**ISO/TC37/SC 4 N412 rev01**
**(N269 rev12)**

Date: 2008-01-08

**ISO/CD 24617-1**

ISO/TC 37/SC 4/WG 2

Secretariat: KATS

# Language resource management — Semantic annotation framework (SemAF) — Part 1: Time and events

*Gestion des ressources linguistiques — Annotation Semantique (SemAF) — Partie 1: le temps et les événements*

Document type: International Standard
Document subtype: Not applicable
Document stage: (30.20) Committee Draft
Document language: E


Reference Number
ISO/CD 24617-1(E)

*ISO copyright office*
*Case postale 56 CH-1211 Geneva 20*
*Tel. +41 22 749 01 11*
*Fax +41 22 749 09 47*
*E-mail copyright@iso.org*
*Web www.iso.org*

*Published in Switzerland*

## Contents

Page

## Foreword

ISO (The International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 3.

Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75% of the member bodies casting a vote.

International Standard 24617-1 entitled *Language resource management - Semantic annotation framework (SemAF) - Part 1: Time and events* was prepared by Technical Committee ISO/TC 37, *Terminology and other language and content resources*, Subcommittee SC 4, *Language resource management*, Working Group WG 2 *Representation schemes* in collaboration with the TimeML Working Group.

NOTE The TimeML Working Group is headed by James Pustejovsky, jampesp@cs.brandeis.edu, Brandeis University.

The main parts of the ISO 24617-1 consist of:

- Scope
- Normative references
- Terms and definitions
- Motivation and requirements
- Basic concepts and metamodel
- Specification of ISO-TimeML, a formal annotation (specification) language for events and temporal expressions in natural language
- Semantics of ISO-TimeML

The proposed international standard provides 10 annexes. One is a normative annex providing core annotation guidelines. The rest are informative annexes. Four annexes provide annotated examples, one being a set of completely annotated examples of English and the other consisting examples from various languages other than English, namely Chinese, Italian and Korean. Two other annexes deal with ISO-TimeML DTD and schema. Information is given on past and current activities on temporal and event annotation and also on tools and templates. In the final annex, editorial or authorship information is provided with a list of editors and contributors along with meeting records.

## Introduction

This standard proposal results from the agreement between the TimeML Working Group and the ISO committee TC 37/SC 4/WG 2 and TDG 3 (ad hoc Thematic Domain Group 3 Semantic Content) that a joint activity should take place to accommodate the two existing documents for annotating temporal information, *TimeML 1.2.1* and *TimeML Annotation Guidelines*, into ISO international standards. This work should lead to the achievement of two objectives:

- Modification of the two documents in conformance to the ISO international standards.
- Verification of the annotation guidelines for a wide coverage of multilingual resources.

It should be noted that this standard provides normative guidelines not just for temporal information, but also for information content in various types of events in English as well as other languages.

# Language resource management — Semantic annotation framework (SemAF) — Part 1: Time and events

## 1 Scope

Temporal information in natural language texts is an increasingly important component to the understanding of those texts. This proposed international standard provides a formal specification language called *ISO-TimeML* for temporal information markup and includes a specific set of guidelines for such markup.

Adopting XML as its formal language, the SemAF/Time annotation standard provides a formalized markup language called ISO-TimeML with a systematic way to extract and represent temporal information, as well as to facilitate the exchange of temporal information, both between operational language processing systems and between different temporal representation schemes. The use of guidelines has been fully attested with examples from the TimeBank corpus, a collection of over 180 documents that have been annotated by TimeML before the current version of ISO-TimeML was formulated. It also provides an appropriate DTD and schema files as annexes.

## 2 Normative references

For this international standard there are five main normative references:

- ISO 8601:2004 Data elements and interchange formats – Information exchange – Representation of dates and times
- ISO 8879:1986 (SGML) as extended by TC2 (ISO/IEC JTC 1/SC 34 N029: 1998-12-06).
- ISO 19757-2, Document schema definition language, part 2.
- ISO 24610-1:2006 Language resource management - Feature structures - Part 1: Feature structure representation.
- ISO CD 24610-2 Language resource management - Feature structures - Part 2: Feature system declaration.

## 3 Terms and definitions

NOTE This clause was provided by Amanda Schiffrin and Harry Bunt of Tilburg Uniersity as part of their 2007 LIRICS project report.

For the purposes of this document, the terms and definitions in ISO 8601:2004 apply. This list is to clarify the terminology relating to the metamodel, specification, and semantics of ISO-TimeML and the normative annex of Core annotation guidelines used throughout this part of ISO 24617-1. Terminology derived from XML and other formal languages as well as from general temporal logics is not defined here.

**3.1**
**event**
something that can be said to obtain or hold true, to happen or to occur. Also referred to as eventuality.

NOTES
1. Adapted from Weiner and Simpson (1996) and Pustejovsky et al. (2004).
2. This is a very broad notion of event, also known in the literature as 'eventuality', which includes all kinds of actions, states, processes, etc. It is not to be confused with the more narrow notion of event as something that happens at a certain point in time (such as the clock striking 2, or waking up) or during a short period of time (such as laughing).

**3.2**
**instant**
point in time with no interior points

NOTE Time is often viewed as a straight line from minus infinity to plus infinity. On this view, time is formed by an infinite sequence of points. An instant can also be seen as an infinitesimally small interval. Cf. OWL-Time Ontology, "instant"; http://www.w3.org/TR/owl-time/.

**3.3**
**temporal interval**
uninterrupted stretch of time, with internal point structure. Cf. OWL-Time Ontology, "interval"; http://www.w3.org/TR/owl-time/.

NOTES
1. Adapted from WordNet.
2. Time is often viewed as a straight line from minus infinity to plus infinity. A temporal interval is a part of that line without any holes, containing all the points between its beginning and its end.
3. In mathematics, an important issue is whether an interval includes its beginning and its end (is 'closed') or not (is 'open' or 'half-open'). In natural language descriptions of intervals this may also be relevant, as when describing an interval in terms of a number of days, but not with the same granularity as in mathematics.

**3.4**
**beginning**
**instant** at which a **temporal interval** begins

NOTE Adapted from Hobbs and Pan (2004).

**3.5**
**end**
**instant** at which a **temporal interval** ends

NOTE Adapted from Hobbs and Pan (2004).

**3.6**
**temporal unit**
element in a time amount that quantifies the length of a **temporal interval** or a set of **temporal intervals**

NOTES
1. Adapted from Bunt (1985).
2. In measurement systems various units are defined for different purposes. Small units such as seconds and minutes are defined to measure small temporal intervals; as one may want to avoid working with big numbers, for larger temporal intervals units such as week, year, decade, and century are defined.
3. The amount of a temporal unit is called a measure.

**3.7**
**period**
another term for **temporal interval**, replaces the notion of **3.8**
**duration**
i

n previous versions

**3.9**
**time amount**
A quantity of time, measured by temporal units over intervals

NOTES
1. Adapted from Bunt (1985).
2. A time amount is a measure of time that can be expressed in terms of a number of temporal units, such as half an hour, or 30 minutes.

**3.10**
**point of speech**
**time unit** at which a given utterance occurs

NOTES
1. Adapted from Reichenbach (1947).
2. The notion of Point of speech is needed. In order to interpret tense, one needs to define anchor points in time of which point of speech is one. For example, in "Arthur smiled" the point of speech is the time that the utterance is made.
3. For a document as a whole, this may be considered to be the same as the document creation time.

**3.11**
**point of event**
**instant** at which the **event** mentioned in a given utterance occurs

NOTES
1. Adapted from Reichenbach (1947).
2. Next to a 'point of speech', a 'point of event' also needs to be defined in order to interpret tense. For example, in "Arthur smiled" the temporal location of the point of event can be defined as being prior to the point of speech.

**3.12**
**point of reference**
**instant** of temporal perspective on the **event** in a given utterance

NOTES
1. Adapted from Reichenbach (1947).
2. To locate certain tenses in time a third anchor point is also required, defined as the point of reference.
3. Example: "Arthur will have gone by then", where the point of speech is now, the point of event is some time.

**3.13**
**point of text**
**instant** at which reported speech is anchored.

NOTE It is the point of time considered in the text of the speech. So for example, when a person is telling a story, it is not enough to know the point of the speech itself (the document creation time), but the point at which the speech in the story is taking place.

**3.14**
**time zone**
area of the Earth that has adopted the same standard time

NOTES
1. Usually referred to as the local time.
2. Most time zones are exactly one hour apart, and by convention compute their local time as an offset from Greenwich Mean Time.
3. There has already been extensive work carried out on the consistent representation of date and time (including time zone information) within ISO: see ISO 8601:2004.

**3.15**
**temporal ordering relation**

relation that determines how objects are ordered in time

NOTES
1. There is a limited number of ways to order objects which are collectively called ordering relations.
2. Examples: precedence, simultaneity.

**3.16**
**tense**
way that languages express the time at which an **event** described by a sentence occurs

NOTE Characterized as a property of a verb form. Noun forms will not be said to exhibit tense but rather temporal markers.

**3.17**
**ALINK**
linking tag that represents a phase relation between an aspectual verb (or morpheme) and a predicate denoting an **eventuality**

**3.18**
**SLINK**
linking tag that represents a subordinating relation between two **eventualities**

**3.19**
**TLINK**
linking tag that represents a temporal relation between two temporal entities: namely, between two eventualities, two temporal expressions, or between a temporal expression and an eventuality.

NOTES
1. Adapted from Pustejovsky et al. (2004).
2. Some ordering relations cannot be expressed by an ordering relation between two eventualities because a signal, like a temporal preposition, complicates the ordering or there is an ordering relation between a temporal signal and an eventuality.

## 4 Overview

An understanding of temporal information is needed to better understand natural language texts in general. Previous work in time stamping is a step in the right direction, but to fully appreciate the complexity of a text with respect to time, the ability to order eventualities and temporal expressions is needed. This standard defines ISO-TimeML, a markup language for time, which has been specifically designed for this task.

ISO-TimeML annotates all expressions having temporal import, broadly categorized as temporal expressions and eventualities (situations, events, states, and activities). Temporal expressions and eventualities participate in temporal relationships (e.g., “before", “simultaneous"), subordinating relationships (e.g., “intensional", “factive"), and aspectual relationships (e.g., “initiates", “continues"). ISO-TimeML provides an additional expressive capability of capturing and representing the complexities of these relationships.

TimeML, the precursor of ISO-TimeML, is already in use in a number of applications focusing on analysis (manual and automatic) of news articles. The TimeBank corpus contains approximately 185 such documents and has been validated against the most recent version of TimeML. The resulting output of a TimeML annotated document is in XML, which allows for general XML validation methods to be used. In addition to supporting interoperability, among different temporal representation schemes, TimeML has been shown adequate to support a mapping from the temporal information in a text to its formal representation in a Web Ontology Language such as OWL-Time.

Unlike prior event annotation schemes, ISO-TimeML’s somewhat unique definition of an event does not limit the standard’s applicability to other natural language genres. An ISO-TimeML event is simply something that can be related to another event or temporal expression using an ISO-TimeML relationship — thus an ISO-

TimeML-compliant representation can be adapted (derived) from the full standard specification, appropriate to different genres, styles, domain, and applications. Future work will involve applying the standard in such different contexts, and formulating guidelines and principles for appropriate use of ISO-TimeML in a variety of language engineering environments.

## 5 Motivation and requirements

The identification of temporal and event expressions in natural language text is a critical component of any robust information retrieval or language understanding system, and recently this has become an area of intense research in computational linguistics and Artificial Intelligence. The importance of temporal awareness to question answering systems has become more obvious as current systems strive to move beyond keyword and simple named entity extraction. Named entity recognition (Chinchor et al, 1999) has moved the fields of information retrieval and information exploitation closer to access by content, by allowing some identification of names, locations, and products in texts. Beyond such metadata tags, however, there is only a limited ability at marking up text for real content. Besides semantic tagging and semantic role labeling, one of the major problems that has not been solved is the recognition of events and their temporal anchorings in text. Events are naturally anchored in time within a narrative. Without a robust ability to identify and extract events and their temporal anchoring from a text, the real aboutness of the article can be missed. Moreover, since entities and their properties change over time, a database of assertions about entities will be incomplete or incorrect if it does not capture how these properties are temporally updated. To this end, event recognition drives basic inferences from text.

As it happens, however, much of the temporal information in an article or narrative is left implicit in the text. The exact temporal designation of events is rarely explicit and many temporal expressions are vague at best. A crucial first step in the automatic extraction of information from such texts, for use in applications such as automatic question answering or summarization, is the capacity to identify what events are being described and to make explicit when these events occurred.

Another important point is that, although most of information on the web is in natural language, it is unlikely that it will ever be marked up for semantic retrieval, if that entails hand annotation. Natural language programs will have to process the contents of web pages to produce annotations. Remarkable progress has been made in the last decade in the use of statistical techniques for analyzing text. However, these techniques for the most part depend on having large amounts of annotated data, and annotations require an annotation scheme. Hence, in addition to developing the necessary tools for temporal analysis, it is important to enable for seamless integration into existing and emerging ontologies (KR standards). For example, OWL and other event ontologies have the potential to become standards in various web-based reasoning and computing environments. The proposed standard and applications based on this work will ensure that such applications will continue to be useful in a variety of temporal analytic contexts..

Interest in temporal analysis and event-based reasoning has spawned a number of important workshops, particularly as applied to IE and QA tasks (cf. at COLING 2000; ACL 2001; LREC 2002; TERQAS 2002; TANGO 2003, Dagstuhl 2005). Significant progress has been made in these meetings, leading to developing a standard for a specification language for events and temporal expressions and their orderings (TimeML). While recent research in the broader community (as indicated, for instance, in the Dagstuhl 2005 seminar) highlights TimeML's status as an emerging standard, this workshop is not intended to focus on TimeML exclusively. Likewise, while the ultimate goal of temporal analysis is to facilitate reasoning about time and events, the formal aspects of this problem are being addressed by other meetings (see, for instance, the TIME 2006 Symposium). Many issues relating to temporal and event identification remain unresolved, however, and these issues ISO-TimeML has been designed to address. Specifically, four basic problems in event-temporal identification have been addressed in the design of ISO-TimeML:

(a) Time anchoring of events (identifying an event and anchoring it in time);

(b) Ordering events with respect to one another (distinguishing lexical from discourse properties of temporal ordering);

(c) Reasoning with contextually underspecified temporal expressions (temporal functions such as *last week* and *two weeks before*);

(d) Reasoning about the persistence of events (how long does an event or the outcome of an event last).

The specification language, ISO-TimeML, is designed to address these issues, in addition to handling basic tense and aspect features.

Reasoning about time is also one of the most important aspects of commonsense reasoning. Work on OWL-Time integrating information from annotated linguistic texts is a particularly important area, and this is explored in Hobbs and Pustejovsky (2003) and Hobbs and Pan (2004). Linking a formal theory for time with an annotation scheme aimed at extracting rich temporal information from natural language text is significant for at least two reasons. It will allow us to use the multitude of temporal facts expressed in text as the ground propositions in a system for reasoning about temporal relations. It will also constitute a forcing function for developing the coverage of a temporal reasoning system, as we encounter phenomena not normally covered by such systems, such as complex descriptions of temporal aggregates.

## 6 Basic concepts and metamodel

Regarding the temporal information in a document, a distinction can be made between (1) the temporal metadata, regarding when the document was created, published, distributed, received, revised, etc., and most importantly (2) the temporal properties of the events and situations that are described in the document. The former type of information is associated with the document as a whole; information of the latter type will be associated in annotations with parts of the text in the document, 'markables' such as phrases and sentences.

Temporal objects and relations have been studied from logical and ontological points of view; well-known studies include those by Allen (1984), Prior (1967), and more recently Hobbs and Pan (2004); see also the collection of papers in Mani et al. (2005). The most common view of time, that underlies most natural languages, is that time is an unbounded linear space running from a metaphorical 'beginning of time' at minus infinity to an equally metaphorical 'end of time' at plus infinity. This linear space can be represented as a straight line, the points of which correspond to moments in time; following Hobbs & Pan (2004) we will also use the term 'instant' to refer to points of time. From a mathematical point of view, the points on the time line are line segments of infinitesimally small size, corresponding to the intuition that a moment in time can in principle be determined with any precision that one may wish.

For linguistic and philosophical reasons, several classifications have been proposed of verbs describing various types of states or events, the Vendler classification being the best known (Vendler, 1967). For the annotation of temporal information in text, not only verbs with their tenses and temporal modifications should be considered, but also nouns, since nouns may also denote events and situations (*The meeting at twelve; The six o'clock news*). In the TimeML annotation language for temporal information, Pustejovsky et al. (2007) have proposed a classification of states and events into seven categories. In the literature a distinction is often made between 'states' and 'events', where the latter are characterized as occurring at a point in time or during a certain definite interval, whereas states may obtain for any indefinite stretch of time (*The Mediterranean Sea separates Europe from Africa*). On a terminological note: the term 'event' will henceforth be used as a generic term that also covers such notions as 'state', 'situation', 'action', 'process', etc.; this broad notion of event has also been termed 'eventuality' (Bach, 1986).

In reality nothing happens in infinitesimally small time; every event or state that occurs in reality (or in someone's mind) requires more than zero time, although natural languages offer speakers the possibility to express themselves as if something occurs at a precise instant (like *I will call you at twelve o'clock*). Since instants are formally a special kind of interval, a consistent approach to modelling the time that an event occurs is to always use intervals, where it may happen that the interval associated with a particular event is regarded as having zero length, and thus being an instant. This is reflected in the metamodel presented in Figure 1, which uniformly relates events with temporal intervals.

The length of an interval can also occur as temporal information in a text, as in *I used twelve hours to read that book* and *It takes seven minutes to walk to the station.* An expression like *seven minutes* does not denote an interval, but the length of an interval. It is the temporal equivalent of spatial distance (*7 kilometres*). To describe the length of a temporal interval one needs a unit of measurement, which may be combined with a numerical expression to obtain an amount of time. The metamodel presented below therefore includes the concept of an amount of time, related to intervals through the function 'length', and the auxiliary concepts of temporal units and real numbers. (Moreover, in the ISO-TimeML semantics, different temporal units are related through a conversion function, stipulating such things as 1 hour = 60 minutes; 1 day = 24 hours, etc. An amount of time can be characterized equivalently by as many pairs <numeral, temporal unit> as there are temporal units, the equivalence being defined through the numerical conversions between units (see Bunt, 1985).)

Regarding the temporal anchoring of events in time, it may be noted that the assocation of a temporal interval with an event does not necessarily mean that the event took place during every moment within that interval. When someone says *"I've been working on my presentation from 8.30 to 12 o'clock"*, that presumably does not mean that the speaker has been working on his presentation for every single moment between 8.30 and 12 o'clock; there must have been interruptions for having some coffee, going to the bathroom, etc. In such a case it is more accurate to anchor the event at the *time span* starting at 8.30 and ending at 12 o'clock, a 'time span' being understood as a period of time that may have 'holes', where the event was interrupted. The metamodel shown in Figure 1 does not distinguish time spans, but reflects the assumption that whether an event occurs during an interval with or without any interruptions can only be decided on a case by case basis, and is best modelled as a property of the temporal anchoring relation applied to a specific event.

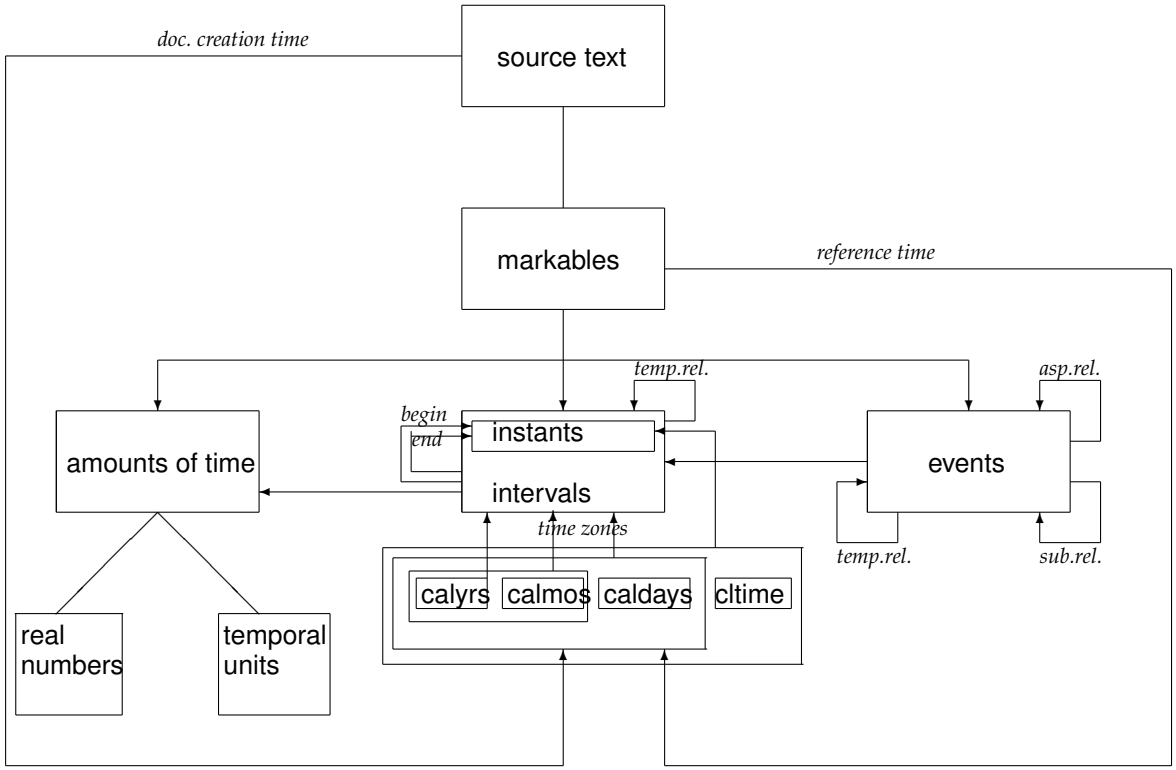


Figure 1. ISO-TimeML Metamodel

The Linguistic Annotation Framework (Ide & et al., 2003), a proposed ISO standard for linguistic annotation, insists on the use of *stand-off annotation*, i.e. the construction of annotations in separate files, separate from the document containing the primary language data, as contrasted with in-line annotation. Stand-off annotations

refer to specific locations in the primary data by addressing byte offsets, linguistic elements such as words, or times associated with recorded data, to which the annotation applies. Compared to in-line annotation, stand-off annotation has the advantages of respecting the integrity of the primary data and of allowing multiple annotations to be layered over a given primary document. Since semantic annotations typically occur at a relatively high level in a layered annotation structure, they do not necessarily refer directly to segments in the primary data, but may also refer to structures in other annotation layers. The generic term *markable* is used to refer to the entities that the annotations are associated with. There are two kinds of markables in ISO-TimeML: *event markables* and *time markables*, corresponding to segments of primary data that describe events, and to those that describe temporal entities or relations, respectively.

Markables are derived from documents, which will have certain metadata that are particularly important for the interpretation of temporal annotations. For interpretating the tenses of verb forms and adverbial temporal deixis in a text, for instance, one must know when the text was produced. This will often be defined by the document creation time, and more precisely by the combination of a creation time and a creation location, since the latter defines the time zone within which the creation time is precisely defined. In many documents the time and place of the document creation will be those of all the markables that may be derived from the document, but it may also happen that the text in a document introduces other times and places relative to which the annotations of the markables should be understood. A reasonable strategy would seem to be to assume that each markable has a time and place (or time zone), which by default is that of the document from which it derives. A time zone, like Greenwich Meat Time (GMT) can be seen as a way of segmenting the time line into named segments of particular lengths, such as (calendar) years, months, days, hours, and minutes. Accordingly, time zones shows up in the metamodel as functions mapping a calendar year ('2008'), a combination of a calendar year and a calendar month ('*May 2008'*), a date ('*May 25, 2008'*), or a date plus a clock time ('*May 25, 2008, 12.30 p.m.'*) onto a temporal interval (in the latter case, an instant).

A markable may refer to more than one, related event, as in *She started to laugh* (two aspectually related events); *John drove to Boston after the concert* (two temporally related events); or *Will you attend the meeting on Tuesday?* (one event having a subordination relation to another). For expressing such relations the metamodel includes the corresponding classes of relations showing up as inter-event links. Temporal relations between events may also be stated between intervals, hence they show up again there in the metamodel.

# 7 Specification of ISO-TimeML

## 7.1 Overview

The Linguistic Annotation Framework (Ide & et al., 2003) makes a fundamental distinction between the concepts of *annotation* and *representation*. The term 'annotation' is used to refer to the process of adding information to segments of language data, or to refer to that information itself. This notion is independent of the format in which this information is represented. The term 'representation' is used to refer to the format in which an annotation is rendered, for instance in XML, independent of its content. According to the Linguistic Annotation Framework (LAF), *annotations* are the proper level of standardization, not representations. The present standard therefore defines a markup language for annotating documents with information about time and events at the level of annotations. This language is called ISO-TimeML.

The distinction between annotations and representations is reflected in the specification of ISO-TimeML given below, where an *abstract syntax* is defined as well as a *concrete syntax*. The abstract syntax specifies the elements making up the information in annotations, and how these elements may be combined to form complex annotation structures; these combinations are defined as set-theoretical structures, independent of any particular representation format. There are infinitely many ways in which these structures can be represented. In line with other ISO TC 37/SC 4 proposals, an XML-based concrete syntax is defined for representing ISO-TimeML annotations. Any other representation that is a faithful rendering of the abstract syntax of ISO-TimeML can readily be converted into this XML representation and vice versa. ISO-TimeML has a semantics associated with its abstract syntax, which defines the meanings of ISO-TimeML annotation structures. The fact that this semantics

is associated with the *abstract* syntax, rather than with a particular concrete syntax, explains why all concrete representations of ISO-TimeML annotations are semantically equivalent.

## 7.2 ISO-TimeML abstract syntax

### 7.2.1 Abstract syntax

The abstract syntax of ISO-TimeML defines the set-theoretical structures that constitute the information about time and events that may be contained in annotations. The definition of the abstract syntax consists of two parts:( a) a specification of the elements from which these structures are built up, called a 'conceptual inventory'; and (b) a set of syntax rules which describe the possible combinations of these elements. What these combinations mean, i.e. which information they capture, is specified by the semantics associated with the abstract syntax.

**a. Conceptual inventory**
The concepts that can be used to build ISO-TimeML annotations fall into five categories, all formed by finite sets, plus the concepts of real number and natural number. Natural numbers are needed for capturing the information expressed in English by *"twice"* and *"three times"*; real numbers are needed for cases such as *'two and a half hours"*.

- finite sets of elements called 'event classes'; 'tenses', 'aspects', 'polarities', and 'set-theoretic types' ;
- finite sets of elements called "temporal relations',duration relations', 'numerical relations, 'event subordination relations', and 'aspectual relations';
- a finite set of elements called 'time zones';
- finite sets of elements called 'calendar years', 'calendar months', 'calendar day numbers'; 'clock times' (natural numbers ranging from 0000 to 0059; from 0100 to 0159; ... from 2300 to 2400);
- a finite set of elements called 'temporal units'.

**b. Syntax rules**
Annotation structures in ISO-TimeML come in two varieties, which we will refer to as *entity structures* and *link structures*. Entity structures contain semantic information about a segment of source text; link structures describe semantic relations between segments of source text by means of links between entity structures.

The simplest kind of ISO-TimeML structures are a single entity structure, which is a pair $< m, a >$ consisting of a markable $m$ and an annotation $a$, or a single link structure which relates two entity structures. More complex annotation structures consist of a set of entity structures and a set of link structures which link the entity structures together through temporal and inter-event relations.

More formally, an ISO-TimeML annotation structure consists of two sets $M$ and $L$, where $M$ is a set of pairs $<markable, entity\ structure>$ and $L$ is a set of triples $<entity\ structure, entity\ structure, link\ structure>$, such that each element of $L$ contains at least one entity structure that occurs in $M$. (This latter condition ensures that the links in an annotation structure relate to the entity structures that it contains; they may in addition also contain links to elements outside the current annotation structure).

