* in September 2000
> (SOURCE: APW_ENG_20030519.0367, sentence 15)

> (S100) In a renewed mediation attempt, Egyptian envoys have been holding *meetings* in Gaza and Damascus in recent days with leaders of the militant groups.
> (SOURCE: APW_ENG_20030519.0367, sentence 22)

> (S101) The latest wave of *attacks* began Saturday evening when a Hamas bomber blew himself up in a square in the West Bank city of Hebron, killing an Israeli settler and his pregnant wife.
> (SOURCE: nw/APW_ENG_20030519.0367, sentence 43)

2.6.4 Unspecified

Whenever the tense of an event is not specified or cannot be determined from the context, the *Unspecified* value is used.

Typical evidence

- This value will often be used for general statements that do not refer to specific events or series of events (S102)
- In other cases, it may not be clear from the sentential context whether an event is ongoing or happened in the past (S103)

Examples

> (S102) Sometimes reporters, like "The Wall Street Journal's Danny Pearl, *die* in covering a different kind of war.

(SOURCE: bc/CNN_IP_20030408.1600.04, sentence 29)

(S103) The _casualty_ count helps put the ongoing debate about journalistic neutrality into a
    sharper light.
(SOURCE: bc/CNN_IP_20030408.1600.04, sentence 31)