*Entity structures:*
There are five types of annotations that may form an entity structure, containing information about (1) events; (2-4) temporal objects (intervals, instants, and amounts of time); (5) frequencies of events; and (6) explicit temporal relations (as for instance expressed in English by temporal prepositions).

a) An *event structure* is a 7-tuple $< C, T, A, \Sigma, N, P_N, V >$ where $C$ is a member of the set of event classes; $T$ and $A$ are a tense and an aspect, respectively; $\Sigma$ is a set-theoretical type (such as *individual object* or *set of individual objects*); $N$ is a natural number (e.g. the number 2 for dealing with such examples as *"John kissed Mary twice"*); $P_N$ is an amount of time (such as two and a half hours, for such examples as *"John called Mary twice every two and a half hours"*), and $V$ is a veracity (claimed truth or falsity, corresponding to positive or negative polarity in natural language).

b) An *instant structure* is either a triple $<$ *time zone*, *date*, *clocktime* $>$, where a *date* is a triple consisting of a calendar year, a calendar month, and a calendar day number; or a triple $<$ *time-amount structure*, *instant structure*, *temporal relation*$>$ (*"half an hour before midnight"*).

c) The following set-theoretical structures are *interval structures*:

   1) a pair $< t_1, t_2 >$ of two instant structures, corresponding to the beginning and end points of the interval;
   2) a triple $<$ *time-amount structure*, *interval structure*, *temporal relation* $>$ ("three weeks before Christmas"; "two years from today");
   3) a triple $< t_1, t_2, R >$ where $t_1$ and $t_2$ are either instant structures or interval structures, and where $R$ is a duration relation (examples: *"from 1992 until 1995"*; *"from 1882 through 1995"*)

d) A *time-amount structure* is a pair $< n, u >$, where $n$ is a real number and $u$ a temporal unit, or a triple $< R, n, u >$, where $R$ is a numerical relation (like *greater than*) and $n$ and $u$ as before;

e) A *temporal relation structure* is just a temporal relation.

*Link structures:*
There are five types of link structures in ISO-TimeML: for temporal *anchoring* of events in time; for temporal *ordering of events and/or intervals or instants* relative to each other; for *measuring* the length of an interval; for *subordination relations* between events, and for *aspectual relations* between events.

a) A *temporal anchoring structure* is a triple $<$ *event structure*, *interval structure*, *temporal anchoring relation* $>$, or a triple $<$ *event structure*, *instant structure*, *temporal anchoring relation* $>$;

b) A *temporal relation structure* is a triple $<$ *event structure*, *event structure*, *temporal relation* $>$, or a triple $<$ *interval* or *instant structure*, *interval* or *instant structure*, *temporal relation* $>$

c) A *time measurement structure* is a pair <eventstructure,time-amountstructure>*orapair*<intervalstructure,time-amountstructur>;

d) A *subordination structure* is a triple $<$ *event structure*, *event structure*, *subordination relation* $>$;

e) An *aspectual structure* is a triple $<$ *event structure*, *event structure*, *aspectual relation* $>$.

### 7.3 Concrete XML-based syntax

A concrete syntax consists of he specification of names for the various sets forming the conceptual vocabulary, plus a listing of specific named elements of these sets, and a specification of how to represent ISO-TimeML annotation structures defined by the syntax rules of the abstract syntax.

A particular XML-based syntax for temporal annotation has been defined in the TimeML effort (Pustejovsky et al., 2003; 2007) and is reproduced below with a few amendments to do justice to the stand-off character of ISO-TImeML annotations and to the conceptual differences between ISO-TImeML and the original TimeML.

### 7.3.1 Naming conventions

Inasmuch as XML is case-sensitive, it is necessary for ISO-TimeML to specify exactly the case of all its elements. This document follows the convention of indicating tag names and attribute values in all upper case (e.g. EVENT, PROGRESSIVE) and attribute names in lower or mixed case (e.g. tense, relatedToTime). Since attribute values are typically atomic (one-word) while attribute names often consist of multiple words, this convention would seem to maximize readability of the annotation. (Multi-word attribute values use the underscore character to separate their component parts.)

This document also follows the attribute naming convention introduced in Setzer (2001). Attributes that range over values of XML datatype ID—a unique index—are short, consisting of one or two characters indicating the name of the element, followed by 'id' (e.g. tid, eiid). Attributes that range over values of XML datatype IDREF—references to IDs—typically consist of the name of the element indexed, followed by 'ID' (e.g. eventID) or a descriptive name (e.g. relatedToTime).

The values of the various ID attributes are specified as beginning with one or two characters, followed by an integer. This scheme is mandated by the syntax of XML. While attribute values of type ID can consist of any sequence of letters, digits, and the hyphen, underscore, and period characters, they must begin with either an underscore or a letter. Therefore "e23" is a valid XML ID; but "23" is not. This naming convention also helps make the examples a bit more readable, especially in the case of link tags, which can contain multiple IDREFs of different kinds.

### 7.3.2 Example annotations

Though this document describes the full ISO-TimeML language, many of the example annotations provided show the result of annotation only through the output of initial automatic tagging combined with human annotation/editing, but do not include elements (e.g. attributes and/or attribute values) that may be introduced by later processing components (e.g. the closure tool). In particular, TIMEX3 tags that are treated as temporal functions typically appear in the examples in an underspecified form. However, those elements that do appear are sufficient for the output of manual annotation.

Finally, note that all examples in this document have been validated against an ISO-TimeML DTD corresponding to the BNF given here, using the oXygen XML editor, version 1.1.

## 7.4 Basic ISO-TimeML tags

### 7.4.1 <EVENT>

The `EVENT` tag is used to annotate those elements in a text that describe what is conventionally referred to as an *eventuality*. Syntactically, events are typically expressed as inflected verbs, although event nominals, such as "crash" in "... killed by the crash", should also be annotated as `EVENT`s.

The `EVENT` tag is also used to annotate a subset of the states in a document, typically expressed as adjectives. This subset of states includes those that are either transient or explicitly marked as participating in a temporal relation. See Annex A, Core annotation guidelines, for more details.

Below is the syntax and definition for the `EVENT` tag:

```
attributes ::= eid eiid class pos tense aspect
           polarity mood [modality] [comment]
eid ::= ID
{eid ::= EventID
 EventID ::= e<integer>}
eiid ::= ID
{eiid ::= EventInstanceID
 EventInstanceID ::= ei<integer>}
class ::='OCCURRENCE' | 'PERCEPTION' | 'REPORTING' |
```

```
         'ASPECTUAL' | 'STATE' | 'I_STATE' | 'I_ACTION'
pos ::= 'ADJECTIVE' | 'NOUN' | 'VERB' | 'PREPOSITION' | 'OTHER'
tense ::= 'FUTURE' | 'PAST' | 'PRESENT' | 'IMPERFECT' | 'NONE'
aspect ::= 'PROGRESSIVE' | 'PERFECTIVE' | 'IMPERFECTIVE'
   | 'PERFECTIVE_PROGRESSIVE'  | 'IMPERFECTIVE_PROGRESSIVE' | 'NONE'
vform ::= 'INFINITIVE' | 'GERUNDIVE' | 'PASTPART' | 'PRESPART' | 'NONE'
polarity ::= 'NEG' | 'POS'
{default, if absent, is 'POS'}
mood ::= 'SUBJUNCTIVE' | 'NONE'
{default, if absent, is 'NONE'}
modality::= CDATA
comment ::= CDATA
```

NOTE The `comment` attribute is a part of all the ISO-TimeML tags, and exists for annotators to add clarifications and other observations about the text being marked.

Each `EVENT` tag represents a unique *instance* of an event, identified as the event instance identification number. If additional instances of an event are needed, a non-consuming `EVENT` tag can be created with the same event ID, but a new event instance ID. One can create as many instances as are motivated by the text. All relations indicated by the other links are stated over these instances.

The `tense` and `aspect` of the event are represented by specific attribute values within this tag. In addition, if the event is modified by a negation, this is indicated by the appropriate value in the `polarity` attribute. The term 'mood' in traditional grammar refers to `SUBJUNCTIVE` or `INDICATIVE`: "If I were (`PRESENT SUBJUNTIVE`) a bird, I would fly?" vs "If I am (`PRESENT INDICATIVE`) a bird, I can fly." "If I had been (`PAST SUBJUNCTIVE`) in the airport, I would have died (`PAST COUNTERFACTUAL CONDITiONAL` sentence)." Here the attribute `mood` has the value of `SUBJUNCTIVE` or `NONE`, and is used when `mood` is expressed by *inflectional morphology* on the verb; `modality`, on the other hand, is reserved for the presence of an explicit modal auxiliary verb, such as *should* or *must*. We expect that the tense and aspect attributes will have their values filled in by a pre-processing program, according to the following paradigm:

**Table 1: Active Voice**

| verb group | tense | aspect |
|---|---|---|
| teaches | PRESENT | NONE |
| is teaching | PRESENT | PROGRESSIVE |
| has taught | PRESENT | PERFECTIVE |
| has been teaching | PRESENT | PERFECTIVE_PROGRESSIVE |
| taught | PAST | NONE |
| was teaching | PAST | PROGRESSIVE |
| had taught | PAST | PERFECTIVE |
| had been teaching | PAST | PERFECTIVE_PROGRESSIVE |
| will teach | FUTURE | NONE |
| will be teaching | FUTURE | PROGRESSIVE |
| will have taught | FUTURE | PERFECTIVE |
| will have been teaching | FUTURE | PERFECTIVE_PROGRESSIVE |

Non-tense forms are encoded with the feature **vform**, which defaults to NONE when not otherwise specified:

**Table 2: Verb Forms**

| verb group | tense | vform |
|---|---|---|
| to teach | NONE | INFINITIVE |
| taught | NONE | PASTPART |
| teaching | NONE | GERUNDIVE |

**Table 3: Passive Voice**

| verb group | tense | aspect |
|---|---|---|
| is taught | PRESENT | NONE |
| is being taught | PRESENT | PROGRESSIVE |
| has been taught | PRESENT | PERFECTIVE |
| has been being taught | PRESENT | PERFECTIVE_PROG |
| was taught | PAST | NONE |
| was being taught | PAST | PROGRESSIVE |
| had been taught | PAST | PERFECTIVE |
| had been being taught | PAST | PERFECTIVE_PROG |
| will be taught | FUTURE | NONE |
| will be being taught | FUTURE | PROGRESSIVE |
| will have been taught | FUTURE | PERFECTIVE |
| will have been being taught | FUTURE | PERFECTIVE_PROG |
| being taught | NONE | PRESPART |

Similarly, non-tense forms are:

**Table 4: Non-tense Forms**

| verb group | tense | vform |
|---|---|---|
| to be taught | NONE | INFINITIVE |
| being taught | NONE | GERUNDIVE |
| having been taught | NONE | PASTPART |

The `pos` ('part of speech') attribute captures distinctions among the grammatical categories of phrases which are marked as events, as not all such phrases contain finite verbs.

The values of polarity and modality are determined by modifiers found near the event in the text. Formally, this information was annotated using a SIGNAL and an SLINK. Fo languages that encode mood in the morphology of the verb, the mood attribute is used. Some examples:

(1) should have bought

```
should have
<EVENT eid="e1" eiid="ei1" class="OCCURRENCE" pos="VERB"
tense="PAST" aspect="PERFECTIVE" modality="SHOULD"
polarity="POS">
bought
</EVENT>
```

(2) did not teach

```
did not
<EVENT eid="e1" eiid="ei1" class="OCCURRENCE" pos="VERB"
tense="PAST" aspect="NONE" polarity="NEG">
teach
</EVENT>
```

(3) must not teach twice

```
must not
<EVENT eid="e1" eiid="ei1" class="OCCURRENCE" pos="VERB"
tense="PRESENT" aspect="NONE" polarity="NEG"
modality="MUST">
teach
</EVENT>
twice
```

NOTE The information provided from the temporal quantifier *twice* will be included via TLINK.

### 7.4.2 <TIMEX3>

The TIMEX3 tag is primarily used to mark up explicit temporal expressions, such as times, dates, durations, etc. It is modeled on Setzer's (2001) TIMEX tag, as well as the TIDES (Ferro, et al. (2002)) TIMEX2 tag. Since it differs both in attribute structure and in use, it seemed best to give it a separate name, which reveals its heritage while at the same time indicating that it is different from its forebears.

```
attributes ::= tid type [functionInDocument] [beginPoint]
               [endPoint][quant] [freq] [temporalFunction]
               (value | valueFromFunction)[mod] [anchorTimeID]
               [comment]
tid ::= ID
{tid ::= TimeID TimeID ::= t<integer>}
type ::= 'DATE' | 'TIME' | 'DURATION' | 'SET'
functionInDocument ::= 'CREATION_TIME' | 'EXPIRATION_TIME' |
     'MODIFICATION_TIME' |'PUBLICATION_TIME' | 'RELEASE_TIME'|
     'RECEPTION_TIME' |'NONE' {default, if absent, is 'NONE'}
beginPoint ::= IDREF
{beginPoint ::= TimeID}
```

```
endPoint ::= IDREF
{endPoint ::= TimeID}
quant ::= CDATA
freq ::= Duration
temporalFunction ::= 'true' | 'false' {default, if absent,
                     is 'false'}
{temporalFunction ::= boolean}
value ::= Duration | Date | Time | WeekDate | WeekTime | Season
     | PartOfYear | PaPrFu
valueFromFunction ::= IDREF
{valueFromFunction ::= TemporalFunctionID
 TemporalFunctionID ::= tf<integer>}
mod ::= 'BEFORE' | 'AFTER' | 'ON_OR_BEFORE' | 'ON_OR_AFTER'
     |'LESS_THAN' | 'MORE_THAN' |'EQUAL_OR_LESS' | 'EQUAL_OR_MORE'
     | 'START' | 'MID' | 'END' | 'APPROX'
anchorTimeID ::= IDREF
{anchorTimeID ::= TimeID}
comment ::= CDATA
```

`functionInDocument`, an optional attribute, indicates the function of the `TIMEX3` in providing a temporal anchor for other temporal expressions in the document. If this attribute is not explicitly supplied, the default value is `NONE`. The non-empty values take their names from the temporal metadata tags in the Prism draft standard, which is available at

http://www.prismstandard.org/techdev/prismspec1.asp,

and are intended to have the same interpretations:

There are several times that mark the major milestones in the life of a news resource: The time the story is published, the time it may be released (if not immediately), the time it is received by a customer, and the time that the story expires (if any). Dates and times should be represented using the W3C-defined profile of ISO 8601:2004 [W3C-NOTE-datetime].

Table 5: Elements for time and date information

| Element | Role |
|---|---|
| prism:creationTime | Date and time the identified resource was first created |
| prism:expirationTime | Date and time when the right to publish material expires |
| prism:modificationTime | Date and time the resource was last modified |
| prism:publicationTime | Date and time when the resource is released to the public |
| prism:releaseTime | Earliest date and time when the resource may be distributed |
| prism:receptionTime | Date and time when the resource was received on current system |

NOTES

1. There can be as many instances of TIMEX3s containing a `functionInDocument` attribute with a non-empty value as there are `TIMEX3`s that express different functions. In practice, there will probably be no more than two, one with `CREATION_TIME` and another with `PUBLICATION_TIME`, since these are likely to be the only attributes that will appear in the text of documents to be annotated.

2. `RELEASE_TIME` does not indicate when the document was actually released. It is a specification of when the document is allowed to be released. This comes up in documents that are syndicated and where the issuing organization wants to delay publication by syndicators, so as not to be scooped.

3. The Prism standard, at least in its temporal indicators, is interested only in the document as an artifact, a piece of intellectual property. This means that the Prism values do not indicate the function of a `TIMEX3` relative to the internal narrative of the document. The specification of the ISO-TimeML language can fill this gap by adding values for the

functionInDocument attribute that capture narrative functions. At present, we leave the specification of possible values as is, and will defer the obvious extension until annotation of existing texts indicates that this is a pressing issue.

temporalFunction, an optional attribute, indicates whether the TIMEX3 is used as a temporal function; e.g. "two weeks ago". If this attribute is not explicitly supplied, the default value is "false". It is used in conjunction with anchorTimeID, which indicates the TIMEX3 to which its denotation is applied. It also appears with valueFromFunction, a pointer to a temporal function that determines its value. As was noted above, TIMEX3 tags that behave as temporal functions are often underspecified in the example annotations below.

The datatypes specified for the value attribute—Duration, Date, Time, WeekDate, WeekTime, Season, PartOfYear, PaPrFu—are XML datatypes based on the 2002 TIDES guideline, which extends the ISO 8601:2004 standard for representing dates, times, and durations.

NOTE See the 2002 TIDES guidelines for details about the value attribute, and see Annex G, ISO-TimeML schema, for complete definitions of each of these datatypes.

The attribute mod is an optional attribute adopted from TIDES. It is used for temporal modifiers that cannot be expressed either within value proper, or via links or temporal functions.

**Examples:**

(4) no more than 60 days

```
<TIMEX3 tid="t1" type="DURATION" value="P60D"
mod="EQUAL_OR_LESS">
no more than 60 days
</TIMEX3>
```

(5) the dawn of 2000

```
<TIMEX3 tid="t2" type="DATE" value="2000"
mod="START">
the dawn of 2000
</TIMEX3>
```

anchorTimeID is used to point to another TIMEX3 in the case of expressions such as "last week", which have a functional interpretation. The value of anchorTimeID provides the reference point to which the functional interpretation applies.

The attributes quant and freq are used to specify sets that that denote quantified times in TIMEX3. The attribute quant is generally a literal from the text that quantifies over the expression. The attribute freq contains an integer value and a time granularity to represent any frequency contained in the set, just as a period of time is represented in a duration.

**Examples:**

(6) twice a month

```
<TIMEX3 tid="t3" type="SET" value="P1M" freq="2X">
twice a month
</TIMEX3>
```

(7) three days every month

```
<TIMEX3 tid="t4" type="SET" value="P1M" quant="EVERY" freq="3D">
three days every month
</TIMEX3>
```

(8) daily

```
<TIMEX3 tid="t5" type="SET" value="P1D quant="EVERY">
daily
</TIMEX3>
```

`beginPoint` and `endPoint` are used to anchor durations to other time expressions in the document. If there is no explicit `tid` to assign to one of these values, then an empty `TIMEX3` tag is created to represent the unspecified point. Conversely, if both the beginning and end points of a duration are explicitly stated in the document, an empty `TIMEX3` tag is created to represent the unspecified duration.

**Examples**:

(9) two weeks from June 7, 2003

```
<TIMEX3 tid="t6" type="DURATION" value="P2W"
beginPoint="t61" endPoint="t62">
two weeks
</TIMEX3>
<SIGNAL sid="s1">
from
</SIGNAL>
<TIMEX3 tid="t61" type="DATE" value="2003-06-07">
June 7, 2003
</TIMEX3>
<TIMEX3 tid="t62" type="DATE" value="2003-06-21"
temporalFunction="true" anchorTimeID="t6"/>
```

(10) 1992 through 1995

```
<TIMEX3 tid="t71" type="DATE" value="1992">
1992
</TIMEX3>
<SIGNAL sid="s1">
through
</SIGNAL>
<TIMEX3 tid="t72" type="DATE" value="1995">
1995
</TIMEX3>
<TIMEX3 tid="t7" type="DURATION" value="P4Y"
beginPoint="t71" endPoint="t72"
temporalFunction="true"/>
```

### 7.4.3 <SIGNAL>

```
attributes ::= sid [comment]
sid ::= ID
{sid ::= SignalID
 SignalID ::= s<integer>}
comment ::= CDATA
```

SIGNAL is used to annotate sections of text, typically function words, that indicate how temporal expressions or eventualities are to be related to each other. The material marked by SIGNAL constitutes several types of linguistic elements: indicators of temporal relations such as temporal prepositions (e.g "on", "during") and other temporal connectives (e.g. "when") and subordinators (e.g. "if"). This functionality of the SIGNAL tag was introduced by Setzer (2001).

## 7.5 The link tags: <TLINK>, <SLINK>, and <ALINK>

### 7.5.1 <TLINK>

TLINK is one of the three ISO-TimeML link tags. Link tags encode the various relations that exist between the temporal elements of a document. The motivations for having multiple types of links are the following:

- To distinguish between event types and event instances, such as those introduced by conjunction, quantification, or negation.
- To adequately handle subordinating contexts involving modality and reported speech.

TLINK is a temporal link. It represents the relation between two temporal elements.

```
attributes ::= [lid] [origin] (eventInstanceID | timeID)
               [signalID](relatedToEventInstance |
               relatedToTime) relType [comment][syntax]
lid ::= ID
{lid ::= LinkID
 LinkID ::= l<integer>}
origin ::= CDATA
eventInstanceID ::= IDREF
{eventInstanceID ::= EventInstanceID}
timeID ::= IDREF
{timeID ::= TimeID}
signalID ::= IDREF
{signalID ::= SignalID}
relatedToEventInstance ::= IDREF
{relatedToEventInstance ::= EventInstanceID}
relatedToTime ::= IDREF
{relatedToTime ::= TimeID}
relType ::= 'BEFORE' | 'AFTER' | 'INCLUDES' | 'IS_INCLUDED'
            | 'DURING' |'SIMULTANEOUS' | 'IAFTER' | 'IBEFORE'
            | 'IDENTITY' |'BEGINS' | 'ENDS' | 'BEGUN_BY'
            | 'ENDED_BY' |'DURING_INV'
comment ::= CDATA
syntax ::= CDATA
```

The value of the optional origin attribute will be supplied by closure. This information and the link ID (lid) are primarily used by the closure algorithm. All links in ISO-TimeML may have these two attributes, but neither will be included in the examples presented here.

**Examples**:

(11) John taught 20 minutes every Monday.

John

```
<EVENT eid="e1" eiid="ei1" class="OCCURRENCE" pos="VERB"
tense="PAST" aspect="NONE" polarity="POS">
taught
</EVENT>
<TIMEX3 tid="t1" type="DURATION" value="P20TM">
20 minutes
</TIMEX3>
<TIMEX3 tid="t2" type="SET" value="xxxx-wxx-1"
quant="EVERY">
every Monday
</TIMEX3>
<TLINK timeID="t1" relatedToTime="t2"
relType="IS_INCLUDED"/>
<TLINK eventInstanceID="ei1" relatedToTime="t1"
relType="SIMULTANEOUS"/>
```

(12) John taught on Monday and on Tuesday too.

```
John
<EVENT eid="e1" class="OCCURRENCE"  eiid="ei1"  tense="PAST"
aspect="NONE" polarity="POS">
taught
</EVENT>
<SIGNAL sid="s1">
on
</SIGNAL>
<TIMEX3 tid="t1" type="DATE" value="xxxx-wxx-1">
Monday
</TIMEX3>
and
<SIGNAL sid="s2">
on
</SIGNAL>
<TIMEX3 tid="t2" type="DATE" value="xxxx-wxx-2">
Tuesday
</TIMEX3>
too
<EVENT eid="e1" class="OCCURRENCE"  eiid="ei2"  tense="PAST"
aspect="NONE" signalID="s2" polarity="POS"/>
<TLINK eventInstanceID="ei1" signalID="s1"
relatedToTime="t1" relType="IS_INCLUDED"/>
<TLINK eventInstanceID="ei2" signalID="s2"
relatedToTime="t2" relType="IS_INCLUDED"/>
```

(13) John taught 5 minutes after the explosion.

```
John
<EVENT eid="e1" eiid="ei1" class="OCCURRENCE"
tense="PAST" aspect="NONE" polarity="POS">
taught
</EVENT>
<TIMEX3 tid="t1" type="DURATION" value="PT5M"
beginPoint="t2" endPoint="t3">
5 minutes
</TIMEX3>
```

```
<SIGNAL sid="s1">
after
</SIGNAL>
the
<EVENT eid="e2" eiid="ei2" class="OCCURRENCE"
pos="NOUN" tense="NONE" aspect="NONE"
polarity="POS">
explosion
</EVENT>
<TIMEX3 tid="t2" type="TIME"
value="xxxx-xx-xx" temporalFunction="true"
anchorTimeID="t1"/>
<TIMEX3 tid="t3" type="TIME" value="xxxx-xx-xx"
temporalFunction="true" anchorTimeID="t1"/>
<TLINK eventInstanceID="ei2" signalID="s1"
relatedToTime="t1" relType="BEGINS"/>
<TLINK eventInstanceID="ei2" relatedToTime="t2"
relType="IS_INCLUDED"/>
<TLINK eventInstanceID="ei1"
relatedToTime="t3" relType="IS_INCLUDED"/>
```

**Treatment of Temporal Functions**:

(14) John taught from September to December last year.

```
John
<EVENT eid="e1" eiid="ei1" class="OCCURRENCE" pos="VERB"
tense="PAST" aspect="NONE" polarity="POS">
taught
</EVENT>
<SIGNAL sid="s1">
from
</SIGNAL>
<TIMEX3 tid="t1" type="DATE" value="xxxx-09">
 September
</TIMEX3>
<SIGNAL sid="s2">
to
</SIGNAL>
<TIMEX3 tid="t2" type="DATE" value="xxxx-12">
December
</TIMEX3>
<TIMEX3 tid="t5" type="DURATION" value="P4M"
beginPoint="t1" endPoint="t2" temporalFunction="true"/>
<TIMEX3 tid="t3" type=DATE" value="1995"
temporalFunction="true" anchorTimeID="t4">
last year
</TIMEX3>
<TIMEX3 tid="t4" type="DATE" value="1996-03-27"
functionInDocument="CREATION_TIME">
03-27-96
</TIMEX3>
<TLINK timeID="t1" signalID="s1"
relatedToTime="t5" relType="BEGINS"/>
```

```
<TLINK timeID="t2" signalID="s2" relatedToTime="t5"
relType="ENDS"/>
<TLINK eventInstanceID="ei1" relatedToTime="t5"
relType="SIMULTANEOUS"/>
```

(15) John taught last week.

```
John
<EVENT eid="e1" eiid="ei1" class="OCCURRENCE" pos="VERB"
tense="PAST" aspect="NONE" polarity="POS">
taught
</EVENT>
<TIMEX3 tid="t1" type="DATE" value="XXXX-WXX"
temporalFunction="true" anchorTimeID="t2">
last week
</TIMEX3>
<TIMEX3 tid="t2" type="DATE" value="1996-03-27"
functionInDocument="CREATION_TIME">
03-27-96
</TIMEX3>
<TLINK eventInstanceID="ei1" relatedToTime="t1"
relType="IS_INCLUDED"/>
```

NOTE The `TLINK` relates `TIMEX3` expressions. This is the only representation that will adequately express the temporal anchoring of this event.

(16) John taught last week on Monday.

```
John
<EVENT eid="e1" eiid="ei1" class="OCCURRENCE" pos="VERB"
tense="PAST" aspect="NONE" polarity="POS">
taught
</EVENT>
<TIMEX3 tid="t1" type="DATE" value="XXXX-WXX"
temporalFunction="true" anchorTimeID="t2">
last week
</TIMEX3>
<SIGNAL sid="s1">
on
</SIGNAL>
<TIMEX3 tid="t3" type="DATE" value="XXXX-WXX-1"
temporalFunction="true" >
Monday
</TIMEX3>
<TIMEX3 tid="t2" type="DATE" value="1996-03-27"
functionInDocument="CREATION_TIME">
03-27-96
</TIMEX3>
<TLINK eventInstanceID="ei1" relatedToTime="t1"
relType="IS_INCLUDED"/>
<TLINK timeID="t3" signalID="s1"
relatedToTime="t2" relType="IS_INCLUDED"/>
```

### 7.5.2 <SLINK>

This is a subordination link that is used for contexts involving modality, evidentials, and factives. An SLINK is used in cases where an event instance subordinates another event instance type. These are cases where a verb takes a complement and subordinates the event instance referred to in this complement.

```
attributes ::= lid eventInstanceID [signalID]
               subordinatedEventInstance relType
               [comment][syntax]
lid ::= ID
{lid ::= LinkID
 LinkID ::= l<integer>}
origin ::= CDATA
eventInstanceID ::= IDREF
{eventInstanceID ::= EventInstanceID}
subordinatedEventInstance ::= IDREF
{subordinatedEventInstance ::= EventInstanceID}
signalID ::= IDREF
{signalID ::= SignalID}
relType ::= 'INTENSIONAL' | 'EVIDENTIAL' | 'NEG_EVIDENTIAL' |
            'FACTIVE' | 'COUNTER_FACTIVE' | 'CONDITIONAL'
comment::= CDATA
syntax ::= CDATA
```

NOTE eventInstanceID is no longer optional (changed from TimeML 1.2.0).

The following EVENT classes interact with SLINK:

REPORTING, I_STATE or I_ACTION

Some lexical notes: Verbs that introduce I_STATE EVENTs that induce SLINK:

```
want, desire, crave, lust
believe, doubt, suspect
hope, aspire, intend
fear, hate, love, enjoy, like,
know
```

Verbs that introduce I_ACTION EVENTs that induce SLINK:

```
attempt, try, persuade, promise, name, swear, vow
```

**Examples**:

(17) If Graham leaves today, he will not hear Sabine.

```
<SIGNAL sid="s1">
if
</SIGNAL>
```

```
Graham
<EVENT eid="e1" eiid="ei1" class="OCCURRENCE" pos="VERB"
tense="PRESENT" aspect="NONE" polarity="POS">
leaves
</EVENT>
<TIMEX3 tid="t1" type="DATE" value="XXXX-XX-XX"
temporalFunction="true" >
today
</TIMEX3>
he will not
<EVENT eid="e2" eiid="ei2" class="OCCURRENCE" pos="VERB"
tense="FUTURE" aspect="NONE"
polarity="NEG">
hear
</EVENT>
Sabine.
<SLINK eventInstanceID="ei1"
subordinatedEventInstance="ei2"
signalID="s1" relType="CONDITIONAL"/>
<TLINK eventInstanceID="ei1"
relatedToEventInstance="ei2" relType="BEFORE"/>
```

(18) Bill denied that John taught on Monday.

```
Bill
<EVENT eid="e1" eiid="ei1" class="I_ACTION" pos="VERB"
tense="PAST" aspect="NONE" polarity="POS">
denied
</EVENT>
that John
<EVENT eid="e2" eiid="ei2" class="OCCURRENCE" pos="VERB"
tense="PAST" aspect="NONE" polarity="POS">
taught
</EVENT>
<SIGNAL sid="s1">
on
</SIGNAL>
<TIMEX3 tid="t1" type="DATE" value="XXXX-WXX-1">
Monday
</TIMEX3>
<TLINK eventInstanceID="ei2" signalID="s1"
relatedToTime="t1" relType="IS_INCLUDED"/>
<SLINK eventInstanceID="ei1"
subordinatedEventInstance="ei2"
relType="NEG_EVIDENTIAL"/>
```

(19) Bill wants to teach on Monday.

```
Bill
<EVENT eid="e1" eiid="ei1" class="I_STATE" pos="VERB"
tense="PRESENT" aspect="NONE" polarity="POS">
wants
</EVENT>
to
<EVENT eid="e2" eiid="ei2" class="OCCURRENCE" pos="VERB"
```

```
aspect="NONE" tense="NONE" vform="INFINITIVE" polarity="POS">
teach
</EVENT>
<SIGNAL sid="s2">
on
</SIGNAL>
<TIMEX3 tid="t1" type="DATE" value="XXXX-WXX-1">
Monday
</TIMEX3>
<TLINK eventInstanceID="ei2" signalID="s2"
relatedToTime="t1" relType="IS_INCLUDED"/>
<SLINK eventInstanceID="ei1"
subordinatedEventInstance="ei2" relType="INTENSIONAL"/>
```

Notice how **vform** and **aspect** interact to account for the following verbal form in the embedded sentence below:

(19′) John is believed *to have lived* in Rome.

```
to have
<EVENT eid="e2" eiid="ei2"class="OCCURRENCE" pos="VERB"
tense="NONE" aspect="PERFECTIVE" vform="INFINITIVE" polarity="POS">
lived
</EVENT>
```

(20) Bill attempted to save her.

```
Bill
<EVENT eid="e1" eiid="ei1" class="I_ACTION" pos="VERB"
tense="PAST" aspect="NONE" polarity="POS">
attempted
</EVENT>
to
<EVENT eid="e2" eiid="ei2"class="OCCURRENCE" pos="VERB"
tense="NONE" aspect="NONE" vform="INFINITIVE" polarity="POS">
save
</EVENT>
her
<SLINK eventInstanceID="ei1"
subordinatedEventInstance="ei2" relType="INTENSIONAL"/>
```

#### 7.5.3 <ALINK>

ALINK is an aspectual link; it indicates an aspectual connection between two events. In some ways, it is like a cross between TLINK and SLINK in that it indicates both a relation between two temporal elements, as well as aspectual subordination

```
attributes ::= lid eventInstanceID [signalID]
           relatedToEventInstance relType [comment] [syntax]
lid ::= ID
{lid ::= LinkID LinkID ::= l<integer>}
eventInstanceID ::= ID
{eventInstanceID ::= EventInstanceID}
signalID ::= IDREF
```

```
{signalID ::= SignalID}
relatedToEventInstance ::= IDREF
{relatedToEventInstance ::= EventInstanceID}
relType ::= 'INITIATES' | 'CULMINATES' | 'TERMINATES'
          | 'CONTINUES' | 'REINITIATES'
comment::= CDATA
syntax ::= CDATA
```

**Some examples:**

(21) The boat began to sink.

```
The boat
<EVENT eid="e1" eiid="ei1" class="ASPECTUAL" pos="VERB"
tense="PAST" aspect="NONE" polarity="POS">
began
</EVENT>
to
<EVENT eid="e2" eiid="ei2" class="OCCURRENCE" pos="VERB"
tense="NONE" aspect="NONE" vform="INFINITIVE" polarity="POS">
sink
</EVENT>
<ALINK eventInstanceID="ei1"
relatedToEventInstance="ei2" relType="INITIATES"/>
```

(22) The search party stopped looking for the survivors.

```
The search party
<EVENT eid="e1" eiid="ei1" class="ASPECTUAL" pos="VERB"
tense="PAST" aspect="NONE" polarity="POS">
stopped
</EVENT>
<EVENT eid="e2" eiid="ei2" class="OCCURRENCE" pos="VERB"
tense="NONE" aspect="NONE" vform="GERUNDIVE" polarity="POS">
looking
</EVENT>
<ALINK eventInstanceID="ei1"
relatedToEventInstance="ei2" relType="TERMINATES"/>
for the survivors
```

### 7.6 Other tags: <CONFIDENCE> and <ISO-TimeML>

#### 7.6.1 <CONFIDENCE>

In various discussions of the full TERQAS groups, the utility of being able to mark confidence values for various aspects of the annotation was pointed out. In general, it would be useful to allow confidence values to be assigned to any tag, and, in fact, to any attribute of any tag.

A convenient way to do this would be to create a confidence tag, which would consume no input, and which would have the following attributes:

```
attributes ::= tagType tagID [attributeName]
```

```
               confidenceValue [comment]
tagType ::= CDATA
tagID ::= IDREF
attributeName ::= CDATA
confidenceValue ::= CDATA
{confidenceValue ::= 0 < x < 1}
comment ::= CDATA
```

`tagType` would range over the names of all the tags of ISO-TimeML. `tagID` would range over the set of actual tag IDs within the current document (XML type IDREF). `attributeName` would range over the names of all the attributes of all the tags of ISO-TimeML. `confidenceValue` would range over the rationals (i.e. would have a floating point value) between 0 and 1. So, for example, given this annotation:

(23) The TWA flight crashlanded on Easter Island two weeks ago.

```
The TWA flight
<EVENT eid="e1" eiid="ei1" class="OCCURRENCE" pos="VERB"
tense="PAST" aspect="NONE" polarity="POS">
crashlanded
</EVENT>
on Easter Island
<TIMEX3 tid="t1" type="DURATION" value="P2W"
beginPoint="t2" endPoint="t3">
two weeks ago
</TIMEX3>
<TIMEX3 tid="t3" type="DATE" value="1999-12-06"
temporalFunction="true" anchorTimeID="t1"/>
.........
<TIMEX3 tid="t2" type="DATE"
functionInDocument="CREATION_TIME"
value="1999-12-20">
12-20-1999
</TIMEX3>
<TLINK eventInstanceID="ei1" relatedToTime="t3"
relType="IS_INCLUDED"/>
```

If we wanted to indicate that we were unsure that we had annotated "two weeks ago" correctly, we could add this annotation:

(23′)

```
<CONFIDENCE tagType="TIMEX3" tagID="t1" confidenceValue="0.50"/>
```

where the lack of the optional attribute, attributeName, indicates that the confidence applies to the whole tag.

On the other hand, if we wanted to indicate that we weren't sure if the tense of "crashlanded" was really "PAST", we could add this annotation:

(23″)

```
<CONFIDENCE tagType="EVENT" tagID="e1" attributeName="TENSE"
confidenceValue="0.75"/>
```

Abstracting confidence measures as a separate tag frees the annotation from having to include a confidence value attribute in every tag and eliminates the problem of uncertainty over the exact attribute of a tag the confidence value applies to.

As for how confidence values should be assigned in manual annotation, we feel that, in a large-scale annotation effort such as TIMEBANK, two conditions should be satisfied:

- Fairly high inter-annotator agreement on the tag assignment in the text.
- Ease of use and habitability of the tool from the annotator's perspective.

Therefore, the annotation of a scalar value such as confidence should have at least two features:

The choice of confidence values should be as clearly defined as possible to cover the options; this relates to the granularity and orders of magnitude as presented by Jerry Hobbs as well. This would suggest a selection from a small set (e.g. low, mid, high; not_sure, sure, abolutely_sure). These could be interpreted or rescaled to a (0,1] range, if need be, for subsequent inference. There should be a default value specified (at high (=1)) so that it is not necessary to annotate all links and attributes for them with a confidence. The constraint on human annotators to a subset of the possible values should be documented in the annotation guidelines and implemented in the annotation tool. And it would probably be best if the annotation tool did not present numbers but rather natural language descriptions such as those suggested above, which would be represented in the underlying annotation numerically. For example, the annotator might pick "moderately certain", which would enter the annotation as 0.5.

Moreover, for manual annotation, it does not seem that the 0 and 1 values will be used/useful. Presumably if the annotator doesn't trust an annotation at all s/he won't add it. And, as was suggested above, 1, at least for manual annotation, should be the default or unmarked value, and so need not be noted, since it would bulk up the files considerably, even if it were used only on entire tags.

#### 7.6.2 <**ISO-TimeML**>

Inasmuch as every well-formed XML document must have a single root node, we supply ISO-TimeML as this node. For example, a sample annotated ISO-TimeML document might look like this:

```
<?xml version="1.0"?>
<!DOCTYPE ISO-TimeML SYSTEM "ISO-TimeML.dtd">
<ISO-TimeML>
FAMILIES SUE OVER AREOFLOT CRASH DEATHS
The Russian airline Aeroflot has been
<EVENT eid="e1" eiid="ei1" class="OCCURRENCE" pos="VERB"
tense="PRESENT" aspect="PERFECTIVE" polarity="POS"/>
hit </EVENT>
with a writ for loss and damages,
<EVENT eid="e2" eiid="ei2" class="OCCURRENCE" pos="VERB"
tense="NONE" aspect="NONE" vform="PASTPART" polarity="POS">
filed
</EVENT>
in Hong Kong by the families of seven passengers
<EVENT eid="e3" eiid="ei3" class="OCCURRENCE" pos="VERB"
tense="NONE" aspect="NONE" vform="PASTPART" polarity="POS">
killed
</EVENT>
<SIGNAL sid="s1">
in
</SIGNAL>
an air
```

```
<EVENT eid="e4" eiid="ei4" class="OCCURRENCE" pos="VERB"
tense="NONE" aspect="NONE" polarity="POS">
crash.
</EVENT>
All 75 people
<EVENT eid="e7" eiid="ei7" class="STATE" pos="PREP"
tense="NONE" aspect="NONE" polarity="POS">
on board
</EVENT>
<TLINK eventInstanceID="ei7" relatedToEvent="ei5"
relType="INCLUDES"/>
the Aeroflot Airbus
<EVENT eid="e5" eiid="ei5" class="OCCURRENCE" pos="VERB"
tense="PAST" aspect="NONE" polarity="POS"/>
died
</EVENT>
<TLINK eventInstanceID="ei5" signalID="s2"
relatedToEvent="ei6" relType="IAFTER"/>
<SIGNAL sid="s2">
when
</SIGNAL>
it
<EVENT eid="e6" eiid="ei6" class="OCCURRENCE" pos="VERB"
tense="PAST" aspect="NONE" polarity="POS">
ploughed
</EVENT>
<TLINK eventInstanceID="ei6" signalID="s3"
relatedToTime="t2" relType="IS_INCLUDED"/>
<TLINK eventInstanceID="ei6" relatedToEvent="ei4"
relType="IDENTITY"/>
into a Siberian mountain
<SIGNAL sid="s3">
in
</SIGNAL>
<TIMEX3 tid="t2" type="DATE" value="1994-03">
March 1994
</TIMEX3>
.. ...
<TIMEX3 tid="t1" type="DATE" value="1996-03-27">
03-27-96
</TIMEX3>
<TLINK eventInstanceID="ei1" relatedToTime="t1"
relType="BEFORE"/>
<TLINK eventInstanceID="ei2" relatedToEvent="ei1"
relType="BEFORE"/>
<TLINK eventInstanceID="ei3" relatedToEvent="ei2"
relType="BEFORE"/>
<TLINK eventInstanceID="ei3" signalID="s1"
relatedToEvent="ei4" relType="IS_INCLUDED"/>
</ISO-TimeML>
```

NOTES

Here are some notes concerning changes from version TimeML 1.2 to TimeML 1.2.1:

1. The `nf_morph` attribute that was part of `MAKEINSTANCE` (now part of `EVENT` in ISO-TimeML) has been changed to `pos` (part of speech), and the `PASTPART, PRESPART, INFINITIVE,` and `GERUNDIVE` elements of `nf_morph` have been redistributed to `vform`.

2. The optional syntax attribute was added to `SLINK, ALINK,` and `TLINK`. Syntax can be used to hold `CDATA`, but is gen-

erally only used by annotation programs to hold the data that led to the creation of the tag.
3. The optional `comment` attribute was added to all ISO-TimeML elements, for the purpose of giving (human) annotators a place to put observations about annotated text.

# 8 Towards a semantics for ISO-TimeML

The syntactic apparatus of ISO-TimeML touches on several distinct areas of linguistic description. These include temporal semantics, as well as issues of modally subordinating contexts and the semantics of Aktionsarten (event classifications). In this clause, we focus on the semantics of the first of these areas, in particular: the representation of predicates as events; the representation of temporal expressions as intervals or quantification over intervals; and the relations between these interval structures.

## 8.1 Tense and Aspect in language

### 8.1.1 Tense

Tense can be defined as "the grammaticalized expression of location in time" (Comrie, 1986). This grammaticalized expression involves marking, via change of form, of particular syntactic elements, e.g., the verb and auxiliaries. For example, in *John ran a marathon*, the past tense morpheme represented as *-ed* (producing the inflected verb form *ran*) is used to indicate that the event occurred at a time earlier than the speech time. In *John will run a marathon*, the modal auxiliary *will* is used to locate the event as occurring at a future time, i.e., later than the speech time. While tense is mainly marked on the verb and auxiliaries associated with the verb group, in some languages, like the North American Indian language Nootka (Comrie, 1986), tense is expressed on the noun phrase.

Tense is not the only mechanism for expressing location in time. In languages such as Mandarin Chinese, which lacks tense morphemes, aspectual markers can be used to express location in time, though sometimes even these may be absent (Lin, 2003). There are also non-grammaticalized expressions of location in time given by temporal adverbials, e.g., *tomorrow*, *yesterday*, *two hours later*, etc. In the case of *tomorrow* or *yesterday*, the temporal location is with respect to the speech time. Temporal locations can also of course be expressed relative to a coordinate system given by a calendar, e.g., *1991 (C.E.)*, or a cyclically occurring event, e.g., *morning*, *spring*, or an arbitrary event, e.g., *the day after he married her*.

The few languages that lack tense altogether are not able to distinguish past from present or future. However, they all have a **realis/irrealis** distinction. In Burmese, for example (Comrie 1986), events that are ongoing or that were observed in the past are expressed by sentence-final realis particles *–te*, *-th*a, *-ta*, and *–hta*. In other cases, i.e., for unreal or hypothetical events (including future events, present events, and hypothetical past events), the sentence-final irrealis particles *–m*e, *-ma*, and *–hma* are used.

ISO-TimeML expresses four values for the attribute tense plus the value `NONE`. These are: `FUTURE`, `PAST`, `PRESENT`, and `IMPERFECT`.

### 8.1.2 Aspect

ISO-TimeML makes the traditional linguistic distinction between tense and aspect. The `tense` attribute values cover all the languages thus far examined by the ISO-TimeML Specification Working Group, but obviously there will be languages not covered by the current set.

While tense allows the speaker to relate the time of an eventuality to a deictic center or some other reference point, grammatical aspect allows the speaker to represent the structure of an eventuality. Here there is a distinction between **perfective aspect**, where an entire eventuality is presented without its internal temporal structure, e.g., *John built a house*, and **imperfective aspect**, where the speaker represents internal phases of the eventuality, e.g., *John is building a house*. Perfective aspect can express termination or completion of an eventuality, while imperfective aspect can express the ongoing nature of an activity. It is important to realize that many of the traditional tenses, e.g., Spanish imperfective, *as* in *Juan leia cuando entre* (*John was reading when I entered*), may combine both tense and aspect, e.g., past and imperfective. The same is true of the 'complex tenses' in English,

such as present progressive, present perfect, etc. Grammatical aspect is expressed in systematic ways across languages, depending on the lexical aspect of the eventuality. The following cross-linguistic account, derived from arguments by Smith (1991), summarizes some of this systematicity.

In English and French, **perfective aspect** is signaled by verbal tense and aspect morphemes. Termination is expressed in activities, completion is expressed in accomplishments and achievements, and statives can either express termination (e.g., French *passé composé* tense morpheme) or not (English, e.g., *I have lived in Paris*). In Mandarin Chinese, which lacks tense markers but which does have the semantic notion of tense (Lin 2003), the perfective is signaled by morphemes *–le* and *–guo*, usually indicating termination for activities, accomplishments and achievements; completion is indicated by a separate resultative morpheme *–wan*. In Russian, the perfective doesn't apply to statives, but is signaled by prefixes *po-* (short duration) and *pro-* (unexpected interval) in activities.

The **imperfective aspect** is signaled in English by the progressive morpheme *-ing*. It occurs in activities, accomplishments and achievements. In French, as in Russian, it is signaled by tense morphemes (e.g., the French *imparfait*). In Mandarin, it is signaled by the progressive morpheme *–zai* and resultative morpheme *–zhe*. The particle *-le* can also have an imperfective use with atelic predicates (Lin 2003).

ISO-TimeML expresses five values for the attribute aspect, plus the value `NONE`. These are: `PROGRESSIVE`, `PERFECTIVE`, `IMPERFECTIVE`, `PERFECTIVE_PROG`, and `IMPERFECTIVE_PROG`. The values not mentioned here are described in the annexes for those languages introducing them.

### 8.2 Temporal relations

ISO-TimeML assumes the general framework of Allen's (1984) interval algebra. In Allen's interval algebra, there are 13 basic (binary) interval relations, where six are inverses of the other six, excluding equality.

(1) a. before (b), after (bi);
b. overlap (o), overlappedBy (oi);
c. start (s), startedBy (si);
d. finish (f), finishedBy (fi);
e. during (d), contains (di);
f. meet (m), metBy (mi);
g. equality (eq).

These are shown schematically in Figure 2 below.

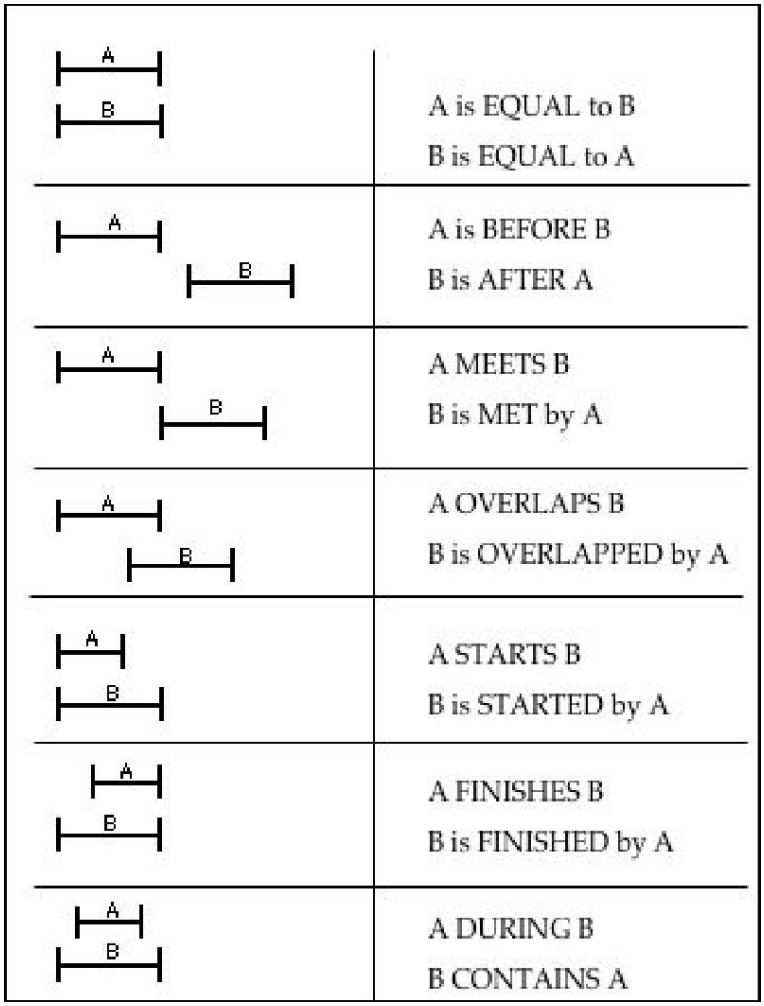


Figure 2: Allen's interval relations

In the clauses that follow, we will follow the syntax for ISO-TimeML by assuming a single type of event, class where there is no distinction made in Aktionsarten. Subsequent work within the ISO-TimeML community will likely extend the specification to account for differing event types in language, but this is not a necessary component to the current scope of the specification.

## 8.3 An Interval-based Semantics for ISO-TimeML

### 8.3.1 Technical preliminaries for interval temporal logic

We assume the usual apparatus and notation of the Simply Typed Lambda Calculus (see, e.g. Hindley and Seldin [?], pp.159–165), with primitive types $i$ and $t$. The type $i$ is identified with the set of non-empty compact, connected subsets of the reals (hereinafter: *intervals*), while the type $t$ is identified with the set $\{\top, \bot\}$ (hereinafter: *Booleans*). The primitive types are assumed to be mutually disjoint. We further assume a variety of temporal and logical constants. For example, the constant $\wedge$ of type $(t\ (t\ t))$ will be used to represent conjunction in the usual way. Likewise, the constant $Q_{\texttt{EVERY}}$ of type $((i\ t)\ (i\ t)\ t)$ will map two properties of intervals to $\top$ just in case every interval satisfying the first satisfies the second. Such constants will be introduced as and when required.

The semantics presented here concerns only a subset of ISO-TimeML. This is partly because some of the information recorded in ISO-TimeML is essentially syntactic rather than semantic in character, and partly because much of the semantic content of natural language texts has, for one reason or another, no clear rendition within the usual apparatus of formal semantics. Specifically, only three varieties of ISO-TimeML tags: `EVENT`-tags,

TIMEX3-tags and TLINK-tags are considered. In addition, only certain slots within these tags are taken to contribute to the semantics.

#### 8.3.2 Basic event-structure

Consider the following text.

(2) After his talk with Mary, John drove to Boston.

The ISO-TimeML-marked-up version of this text contains a pair of EVENT-tags and a TLINK-tag:

(3) After his <EVENT eid=e2 eiid=ei2> talk </EVENT> with Mary,
John <EVENT eid=e1 eiid=ei1> drove </EVENT> to Boston.
<TLINK eventInstanceID=ei1 relatedToEventInstance=ei2
relType=AFTER />.

At the same time, the event-structure of (**??**) can be captured by the first-order formula

(4) $p_{\mathtt{e1}}(I_{\mathtt{ei1}}) \wedge p_{\mathtt{e2}}(I_{\mathtt{ei2}}) \wedge R_{\mathtt{AFTER}}(I_{\mathtt{ei1}}, I_{\mathtt{ei2}})$,

where the variables $I_{\mathtt{ei1}}, I_{\mathtt{ei2}}$ range over intervals, the unary predicates $p_{\mathtt{e1}}, p_{\mathtt{e2}}$ are interpreted as the sets of such intervals over which John drove to Boston and John talked to Mary, respectively, and the binary predicate $R_{\mathtt{AFTER}}$ is interpreted as the relation which holds between intervals $[a,b]$ and $[c,d]$ just in case $d < a$. The variables $I_{\mathtt{ei1}}, I_{\mathtt{ei2}}$ may be assumed by default to be bound by existential quantifiers that take scope over the whole text; however, this default assumption may be overridden by further ISO-TimeML tags, as explained below. Making these global existential quantifiers explicit would add no information, and we do not do so.

Comparing (**??**) and (**??**), we see that each eid-value $e$ gives rise to a unary predicate $p_e$, each eiid-value $ei$ gives rise to a variable $I_{ei}$, and each relType-value $r$ gives rise to a binary predicate $R_r$.

For compatibility with the more complicated cases treated below, we re-formulate the semantics proposed in (**??**) using the syntax of higher-order logic. Under this régime, all constituents—logical constants, nonlogical constants and variables—are regarded as functions of (at most) one argument, and the application of a function $f$ to an argument $a$ is denoted $(f\ a)$. In addition, where convenient we split a conjunction $\phi \wedge \psi$ (in higher-order-logic notation: $((\wedge\ \phi)\ \psi)$) into the the pair of formulas $\phi$, $\psi$. Making use of this facility for one of the conjuncts in (**??**) yields the pair of higher-order logic formulas

(5) $(p_{\mathtt{e1}}\ I_{\mathtt{ei1}})$, $((\wedge\ (p_{\mathtt{e2}}\ I_{\mathtt{ei2}}))\ ((R_{\mathtt{AFTER}}\ I_{\mathtt{ei1}})\ I_{\mathtt{ei2}}))$.

Representations in the style of (**??**) are easier to generate from ISO-TimeML-tags than those in the style of (**??**); for the sake of readability, however, we continue to give both forms.

We can generate the formulas (**??**) as follows. Consider any EVENT-tag with eiid-value *ei* and eid-value *e*. We first create a new variable $I_{ei}$, of type $i$, and a new constant $p_e$, of type $(i\ t)$. In addition, we assign to *ei* a *semantic value* $\sigma_{ei}$, which is an expression of higher-order logic featuring $I_{ei}$ and $p_e$. (In the cases encountered in (**??**), all semantics values have Boolean type, so we can think of them as ordinary formulas.) The interpretation rule for simple EVENT-tags such as those encountered in (**??**) may be given as

(6) <EVENT eid=*e* eiid=*ei*> $\Rightarrow$ $\sigma_{ei} := (p_e\ I_{ei})$,

Since any eiid-value *ei* is allowed to occur in at most one EVENT-tag, these assignments cannot clash.

Suppose, then, that *ei* is the value of an `eiid`-slot in some `EVENT`-tag, so that $\sigma_{ei}$ has been assigned. We take any `TLINK`-tag with `relatedToEventInstance`-value *ei* to *modify* the expression assigned to $\sigma_{ei}$. In the case encountered in (**??**), this modification results simply in the addition of a conjunct expressing a temporal relation between two eventualities, thus:

(7) `<TLINK eventInstanceID=`$ei'$ `relatedToEventInstance=`$ei$ `relType=`$r$`>` $\Rightarrow \quad \sigma_{ei} := ((\wedge\, \sigma_{ei})\ ((R_r\ I_{ei'})\ I_{ei}))$,

where the $R_r$ is the constant of type $(i\,(i\,t))$ corresponding to $r$.

By applying rule (**??**) and subsequently rule (**??**) to example (**??**), we obtain the assignments

$$\begin{aligned} \sigma_{\mathrm{ei1}} &= (p_{\mathtt{e1}}\ I_{\mathtt{ei1}}) \\ \sigma_{\mathrm{ei2}} &= ((\wedge\,(p_{\mathtt{e2}}\ I_{\mathtt{ei2}}))\ ((R_{\mathtt{AFTER}}\ I_{\mathtt{ei1}})\ I_{\mathtt{ei2}})). \end{aligned}$$

Collecting together these assignments yields the two formulas (**??**), as required.

In general, *ei* may be the `relatedToEventInstance`-value of many `TLINK`-tags such as the one encountered in (**??**). In that case, the corresponding instances of rule (**??**) may be applied in any order; the chosen order will not affect the representation of event-structure, modulo logical equivalence.

We turn now to negated events. Consider

(8) John did not drive to Boston,

which has ISO-TimeML-mark-up

(9) John `<EVENT eid=e1 eiid=ei1 polarity=NEG>` drove `</EVENT>` to Boston.

We take (**??**) to assert that, within some contextually determined interval, no event of John's driving to Boston took place. If we represent our contextually determined interval using the variable $I_{\mathtt{ei1}}$ (corresponding to the `eiid`-value of the relevant `EVENT`-tag), then we can express these truth-conditions using the formula

(10) $\neg\exists I_1(R_{\mathtt{DURING}}(I_1, I_{\mathtt{ei1}}) \wedge p_{\mathtt{e1}}(I_1))$,

where the binary predicate $R_{\mathtt{DURING}}$ is interpreted as the relation which holds between intervals $[a,b]$ and $[c,d]$ just in case $c < a$ and $b < d$. Using the notation of higher-order logic, (**??**) may be re-written

(11) $(Q_{=0\mathtt{X}}\ \lambda I_1.((\wedge\ ((R_{\mathtt{DURING}}\ I_1)\ I_{\mathtt{ei1}}))\ (p_{\mathtt{e1}}\ I_1)))$,

where the unary quantifier $Q_{=0\mathtt{X}}$, of type $((i\,t)\,t)$, has the obvious interpretation: it maps a property of intervals to $\top$ just in case that property has no instances. The formula (**??**) can be generated from (**??**) by the interpretation rule

(12) `<EVENT eid =`$e$ `eiid=`$ei$ `polarity=NEG>` $\Rightarrow \quad \sigma_{ei} := (Q_{=0\mathtt{X}}\ \lambda I_1.((\wedge\ ((R_{\mathtt{DURING}}\ I_1)\ I_{ei}))\ (p_e\ I_1)))$,

where $I_1$ is a variable of type $i$, and $p_e$ is the constant of type $(i\,t)$ corresponding to $e$.

#### 8.3.3 The interpretation of TIMEX3-tags

TIMEX3-tags come in many varieties, for example:

```
<TIMEX3 tid=t61 type=DATE value=2003-06-07>
June 7, 2003 </TIMEX3>

<TIMEX3 tid=t1 type=SET value=XXXX-WXX-1 quant=EVERY>
every Monday </TIMEX3>

<TIMEX3 tid=t3 type=SET value=P1M freq=2X>
twice a month </TIMEX3>.
```

Clearly, we cannot expect a uniform semantic treatment of these tags.

##### 8.3.3.1 Unquantified TIMEX3-tags

Let us begin with the simplest case. Sometimes, TIMEX3-tags can be treated analogously to EVENT-tags. Consider, for example,

(13) John drove to Boston on Saturday, 31st January, 2004.

This text may be marked up as

(14)
```
John <EVENT eid=e1 eiid= ei1> drove </EVENT> to Boston
on <TIMEX3 tid=t1 value=2004-01-31>
Saturday, 31st January, 2004 </TIMEX3>
<TLINK eventInstanceID=ei1 relatedToTime= t1 relType=DURING />,
```

and its event-structure may be represented by the formula

(15) $p_{\texttt{2004-01-31}}(I_{\texttt{t1}}) \wedge p_{\texttt{e1}}(I_{\texttt{ei1}}) \wedge R_{\texttt{DURING}}(I_{\texttt{ei1}}, I_{\texttt{t1}})$,

where the variable $I_{\texttt{t1}}$ ranges over intervals, and the unary predicate $p_{\texttt{2004-01-31}}$ is interpreted as the (singleton) set of time-intervals coinciding with the 31st January, 2004. Re-writing (**??**) in the style of (**??**), we obtain

(16) $(p_{\texttt{2004-01-31}}\ I_{\texttt{t1}}),\ \ ((\wedge\ (p_{\texttt{e1}}\ I_{\texttt{ei1}}))\ ((R_{\texttt{DURING}}\ I_{\texttt{ei1}})\ I_{\texttt{t1}}))$.

This example suggests the following interpretation process. Given a TIMEX3-tag with tid-value $t$, we create a new variable $I_t$ of type $i$, and we assign the semantic value $\sigma_t$ according to the rule

(17) <TIMEX3 tid=$t$ value=$v$> $\Rightarrow$ $\sigma_t := (p_v\ I_t)$,

where $p_v$ is the constant, of type $(i\ t)$, corresponding to $v$.

In addition, we interpret the EVENT-tag in (**??**) using the rule (**??**), and the TLINK-tag using the following generalization of rule (**??**):

(18)
```
<TLINK
 {eventInstanceID|timeID}=x
 {relatedToEventInstance|
  relatedToTime}=y
 relType=r>
```
$\Rightarrow$ $\sigma_x := ((\wedge\ \sigma_x)\ ((R_r\ I_x)\ I_y))$.

These rules generate the representation of event-structure (**??**) from (**??**), exactly as for example (**??**).

#### 8.3.3.2 Quantifying TIMEX3-tags

Other TIMEX3-tags make a more complex semantic contribution. Consider, for example,

(19) John drove to Boston twice.

which has ISO-TimeML-mark-up

(20)
```
John <EVENT eid=e1 eiid=ei1> drove </EVENT> to Boston
<TIMEX3 tid=t1 freq=2X> twice </TIMEX3>
<TLINK eventInstanceID=ei1 relatedToTime=t1
 relType= DURING />
```

Assuming that twice here has the reading 'at least twice', we take (**??**) to assert that, within some contextually determined interval, at least two events of John's driving to Boston took place. If we represent this interval using the variable $I_{\texttt{t1}}$ (corresponding to the tid-value of the relevant TIMEX3-tag), then we can express these truth-conditions using the formula

(21) $\exists_{\geq 2} I_{\texttt{ei1}}(R_{\textsc{during}}(I_{\texttt{ei1}}, I_{\texttt{t1}}) \wedge p_{\texttt{e}}(I_{\texttt{ei1}}))$,

where, as before, the predicate $p_{\texttt{e}}$ is true of precisely those intervals over which John drives to Boston. Note that the variable $I_{\texttt{ei1}}$ arising from the EVENT-tag is bound by a quantifier in (**??**).

In the notation of higher-order logic, (**??**) becomes

(22) $(Q_{2\texttt{X}}\ \lambda I_{\texttt{ei1}}.((\wedge\ ((R_{\textsc{during}}\ I_{\texttt{ei1}})\ I_{\texttt{t1}}))\ (p_{\texttt{e}}\ I_{\texttt{ei1}})))$,

where $Q_{2\texttt{X}}$ is the function mapping a property of intervals to $\top$ just in case it has at least two instances.

To generate (**??**) from the ISO-TimeML-tags in (**??**), we adopt the rule

(23)
```
<TIMEX3
 tid=t
 freq=q>
```
$\Rightarrow \quad \lambda R \lambda P.(Q_q\ \lambda I_1.((\wedge\ ((R\ I_1)\ I_t))\ (P\ I_1)))$,

where $R$ is a variable of type $(i\ (i\ t))$, $P$ a variable of type $(i\ t)$, $I_1$ a variable of type $i$, $Q_q$ the constant of type $((i\ t)\ t)$ corresponding to $q$, and $I_t$ the variable of type $i$ corresponding to $t$. We also adopt the rule

(24)
```
<TLINK
 {eventInstanceID|timeID}=x
 {relatedToEventInstance|
  relatedToTime}=y
 relType=r>
```
$\Rightarrow \quad \sigma_y := ((\sigma_y\ R_r)\ \lambda I_x.\sigma_x)$
$\sigma_x :=$ undefined,

where $R_r$ is the constant of type $(i\ (i\ t))$ corresponding to $r$, and $I_x$ the variable of type $i$ corresponding to $x$.

Let us work through the example (**??**) to see how these rules produce the event-structure representation (**??**). First,

(25) John <EVENT eid=e1 eiid=ei1> drove </EVENT> to Boston

is processed by rule (**??**), yielding the assignment

(26) $\sigma_{\texttt{ei1}} := (p_{\texttt{e1}}\ I_{\texttt{ei1}})$,

while

(27) <TIMEX3 tid=t1 freq=2X> twice </TIMEX3>

is processed by rule (**??**), yielding the assignment

(28) $\sigma_{\texttt{t1}} := \lambda R \lambda P.(Q_{\texttt{2X}}\ \lambda I_1.((\wedge\ ((R\ I_1)\ I_{\texttt{t1}}))\ (P\ I_1)))$.

Note that (**??**) is a function which takes a relation between intervals as argument, and returns a function which itself takes a property of intervals as argument. These assignments having been made,

(29) <TLINK eventInstanceID=ei1 relatedToTime=t1 relType=DURING />,

is processed by rule (**??**) to yield the *re-assignment*

(30) $\sigma_{\texttt{t1}} := ((\lambda R \lambda P.(Q_{\texttt{2X}}\ \lambda I_1.((\wedge\ ((R\ I_1)\ I_{\texttt{t1}}))\ (P\ I_1)))\ R_{\texttt{DURING}})\ \lambda I_{\texttt{ei1}}.(p_{\texttt{e1}}\ I_{\texttt{ei1}}))$,

and the *de-assignment* of $\sigma_{\texttt{ei1}}$. Thus, rule (**??**) supplies arguments for the function in (**??**), and gets rid of the assignment in (**??**) altogether. Routine calculation shows that (**??**) normalizes to (**??**); and since, following the de-assignment of $\sigma_{\texttt{ei1}}$, there are no other semantic values to consider, this is the final semantics for (**??**), as required.

There is one further matter to discuss in relation to this example. As part of the process of de-assigning $\sigma_{\texttt{ei1}}$, we assume that `ei1` is made *inaccessible* to any other `TLINK`-tags than the one which caused rule (**??**) to be applied. That is, no other `TLINK`-tags may have `ei1` as the value of any of the slots `eventInstanceID`, `timeID`, `relatedToEventInstance` or `relatedToTime`. This inaccessibility reflects the fact that we cannot meaningfully use a `TLINK`-tag to relate `ei1` to any *other* `eiid`-value. After all, if such a `TLINK`-tag did involve `ei1`, *which* of the two (or more) events of John's driving to Boston would the relation involve? ISO-TimeML-marked-up text violating this accessibility constraint is considered uninterpretable.

#### 8.3.3.3 Quantifying `TIMEX3`-tags with unquantified complements

Moving on to a different kind of `TIMEX3`-tag, consider

(31) John drove to Boston every Monday.

This text may be marked up as

(32)
```
John <EVENT eid=e1 eiid=ei1 /> drove </EVENT> to Boston
<TIMEX3 tid=t1 type=SET value=XXXX-WXX-1 quant=EVERY>
every Monday </TIMEX3>
<TLINK eventInstanceID=ei1 relatedToTime=t1
 relType=DURING />,
```

and its event-structure may be represented by the formula

$$(33)\quad \begin{array}{l}\forall I_1(p_{\texttt{XXXX-WXX-1}}(I_1)\wedge R_{\texttt{DURING}}(I_1,I_{\texttt{t1}})\rightarrow\\ \quad\exists I_{\texttt{ei1}}(p_{\texttt{e1}}(I_{\texttt{ei1}})\wedge R_{\texttt{DURING}}(I_{\texttt{ei1}},I_1))),\end{array}$$

where the free variable $I_{\texttt{t1}}$ represents some contextually given interval to which the universal quantification in (**??**) is confined, the unary predicate $p_{\texttt{XXXX-WXX-1}}$ is satisfied by precisely those intervals which coincide with Mondays, and the predicates $p_{\texttt{e1}}$ and $R_{\texttt{DURING}}$ are as above. Thus, we take (**??**) to exhibit two levels of quantification: *explicit* universal quantification over Mondays, and *implicit* existential quantification over events within each of those Mondays.

In the notation of higher-order logic, (**??**) may be written

$$(34)\quad \begin{array}{l}((Q_{\texttt{EVERY}}\ \lambda I_1.((\wedge(p_{\texttt{XXXX-WXX-1}}\ I_1))\ ((R_{\texttt{DURING}}\ I_1)\ I_{\texttt{t1}})))\\ \ \lambda I_2.(Q_{\texttt{1X}}\ \lambda I_{\texttt{ei1}}.((\wedge\ (p_{\texttt{e1}}\ I_{\texttt{ei1}}))\ ((R_{\texttt{DURING}}\ I_{\texttt{ei1}})\ I_2)))).\end{array}$$

Here, $Q_{\texttt{EVERY}}$ is the constant of type $((i\,t)\ ((i\,t)\,t))$ expressing the binary universal quantifier, and $Q_{\texttt{1X}}$ the constant of type $((i\,t)\,t)$ expressing the unary existential quantifier. That is, for all $P$, $Q$ of type $(i\,t)$: $((Q_{\texttt{EVERY}}\ P)\ Q) = \top$ if and only if, for every interval $I$, $(P\,I) = \top$ implies $(Q\,I) = \top$; and $(Q_{\texttt{1X}}\ P) = \top$ if and only if, for some interval $I$, $(P\,I) = \top$.

The question now is how to recover (**??**) from the ISO-TimeML tags in (**??**). We propose the rule

$$(35)\quad \begin{array}{l}\texttt{<TIMEX3}\\ \texttt{ tid=}t\\ \texttt{ type=SET}\\ \texttt{ value=}R_v\\ \texttt{ quant=}q\ \texttt{>}\end{array}\Rightarrow \begin{array}{l}\sigma_t := \lambda R\lambda P.((Q_q\ \lambda I_1.((\wedge\ ((R_{\texttt{DURING}}\ I_1)\ I_t))\ (v\ I_1)))\\ \qquad\qquad \lambda I_2.(Q_{\texttt{1X}}\ \lambda I_3.((\wedge\ ((R\ I_3)\ I_2))\ (P\ I_3)))).\end{array}$$

With the above rules at our disposal,

(36) John <EVENT eid=e1 eiid=ei1 /> drove </EVENT> to Boston

is processed by rule (**??**), yielding the assignment

$$(37)\quad \sigma_{\texttt{ei1}} := (p_{\texttt{e1}}\ I_{\texttt{ei1}}),$$

while

(38) <TIMEX3 tid=t1 type=SET value=XXXX-WXX-1 quant=EVERY>
every Monday </TIMEX3>

is processed by rule (**??**), yielding the assignment

$$(39)\quad \begin{array}{l}\sigma_{\texttt{t1}} := \lambda R\lambda P.((Q_{\texttt{EVERY}}\ \lambda I_1.((\wedge\ ((R_{\texttt{DURING}}\ I_1)\ I_{\texttt{t1}}))\ (p_{\texttt{XXXX-WXX-1}}\ I_1)))\\ \qquad\qquad \lambda I_2.(Q_{\texttt{1X}}\ \lambda I_3.((\wedge\ ((R\ I_3)\ I_2))\ (P\ I_3)))).\end{array}$$

These assignments having been made,

(40) <TLINK eventInstanceID=ei1 relatedToTime=t1 relType=DURING />

is processed by rule (**??**), yielding the re-assignment

$$\begin{array}{ll}(41) & \sigma_{\texttt{t1}} := ((\lambda R \lambda P.((Q_{\texttt{EVERY}}\ \lambda I_1.((\wedge\ ((R_{\texttt{DURING}}\ I_1)\ I_{\texttt{t1}}))\ (p_{\texttt{XXXX-WXX-1}}\ I_1)))\\ & \qquad\qquad \lambda I_2.(Q_{\texttt{1X}}\ \lambda I_3.((\wedge\ ((R\ I_3)\ I_2))\ (P\ I_3))))\\ & \quad R_{\texttt{DURING}})\\ & \lambda I_{\texttt{ei1}}.(p_{\texttt{e1}}\ I_{\texttt{ei1}})),\end{array}$$

and the de-assignment of $\sigma_{\texttt{ei1}}$. But (**??**) normalizes to (**??**); and since there are no other semantic values to consider, this is the final semantics for (**??**), as required.

In some texts, the implicit existential quantification over events of the kind exhibited in (**??**) is made explicit. Consider, for example,

(42) John drove to Boston twice a month.

(Again, we assume that twice here has the reading 'at least twice'.) This text may be marked up as

(43)
```
John <EVENT eid=e1 eiid=ei1 /> drove </EVENT> to Boston
<TIMEX3 tid=t1 type=SET value= month quant=EVERY freq=2X>
twice a month </TIMEX3>
<TLINK eventInstanceID=ei1 relatedToTime=t1 relType=DURING />,
```

and its event-structure may be represented by the formula

$$\begin{array}{ll}(44) & \forall J(p_{\texttt{month}}(J) \wedge R_{\texttt{DURING}}(J, I_{\texttt{t1}}) \rightarrow\\ & \quad \exists_{\geq 2} I_{\texttt{ei1}}(p_{\texttt{e1}}(I_{\texttt{ei1}}) \wedge R_{\texttt{DURING}}(I_{\texttt{ei1}}, J))),\end{array}$$

where the variable $I_{\texttt{t1}}$ represents the interval to which universal quantification over months is restricted, and the unary predicate $p_{\texttt{month}}$ is true of exactly those intervals which are (calendar) months. Here we have two levels of quantification: universal quantification restricted to $I_{\texttt{t1}}$, and 'at-least-twice' quantification restricted to calendar months included in $I_{\texttt{t1}}$. The universal quantification is somewhat obliquely signalled by the English a month, but is encoded directly in the `quant`-value of the `TIMEX3`-tag.

We may re-write (**??**) using the notation of higher-order logic as

$$\begin{array}{ll}(45) & ((Q_{\texttt{EVERY}}\ \lambda I_1.((\wedge\ (p_{\texttt{month}}\ I_1))\ ((R_{\texttt{DURING}}\ I_1)\ I_{\texttt{t1}})))\\ & \ \lambda I_2.(Q_{\texttt{2X}}\ \lambda I_{\texttt{ei1}}.((\wedge\ (p_{\texttt{e1}}\ I_{\texttt{ei1}}))\ ((R_{\texttt{DURING}}\ I_{\texttt{ei1}})\ I_2)))).\end{array}$$

This suggests the interpretation rule

$$\begin{array}{l}\texttt{<TIMEX3}\\ \texttt{\ tid=}t\\ \texttt{\ type=SET}\\ \texttt{\ value=}R_v\\ \texttt{\ quant=}q\\ \texttt{\ freq=}q'\texttt{>}\end{array} \quad \Rightarrow \quad \begin{array}{l}\sigma_t := \lambda R \lambda P.((Q_q\ \lambda I_1.((\wedge\ ((R_{\texttt{DURING}}\ I_1)\ I_t))\ (v\ I_1)))\\ \qquad\qquad \lambda I_2.(Q_{q'}\ \lambda I_3.((\wedge\ ((R\ I_3)\ I_2))\ (P\ I_3)))).\end{array}$$

with other tags interpreted as before.

It may be verified that, after applying these rules and normalizing, $\sigma_{\texttt{t1}}$ is assigned the expression (**??**), while $\sigma_{\texttt{ei1}}$ is de-assigned. Since there are no other semantic values to consider, (**??**)—or its more readable equivalent, (**??**)—is the meaning of the ISO-TimeML tags in (**??**).

#### 8.3.3.4 Quantifying TIMEX3-tags with quantified complements

In example (**??**), the EVENT-tag, corresponding to John drove to Boston contributes an *unquantified* expression $(p_{\texttt{e1}}\ I_{\texttt{ei1}})$, while the TIMEX3-tag, corresponding to the adverbial twice, contributes a quantifier $Q_{\texttt{2X}}$, binding the variable $I_{\texttt{ei1}}$. Similarly, in example (**??**), the EVENT-tag again contributes the unquantified expression $(p_{\texttt{e1}}\ I_{\texttt{ei1}})$, while the TIMEX3-tag, corresponding to the adverbial every Monday, contributes an (implicit) existential quantifier $Q_{\texttt{1X}}$, again binding the variable $I_{\texttt{ei1}}$. In both these cases, a single temporal adverbial provides a quantifier binding the variable in an unquantified event-reporting expression.

It can also happen, however, that a temporal adverbial modifies an expression which has already been *explicitly quantified* by another adverbial. Since both adverbials will be marked with TIMEX3-tags, additional rules are required to interpret these tags. Consider, for example,

(46) One Monday, John drove to Boston twice,

where the adverbial One Monday modifies the explicitly quantified John drove to Boston twice. This text may be marked up as

(47)
```
<TIMEX3 tid=t2 type=SET value=XXXX-WXX-1 quant=SOME>
One Monday </TIMEX3>
John <EVENT eid=e1 eiid=ei1> drove </EVENT> to Boston
<TIMEX3 tid=t1 freq=2X> twice </TIMEX3>
<TLINK eventInstanceID=ei1 relatedToTime=t1 relType=DURING />
<TLINK tid=t1 relatedToTime=t2 relType=DURING />,
```

with the two TLINK-tags forming, so to speak, a cascade: ei1 → t1 → t2. The event-structure of (**??**) may be represented by the formula

(48) $\exists I_1(p_{\texttt{XXXX-WXX-1}}(I_1) \wedge R_{\texttt{DURING}}(I_1, I_{\texttt{t2}}) \wedge \exists_{\geq 2} I_{\texttt{ei1}}(p_{\texttt{e1}}(I_{\texttt{ei1}}) \wedge R_{\texttt{DURING}}(I_{\texttt{ei1}}, I_1))).$

or, in higher-order logic notation,

(49) $((Q_{\texttt{SOME}}\ \lambda I_1.((\wedge\ (p_{\texttt{XXXX-WXX-1}}\ I_1))\ ((R_{\texttt{DURING}}\ I_1)\ I_{\texttt{t2}})))\ \lambda I_2.(Q_{\texttt{2X}}\ \lambda I_{\texttt{ei1}}.((\wedge\ (p_{\texttt{e1}}\ I_{\texttt{ei1}}))\ ((R_{\texttt{DURING}}\ I_{\texttt{ei1}})\ I_2))))),$

with the binary quantifier $Q_{\texttt{SOME}}$ interpreted in the obvious way.

To see how (**??**) may be generated from (**??**), let us split the latter into

(50)
```
John <EVENT eid=e1 eiid=ei1> drove </EVENT> to Boston
<TIMEX3 tid=t1 freq=2X> twice </TIMEX3>
<TLINK eventInstanceID=ei1 relatedToTime=t1
 relType=DURING />
```

and

(51)
```
<TIMEX3 tid=t2 type=SET value=XXXX-WXX-1 quant=SOME>
One Monday </TIMEX3>
<TLINK tid=t1 relatedToTime=t2 relType=DURING />.
```

Now, (**??**) is identical to (**??**), and will therefore be given the same semantics. Crucially, the expression (**??**) thereby assigned to $\sigma_{\texttt{t1}}$ already quantifies over events; therefore, we should not proceed by analogy with (**??**), where the `TIMEX3`-tag is mapped to an expression providing implicit existential quantification. Instead, we propose the additional interpretation rules

(52)
```
<TIMEX3
 tid=t
 type=SET ⇒ σt := (Qq λI1.((∧ ((RDURING I1) It)) (v I1)))
 value=Rv
 quant=q >
```

and

(53)
```
<TLINK
 {eventInstanceID|timeID}=x
 {relatedToEventInstance|          ⇒   σy := (σy λIx.σx)
  relatedToTime}=y                      σx := undefined.
 relType=DURING>
```

Let us work through the example (**??**) to see how these rules produce the event-structure representation (**??**). We have already seen how the tags in (**??**) generate the assignments:

(54) $\sigma_{\texttt{t1}} = (Q_{\texttt{2X}}\ \lambda I_{\texttt{ei1}}.((\wedge\ ((R_{\texttt{DURING}}\ I_{\texttt{ei1}})\ I_{\texttt{t1}}))\ (p_{\texttt{e}}\ I_{\texttt{ei1}})))$
$\sigma_{\texttt{ei1}} = \text{undefined}.$

Turning therefore to the tags in (**??**), rule (**??**) results in the assignment

(55) $\sigma_{\texttt{t2}} := (Q_{\texttt{SOME}}\ \lambda I_1.((\wedge\ ((R_{\texttt{DURING}}\ I_1)\ I_{\texttt{t2}}))\ (p_{\texttt{XXXX-WXX-1}}\ I_1)));$

and rule (**??**) results in the re-assignment

(56) $\sigma_{\texttt{t2}} := ((Q_{\texttt{SOME}}\ \lambda I_1.((\wedge\ ((R_{\texttt{DURING}}\ I_1)\ I_{\texttt{t2}}))\ (p_{\texttt{XXXX-WXX-1}}\ I_1)))$
$\lambda I_{\texttt{t1}}.(Q_{\texttt{2X}}\ \lambda I_{\texttt{ei1}}.((\wedge\ ((R_{\texttt{DURING}}\ I_{\texttt{ei1}})\ I_{\texttt{t1}}))\ (p_{\texttt{e}}\ I_{\texttt{ei1}})))),$

and the de-assignment of $\sigma_{\texttt{t1}}$. But (**??**) is identical to (**??**) up to variable renaming. And since there are no other semantic values to consider, this is the final semantics for (**??**), as required.

#### 8.3.4 Interpretive rule summary

We summarize the interpretation rules employed above.

<EVENT eid =*e* eiid=*ei*> $\Rightarrow$ $\sigma_{ei} := (p_e\ I_{ei})$,

<EVENT
eid =*e*
eiid=*ei*
polarity=NEG> $\Rightarrow$ $\sigma_{ei} := (Q_{=0\mathrm{X}}\ \lambda I_1.((\wedge\ ((R_{\mathrm{DURING}}\ I_1)\ I_{ei}))\ (p_e\ I_1)))$,

<TIMEX3
tid=*t*
freq=*q*> $\Rightarrow$ $\lambda R \lambda P.(Q_q\ \lambda I_1.((\wedge\ ((R\ I_1)\ I_t))\ (P\ I_1)))$,

<TIMEX3
tid=*t*
type=SET
value=$R_v$
quant=*q* > $\Rightarrow$ $\sigma_t := \lambda R \lambda P.((Q_q\ \lambda I_1.((\wedge\ ((R_{\mathrm{DURING}}\ I_1)\ I_t))\ (v\ I_1)))\ \lambda I_2.(Q_{1\mathrm{X}}\ \lambda I_3.((\wedge\ ((R\ I_3)\ I_2))\ (P\ I_3)))))$.

<TIMEX3
tid=*t*
type=SET
value=$R_v$
quant=*q* > $\Rightarrow$ $\sigma_t := (Q_q\ \lambda I_1.((\wedge\ ((R_{\mathrm{DURING}}\ I_1)\ I_t))\ (v\ I_1)))$

<TLINK
{eventInstanceID|timeID}=*x*
{relatedToEventInstance|
relatedToTime}=*y*
relType=*r*> $\Rightarrow$ $\sigma_x := ((\wedge\ \sigma_x)\ ((R_r\ I_x)\ I_y))$.

<TLINK
{eventInstanceID|timeID}=*x*
{relatedToEventInstance|
relatedToTime}=*y*
relType=*r*> $\Rightarrow$ $\sigma_y := ((\sigma_y\ R_r)\ \lambda I_x.\sigma_x)$
$\sigma_x :=$ undefined,

<TLINK
{eventInstanceID|timeID}=*x*
{relatedToEventInstance|
relatedToTime}=*y*
relType=DURING> $\Rightarrow$ $\sigma_y := (\sigma_y\ \lambda I_x.\sigma_x)$
$\sigma_x :=$ undefined.

There is a general assumption that rules may only apply if the relevant types match.

## 8.4 An Event-based Semantics for ISO-TimeML

### 8.4.1 Introduction

The ISO-TimeML language has a semantics associated with its abstract syntax. A simple way to describe this semantics is in the form of a mapping from the set-theoretical structures defined by the abstact syntax to the language of first-order predicate logic with lambda abstraction. This semantics has been inspired by the compositional translation, defined in Bunt & Overbeeke (2008), from a concrete syntax of ISO-TimeML in the style of the underlying TimeML language (Pustejovsky et al., 2003; 2007) to first-order logic. Given that first-order logic has a well-established compositional formal semantics, this approach defines a compositional semantics for ISO-TimeML. The details of this formal semantics can be found in Annex 1. Here we illustrate the way the semantics works with an example.

Consider the sentence *"John started to read at half past ten"*. The ISO-TimeML annotation structure for this sentence would be as follows, where we identify markables by the corresponding text segments in the sentence, and use self-explanatory names for particular elements of the conceptual inventory.

*Entity structures:*
- event structure for the start event:
$<$ "*started*", $<$ *process*, *past*, *inchoative*, *individual*, *positive* $>>$
- event structure for the read event:
$<$ "*read*", $<$ *process*, $--$, $--$, *individual*, *positive* $>>$
- instant structure for *"half past ten*:
$<$ "10 : 30", $<$CET, $<$ 2008, $--$, $--$ $>$, 1030$>>$

*Link structures*:
- subordination structure relating the two events:
$<<$"*started*",$<$*process*, *past*, *inchoative*, *individual*, *positive* $>>$, $<$ "*read*", $<$ *process*, $--$, $--$, *individual*, *positive* $>>$, INITIATES$>$
- temporal anchoring structure relating the start event to the temporal instant:
$<<$ "*started*", $<$ *process*, *past*, *inchoative*, *individual*, *positive* $>>$, $<$ "10 : 45", $<$CET, $<$ 2008, $--$, $--$ $>$, 1045 $>>$, AT$>$

The semantics of this annotation structure can be computed by combining the first-order logic representations of the subordination structure and the temporal anchoring structure (see Bunt & Overbeeke, 2008 for how to do this), with the result:

$\exists e_1$ : PROCESS($e_1$) $\wedge$ $\exists t_1$ : calyear(CET, $t_1$)=2008 $\wedge$ clocktime(CET($t_1$))=1045 $\wedge$ EV-TIME(CET($e_1$))=$t_1$ $\wedge$
$\exists e_2$ : INITIATES($e_1, e_2$)

### 8.4.2 Defining an Event-based Semantics

The following clauses define an interpretation function *Infl* as a mapping from annotations to first-order logic with lambda abstraction (and real numbers). In the left hand s&s of the definition clauses, capital letters are used to designate elements form the conceptual inventory, and small letters to designate annotation structures.

- Elements from the conceptual inventory:
  - to every event class $C_i$, tense $T_j$, and aspect $A_m$ the interpretation function assigns a predicate constant which is indicated by an accent "′", e.g. $Infl(T_j) = T'_j$;
  - polarities and set-theoretic types as such are not represented in the first-order logic representations, but are interpreted through the stipulation of different interpretations of event annotations, depending on the values of these elements;
  - to every temporal relation, temporal measurement function, duration relation, numerical relation, event subordination relation, and aspectual relation the interpretation function assigns a predicate constant which is indicated by an accent "′";
  - to every time zone $Z$ a function constant $Z'$ is assigned; (Such functionsthat map the time line onto pairs consisting of a date and a clock time)
  - to all calendar years, calendar months, calendar day numbers, and clock times the interpretation function assigns their usual numerical string name (like '20090229' and '1245');
  - for every real number $N$, $Infl(N)$ will be its usual string name (like '5');
  - to every temporal unit the interpretation function assigns an individual constant which is indicated by an accent "′": $Infl(u) = u'$.
- Event annotations:
  $Infl(< C, T, A, indiv, pos >) = \lambda P.\exists e : C'(e) \wedge T'(e) \wedge A'(e) \wedge P(e)$
  $Infl(< C, T, A, indiv, neg >) = \lambda P.\neg\exists e : C'(e) \wedge T'(e) \wedge A'(e) \wedge P(e)$
  $Infl(< C, T, A, set, pos >) = \lambda P.\exists E : \forall e \in E : C'(e) \wedge T'(e) \wedge A'(e) \wedge P(e)$
  $Infl(< C, T, A, set, neg >) = \lambda P.\neg\exists E : \forall e \in E : C'(e) \wedge T'(e) \wedge A'(e) \wedge P(e)$

$Infl(< C,T,A,set,N,pos >) = \lambda P.\exists_N e : C'(e) \wedge T'(e) \wedge A'(e) \wedge P(e)$
$Infl(< C,T,A,set,N,P_N,pos >) = \lambda P.\forall I : Infl(P_N)(I) \rightarrow \exists_N e : C'(e) \wedge T'(e) \wedge A'(e) \wedge P(e)$
$Infl(< C,T,A,set,N,neg >) = \lambda P.\neg\exists_N e : C'(e) \wedge T'(e) \wedge A'(e) \wedge P(e)$
$Infl(< C,T,A,set,N,P_N,neg >) = \lambda P.\neg\forall I : Infl(P_N)(I) \rightarrow \exists_N e : C'(e) \wedge T'(e) \wedge A'(e) \wedge P(e)$

- Interval annotations:
  1. $Infl(< t_1,t_2 >) = \lambda P.\ \exists T : Infl(t_1)(\lambda x.\ \text{start}(T){=}x) \wedge Infl(t_2)(\lambda y.\ \text{end}(T){=}\ y) \wedge P(T)$
  2. $Infl(<< N,U >, T_1,R >) = \lambda P.\exists T : \exists T' : \text{end}(T){=}\text{start}(T') \wedge \text{end}(T'){=}\text{start}(T_1) \wedge \text{length}(T',U') = N' \wedge P(T)$
  3. $Infl(< t_1,t_2,R >) = \lambda P.\ \exists T : R'(T,T_1,T_2) \wedge P(T)$
- Instant annotations:
  $Infl(< Z,d,T >) = \lambda P.\ \exists t : \text{date}(Z'(t)) = Infl(d) \wedge \text{clocktime}(Z'(t)) = T' \wedge P(t)$
  $Infl(<< N,U >,t_1 >) = \lambda P.\exists t : \exists T : t = \text{start}(T) \wedge t_1 = \text{end}(T) \wedge \text{length}(T,U') = N' \wedge P(t)$
- Time-amount annotations:
  $Infl(< N,U >) = \lambda x.\ \text{length}(x,U') = N'$
  $Infl(< R,N,U >) = \lambda x.\ RF'(\text{length}(x,U'),N'$
- Temporal relation annotations: $Infl(R) = R'$
- Temporal anchoring structures: $Infl(< e,t,R >) = \lambda e.\lambda T.R'(\text{EV-TIME}(e),T)$
- Temporal relation structures:
  $Infl(< e_1,e_2,R >) = \lambda e_1.\lambda e_2.R'(e_1,e_2,T)$
  $Infl(< t_1,t_2,R >) = \lambda t_1.\lambda t_2.R'(t_1,t_2)$
- Time measurement structures:
  $Infl(< t,< N,U >>) = \lambda T.\text{length}(T,U') = N'$
  $Infl(< e,< N.U >>) = \lambda e.\text{length}(\text{EV-TIME}(e),\ U') = N'$
- Subordination structures: $Infl(< e_1,e_2,R >) = \lambda e_1.\lambda e_2.R'(e_1,e_2)$
- Aspectual structures: $Infl(< e_1,e_2,A >) = \lambda e_1.\lambda e_2.A'(e_1,e_2)$

These first-order logic representations have a formal semantics which makes use of a model $M =< D,F >$ where:

- $D$ is the model structure, which is an octet $D = < C_E,T,F_T,U,C_U,Z_T,date,clocktime >$, where
  - $C_E$ is a set of event classes (such as *Process, Transition, State*);
  - $T$ is an 8-tuple ($\Pi$, $\leq_T$,start, end, $cY,cM,cD,cT >$ where $\Pi$ is an infinite a set of time points with a total ordering $\leq_T$; 'start' and 'end' are functions from temporal intervals to the time points defining their beginning and end; $cY$, $cM$, $cD$ and $cT$ are sets of intervals in $T$ (calendar years, calendar months, calendar days, and clocktimes respectively); $cT$ is a structure $cT = < clH,clM,.. >$ (of clock hours and minutes, extended, if necessary, with seconds, milliseconds, etc.)
  - $U$ is a finite set of units for temporal measurement;
  - $F_T$ is a set of temporal functions
  - $C_U$ is a pair consisting of a function from pairs of temporal units to real numbers, i.e. a conversion function between temporal units, and a function 'length' that computes the length of a temporal interval, given a temporal unit (i.e. a function from pairs consisting of a temporal interval and a temporal unit to real numbers);
  - *length* is a function assigning a numerical value to a pair, consisting of a temporal interval and a temporal unit
  - $Z_T$ (time zones) is a set of functions from $T$ to pairs of dates and times;
  - *date* and *clocktime* are functions projecting each temporal value, defined by a time zone, onto its date and its clock time, respectively
- $F$ is the interpretation function, assigning to the predicates in the first-order representations of annotation structures, elements of the model structure. The interpretation rules for first-order logic expressions are as usual.

 

# Annex A
(normative)
# Core annotation guidelines
# Language-independent ISO-TimeML annotation guidelines

## A.1 Introduction

This annex describes the annotation guidelines for marking up text according to the ISO-TimeML language. It is organized as follows. The first clause explains what the ISO-TimeML tags (XML elements) are and how to annotate them. It also specifies for each tag what its attributes are and provides a BNF definition for the tag and its attributes. While this exposition contains many examples illustrating what and how to tag, the examples focus, for clarity's sake, on the tag under discussion at any given point. The following Annex B provides a set of fully annotated examples, illustrating all of the interactions between the various entity and relational tags.

For the sake of convenience, I&P(02) will be used to refer to [?] and TIDES(02) throughout the whole Annex.

## A.2 ISO-TimeML tags and their attributes

### A.2.1 The tag <EVENT>

*Event* is a cover term for expressions denoting:

- Situations that *happen* or *occur*, which can both be punctual (**??**) or last for a period of time (**??**).

  (57) a. Ferdinand Magellan, a Portuguese explorer, first **reached** the islands in search of spices.
  b. A fresh flow of lava, gas and debris **erupted** there Saturday.

  (58) a. *11,024 people, including local Aeta aborigines,* **were evacuated** *to 18 disaster relief centers.*
  b. *"We're* **expecting** *a major eruption," he said in a telephone interview early today.*

- *States* or *circumstances* in which something obtains or holds true (**??**).

  (59) Israel has been scrambling to buy more masks abroad, after a **shortage** of several hundred thousand gas masks.

#### A.2.1.1 How to annotate EVENTs

The types of expressions denoting events vary cross-linguistically. Chinese for instance does not use tensed forms of verbs, whereas this is the most common way of conveying events in Germanic and Romance languages. In general, however, events can be expressed by means of (at least some of) the following phrase types: VPs headed by either tensed or untensed verbs (**??**-**??**), NPs (**??**), APs (**??**), or PPs (**??**).

Note that in the above sentences not all markable elements are tagged –only those that are relevant for the discussion. In the first example, for instance, *flow* was not marked.

(60) *A fresh flow of lava, gas and debris* **erupted** *there Saturday.*

(61) *Prime Minister Benjamin Netanyahu called the prime minister of the Netherlands* **to thank** *him for thousands of gas masks his country has already contributed.*

(62) *Israel will ask the United States to delay a military* **strike** *against Iraq until the Jewish state is fully prepared for a possible Iraqi* **attack**.

(63) *A Philippine volcano,* **dormant** *for six centuries, began exploding with searing gases, thick ash and deadly debris.*

(64) *All 75 people* **on board** *the Aeroflot Airbus died.*

**A. Event identification:**
Event identification is based on the notion of *minimal chunk*, because higher constituents (i.e., phrases) may contain more than one event-denoting expression. For example: VPs headed by an aspectual verb (**??**), light verb constructions (**??**), causative constructions (**??**) or other VPs whose complement expresses an additional event. In ISO-TimeML, both event expressions will be annotated with independent tags because both verbal and nominal heads are relevant to different kind of event information.

NOTE The two tagged events will be related by means of a temporal or aspectual link (`TLINK` and `ALINK`, respectively). See clause A.3.

In the examples below, the phrase and minimal chunk constituency levels are marked in square and regular brackets, respectively.

(65) They **[probably (would have began) (responding) to President Reagan's 600 ships plan with new construction]**.

NOTE The VP contains two additional event expressions not signaled here: *plan* and *construction*.

(66) They **[(will definitely take) it into (consideration)]**.

(67) a. The rains **[(caused) (the flooding)]**.

b. John **[(caused) (the fire)]**.

NOTE The subject *the rains* denotes here an additional event.

If the event is denoted by means of a predicative construction, only the predicative element (the adjective, nominal, or prepositional complement) will be marked, disregarding the copular element (expressed by copulas equivalent to English form *be*).

(68) a. There is no reason why we **[would not be (fully prepared)]**.

b. If, in spite of everything, we **[will not be (ready)]**, we will ask the United States to delay the operation.

c. James Pustejovsky **[was (CTO of LingoMotors)]** for several years.

**B. Event tag span:**
In most cases, the event tag will span over only one word: the head of the minimal chunk expressing the event. The following cases are contemplated:

- If the event is expressed by a verb, noun, or adjective chunk, the tag will cover only the head of the chunk. We will therefore disregard: auxiliaries, clitics, prepositions, polarity markers, and other particles, in the case of verbal chunks (**??**); specifiers, complements, and modifiers, in the case of noun chunks (**??**); or modifiers, in the case of adjective chunks (**??**). In the examples below, the event-denoting chunk is in bold face, whereas the tagged head is underlined.

  (69) a. A fresh flow of lava, gas and debris **erupted** there Saturday.
  b. Israel **has been scrambling to buy** more masks abroad.
  c. The private sector **could establish** a private agency.
  d. Kaufman **did not disclose** details of the deal.
  e. Additional distribution centers would be **set up** next week.

  (70) a. **A fresh flow of lava, gas and debris** erupted there Saturday.
  b. Israel will ask the United States to delay **a military strike** against Iraq until the Jewish state is fully prepared for **a possible Iraqi attack**.

  (71) a. A Philippine volcano, **dormant** for six centuries, began exploding with searing gases, thick ash and deadly debris.
  b. There is no reason why we would not be **fully prepared.**

- If the event is expressed by a prepositional chunk, we will annotate the noun head of the embedded NP in case it expresses an event (**??**). Otherwise, we will annotate the whole PP (**??**):

  (72) a. Prof. Abramovitz, **on sabatical** in Heidelberg University, declared the discovery is of remarkable relevance.
  b. All 75 people **on board** the Aeroflot Airbus died

**C. What *NOT* to tag:**
Events will not be tagged in the following two situations:

- When they express states that are not temporally relevant; that is, states that (a) are not directly related to a temporal expression, or (b) are not identifiably changed over the course of the text being marked up. See subclause **??**, on events belonging to the class STATE.
- When the event reading of a logically polysemous nominalization is not exploited in the predication; For example, in (**??**) *reports* is not tagged as an event because the sense of the nominalization relevant here is the "information" or "content" interpretation of the noun.

  (73) Newspaper **reports** have **said** Amir was infatuated with Har-Shefi.

#### A.2.1.2 BNF for the EVENT tag

```
attributes ::= eid eiid class tense aspect pos polarity mood [modality]
eid ::= e<integer>
eiid ::= ei<integer>
type ::= 'STATE' | ''PROCESS' | 'TRANSITION'
class ::= 'REPORTING'|'PERCEPTION'|'ASPECTUAL'|'I_ACTION'|'I_STATE' |
      'OCCURRENCE'
pos ::= 'ADJECTIVE'|'NOUN'|'VERB'|'PREPOSITION'|'OTHER'
tense ::= 'FUTURE' | 'PAST' | 'PRESENT' | 'IMPERFECT' | 'NONE'
aspect ::= 'PROGRESSIVE' | 'PERFECTIVE' | 'IMPERFECTIVE'
   | 'PERFECTIVE_PROGRESSIVE'  | 'IMPERFECTIVE_PROGRESSIVE' | 'NONE'
vform ::= 'INFINITIVE' | 'GERUNDIVE' | 'PASTPART' | 'PRESPART' | 'NONE'
polarity ::= 'NEG'|'POS' {default, if absent, is 'POS'}
mood ::= 'SUBJUNCTIVE'|'NONE' {default, if absent, is 'NONE'}
modality ::= CDATA
comment ::= CDATA
```

#### A.2.1.3 Attributes for EVENT

**A. Attribute** `eid` **(event ID number):**
Required attribute. Each event must be identified by a unique ID number.

**B. Attribute** `class`**:**
Required attribute. Each event belongs to one of the following classes:

NOTE The verbs provided as examples of each class may have multiple senses, some of which may not belong to that particular class.

- `REPORTING`: Reporting events describe the action of a person or an organization declaring something, narrating an event, informing about an event, etc. Some examples in English: *say, report, tell, explain, state*:

  (74) a. Punongbayan **said** that the 4,795-foot-high volcano was spewing gases up to 1,800 degrees.
  b. No injuries were **reported** over the weekend.
  c. **Citing** an example, ...

- `PERCEPTION`: Events involving the physical perception of another event. In English, such events are typically expressed by verbs like: *see, watch, glimpse, behold, view, hear, listen, overhear.*

  (75) a. Witnesses tell Birmingham police they **saw** a man running.
  b. "You can **hear** the thousands of small explosions down there," a witness said.

- `ASPECTUAL`: In languages such as English and French, there is a grammatical device of aspectual predication, which focuses on different facets of event history:

  a) Initiation: *begin*, *start*, *commence, set out*, *originate, initiate*.
  b) Reinitiation: *restart, reinitiate, reignite* (metaphorically)
  c) Termination: *stop*, *terminate*, *cease*, *discontinue*, *interrupt*, *quit*.
  d) Culmination: *finish*, *complete*.
  e) Continuation: *continue*, *keep, go on*.

  A couple of examples:

  (76) a. The volcano **began** showing signs of activity in April for the first time in 600 years,...
  b. All non-essential personnel should **begin** evacuating the sprawling base.

- `I_ACTION`: `I_ACTION` stands for intensional action. `I_ACTION`s describe an action or situation which introduces another event as its argument, which must be in the text explicitly. Explicit performative predicates (like those in (e)-(i), below) are also included here.

  NOTES
  1. The `I_ACTION` class does not cover states (but see the description of `I_STATES` below).
  2. Note the distinction between "intensional" and "intentional" or purposeful. This class includes but is broader than actions with intended consequences.

  The following list of English predicates is representative (not exhaustive) of the types of events included in this class. In the examples, `I_ACTION`s are in bold face and their event arguments, underlined.

  a) **attempt, try, scramble:**

  (77) Companies such as Microsoft are **trying** to monopolize Internet access.

  b) **investigate, investigation, look at, delve:**

  (78) A new Essex County task force began **delving** Thursday into the slayings of 14 black women.

  c) **delay, postpone, defer, hinder, set back:**

  (79) Israel will ask the United States to **delay** a military strike against Iraq.

d) **avoid, prevent, cancel:**

(80) Palestinian police **prevented** a planned pro-Iraq rally by the Palestinian Professionals' Union.

e) **ask, order, persuade, request, beg, command, urge, authorize:**

(81) Iraqi military authorities **ordered** all Americans and Britons in Kuwait to assemble at a hotel.

f) **promise, offer, assure, propose, agree, decide:**

(82) Germany has **agreed** to lend Israel 180,000 protective kits against chemical and biological weapons, and Switzerland **offered** to lend Israel another 25,000 masks.

g) **swear, vow.**

h) **name, nominate, appoint, declare, proclaim.**

i) **claim, allege, suggest.**

- I_STATE: They are similar to the events in the previous class. I_STATEs also select for another event as their argument, but contrary to I_ACTIONs, they denote stative situations. As above, the I_STATE is in bold face, whereas the embedded argument is underlined. The following list is not exhaustive, but only representative.

a) **believe, think, suspect, imagine, doubt, feel, be conceivable, be sure**:

(83) "We **believe** that his words cannot distract the world from the facts of Iraqi aggression."

b) **want, love, like, desire, crave, lust**:

NOTE The verb *love* as in *John loves Paul's cousin* is not considered an I_STATE. Similarly for *like*. An I_STATE must govern another event.

(84) We aim at triggering associations that will generate **lust** for change.

c) **hope, expect, aspire, plan**:

(85) We **aim** at triggering associations that will generate lust for change.

d) **fear, hate, dread, worry, be afraid**:

(86) The agencies **fear** they will be unable to crack those codes to eavesdrop on spies and crooks.

e) **need, require, demand**

f) **be ready, be eager, be prepared**

(87) The young industry's rapid growth also is attracting regulators **eager** to police its many facets.

g) **be able, be unable**

(88) The agencies fear they will be **unable** to crack those codes to eavesdrop on spies and crooks.

- STATE: States describe *circumstances* in which something obtains or holds true. However, we will only annotate temporally relevant states; that is:

a) States that are identifiably changed over the course of the document being marked up.
In these and the following examples the markable state is in bold face.

(89) a. All 75 people **on board** the Aeroflot Airbus died.
b. Israel has been scrambling to buy more masks abroad, after a **shortage** of several hundred thousand gas masks was discovered.
c. No **injuries** were reported over the weekend.

b) States that are directly related to a temporal expression.
This criterion includes all states that are linked to a TIMEX3 markable by means of a TLINK (see clauses **??** and **??**). Two examples are given here, where the state is in bold face and the temporal expression associated with it is underlined.

(90) a. James Pustejovsky was **CTO** for several years.
b. They **lived** in U.N.-run refugee camps for 2 1/2 years.

c) States that are introduced by an I_ACTION, an I_STATE, or a REPORTING event. States are in bold face, the introducing event underlined.

(91) a. He mediated the **crisis**.
b. Saddam Hussein sought **peace** on another front.
c. Har-Shefi told police that Rabin was a **traitor**.

d) Predicative states the validity of which is dependent on the document creation time
In spite of not being explicitely related to any TIMEX3 expression, the states underlined in the examples below will be tagged because their validity is relative to the point in time they have been asserted (the DCT).

(92) a. A total of about 3,000 Americans, 3,000 Britons and more than 450 Japanese are **in Iraq**.
b. Overall, more than 2 million foreigners are **in both countries.**

This criterion also includes quantitative statements such as those from financial journals:

(93) Gas prices fell from a **twenty-two dollar barrel level** down to the **fourteen dollars** we're seeing today.

However, it applies only to predicative states. Sortal states (like *President, CTO*, etc.) will not be marked up.

NOTE The current class, STATE, does not contain states that have been tagged as I_STATEs.

- OCCURRENCE: This class includes all the many other kinds of events describing situations that happens or occurs in the world.

(94) a. The Defense Ministry said 16 planes have **landed** so far with protective equipment against biological and chemical warfare.
b. Two moderate **eruptions** shortly before 3 p.m. Sunday appeared to signal a larger **explosion**.
c. RMS said it **had** a **loss** of $158,666, or 10 cents a share, in the third quarter, compared with a year-earlier **loss** of $29,956, or two cents a share.
d. Ralston said its restructuring **costs** include the **phase-out** of a battery facility in Greenville, N.C....

**C. Attribute** pos**:**
Required attribute. It captures distinctions among the grammatical categories of elements which are marked as events. This attribute can have the following values: ADJECTIVE, NOUN, VERB, PREPOSITION, OTHER. Sentences (**??**-**??**), repeated below, illustrate the use of attributes VERB (**??**), NOUN (**??**), ADJECTIVE (**??**), and PREPOSITION (**??**).

(95) a. A fresh flow of lava, gas and debris **erupted** there Saturday.
b. Israel will ask the United States to delay a military **strike** against Iraq until the Jewish state is fully prepared for a possible Iraqi **attack**.
c. A Philippine volcano, **dormant** for six centuries, began exploding with searing gases, thick ash and deadly debris.
d. All 75 people **on board** the Aeroflot Airbus died.

**D. Attribute** tense**:**
Required. Capturing standard distinctions in the grammatical category of verbal tense. It can have values PRESENT, PAST, FUTURE, IMPERFECT, or NONE.

For languages in which tense distinctions do not apply, the value NONE will be used as default. This value can however be overwritten if there is a non-verbal element (e.g., adverbs of time) in the sentence conveying a value equivalent to tense for a particular event.

Among languages with tense distinctions, tenses are not easily mapped. The annotators should develop a specific ISO-TimeML spec for annotating tense (and also aspect) in the particular language they annotate. Basic guides for deciding among the different values are:

- PRESENT: It is the tense generally used to express action at the present time and states, but also, possibly: habitual events, occurrences in the near future, or actions that started in the past and still hold in the present.
- PAST: Expressing action and states of being in a past time.
- FUTURE: Used for describing events as not having happened yet, but expected to in the future.
- IMPERFECT: Assigned to finite forms, the imperfect is a descriptive past tense which indicates an ongoing state of being or a repeated or incomplete action. The beginning and end of the state of being or action are not indicated, and the imparfait is very often translated in English as "was" or "was ____-ing."
- NONE: No tense value is found (relevant for non-verbal event-denoting expressions: nouns, adjectives, and prepositions.

ISO-TimeML is conceived as a highly surface-based language, in the sense that it does not aim at annotating meaning but at providing a way to normalize temporally relevant expressions. The annotators for each language are encouraged to develop language-specific ISO-TimeML specs having that in mind.

**E. Attribute** `aspect`**:**
Required. Similarly to `tense`, it captures standard distinctions in the grammatical category of verbal aspect. It can have values `PROGRESSIVE, IMPERFECTIVE, PERFECTIVE, IMPERFECTIVE_PROGRESSIVE, PERFECTIVE_PROGRESSIVE`, or `NONE`.

As with `tense`, languages without aspect distinctions in the verbal system will assign the value `NONE` as default. This value can however be overwritten if there is a non-verbal element (e.g., adverbs of time, phrases, etc.) in the sentence conveying a value equivalent to aspect for a particular event.

For languages in which aspect distinctions apply, the basic guides for deciding among different aspect values are:

- PROGRESSIVE: Expressing, among other possibilities: actions in progress and outgoing activities; durative activities and continuous states; activities posing the background for other activities; simultaneous activities; etc. (e.g., *Prof. Abramovitz was teaching that day/could be teaching on Friday.*)
- PERFECTIVE: Generally expressing states and activities which were ended (e.g., *Prof. Abramovitz has conducted experiments in different countries around the world.*)
- IMPERFECTIVE: Generally expressing states and activities that are seen from a particular viewpoint as ongoing, habitual, repeated, or generally containing internal structure. This is distinct from the progressive. English does not have a proper imperfective aspect.
- PERFECTIVE_PROGRESSIVE: Combining the meanings of progressive and perfective (e.g., *Prof. Abramovitz has been teaching for his whole life.*)
- IMPERFECTIVE_PROGRESSIVE: Combining the meanings of progressive and imperfective in languages that have imperfective aspect.
- NONE: No aspect value is found. Relevant for non-verbal event-denoting expressions (nouns, adjectives, and prepositions).

**F. Attribute** `polarity`**:**
Required. Boolean attribute that conveys the polarity of the event in question. If it is set to `NEG`, the event instance is negated. If it is set to `POS`, the event instance is not negated.

**G. Attribute** `mood`**:**
Required. Captures the mood of the event. Presently, it can have either `SUBJUNCTIVE` or `NONE`. If no inflectional morphology is present to indicate mood, then the default value is `NONE`.

**H. Attribute** `modality`**:**
Optional. Conveying the modality nature of the event: different degrees of epistemic modality, deontic modality, etc. The particular values for this attribute will be language specific.

**I. Attribute** `vform`**:**
Required. Capturing standard distinctions in the grammatical category of non-tensed verbal forms . It can have values `INFINITIVE`, `PRESPART`, `PASTPART`, `GERUNDIVE`, or `NONE`.

- `INFINITIVE`: Assigned to the basic form of a verb (non-finite) when used in an embedded section.
- `PRESPART`: Assigned to non-finite forms of the verb with an active meaning. PRESPARTs can perform different functions, among others: (a) modify nouns (**??**) or (b) head a predicate (**??**).

  (96) The woman **wearing** the long white dress is the hostess.

  (97) **Being awfully tired**, she sat down.
- `PASTPART`: Assigned to non-finite forms of the verb with a passive (or also perfective) meaning.
- `GERUNDIVE`: Assigned to non-finite forms of a gerund verbal form (e.g., *fishing*, *teaching*), where they function as nouns (**??**).

  (98) **Walking** is good for the heart.

### A.2.2 The tag `<TIMEX3>`

The `<TIMEX3>` tag annotate any temporal expression (a.k.a. *timex*) referring to:

- Day times (*noon, 3p.m., the evening, ...*).
- Dates of different granularity: days (*yesterday, Jan 8 2001, last Friday*, etc.), weeks (*next week, the second week of July*, etc.), months (*in two months, August 1971*), seasons or business quarters (*last spring, the third quarter*, etc.), years (*1978, the previous year*), centuries, etc.
- Durations (*two months, five hours*).
- Sets (*every Thursday, the first Sunday of the month*).

Previous to the TimeML initiative, there is already some research devoted to the annotation of temporal expressions; most remarkably STAG (Sheffield Temporal Annotation Guidelines), [?]), which uses the tag `TIMEX` for annotating temporal expressions in the context of newswire articles, and TIDES ([?], [?]), which introduces the `TIMEX2` tag.

The specifics of the TimeML tagset for annotating temporal expressions differ in detail from both the `TIMEX` tag in STAG and the `TIMEX2` tag in TIDES. Because of that, we adopt the tag name `TIMEX3`.

#### A.2.2.1 How to annotate TIMEX3s

**A. Timex identification:**
In order to be as compliant as possible with TIDES `TIMEX2` annotation, the `TIMEX3` tag will, for the most part, be applied to the same TIDES `TIMEX2` markable expressions (refer to TIDES(02), section 2). However, ISO-TimeML will differ from TIDES with regard to the tag span, as detailed in the following subclause.

**B. Timex tag span:**
As already seen, ISO-TimeML aims at a surface-oriented approach to the tagging of expressions in text. Because of that, it is conceived from a highly compositional view. The `TIMEX3` span will be based on the constituent structure of each particular language, but it will also make use of the classification of temporal units shown in Table 6, and the type of relations holding among two different temporal expressions (**??**).

**Table 6: Time units classification**

| *t* <**day** | *t*=**day** | **day**<*t*<**year** | *t*=**year** | *t* >**year** |
|---|---|---|---|---|
| *twelve o'clock*<br>*midnight*<br>*morning*<br>*ten minutes to two* | *Wednesday*<br>*tomorrow*<br>*Jan. 2*<br>*the 8th*<br>*Christmas Day*<br>*Worker's Day* | *first week*<br>*month*<br>*semester*<br>*January*<br>*season*<br>*Fall* | *1984*<br>*next year* | *last century* |

(99) a. **Specification relation:** Involving two temporal expressions, one of which is helping to further specify the other (e.g., *[twelve o'clock] [midnight], [four] in [the afternoon], [Tuesday] [Jan. the 18th], [this year's] [summer], [some Thursdays] in [1984]*).

b. **Anchoring relation:** Involving two temporal expressions, one of which is ordered, or anchored, relative to the other. In English, they generally involve the use of temporal prepositions and conjunctions like *from, before, after, following, prior to*, etc. For instance, *[two weeks] from [next Tuesday], [2 days] before [yesterday], [ten minutes] to [four], [three years ago] [today]*. These expressions are also known as *anchored durations*.

c. **Conjunction relation:** Involving two temporal expressions related by a coordination conjunction – mainly, *and* and *or* (e.g., *[six months] or [a year]*).

The `TIMEX3` span need be compliant with the following general rules:

a) **The full extent of the tag** must correspond to one of the following categories:

- Noun phrase (*the afternoon, last summer, yesterday, Sunday*).
- Adjective phrase (*half an hour long, half-hour* –as in *a half-hour trip*).
- Adverbial phrase (*fairly recently*).

Therefore, any preposition preceding a temporal expression (as in *in the afternoon, before yesterday, in half an hour*,...) will not be included as part of the tag:

NOTE Temporally relevant prepositions will be annotated as signals. See clause **??**.

(100) a. in the afternoon

b.
```
in
<TIMEX3 tid="t1">
the afternoon </TIMEX3>
```

On the other hand, adverbial postmodifiers (*ago, ever*) will be considered part of the `TIMEX3` markable expression (**??-??**), but not postmodifiers that express an event (**??-??**).

(101) a. the best second quarter **ever**

b.
```
<TIMEX3 tid="t1">
the best second quarter ever </TIMEX3>
```

(102) a. three years **ago**

b.
```
<TIMEX3 tid="t1">
three years ago </TIMEX3>
```

(103) a. five days after he came back

b.
```
<TIMEX3 tid="t1">
five days
</TIMEX3>
after he
<EVENT eid="e1" class="OCCURRENCE">
came
</EVENT>
back
```

(104) a. nearly four decades of experience

b.
```
<TIMEX3 tid="t1">
nearly four decades
</TIMEX3>
of
<EVENT eid="e2" class="STATE">
experience
</EVENT>
```

(105) a. months of renewed hostility

b.
```
<TIMEX3 tid="t1">
months
</TIMEX3>
of renewed
<EVENT eid="e2" class="STATE">
hostility
</EVENT>
```

b) **Two temporal expressions in a specification relation** will be marked up with a single tag if:

- The two expressions belong to the same class in Table 6 (e.g., *12 o'clock midnight, Tuesday, Jan 18th, eleven in the morning*).
- The two expressions belong to the same syntactic constituent. Syntactic constituency can be checked using the fronting or clefting tests (examples b-c and d-e in the examples below, respectively). Different constituents allow fronting and clefting, but not two parts of the same constituent. For example, the two temporal expressions in (**??**) are part of the same constituent, while those in (**??**) are not.

(106) a. The different groups will meet at 11a.m. Jan. 3, 2005.
b. *Jan. 3, 2005, the different groups will meet at 11a.m.
c. *At 11a.m., the different groups will meet Jan. 3, 2005.
d. *It will be Jan. 3, 2005, that the different groups will meet at 11a.m..
e. *It will be at 11a.m, that the different groups will meet Jan. 3, 2005..

(107) a. The different groups will meet at 11a.m. on Jan. 3, 2005.
b. On Jan. 3, 2005, the different groups will meet at 11a.m.
c. At 11a.m., the different groups will meet on Jan. 3, 2005.
d. It will be on Jan. 3, 2005, that the different groups will meet at 11a..m.
e. It will be at 11a.m, that the different groups will meet on Jan. 3, 2005.

Some other examples of one-constituent temporal expression are: *Friday evening, Tuesday the 18th, twelve o'clock January 3, 1984, the second of December, October of 1963, last year's summer*.

(108) a. **This year's summer** was unusually hot.

b.
```
<TIMEX3 tid="t1">
This year's summer
</TIMEX3>
```

c) **Temporal expressions in an anchoring relation** will be generally marked up with two independent tags. For instance: *[two weeks] from [next Tuesday]*, *[2 days] before [yesterday]*, *[three years ago] [today]* (**??-??**).

NOTE Links will be used to express the relative ordering of the two temporal expressions – see clause **??**.In addition, these types of expression are considered **anchored durations** and can be annotated as such in the newest version of ISO-TimeML – see subclause **??**.

The only exception will be those temporal expressions denoting day time, such as *[ten minutes] to [four]*, which will be annotated with one single tag (**??**).

(109) a. I'm leaving on vacation **two weeks from next Tuesday**.

b.
```
<TIMEX3 tid="t1">
two weeks
</TIMEX3>
from
<TIMEX3 tid="t2">
next Tuesday
</TIMEX3>
```

(110) a. John left **2 days before yesterday.**

b.
```
<TIMEX3 tid="t1">
2 days
</TIMEX3>
before
<TIMEX3 tid="t2">
yesterday.
</TIMEX3>
```

(111) a. A major earthquake struck Los Angeles **three years ago today**.

b.
```
<TIMEX3 tid="t1">
three years
</TIMEX3>
ago
<TIMEX3 tid="t2">
today
</TIMEX3>
```

(112) a. I'm leaving at **ten minutes to four**.

b.
```
<TIMEX3 tid="t1">
ten minutes to four
</TIMEX3>
```

d) **Two temporal expressions in a conjunction relation** will be marked up as two different tags.

(113) a. Saddam might play the whole game again **six months or a year from now**

b.
```
<TIMEX3 tid="t1">
six months
</TIMEX3>
or
<TIMEX3 tid="t2">
a year
</TIMEX3>
from <TIMEX3 tid="t3">
now
</TIMEX3>
```

NOTE In the example above, the expression *(from) now* is in an anchoring relation –with both *six months* and *a year*. Hence, it needs to be marked up with an independent tag.

#### A.2.2.2 BNF for the TIMEX3 tag

```
attributes ::= tid type [functionInDocument][beginPoint]
          [endPoint] [quant][freq][temporalFunction]
          (value|valueFromFunction) [mod][anchorTimeID]
tid ::= ID
{tid ::= TimeID
TimeID ::= t<integer>}
type ::= 'DATE'|'TIME'|'DURATION'|'SET'
beginPoint ::= IDREF
{beginPoint ::= TimeID}
endPoint ::= IDREF
{endPoint ::= TimeID}
quant ::= CDATA
freq ::= CDATA
functionInDocument ::= 'CREATION_TIME' | 'EXPIRATION_TIME' |
             'MODIFICATION_TIME' | 'PUBLICATION_TIME' |
             'RELEASE_TIME'| 'RECEPTION_TIME' | 'NONE'
             {default, if absent, is 'NONE'}
temporalFunction ::= 'true' | 'false' {default, if absent, is 'false'}
{temporalFunction ::= boolean}
value ::= CDATA
{value ::= duration | dateTime | time | date | gYearMonth |
             gYear | gMonthDay | gDay | gMonth}
valueFromFunction ::= IDREF
{valueFromFunction ::= TemporalFunctionID
TemporalFunctionID ::= tf<integer>}
mod ::= 'BEFORE' | 'AFTER' | 'ON_OR_BEFORE' | 'ON_OR_AFTER' |
             'LESS_THAN' | 'MORE_THAN' | 'EQUAL_OR_LESS' |
             'EQUAL_OR_MORE' | 'START' | 'MID' | 'END' | 'APPROX'
anchorTimeID ::= IDREF
{anchorTimeID ::= TimeID}
```

#### A.2.2.3 Attributes for TIMEX3

**A. Attribute** `tid` **(`TIMEX` ID number)**
Required attribute. Each `TIMEX3` expression has to be identified by a unique ID number.

**B. Attribute** `type`
Required attribute. Each `TIMEX3` is assigned one of the following types: `DATE`, `TIME`, `DURATION`, or `SET`.

- `DATE`: The expression describes a calendar time.

  (114)Mr. Smith left **Friday, October 1, 1999**
  **the second of December**
  **yesterday**
  in **October of 1963**
  in **the summer of 1964**
  on **Tuesday 18th**
  in **November 1943**
  **this year's summer**
  **two weeks from next Tuesday**
  **last week**

  `DATE` can also be the value for the `type` attribute of each of the two `TIMEX3` markable expressions constituting a range, as long as they describe a calendar time.

(115) a. John left between **Monday** and **Wednesday**

b.
```
John left between
<TIMEX3 tid="t1" type="DATE">
Monday
</TIMEX3>
and
<TIMEX3 tid="t2" type="DATE">
Wednesday
</TIMEX3>
```

- TIME: The expression refers to a time of the day, even if in a very indefinite way (as in the two last examples below):

(116)Mr. Smith left **ten minutes to three**
at **five to eight**
at **twenty after twelve**
at **half past noon**
at **eleven in the morning**
at **9 a.m. Friday, October 1, 1999**
**the morning of January 31**
**late last night**
**last night**

As before, TIME can also be the type value for each of two TIMEX3 markable expressions that together refer to a temporal range (e.g., *Mr. Smith left between* **8:00 a.m.** *and* **10:00 a.m.**).

- DURATION: The expression describes a duration. This value is assigned only to explicit durations like the following:

(117)Mr. Smith stayed **2 months** in Boston
**48 hours**
**three weeks**
**all last night**
**20 days** in July
**3 hours** last Monday.

As a rule, if any specific calendar information is supplied in the temporal expression, then the type of the TIMEX3 must be either DATE or TIME. Some annotators may be tempted to make something like "1985" a DURATION if the context suggests that an event holds throughout that year. However, temporal expression like the one described here must be of type DATE, since they refer to a particular area in the temporal axis –even though that area spans over a period of time. By contrast, durations are periods of time not pointing at any specific area in the temporal axis.

- SET: The expression describes a set of times. This value is assigned to expressions such as those in section 3.5 of TIDES(02). For example:

(118)John swims **twice a week**.
**every 2 days**.

**C. Attribute** value**:**
The attribute value (equivalent to VAL in TIMEX2) will be annotated exactly as specified in TIDES(02) sections 3.2.and 3.3. Note however that these sections also include the use of two additional attributes, ANCHOR_VAL and ANCHOR_DIR, which are not used in ISO-TimeML.

The format of this attribute value is determined by the type attribute. For instance, a DURATION must have a value that begins with the letter 'P' (standing for period of time), and a TIME witht the letter 'T' (standing for a time) that includes times of the day.

The following examples, from previous clauses, partially illustrate the use of the value attribute for times of the day, dates, durations, and sets:

- Times of the day, and dates:

  (119) <TIMEX3 tid="t1" type="TIME" value="T16:00">
  4 p.m.</TIMEX3>

  The annotator will introduce as much information as is available both in the time expression and from the context. In case the text would include some reference to the specific date in which the time is anchored. For instance, given the sentence *Last Friday's meeting didn't start until 4 p.m.*, assuming that the document creation time is Friday, July 12, 2002, then the `value` attribute must specify the full date that can be computed from the document creation time, and be:

  (120) <TIMEX3 tid="t1" type="TIME" value="2002-07-05T16:00">
  4 p.m.</TIMEX3>

  Unknown information is left underspecified by means of the placeholder 'X'. In the next example, for instance, the year is unknown.

  (121) <TIMEX3 tid="t5" type="DATE" value="XXXX-12-02">
  the second of December
  </TIMEX3>

- Durations:

  (122) <TIMEX3 tid="t1" type="DURATION" value="P4M">
  4 months</TIMEX3>

  (123) during
  <TIMEX3 tid="t1" type="DURATION" value="P2D">
  two entire days
  </TIMEX3>
  on
  <TIMEX3 tid="t2" type="DATE" value="1999-SU">
  the summer of 1999
  </TIMEX3>

- Sets:

  To fully annotate sets, the `TIMEX3` must also include either the `quant` or `freq` attributes, if not both. The following examples begin the annotation of a `TIMEX3 set` as pertains to the value attribute:

  (124) <TIMEX3 tid="t1" type="SET" value="P1W">
  twice a week</TIMEX3>

  (125) <TIMEX3 tid="t1" type="SET" value="P2D">
  every 2 days </TIMEX3>

  (126) <TIMEX3 tid="t1" type="SET" value="P1W">
  3 days each week </TIMEX3>

  (127) <TIMEX3 tid="t1" type="SET" value="XXXX-10">
  every October </TIMEX>

**D. Attribute** `mod`**:**
Optional attribute, inherited directly from the TIMEX2 MOD attribute. Its value is as specified in TIDES(02), section 3.4.

**E. Attribute** `temporalFunction`**:**
Binary attribute which expresses whether the value of the temporal expression needs to be determined via evaluation of a temporal function. Temporal functions will be applied as a postprocess.

The value for this attribute will be positive for those cases that do not contain all the information necessary to fill the higher-order (left-hand) positions in the `value` attribute (**??**). This will apply even if `value` can be completely filled, given additional information provided by the context.

(128) a. *eleven in the morning:* missing the particular day.
b. *January, 31:* missing the year.
c. *last week:* missing the month and year.

On the other hand, for cases in which the higher-order position of `value` are filled from the information provided by the tagged temporal expression, `temporalFunction` should be assigned a negative value. Such cases include:

(129) a. *twelve o'clock January 3, 1984*
b. *summer of 1964*
c. *Friday, October 1, 1999*
d. *the morning of January 31, 1999*

Durations whose length is underspecified will receive `true` as the value of `temporalFunction`. Compare (**??**) with (**??**), which indicates a specific length.

(130) a. in **recent months**

b.
```
<TIMEX3 tid="t1" type="DURATION" value="PXM"
temporalFunction="true">
```

(131) a. for **two months**

b.
```
<TIMEX3 tid="t1" type="DURATION" value="P2M"
temporalFunction="false">
```

**F. Attribute** `anchorTimeID`**:**
Optional attribute. It introduces the `tid` value of the time expression to which the `TIMEX3` markable is temporally anchored. That is, the time expression needed in order to compute its `value` attribute value.

In (**??**), both `TIMEX3` expressions require the application of a temporal function: *Monday, July 15* needs to be evaluated with respect to the document creation (`t0`) in order to compute the particular year, and *9:00 a.m.* relative to `t1`.

(132) a. The TERQAS Workshop will resume **Monday, July 15**. The session will start at **9:00 a.m.**

b.
```
The TERQAS Workshop will resume
<TIMEX3 tid="t1" type="DATE" value="2002-07-15"
temporalFunction="true" anchorTimeID="t0">
Monday, July 15
</TIMEX3>
.. The session will start at
<TIMEX3 tid="t2" type="TIME" value="T9:00"
temporalFunction="true" anchorTimeID="t1">
9:00 a.m.
</TIMEX3>
```

NOTE The presence of the `anchorTimeID` attribute appears together with `temporalAnchor='true'`. However, this is not always the case.

**G. Attribute** `valueFromFunction`**:**
This attribute is not relevant for the purposes of manual annotation, but only for the postprocess. The human annotator should ignore it.

**H. Attribute** `functionInDocument`**:**
Optional. It indicates the function of a `TIMEX3` in providing a temporal anchor for other temporal expressions in the document. There are several times that mark the major milestones in the life of a textual document:

- the time the text is created
- the time the text is modified
- the time the text is published
- the time it may be released (if not immediately)
- the time it is received by a reader
- the time that the text expires (if any)

The possible values for this attribute are then: `'CREATION_TIME', 'MODIFICATION_TIME', 'PUBLICATION_TIME` `'RELEASE_TIME', 'RECEPTION_TIME', 'EXPIRATION_TIME', 'NONE'`. If this attribute is not explicitly supplied, the default value is `'NONE'`.

The following attributes are used to strengthen the annotation of durations and sets in ISO-TimeML. Each is optional and can be used at the annotator's discretion. It is important to note, however, that `beginPoint` and `endPoint` should only be used when the `type` of the expression is `DURATION`, and that `quant` and `freq` should only be used when the expression is a `set`.

**I. Attributes** `beginPoint` **and** `endPoint`**:**
Used when a duration is anchored by one or two time expressions indicating its begin and/or end points. If only one of these points is provided, the annotator can create an empty `TIMEX3` to represent the missing point.

NOTE The values stored in these attributes can be used by temporal functions to compute the missing points and create a tag for them.

In some ways, the `beginPoint` and `endPoint` attributes are similar to `anchorTimeID`. In (**??**), for instance, the new `tid` introduced by the empty `TIMEX3` can be used to link the teaching event directly to the time at which it takes place.

(133)
```
John begins teaching
<TIMEX3 tid="t1" type="DURATION" value="P1W"
beginPoint="t2" endPoint="t3">
one week </TIMEX3>
from
<TIMEX3 tid="t2" type="DATE" value="XXXX-9-15">
September 15 </TIMEX3>
<TIMEX3 tid="t3" type="DATE" value="XXXX-9-22"
temporalFunction="TRUE" anchorTimeID="t1" />
```

**J. Attributes** `quant` **and** `freq`**:**
Used when a temporal expression is of the type `SET`. `quant` is generally a literal from the text that quantifies over the expression. `freq` contains an integer value and a time granularity that represent the frequency at which the temporal expression regularly reoccurs. These attributes are only used if their values are supplied by the temporal expression (or by a temporal anchor). Though it seems on occasion that values for these attributes can be inferred, they will not be for purposes of manual annotation. Although, if there is no specified `quant`, one imagines that the set is universally quantified. The following examples complete the annotations of the sets listed earlier in this section:

(134) <TIMEX3 tid="t1" type="SET" value="P1W" freq="2X">
twice a week</TIMEX3>

(135) <TIMEX3 tid="t1" type="SET" value="P2D" quant="EVERY">
every 2 days</TIMEX3>

(136) TIMEX3 tid="t1" type="SET" value="P1W" quant="EACH" freq="3d">
3 days each week</TIMEX3>

(137) <TIMEX3 tid="t1" type="SET" value="XXXX-10" quant="EVERY">
every October</TIMEX>

#### A.2.3 The tag <SIGNAL>

A signal is a textual element that makes explicit the relation holding between two entities (timex and event, timex and timex, or event and event). Signals are generally:

- **Temporal prepositions:** *on, in, at, from, to, before, after, during,* etc.
- **Temporal conjunctions:** *before, after, while, when,* etc.
- **Special characters:** "-" and "/", in temporal expressions denoting ranges (*September 4-6, April 1999/July 1999*, etc.).

##### A.2.3.1 How to annotate SIGNALs

Generally, the SIGNAL tag spans over one word or element:

(138) a. John taught <SIGNAL sid="s1">on</SIGNAL> Monday

b. All passengers died <SIGNAL sid="s1">when</SIGNAL> the plane crashed into the mountain.

When two distinct signals appear side by side, they can be annotated separately, if they belong to different signal classes as listed above. However, some situations require that they must be annotated as a single SIGNAL. For example, in (**??**) the three temporal prepositions need to be collapsed into a single SIGNAL in order to properly recover the IS_INCLUDED relation of the TLINK between the *genocide* and the *role* events (see section **??**).

(139) They will investigate the role of the US <SIGNAL sid="s1">before, during and after</SIGNAL> the genocide

##### A.2.3.2 BNF for the SIGNAL tag

```
attributes ::= sid
sid ::= s<integer>
```

**Attributes for SIGNAL**

SIGNAL has only one, non-optional, attribute: sid, the signal's unique id.

### A.3 The link tags: <TLINK>, <SLINK>, and <ALINK>

There are three types of link tags. The function of each will be introduced here, before we move on to explaining in detail how links are annotated.

### A.3.1 The tag <TLINK>

A TLINK (or temporal link) represents the temporal relationship holding between two events, two times, or between an event and a time, and indicates how they are related. The possible **relType** values for the TLINK are:

a) **Simultaneous:** Two events are judged simultaneous if they happen at the same time, or are temporally indistinguishable in context, i.e. they occur close enough so that further distinguishing their times makes no difference to the temporal interpretation of the text. This is also used for expressing the duration of an ongoing event, as in

(140) Mary **taught** for 20 minutes.

Here, the event of teaching is SIMULTANEOUS to the duration interval expressed by the TIMEX3 "20 minutes."

b) **One before the other** (BEFORE): As in the following example between the events *slayings* and *arrested*:

(141) The police looked into the **slayings** of 14 women. In six of the cases suspects have already been **arrested**.

c) **One after the other** (AFTER):

This is just the inverse of the preceding relation. So the two events of the previous example can alternatively be annotated as expressing an **after** relation, if the directionality is changed.

d) **One immediately before the other** (IBEFORE):

As in the following sentence between *crash* and *died*.

(142) All passengers **died** when the plane **crashed** into the mountain

e) **One immediately after than the other** (IAFTER): This is the inverse of the preceding relation.

f) **One including the other** (INCLUDES): As is the case between the temporal expression and the event in the following example:

(143) John **arrived** in Boston **last Thursday**.

g) **One being included in the other** (IS_INCLUDED): The inverse relation to the preceding one.

h) **One holds during the other** (DURING):

Similar to INCLUDES, but used to relate one event included within another event.

(144) Mary **sneezed** while **running**.

i) **One being the beginning of the other** (BEGINS): As it holds between the first of the temporal expressions and the event in the following example:

(145) John was **in the gym** between **6:00 p.m.** and 7:00 p.m.

j) **One being begun by the other** (BEGUN_BY): The inverse relation to the one just introduced.

k) **One being the ending of the other** (ENDS) **:**

As in:

(146) John was **in the gym** between 6:00 p.m. and **7:00 p.m.**.

l) **One being ended by the other** (ENDED_BY) **:** The inverse relation to the one just introduced.

In addition, TLINKs are also used in the following situations:

a) **Event identity** (IDENTITY) **:**

Event identity is annotated via the TLINK (e.g., *John* **drove** *to Boston. During his* **drive** *he ate a donut*). Furthermore, TLINK will be used in order to relate the events in:

- **Causative constructions.** Triggered by verbs like the following, in their causative sense: *cause, stem from, lead to, breed, engender, hatch, induce, occasion, produce, bring about, produce, secure*. Two cases can be distinguished

  **Case 1** EVENT cause EVENT

  *The* **[rains] [caused]** *the* **[flooding]**.

  **Case 2** ENTITY cause EVENT

  *John* **[caused]** *the* **[fire]**.

  The event introduced by the subject in Case 1 is related to the verbal predicate by a TLINK expressing event identity (we will see later that this means setting the TLINK relType attribute as IDENTITY); on the other hand, the relation between the verbal event and that expressed by the object is represented by a TLINK of type BEFORE).

  Case 2 exhibits what is called event metonymy ([?], [?]). Such constructions will not be annotated in the current specification, but will be included in the next release, where we will introduce a skolemized event instance, *ei*1, to act as the proxy in the causation relation.
- **Light verb constructions:** Similar to the case above, the verbal and nominal events will be related by means of a TLINK expressing event identity.

Event identity is a very important relationship, which will not be picked up during the closure part of the annotation. So it is extremely important to make sure that all identity links are annotated.

b) **When a set/subset relationship occurs in the text:**

An example is:

(147) The police looked into the **slayings** of 14 women. In six of the **cases** suspects have already been arrested.

Two EVENT tags are created, for each of the event sets: a first one marking up *slayings*, with cardinality 14, and a second one for *cases*, with cardinality 6. The two events will be related via a TLINK with the temporal relation IS_INCLUDED (or INCLUDES, depending on the directionality).

#### A.3.1.1 How to annotate TLINKs

Here and in the following clauses for SLINKs and ALINKs, we use examples to demonstrate how to create each of the link types. In these examples we do not give detailed annotation of events, times and signals – please refer to the appropriate clauses for instructions on annotating these. Also, we only show the mark-up for those entities which are relevant to the examples.

A TLINK has to be created each time a temporal relationship holding between events or an event and a time needs to be annotated. This includes the important relationship of event identity. Examples:

a) *John taught on Monday*

The temporal relationship holding between the event and the time expression, as indicated by the signal *on*, is marked up by introducing the following TLINK:

```
<TLINK eventInstanceID="ei1" relatedToTime="t1" signalID="s1"
        relType="IS_INCLUDED"/>
```

b) *John taught every Monday*

The TIMEX3 representing the multiple instances of *Monday* looks as follows:

```
<TIMEX3 tid="t1" type="SET" value="XXXX-WXX-1" quant="EVERY">
every Monday
</TIMEX3>
```

The TLINK representing the temporal relation holding between the event and the temporal expression looks like this:

```
<TLINK eventInstanceID="ei9" relatedToTime="t1"
  relType="IS_INCLUDED"/>
```

c) *John taught for 20 minutes on Monday.*

The EVENT tag representing the event *taught* looks as follows:

```
<EVENT eid="e4"> eiid="ei4" taught</EVENT>
```

Two TLINKs have to be introduced.

NOTE As is often the case, these are not the only TLINKs that can be drawn from this text. However, post-processing algorithms should be able to fill in the links the annotator leaves out.

One TLINK captures the fact that the *taught* event holds throughout the *20 minutes*, and one TLINK captures the fact that the *20 minutes* temporal expression is included in *Monday*.

```
<TLINK eventInstanceID="ei4" relatedToTime="t2" signalID="s5"
        relType="SIMULTANEOUS"/>
<TLINK tid="t2" relatedToTime="t3" signalID="s6"
    relType="IS_INCLUDED"/>
```

d) *John drove to Boston. During his drive he ate a donut.*

The EVENT tags presenting the events *drove* and *drive* look as follows:

```
<EVENT eid="e1" eiid="ei1">drove</EVENT>
<EVENT eid="e2" eiid="ei2" >drive</EVENT>
```

The TLINK that has to be introduced to represent the identity of these two events looks as follows:

```
<TLINK eventInstanceID="ei1" relatedToEvent="e2" relType="IDENTITY"/>
```

Please note that we did not include the TLINK that would represent the temporal relationship between *ate* and *drive*.

#### A.3.1.2 BNF for the TLINK tag

```
attributes ::= [lid] [origin] (eventInstanceID | timeID) [signalID]
               (relatedToEventInstance | relatedToTime) relType
lid ::= ID
{lid ::= LinkID
LinkID ::= l<integer>}
origin ::= CDATA
eventInstanceID ::= IDREF
{eventInstanceID ::= EventInstanceID}
timeID ::= IDREF
{timeID ::= TimeID}
signalID ::= IDREF
{signalID ::= SignalID}
relatedToEventInstance ::= IDREF
```

```
{relatedToEventInstance ::= EventInstanceID}
relatedToTime ::= IDREF
{relatedToTime ::= TimeID}
relType ::= 'BEFORE' | 'AFTER' | 'INCLUDES' | 'IS_INCLUDED'
            | 'DURING' | 'DURING_INV' | 'SIMULTANEOUS'
            | 'IAFTER' | 'IBEFORE' | 'IDENTITY' | 'BEGINS'
            | 'ENDS' | 'BEGUN_BY' | 'ENDED_BY'
```

#### A.3.1.3 Attributes for TLINKs

a) **Attributes** `eventInstanceID` **or** `timeID`**:**

Obligatory attribute (one or the other of these needs to be present). This is the ID of the `event` or the `time` involved in the temporal link.

b) **Attribute** `signalID`**:**

Optional attribute. If the temporal relation holding between the entities is explicitly signaled in the text, then the ID of that signal needs to be supplied here.

c) **Attributes** `relatedToEventInstance` **or** `relatedToTime`**:**

Obligatory attribute (one or the other of these needs to be present). This is the ID of the entity that is being related to the event instance with `ID=eventInstanceID` or time expression with `ID=timeID`.

d) **Attribute** `relType`**:**

Obligatory attribute. This is the temporal relation holding between the entities. Possible values are: `BEFORE`, `AFTER`, `INCLUDES`, `IS_INCLUDED`, `DURING`, `DURING_INV`, `SIMULTANEOUS`, `IAFTER`, `IBEFORE`, `IDENTITY`, `BEGINS`, `ENDS`, `BEGUN_BY`, `ENDED_BY` . They are assigned according to the instructions given at the beginning of clause **??**, on `TLINK`s. There will be only one relation assigned per `TLINK`.

### A.3.2 The tag `<SLINK>`

An `SLINK` (or subordination link) is used for contexts introducing relations between two events. `SLINKs` are of one of the following sorts:

`INTENSIONAL:`

This relation is brought up by events introducing a reference to a possible world – mainly `I_ACTION`s and `I_STATE`s:

*John* **promised** *Mary to buy some beer.*
*Mary* **wanted** *John to buy some wine.*
*The police* **attempted** *to arrest the robber.*

`Factive:`

Certain verbs presuppose or entail the veradicity (or factuality) of their event argument. They include *forget* (with a tensed complement), *regret*, or *manage* (in positive contexts):

*John* **forgot** *that he was in Boston last year.*
*Mary* **regrets** *that she didn't marry John.*
*John* **managed** *to leave the party*

`Counter-factive:`

Contrary to the previous relation, in this case the event presupposes the non-veracity of its argument; e.g., *forget (to), unable to* (in past tense), *prevent, cancel, avoid, decline*, etc.

*John* **forgot** *to buy some wine.*
*Mary was* **unable** *to marry John.*
*John* **prevented** *the divorce.*

`Evidential:`

Evidential relations are typically introduced by `REPORTING` or `PERCEPTION` events:

*John* **said** *he bought some wine.*
*Mary* **saw** *John carrying only beer.*

`Negative evidential:`

Introduced by `REPORTING` and `PERCEPTION` events conveying negative polarity:

*John* **denied** *he bought only beer.*

`SLINKs` can be of the following nature:

a) **Lexically-based:**

They are triggered by an event of class `I_ACTION`, `I_STATE`, `PERCEPTION`, or `REPORTING`, which are events that generally take a clausal complement or an NP headed by an event-denoting nominal. The `SLINK` is established between those events and the one denoted by the complement.

For each `REPORTING` or `PERCEPTION` event, an `SLINK` has to be introduced. In the following example, the `REPORTING` and `PERCEPTION` events are in bold face, whereas the subordinated events are underlined:

*15 minutes later I* **saw** *the other plane just slam into the Worl Trade Center.*
*"It sounded like a jet or rocket,"* **said** *Eddie Gonzalez.*

Similarly, for each `I_ACTION` or `I_STATE`, an `SLINK` is introduced, which expresses the relation between the intensional event (in bold face) and its subordinated event (underlined):

*Subcomandante Marcos* **attempted** *to explain this difference in a letter in 1995.*
*We* **want** *to participate directly in the decisions which concern us, to control those who govern us.*

The subordinating event class constrains the `SLINK` relation type in the following way:

i. `PERCEPTION` events:
They will always introduce `SLINKS` of type `EVIDENTIAL` or `NEG_EVIDENTIAL`.

ii. `I_ACTION, I_STATE` events:
They can introduce `SLINKs` of type `INTENSIONAL`, `FACTIVE`, and `COUNTER_FACTIVE`.

iii. `REPORTING` events:
They can introduce `SLINKs` of any type.

b) **Structurally-based:**

1) **Purpose clauses:**
In a sentence involving a purpose clause, an `SLINK` relates the event in the main clause (bold face) and the one in the purpose clause modifying it (underlined).

*The environmental commission must* **adopt** *regulations to ensure people are not exposed to radioactive waste.*

2) **Conditional contructions:**
In a conditional construction, an `SLINK` relates the event in the antecedent section and the one in the consequent section.

*On Dec. 2 Marcos promised to* **return** *to the negotiating table if the conflict zone was demilitarized.*

#### A.3.2.1 How to annotate SLINKs

a) **Lexically-based SLINKS:**

The annotator should put an SLINK relating the subordinating event to each event that is subordinated by it:

- *John said that he taught on Monday.*
  To express the fact that the *taught* event is reported by the *said* event, the following SLINK is created:

```
<SLINK eventInstanceID="e2" subordinatedEventInstance="e3"
  relType="EVIDENTIAL"/>
```

- *John denied that he taught on Monday.*
  To express the fact that the *taught* event is being reported by the *denied* event, the following SLINK is created:

```
<SLINK eventInstanceID="e1" subordinatedEventInstance="e2"
    relType="NEG_EVIDENTIAL"/>
```

In some cases the same subordinating event will introduce more than one SLINK. For instance, in the example below the event *said* is slink-ed to two events: *listed* and *gave*.

- *Rita* **said** *they correctly listed his name but gave a false address for him.*

```
<SLINK eventInstanceID="ei1" subordinatedEventInstance="ei2"
    relType="EVIDENTIAL"/>
<SLINK eventInstanceID="ei1" subordinatedEventInstance="ei3"
    relType="EVIDENTIAL"/>
```

b) **Structurally-based SLINKS:**

1) **Purpose clauses:**
The event in the main clause will correspond to the value of the attribute eventInstanceID. The event in the purpose clause will be taken as the subordinatedEvent value. These SLINKs will always receive relType="INTENSIONAL". The preposition *to* will be taken as the value of the signalID attribute.

   - *The environmental commission must* **adopt** *regulations* to *ensure people are not exposed to radioactive waste.*

```
<SLINK eventInstanceID="ei1" subordinatedEventInstance="ei2"
    signalID="s1" relType="INTENSIONAL"/>
```

2) **Conditional contructions:**
The event in the antecedent clause corresponds to the value in the eventInstanceID attribute. The one in the consequent, to the value of the subordinatedEvent. The conditional conjunction (*if, when*) will be taken as the value of the signalID attribute. These SLINKs will always receive relType="CONDITIONAL".

   - *Mexico pledged to* **support** *an inquiry into Guantanamo* if *it is put to the vote at the UN Human Rights Commission.*

```
<SLINK eventInstanceID="ei1" subordinatedEventInstance="ei2"
    signalID="s1" relType="CONDITIONAL"/>
```

The presence of the same event in several SLINKs is also possible in structurally-based SLINKs, as for instance in the following conditional construction, where the antecent is a coordination. In this case, the repeated event is the subordinated one (*return*), since it is the event in the consequent section.

- *On Dec. 2 Marcos promised to* **return** *to the negotiating table if the conflict zone was demilitarized, Congress passed a bill on indigenous rights and culture, and around 100 Zapatista prisoners were released.*

```
<SLINK eventInstanceID="ei2" subordinatedEventInstance="ei1"
    signalID="s1" relType="CONDITIONAL"/>
<SLINK eventInstanceID="ei3" subordinatedEventInstance="ei1"
    signalID="s1" relType="CONDITIONAL"/>
<SLINK eventInstanceID="ei4" subordinatedEventInstance="ei1"
    signalID="s1" relType="CONDITIONAL"/>
```

#### A.3.2.2 BNF for the SLINK tag

```
attributes ::= [lid] [origin] eventInstanceID
               [signalID] subordinatedEventInstance relType
lid ::= ID
{lid ::= LinkID
LinkID ::= l<integer>}
origin ::= CDATA
eventInstanceID ::= IDREF
{eventInstanceID ::= EventInstanceID}
subordinatedEventInstance ::= IDREF
{subordinatedEventInstance ::= EventInstanceID}
signalID ::= IDREF
{signalID ::= SignalID}
relType ::= 'INTENSIONAL' | 'EVIDENTIAL' | 'NEG_EVIDENTIAL'
               | 'FACTIVE' | 'COUNTER_FACTIVE' | 'CONDITIONAL'
```

#### A.3.2.3 Attributes for SLINKs

a) **Attribute** `eventInstanceID`**:**

Required attribute. It conveys the ID of the `event` involved in the subordination link.

b) **Attribute** `subordinatedEventInstance`**:**

Required attribute. It takes as value the ID of the subordinated event that the event with `ID=eventInstanceID` is related to.

c) **Attribute** `signalID`**:**

Optional attribute. If the subordination relation holding between the events is explicitly signalled in the text, then the ID of that signal needs to be filled in here.

d) **Attribute** `relType`**:**

Obligatory attribute. Expressing the kind of subordintaion relation holding between the two events. Possible values are: `INTENSIONAL`, `EVIDENTIAL`, `NEG_EVIDENTIAL`, `FACTIVE`, `COUNTER_FACTIVE`, `CONDITIONAL`. They are assigned according to the instructions given at the beginning of section **??**, on `SLINK`s.

### A.3.3 The tag <ALINK>

An `ALINK` or (aspectual link) represents relations between aspectual events and their event arguments. Types of aspectual relations to be encoded are:

**Initiation**:

*John* **started** *to read*

**Culmination**:

*John* **finished** *assembling the table.*

**Termination**:

*John* **stopped** *talking*.

**Continuation**:

*John* **kept** *talking*.

#### A.3.3.1 How to annotate ALINKs

Some annotation examples are:

a) *John started to read*

The two EVENT tags for the two events are the following:

```
<EVENT eid="e5" eiid="ei1">started</EVENT>
<EVENT eid="e6" eiid="ei2">read</EVENT>
```

The ALINK that has to be created between the aspectual verb *started* and the event *read* is the following:

```
<ALINK eventInstanceID="ei1" relatedToEventInstance="ei2"
  relType="INITIATES"/>
```

b) *John finished reading*

The two EVENT tags for the two events are the following:

```
<EVENT eid="e5" eiid="ei1">finished</EVENT>
<EVENT eid="e6" eiid="ei2">reading</EVENT>
```

The ALINK that has to be created between the aspectual verb and its argument is the following:

```
<ALINK eventInstanceID="ei1" relatedToEventInstance="ei2"
  relType="TERMINATES"/>
```

#### A.3.3.2 BNF for the ALINK tag

```
attributes ::= [lid] [origin] eventInstanceID [signalID]
               relatedToEventInstance relType
lid ::= ID
{lid ::= LinkID
LinkID ::= l<integer>}
origin ::= CDATA
eventInstanceID ::= ID
{eventInstanceID ::= EventInstanceID}
signalID ::= IDREF
{signalID ::= SignalID}
relatedToEventInstance ::= IDREF
{relatedToEventInstance ::= EventInstanceID}
relType ::= 'INITIATES' | 'CULMINATES' | 'TERMINATES'
               | 'CONTINUES' | 'REINITIATES'
```

#### A.3.3.3 Attributes for ALINKs

a) **Attribute** eventInstanceID**:**

Obligatory attribute. This is the ID of the (aspectual) eventInstance involved in the link.

b) **Attribute** signalID**:**

Optional attribute. If the aspectual relation holding between the events is explicitly signalled in the text, then the ID of that signal needs to be filled in here.

c) **Attribute** relatedToEventInstance**:**

Obligatory attribute. This is the ID of the event instance related to the aspectual event.

d) **Attribute** relType**:**

Obligatory attribute. This is the temporal relation holding between the events. Possible values are: INITIATES, CULMINATES, TERMINATES, CONTINUES, REINITIATES.

**Additional attributes for all LINK tags**

**Attributes** lid **(link ID) and** origin**:**
The lid and origin attributes are optional in all ISO-TimeML links. They are used solely for post-processing efforts such as closure and can be ignored by the human annotator.

 

# Annex B
(informative)
# Completely annotated examples

Assume for all the examples that the document creation time (DCT) is marked up as a TIMEX3 expression with tid="t0".

For simplicity, the following annotations use the event ID to show link participation instead of the event instance ID.

## B.1 Complex TIMEX examples

a) *John left 2 days before yesterday.*

```
<EVENT  eid="e1" eiid="ei1" type="TRANSITION" class="OCCURRENCE" pos ="VERB" tense="PAST"
markable="left" />

<TIMEX3 tid="t1" type="DURATION" value="P2D" beginPoint="t2"
endPoint="t3" markable ="2 days"/>

<SIGNAL sid="s1" markable ="before"/>

<TIMEX3 tid="t2" type="DATE" value="2002-07-10" temporalFunction="true"
anchorTimeID="t0" markable ="yesterday"/>

<TIMEX3 tid="t3" type="DATE" value="2002-07-08" temporalFuntion="true"
anchorTimeID="t1"/>

<TLINK timeID="t1" relatedToTime="t2" signalID="s1" relType="BEFORE"/>
<TLINK eventInstanceID="ei1" relatedToTime="t3" relType="IS_INCLUDED"/>
```

The TIMEX3 that is annotated as a DURATION includes begin and endpoint information. The annotator also could have included additional TLINKs with this information.

The type attribute of the TIMEX3 for "yesterday" denotes a DATE, which can be computed by a temporal function relative to the temporal anchor "t0" (the DCT). Similarly, the final TIMEX3 can be computed by a temporal function relative to the initial DURATION, annotated as "t1".

b) *I'm leaving on vacation two weeks from next Tuesday*.

```
<EVENT  eid="e1" eiid="ei1" class= "OCCURRENCE" pos="VERB" tense="PRESENT" type= "PROCESS"
aspect="PROGRESSIVE" polarity="POS" markable ="leaving" />

<TIMEX3 tid="t1" type="DURATION" value="P2W" beginPoint="t2"
endPoint="t3" markable="two weeks" />

<SIGNAL sid="s1" markable="from" />

<TIMEX3 tid="t2" type="DATE" value="2002-07-02" temporalFunction="true"
anchorTimeID="t0" markable="next Tuesday" />
```

```
<TIMEX3 tid="t3" type="DATE" value="2002-07-23" temporalFunction="true"
anchorTimeID="t1"/>

<TLINK timeID="t1" relatedToTime="t2" signalID="s1" relType="AFTER"/>
<TLINK eventInstanceID="ei1" relatedToTime="t3" relType="IS_INCLUDED"/>
```

c) *A major earthquake struck Los Angeles three years ago today.*

```
<EVENT  eid="e1"  eiid="ei1" type="PROCESS" class="OCCURRENCE" pos="NOUN" tense="NON

<EVENT  eid="e2"  eiid="ei2" type="TRANSITION" class="OCCURRENCE" pos="VERB" tense="
polarity="POS" markable="struck"/>

<TIMEX3 tid="t1" type="DURATION" value="P3Y" beginPoint="t2"
endPoint="t3" markable ="three years ago" />

<TIMEX3 tid="t2" type="DATE" value="2002-07-12" temporalFunction="true"
anchorTimeID="t0" markable="today" />

<TIMEX3 tid="t3" type="DATE" value="1999-07-12" temporalFunction="true"
anchorTimeID="t1"/>


<TLINK eventInstanceID="ei1" relatedToEventInstance="ei2"
relType="IBEFORE"/>

<TLINK eventInstanceID="ei1" relatedToTime="t3" relType="IS_INCLUDED"/>
```

d) *John left 2 days ago.*

```
<EVENT  eid="e1"  eiid="ei1" type="TRANSITION" class="OCCURRENCE" pos="VERB" tense="
 polarity="POS" markable="left" />

<TIMEX3 tid="t1" type="DATE" value="2002-07-08" temporalFunction="true"
anchorTimeID="t0" markable ="2 days ago"/>


<TLINK eventInstanceID="ei1" relatedToTime="t1" relType="IS_INCLUDED"/>
```

NOTE *ago* is NOT a signal but a part of the TIMEX3 expression.

The TIMEX3 expression returns a DATE (not a DURATION), which needs to be computed by a temporal function relative to the DCT or the Speech time. *2 days ago* is ALWAYS a DATE computed relative to the DCT, in contrast to expressions like "2 days before", which necessarily relate two events and thus introduce a TLINK with the magnitude attribute. This can be observed in the 3 following examples.

e) *John left 2 days before the attack.*

```
<EVENT  eid="e1" eiid="ei1" type="TRANSITION" class="OCCURRENCE" pos ="VERB" tense="
markable="left" />

<TIMEX3 tid="t1" type="DURATION" value="P2D" temporalFunction="false"
markable ="2 days"/>

<SIGNAL sid="s1" markable ="before"/>
```

```
<EVENT  eid="e2"  eiid="ei2" type="TRANSITION" class="OCCURRENCE" pos="NOUN" tense="NONE"


<TLINK  eventInstanceID="ei1" signalID="s1" relatedToEvent="ei2"
relType="BEFORE" magnitude="t1"/>
```

f) *5 days after he came back Mary got sick.*

```
<TIMEX3 tid="t1" type="DURATION" value="P5D" temporalFunction="false"
markable="5 days" />

<SIGNAL sid="s1" markable="after" />

<EVENT  eid="e1"  eiid="ei1" type= "TRANSITION"  class="OCCURRENCE" pos="VERB" tense="PAS
polarity="POS" markable="came" />

<EVENT  eid="e2"  eiid="ei2" type ="TRANSITION" class="OCCURRENCE" pos="VERB" tense="PAST
 polarity="POS" markable="got" />


<TLINK  eventInstanceID="ei1" signalID="s1"
relatedToEvent="ei2" relType="BEFORE" magnitude="t1"/>
```

g) *Two months before the attack, a report was sent.*

```
<TIMEX3 tid="t1" type="DURATION" value="P2M" temporalFunction="false"
markavke="Two months" />

<SIGNAL sid="s1" markable="before" />

<EVENT  eid="e1"  eiid="ei1" type="TRANSITION" class="OCCURRENCE" pos="NOUN" tense="NONE"

<EVENT  eid="e2"  eiid="ei2" type="TRANSITION" class="OCCURRENCE" pos="VERB" tense="PAST"
markable="sent" />

<TLINK eventInstanceID="ei1" signalID="s1"
relatedToEvent="ei2" relType="AFTER" magnitude="t1"/>
```

The `TIMEX3` expression here is considered here to be of `type=DURATION`, since it establishes the length of the interval separating the 2 events. As such, the value for the `value` attribute is already known (`P2D`, `P5M`, etc.) and therefore the `temporalFunction` attribute returns `false` as its value.

There is only one `TLINK` relating the two events, which introduces both the `magnitude` attribute (pointing to the ID of the TIMEX3 expression) and the `signalID` attribute.

## B.2 Complex `TLINK` and `SLINK` examples

a) *The attack was not expected at all, although a report had been sent 2 months before.*

```
The
<EVENT eid="e1"  eiid="ei1" type="TRANSITION" class="OCCURRENCE" pos="NOUN" tense="NONE" a
polarity="POS" markable="attack" />

<EVENT eid="e2"  eiid="ei2" type="TRANSITION" class="I_STATE"
pos="VERB" tense="PAST" aspect="NONE"
polarity="NEG"markable="expected" />
```

```
<EVENT eid="e3"  eiid="ei3" type="TRANSITION" class="OCCURRENCE" pos="VERB" tense="P
polarity="POS" markable="sent" />

<TIMEX3 tid="t1" type="DURATION" value="P2M" beginPoint="ei2"
endPoint="ei1" markable="two months" />

<SIGNAL sid=s2 markable="before" />


<SLINK eventInstanceID="ei2" subordinatedEvent="ei1"
relType="INTENSIONAL"/>
<TLINK eventInstanceID="ei1" relatedToEvent="ei3" relType="AFTER"
signalID="s2"/> % JL
```

b) *Mary arrived yesterday but John left 2 days before.*

```
Mary
<EVENT  eid="e1"  eiid="ei1" class="OCCURRENCE" pos="VERB"
tense="PAST" type="TRANSITION" aspect="NONE" polarity="POS"
markable="arrived" />

<TIMEX3  tid="t1" type="DATE" value="2002-07-09"
temporalFunction="true" anchorTimeID="t0"
valueFromFunction="tf1" markable="yesterday" />

<EVENT  eid="e2"  eiid="ei2" class="OCCURRENCE"
pos="VERB" tense="PAST" type = "TRANSITION" aspect="NONE"
 polarity="POS" markable="left" />

<TIMEX3 tid="t2" type="DURATION" value="P2D"
temporalFunction="false" beginPoint="ei2" endPoint="ei1"
markable="2 days" />

<SIGNAL sid="s1" markable="before" />


<TLINK  eventInstanceID="ei1" signalID="s1" relatedToEvent="ei2"
relType="AFTER" />
<TLINK  eventInstanceID="ei1" relatedToTime="t1"
relType="IS_INCLUDED"/>
```

The two events are related by means of a `TLINK`. In addition there is a second `TLINK` relating the event linked to the date (*arrived*) and this date (`yesterday`).

c) *She was sick after the play.*

```
<EVENT eid="e1"  eiid="ei1" type= "STATE"  class="STATE" pos="ADJECTIVE"
tense="NONE" aspect="NONE" polarity="POS"
markable="sick" />

<SIGNAL sid="s1" markable="after" />


<EVENT  eid="e2"  eiid="ei2" type="TRANSITION" class="OCCURRENCE"  pos="NOUN" tense=

<TLINK  eventInstanceID="ei1" signalID="s1"
relatedToEvent="ei2" relType="AFTER"/>
```

d) *She was sick for 2 hours after the play.*

```
<EVENT eid="e1"  eiid="ei1" type= "STATE"  class="STATE" pos="ADJECTIVE"
tense="NONE" aspect="NONE" polarity="POS"
markable="sick" />

<SIGNAL sid="s1" markable="for" />

<TIMEX3 tid="t1" type="DURATION" value="P2H" temporalFunction="false"
markable= "2 hours" />

<SIGNAL sid="s2" markable="after" />

<EVENT  eid="e2"  eiid="ei2" type="TRANSITION" class="OCCURRENCE"  pos="NOUN" tense="NONE"

<TLINK  eventInstanceID="ei1" signalID="s1" relatedToTime="t1"
relType="SIMULTANEOUS"/>
<TLINK  eventInstanceID="ei1" signalID="s2" relatedToEvent="ei2"
relType="AFTER"/>
```

There are two `TLINK`s: The first one introduces the holding relation between the state of being sick and the time it took (2 hours). The second one states the ordering of the two events.

e) *John taught for 20 minutes every Monday.*

```
<EVENT eid="e4"    eiid="ei4" class="OCCURRENCE" pos="VERB"
tense="PAST" aspect="NONE" type="TRANSITION" polarity="POS"
markable="taught" />

<SIGNAL sid="s5" markable="for" />

<TIMEX3 tid="t2" type="DURATION" value="PT20M"
markable="20 minutes" />

<TIMEX3 tid="t3" type="SET" value="XXXX-WXX-1"
markable="every Monday" />


<TLINK eventInstanceID="ei4" relatedToTime="t2" signalID="s5"
relType="SIMULTANEOUS"/>
<TLINK timeID="t2" relatedToTime="t3" relType="IS_INCLUDED"/>
```

f) *John left between Monday and Wednesday*

```
<EVENT eid="e1"  eiid="ei1" type="TRANSITION" class="OCCURRENCE" tense="PAST" aspect="NONE

<SIGNAL sid="s1" markable= "between" />

<TIMEX3 tid="t1" type="DATE" value="2002-07-15" temporalFunction="true"
anchorTimeID="t0" valueFromFunction="tf3" markable="Monday" />

<TIMEX3 tid="t2" type="DATE" value="2002-07-17" temporalFunction="true"
anchorTimeID="t0" valueFromFunction="tf3"markable="Wednesday" />

<TLINK eventInstanceID="ei1" relatedToTime="t1" signalID="s1"
relType="IAFTER"/>
```

```
<TLINK eventInstanceID="ei1" relatedToTime="t2" signalID="s1"
relType="IBEFORE"/>
```

NOTE This current solution is not completely adequate, but we will keep it temporarily.

g) *John taught from 1994 through 1999.*

In this case, one EVENT and the two TIMEX3s need to be created. In addition, the following tags are needed:

1) One TLINK to capture the fact that the event started in 1994.
2) One TLINK to capture the fact that the event finished in 1999.

This then should lead to a duration, which is automatically created by the closure part of the tool.

```
<EVENT eid="e4"    eiid="ei4" class="OCCURRENCE" tense="PAST"
aspect="NONE" type="TRANSITION" polarity="POS"
markable="taught" />

<SIGNAL sid="s5" markable="from" />

<TIMEX3 tid="t2" type="DATE" value="1994" markable="1994" />

<SIGNAL sid="s6" markable="through" />

<TIMEX3 tid="t3" type="DATE" value="1999" markable="1999" />

<TIMEX3 tid="t1" type="DURATION" value="P2Y" beginPoint="t2"
endPoint="t3"/>
<TLINK eventInstanceID="ei4" relatedToTime="t2" signalID="s5"
relType="BEGUN_BY"/>
<TLINK eventInstanceID="ei4" relatedToTime="t3" signalID="s6"
relType="ENDED_BY"/>
```

h) *John did not leave on Monday but on Tuesday.*

One EVENT, **two event instances**, and three SIGNALs need to be created.

```
<EVENT eid="e4"    eiid="ei4" type="TRANSITION" class="OCCURRENCE"
tense="PAST" aspect="NONE" polarity="NEG" markable="leave" />

<SIGNAL sid="s6" markable="on" />

<TIMEX3 tid="t3" type="DATE" value="XXXX-WXX-1"
markable="Monday" />

<SIGNAL sid="s7" markable="on" />

<TIMEX3 tid="t4" type="DATE" value="XXXX-WXX-2"
markable="Tuesday" />

<TLINK eventInstanceID="ei1" relatedToTime="t3" signalID="s6"
relType="IS_INCLUDED"/>
<TLINK eventInstanceID="ei2" relatedToTime="t4" signalID="s7"
relType="IS_INCLUDED"/>
```

### B.3 Causative examples

a) *The rains caused the flooding.*

```
The
<EVENT eid="e1" eiid="ei1" type="PROCESS" class="OCCURRENCE" pos="NOUN" markable="rains" /

<EVENT eid="e2" eiid="ei2" type="TRANSITION" class="OCCURRENCE" pos="VERB"tense="PAST" asp
polarity="POS" markable="caused" />

<EVENT eid="e3" eiid="ei3" type="PROCESS" class="OCCURRENCE" pos="NOUN" markable="floodin


<TLINK eventInstanceID=ei1 relatedtoEvent=ei3  relType="BEFORE" />
<TLINK eventInstanceID=ei1 relatedtoEvent=ei2  relType="IDENTITY" />
```

b) *John caused the fire.*

```
<EVENT eid="e1" eiid="ei1" class="OCCURRENCE" pos="VERB"
tense="PAST" aspect="NONE"  type="TRANSITION"
polarity="POS" markable="caused" />

<EVENT eid="e2" eiid="ei2" type="PROCESS" class="OCCURRENCE" pos="NOUN" markable="fire" />

<TLINK eventInstanceID=ei1 relatedtoEvent=ei2  relType="BEFORE" />
```

c) *Kissinger secured the peace at great cost.*

```
<EVENT eid="e1" eiid="ei1"  class="OCCURRENCE" pos="VERB"
tense="PAST" aspect="NONE"  type= "TRANSITION" polarity="POS"
markable="secured" />

<EVENT eid="e2" eiid="ei2" class="OCCURRENCE"
type="STATE" pos="NOUN" markable="peace" />

<TLINK eventInstanceID=ei1 relatedtoEvent=ei2  relType="BEFORE" />
```

d) *He kicked the ball, and it rose into the air.*

Discourse relations acting as a causative will be handled in the next release as causatives in a Causative Link (CLINK), but this is out of the scope of our current discussion.

```
<EVENT eid="e1" eiid="ei1"  class="OCCURRENCE" pos="VERB"
tense="PAST" aspect="NONE"  type="TRANSITION"
polarity="POS" markable="kicked" />

<EVENT eid="e2"  eiid="ei2" class="OCCURRENCE" pos="VERB"
tense="PAST" aspect="NONE" type="PROCESS"
polarity="POS" markable="rose" />

<TLINK eventInstanceID=ei1 relatedtoEvent=ei2  relType="BEFORE" />
```

# Annex C
(informative)
# Event and temporal annotations for Chinese

Example (1): 他是中国人。
He is a Chinese.

Example (2): 花是红色的。
Flowers are red.

Example (1) and (2) are not annotated in ISO-TimeML, as stated in the ISO document, “if a STATE is deemed persistent throughout the event line of the document, it is factored out and not annotated."

Example (3):

他在跑步。
He is running.

他在
<EVENT eid="e1" eiid="ei1" romanization="PAOBU" content="RUN”
class="OCCURRENCE" pos= “VERB” tense="NONE" aspect="PROGRESSIVE"
polarity="POS">
跑步
</EVENT>
。

Example (4):

很多人在公园里跳健美操。
Many people are practising aerobics in the park.

很多人在公园里
<EVENT eid="e1" eiid="eil" romanization="TIAO" content="PRACTISE"
class="OCCURRENCE" pos="VERB" tense="NONE" aspect="PROGRESSIVE"
polarity="POS">
跳
</EVENT>
健美操。
<CONFIDENCE tagType="EVENT" tagID="e1" attributeName="ASPECT"
confidenceValue="0.50"/>

Example (5):

这部电影拍了三年的时间终告竣工。
The film took three years to shoot.

这部电影
<EVENT eid="e1" romanization="PAI LE" content="SHOOT"
class="OCCURRENCE" pos="VERB" tense="NONE" aspect="PERFECTIVE"
polarity="POS">

拍
</EVENT>
了
<TIMEX3 tid="t1" type="DURATION" value="P3Y">
三年
</TIMEX3>
的时间终告
<EVENT eid="e2" eiid="ei2" romanization="JUNGONG" content="COMPLETE"
class="ASPECTUAL" pos="VERB" tense="NONE" aspect="PERFECTIVE"
polarity="POS">
竣工
</EVENT>。
<TLINK eventInstancdD="ei1" relatedToTime="t1" relType="IS_INCLUDED"/>
<TLINK eventInstanceID="ei2" relatedToEventInstance="ei1"
relType="IAFTER"/>

Example (6):

我花了一周时间才读完这本书。
I spent one week and finished reading the book.

我
<EVENT eid="e1" eiid="ei1" romanization="HUA LE" content="SPEND"
class="OCCURRENCE" pos="VERB" tense="NONE" aspect="PERFECTIVE"
polarity="POS">
花了
</EVENT>
<TIMEX3 tid="t1" type="DURATION" value="P1W">
一周时间
</TIMEX3>
才
<EVENT eid="e2" eiid="ei2" romanization="DU WAN" content="READ"
class="OCCURRENCE" pos="VERB" tense="NONE" aspect="PERFECTIVE"
polarity="POS">
读完
</EVENT>
这本书。
<TLINK eventInstanceID="ei1" relatedToTime="t1" relType="IS_INCLUDED"/>
<TLINK eventInstanceID="ei2" signaledID="s1"
relatedToEventInstance="ei1" relType="IAFTER"/>

Example (7):

钥匙掉到地上去了。
The key dropped to the floor.

钥匙
<EVENT eid="e1" eiid="ei1" romanization="DIAO" content="DROP"
class="OCCURRENCE" pos="VERB" tense="NONE" aspect="NONE"
polarity="POS">
掉
</EVENT>
到地上去了。

Example (8):

门突然被打开。

The door was suddenly pushed open.

门突然被
<EVENT eid="e1" eiid="ei1" romanization="DAKAI" content="OPEN"
class="OCCURRENCE" pos="VERB" tense="NONE" aspect="NONE"
voice="PASSIVE" polarity="POS">
打开
</EVENT>。

Example (9):

他在做作业。
He is doing homework.

他在
<EVENT eid="e1" eiid="ei1" romanization="ZUO" content="DO"
class="OCCURRENCE" pos="VERB" tense="NONE" aspect="PROGRESSIVE"
polarity="POS">
做
</EVENT>
作业。

Example (10):

他做了三个小时的作业才做完。
He spent three hours and finished doing his homework.

他
<EVENT eid="e1" eiid="ei1" romanization="ZUO LE" content="DO"
class="OCCURRENCE" pos="VERB" tense="NONE" aspect="PERFECTIVE"
polarity="POS">
做了
</EVENT>
<TIMEX3 tid="t1" type="DURATION" value="P3H">
三个小时
</TIMEX3>
的作业
<SIGNAL sid="s1">
才
</SIGNAL>
<EVENT eid="e2" eiid="ei2" romanization="ZUO WAN" content="DO"
pos="VERB" class="OCCURRENCE" tense="NONE" aspect="PERFECTIVE"
polarity="POS">
做完
</EVENT>
。

<TLINK eventInstanceID="ei1" relatedToTime="t1"
relType="IS_INCLUDED"/>
<TLINK eventInstanceID="ei2" signalID="s1"
relatedToEventInstance="ei1" relType="IAFTER"/>

Example (11):

他做了作业。
He has done/did his homework.

他
<EVENT eid="e1" eiid="ei1" romanization="ZUO LE" content="DO"
class="OCCURRENCE" pos="VERB" tense="NONE" aspect="PERFECTIVE"
polarity="POS">
做了
</EVENT>
作业。

Example (12):

9 月 11 日，这是一个让全美国人民陷于悲痛的日子。
9/11 is a day that made all the Americans saddened.

<TIMEX3 tid="t1" type="DATE" value="XXXX-09-11">
9 月 11 日
</TIMEX3>
，
这
<EVENT eid="e1" eiid="ei1" romanization="SHI" content="DO"
class="OCCURRENCE" pos="VERB" tense="NONE" aspect="NONE"
polarity="POS">
是
</EVENT>
一个让全美国人民
<EVENT eid="e2"  eiid="ei2" romanization="XIAN" content="DO"
class="OCCURRENCE" pos="VERB" tense="NONE" aspect="NONE"
polarity="POS">
陷
</EVENT>
于
<EVENT eid="e3" eiid="ei3" romanization="BEITONG" content="SAD"
class="STATE" pos="ADJECTIVE" tense="NONE" aspect="NONE"
polarity="POS">
悲痛
</EVENT>
的日子。
<TLINK eventInstanceID="e1" relatedToTime="t1"
relType="IS_INCLUDED"/>
<TLINK eventInstanceID="e2" relatedToEventInstance="e3"
relType="IDNETITY"/>

Example (13):

他一直都被蒙在鼓里。
He has been kept away from the truth.

他一直都被
<EVENT eid="e1" eiid="ei1" romanization= “MENG”
content= “KEEP” class="OCCURRENCE" pos="VERB"
tense="NONE" aspect="NONE" polarity="POS">
蒙
</EVENT>
在鼓里。

Example (14):

自他从北京回来后，脸上的笑容逐渐多了起来。

His face has shown more and more smiles since he got back from Beijing.

自他从北京
<EVENT eid="e1" eiid="ei1" romanization="HUILAI"
content="GET_BACK" class="OCCURRENCE" pos="VERB"
tense="NONE" aspect="NONE" polarity="POS" >
回来
</EVENT>
<SIGNAL sid="s1">
后
</SIGNAL>
，
脸上的笑容逐渐
<EVENT eid="e2" eiid="ei2" romanization="DUO"
content="MORE" class="OCCURRENCE" pos="ADJECTIVE"
tense="NONE" aspect="NONE" polarity="POS">
多
</EVENT>
了起来。
<TLINK eventInstanceID="e2" signalID="s1"
relatedToEventInstance="ei1" relType="AFTER"/>

Example (15):

我吃过饭了。
I have had/had my dinner.

我
<EVENT eid="e1" eiid="ei1" romanization="CHI GUO"
content="EAT" class="OCCURRENCE" pos="VERB"
tense="NONE" aspect="PERFECTIVE" polarity="POS">
吃过
</EVENT>
饭了。

 

# Annex D
(informative)
# Annotation for Italian fragment

## D.1 Introduction

This annex describes the annotation guidelines for marking up Italian text according to the ISO-TimeML language. It is organised as follows. The first part explains how the ISO-TimeML tags are realized in Italian and how to annotate them. The second part is more informative and contains a fully annotated example, illustrating all of the interactions between the various entities and relational tags. For the sake of conveninece Annex A will be referred to throughout the annex.

## D.2 Basic references

This Annex relies on the following references:

- Bertnetto, P.M. (2001) "Sulle proprietà tempo-aspettuali dell'Infinito italiano", available at http://alphalinguistica.sns.it/QLL/QLL01/PMB.Infinito.pdf
- Bertnetto, P.M. (1991) "Il Verbo" in *Grande Grammatica Italiana di Consultazione*, L.Renzi-G.Salvi (a cura di), Il Mulino: 13-163
- Bertinetto, P.M. (1986) *Tempo, Aspetto e Azione nel verbo Italiano*, Accademia della Crusca, Firenze.
- Lavelli A., Magnini B., Negri M., Pianta E., Speranza M., and Sprugnoli R.(2005) *Italian Content Annotation Bank (I-CAB): Temporal Expressions (V. 1.0)*, ITC-irst Technical Report.

## D.3 ISO-TimeML tags and their attributes

### D.3.1 The tag `<EVENT>`

The definition of *event* is the same as that proposed in Annex A.2.1.

### D.3.2 How to annotate EVENTs

The types of expressions detonting an event in Italian are much the same as those in Annex A.2.1.1. For clarity's sake the following phrase types are some examples. Event-denoting expressions are in bold face.

a) Verbs (finite or non-finite form) e.g.:
   - *I pompieri hanno **isolato** la sala.*
   - *Fim-Cisl e Uilm-Uil hanno **annunciato** oggi una conferenza stampa.*
   - *La città mostra i segni della battaglia: cassonetti **incendiati** o **rivoltati**.*

b) Nominalizations, e.g.:
   - *La **caduta** della base aerea di Ubdina allontana il fronte di 120 km.*

c) Adjectives e.g.:
   - *La coppia, **residente** a Milano, stava trascorrendo un periodo di vacanza in Sicilia.*

d) Predicative sections e.g. :

- *Al Sayed è il nuovo **presidente** della Fermenta.*

e) Prepositional phrases, e.g.:

- *Una giovane turista **in vacanza** nel villaggio "Katibubbo" è morta.*
- *Un centinaio di giovani è tuttora **agli arresti**.*

f) Nouns with an event-like reference, e.g.:

- *Queste le principali indicazioni strategiche fatte da Cabassi al termine dell'**assemblea**.*
- ***Allarme** inconsueto alla Tate Gallery..*
- *Rota auspica un ritorno alla **pace**..*
- *Questo consentirebbe di discutere con serenità e fermezza i **problemi** della minoranza italiana.*

### D.3.3 Event identification and tag span

As proposed in Annex A, the annotation of Italian texts is based on the notion of *minimal chunk*. This means that only the head of the chunk will be covered by the tag and thus that auxiliaries, clitics, polarity markers, particles, modifiers, complements and specifiers will be disregarded. In the following examples, the event-denoting chunk is in bold face and the tagged head is underlined.

- *I pompieri **hanno isolato** la sala.*
- ***Accusandoli** di omicidio...*
- *La riunione **sta per chiudersi**.*
- *Il PIL Italiano **non è cresciuto** nell'ultimo trimestre.*
- ***La caduta** della base aerea di Ubdina allontana il fronte di 120 km.*
- *Al Sayed è **il nuovo presidente** della Fermenta.*
- *La coppia, **residente** a Milano, stava trascorrendo un periodo di vacanza in Sicilia.*

As far as prepositions are concerned, if the prepositional chunk represents a fixed expression denoting an event, then the preposition must be included into the tag; otherwise, only the noun head of the embedded NP must be annotated.

- *Le strade mostrano ancora i segni **della battaglia**.*
- *Un centinaio di giovani è tuttora **agli arresti**.*

Most event tags will span over only one word, i.e. the minimal chunk. However, an important question emerges with the nature of the textual extent of idioms, verbal collocations, metaphorical uses, light verb constructions, causative constructions and complex NPs of the kind "NP + PP". The following cases are contemplated:

a) those realizations whose entire event extent is annotated and then classified. They comprise all occurrences of idiomatic expressions, metaphors, light verb constructions of the form "verb + non-deverbal noun", constructions with "*FARE + indefinite article + non-deverbal noun*", and causative constructions with "*FARE + abstract noun*". All instances are in bold characters and the square brackets delimit the tag span; e.g.:

- *I punk hanno **[messo a ferro e fuoco]** la città.*
- *Tocca a Baker **[tirare le somme]** su questo incontro...*
- *Tutte le questioni principali sono **[rimaste sul tappeto]**.*

- *La ragazza è morta mentre [**faceva una doccia**].*
- *Daubre ha deciso di [**tenere sulla corda**] il sindacato.*
- *Mi [**fa paura**].*

NOTE This holds also for the variants with DARE or METTERE

b) those realizations which split in two the extent of the annotation. Such cases involve light verb constructions of the form "verb + deverbal noun", constructions with "*FARE + indefinte article + deverbal noun*", constructions where "FARE" is used to substitute entire VPs (instances of "*fare*" as "*verbo vicario*"), causative constructions, all event-denoting constructions whose meaning is compositional, e.g.:

- *I guardiani hanno [**fatto**] [**scattare**] l'allarme.*
- *Gli Usa hanno[**fatto**][**sapere**] che non sono disponibili.*
- *L'assemblea ha[**preso**][**visione**] del bilancio consolidato.*
- *Marco ha [**fatto**] una[**passeggiata**].*

#### D.3.3.1 Modal verbs

Modal verbs in Romance languages are very different from the English ones. In Italian, modal verbs are to be considered similar to other lexical verbs in that it is possible to assign them values for tense and aspect. Consequently, each instance of Italian modal verbs ("*dovere*", "*potere*", "*volere*", "*sapere*") will be annotated with the tag `<EVENT>`; e.g.:

- *L'assemblea [**deve**] [**prendere**] una decisione...*
- *Non ho [**potuto**] [**chiamare**] l'ufficio cambi.*

#### D.3.3.2 Verbal periphrases

In Italian it is possible to identify different instances of verbal periphrases. We accept here the proposal of Bertinetto (1991) to identify a hierarchy of verbal periphrases:

- aspectual periphrases: to code progressive or habitual aspect, e.g. *sta mangiando*, *è solito riposare*...
- modal periphrases: to code modality not realized by proper modal verbs; we consider expressions like "*essere in grado di + INF*", "*c'è da + INF*" as belonging to this class, e.g. *va fatto, c'è da dire, ho da fare*...
- aspectual (phasal) periphrases: to code information on a particular phase or aspect in the description of a particular event; phasal verbs instances are in bold character, e.g. ***iniziare** a mangiare,* ***continuare** a dormire*...

NOTE Phasal verbs are called aspectual verbs in ISO-TimeML. The ISO-TimeML term will be used to refer to these kinds of verbs.

Due to the treatment proposed in Bertinetto (1991), we claim that only in the last two cases, i.e. modal periphrases and phasal/aspectual periphrases, both elements involved should be annotated, while in the case of the aspectual periphrasis only the main verb has to be marked, as suggested in Annex A, clause B. In the examples below, the event denoting chunk is in bold face, whereas the tagged head(s) is underlined.

- La borsa **stava perdendo** l'1,1% in prima mattinata.(*Progessive periphrasis*)
- A oggi **siamo in grado di dire** che l'accordo non è stato raggiunto. (*Modal periphrasis*)
- **C'è da dire** che questo trattamento non è soddisfacente.(*Modal periphrasis*)
- Il magistrato **ha iniziato a condurre** le indagini sulla mortedi Calipari.(*Aspectual periphrasis*)

#### D.3.4 What *NOT* to tag

Events are not to be tagged in the situations described in Annnex A, clause C.

### D.4 BNF values for <EVENT> in Italian

```
attributes ::= eid class tense aspect pos [polarity] mood [modality]
eid ::= e<integer>
class ::= 'REPORTING'|'PERCEPTION'|'ASPECTUAL'|'I_ACTION'|'I_STATE' |
      'STATE'|'OCCURRENCE'
pos ::= 'ADJECTIVE'|'NOUN'|'VERB'|'PREPOSITION'|'NONE'
tense ::= 'FUTURE'|'PAST'|'PRESENT'|'IMPERFECT'| 'NONE'
aspect ::= 'PROGRESSIVE'|'PERFECTIVE'|'IMPERFECTIVE'|'NONE'
vform::='NONE'|'PASTPART'|'PRESPART'|'INFINITIVE'|'GERUNDIVE'
(default, if absent 'NONE')
polarity ::= 'NEG'|'POS' {default, if absent, is 'POS'}
mood ::= 'SUBJUNC'|'COND'|'NONE' {default, if absent, is 'NONE'}
modality ::= CDATA
```

### D.5 Attributes for EVENT

Annex A, clause A.2.1.3 is referred to for attribute definitions. In this section we illustrate the `class` attribute, for informative purposes, and the language specific values of the other attributes.

#### D.5.1 Attribute `class`:

Required attribute. Each event belongs to one of the following classes, as definied in Annex A.2.1.3, clause B.

NOTE The verbs provided as examples of each class may have multiple senses, some of which may not belong to that particular class.

- `REPORTING`: *dire, spiegare, raccontare, affermare*.

  (148) a. Punongbayan ha **detto** che dal vulcano fuoriuscivano gas con temperature fino a 1.800 gradi.
  b. **Citando** l'esempio di...

- `PERCEPTION`: *vedere, guardare, osservare, ascoltare, sentire*.

  (149) a. Dei testimoni hanno dichiarato alla polizia di aver **visto** delle persone fuggire.
  b. "Puoi **sentire** le migliaia di piccole esplosioni da laggiù", ha detto un testimone.

- `ASPECTUAL`:

  a) Initiation: *iniziare*, *incominciare*.
  b) Reinitiation: *rincominciare*.
  c) Termination: *smettere*, *terminare*, *cessare*, *interrompere*.
  d) Culmination: *finire*, *completare*.
  e) Continuation: *continuare*, *andare avanti*.

  A couple of examples:

  (150) a. Il vulcano ha **iniziato** a mostrare segni di attività in Aprile.
  b. Ho **continuato** a leggere quell'articolo per tutto il giorno.

- `I_ACTION`. In the examples, we report in bold face the I_ACTION events and their arguments, underlined.

a) **cercare, provare**:
   (151) Compagnie coma la Microsft stanno **cercando** di monopolizzare.
b) **investigare, indagare, ricercare**:
   (152) Una nuova task force ha inizato a **indagare** sull'uccisione di 14 donne.
c) **ritardare, postporre, ostacolare**:
   (153) IsraeIe chiederà agli Stai Uniti di **ritardare** l'attacco contro l'Iraq.
d) **evitare, impedire, prevenire, cancellare, disdire**:
   (154) La Questura di Livorno ha **impedito** lo svolgimento della manifestazione di Forza Nuova indetta per il 10 Febbraio.
e) **chiedere, ordinare, persuadere, comandare, richiedere, autorizzare**:
   (155) Le autorità hanno **richiesto** la massima collaborazione da parte dei mezzi di informazione.
f) **promettere, offrire, assicurare, proporre, accordarsi**.
g) **giurare**.
h) **nominare, eleggere, dichiarare, proclamare**.

- `I_STATE`. As above, the I_STATE events are in bold face, whereas the embedded argument is underlined.

a) **credere, pensare, immaginare, essere sicuro, sospettare.**
   (156) "**Crediamo** che le sue parole non abbiano distratto il pubblico da quello che è accaduto".
b) **sembrare, desiderare, bramare, auspicare**.
   (157) Il governo italiano ha **auspicato** un'intesa in tempi rapidi.
c) **sperare, aspirare, decidere.**
   (158) **Sperano** che i residenti rientreranno nelle loro case una volta cessato l'allarme.
d) **temere, odiare, essere preoccupato, aver paura, spaventarsi.**
   (159) **Temevano** per la loro incolumità.
e) **aver bisogno, necessitare.**
f) **dovere, potere, volere, sapere, essere in grado di, riuscire**
   (160) I soldati **devono** essere ritirati dall'Iraq.

- `STATE`:

a) States that are identifiably changed over the course of the document being marked up. In these and the following examples the markable state is in bold face.
   (161) a. Numerosi punk sono tutt'ora **agli arresti**.
   b. Il numero di **feriti** negli scontri è imprescisato.
   c. Si deve guardare agli andamenti economici, in modo da portere correzzioni dove **necessario**.
b) States that are directly related to a temporal expression. This criterion includes all states that are linked to a `TIMEX3` markable by means of a `TLINK` (see clauses **??** and **??**). An example is presented here, where the state is in bold face and the temporal expression associated with it is underlined.
   (162) Silvio Berlusconi è stato il **Presidente del Consiglio** negli ultimi 5 anni.
   NOTE In a sentece like "*Silvio Berlusconi è il proprietario di Mediaset*" the predicative noun "**proprietario**" must not be marked as a STATE because it is not temporally relavant.
c) States that are introduced by an `I_ACTION`, an `I_STATE`, or a `REPORTING` event. States are in bold face, the introducing event underlined.
   (163) a. Una partecipazione garantita dalla **presenza** dei nostri ministri.
   b. Ha dichiarato che è un **bugiardo**.
d) Predicative states the validity of which is dependent on the document creation time.
   (164) Più di 2.000 soldati italiani sono **in Afghanistan**.

(165) Le quote di circolazione sono salite dai **3,6 miliardi** di agosto ai **3,7** di settembre.

- `OCCURRENCE`:

  (166) Il patrimonio dell'Assofondi è **cresciuto**.

  (167) I ministri dei 150 Paesi se ne **tornano** in patria.

#### D.5.2 Attribute `tense`

Required. Capturing standard distinctions in the grammatical category of verbal tense. It can have values `PRESENT`, `PAST`, `FUTURE`, `IMPERFECT`, or `NONE`.
The values assigned to this attribute mirror the higly-surface based character of ISO-TimeML. The values presented are based on classical tense distinctions in Italian. It is imporatant to stress the fact that on the level of general temporal reference there are no major differences between Italian and English and also among other Indo-European languages. In the following Table (Tense classification), correspondences between the classical grammatical tense classification system and the ISO-TimeML values are presented:

Table: Tense classification

| Classical Grammatical Tense Classification | ISO-TimeML values |
|---|---|
| Presente Semplice | PRESENT |
| Passato Composto | PAST |
| Imperfetto | IMPERFECT |
| Passato Semplice | PAST |
| Trapassato | PAST |
| Piucchepperfetto (or Trapassato Prossimo | PAST |
| Futuro Semplice | FUTURE |
| Futuro Composto | FUTURE |

#### D.5.3 Attribute `aspect`

Required. Similar to `tense`, it captures standard distinctions in the grammatical category of semantic aspect. It can have values `PROGRESSIVE, PERFECTIVE, IMPERFECTIVE`, or `NONE`. With respect to English, Italian has not a clearcut morphological distinction to code semantic aspect. It is recognized and determined more on a sort of pragmatic level. In section D.11 some examples on aspect annotation for Italian are presented. Note that due to language specific issues and in the perspective of an automatic annotatoin process we did not propose to use fine-grained values like aorist, perfect, continuous or habitual but general cover term like PERFECTIVE and IMPERFECTIVE. The PROGRESSIVE value, which is a specification of the IMPERFECTIVE aspect, is restricted to explicit cases realised in Italian by an aspectual periphrasis.

#### D.5.4 Attribute `mood`

Required. Captures the mood of the event. It can have the following values:

- `COND`: it signals the conditional mood which in Italian is realized by the morphological inflection on the verb. It is used to speak of an event whose realization is dependent on a certain condition, particularly, but not exclusively, in conditional clauses.

  (168) **Mangerei** del pesce.

- `SUBJUNC(tive)`: has several uses in independent clauses. This mood is required for certain types of dependent clauses.

  (169) Voglio che tu te ne **vada**

- NONE.

If no inflectional morphology is present to indicate mood, then the default value is NONE

### D.6 Attribute vForm

Required. It encodes information for non-finite verb forms. Its values are INFINITIVE, PASTPART, PRESPART, GERUNDIVE and NONE.

### D.7 Attribute modality

Optional. It is used to convey the different degrees of modality nature of an event, mainly epistemic and deontic. Due to the fact that it is not an easy task to recognize the correct modality value of an event, it has been decided that this attribute is to be fulfilled in presence of the modal verbs *dovere*, *potere* and their paraphrases. Its values will be represented by the modal verb itself as in the following example.

I profughi **devono** abbandonare le loro case.

```
<EVENT eid="01" eiid="01" class="I_STATE" pos="VERB"
tense="PRESENT" aspect="NONE" mood="NONE" vForm="NONE"
modality=''DOVERE''>
devono
</EVENT>
```

### D.8 The tag <TIMEX3>

The <TIMEX3> tag annotate any temporal expression (a.k.a. *timex*) referring to:

- Day times (*mezzogiorno, 3, la sera, la mattina...*).
- Dates of different granularity: days (*ieri, 8 Gennaio 1980, venerdì scorso, sabato*, etc.), weeks (*la prossima settimana, la seconda settimana del mese*, etc.), months (*tra due mesi, il mese prossimo, l' Agosto del 1980*), seasons or business quarters (*la scorsa primavera, lo scorso semestre, il primo trimestre, il bimestre*, etc.), years (*1980, l'anno scorso*), centuries, etc.
- Durations (*due mesi, cinque ore, nei prossimi anni, periodo*).
- Sets (*una volta al mese, ogni martedì*).

No major changes are needed for this tag. Readers are referred to Annex A.2.2.2 for the BNF values and A.2.2.3 for the description of attributes. For clarity's sake, instructions on TIMEX3 will be presented here.

#### D.8.1 How to Annotate TIMEX3 in Italian

As stated in Annex A.2.2, the TIMEX3 tag relies on and is as much compliant as possible with the TIDES TIMEX2 annotation. The Italian adaptation of this annotation scheme is presented in Lavelli et al. (2005), to which we refer for the identification of the markable expressions. In the following paragraphs, differences with respect to the TIDES scheme are presented and illustrated.

##### D.8.1.1 Tag span

The surface-oriented approach to the tagging of expressions in ISO-TimeML implies that temporal expression annotation is based on costituent structure and the time unit classification presented in Annex A.2.2.1, B, Table 6.

#### D.8.1.1.1 Syntactic constituents and criteria for the annotation

The full extent of annotation must correspond to one of the following syntactic categories:

- **Noun phrase** (l*a scorsa estate, ieri, lo stesso periodo, maggio*).
- **Adjective phrase** (*primaverile, estivo, settimanale, giornaliero*).
- **Adverbial phrase** (i*eri, ora, recentemente, mensilemente*).

Prepositions preceding a temporal expression are to be excluded from the tag. Due to the specifications of ISO-TimeML, complex or contracted prepositions must be excluded from tagging:

(170) a. nel pomeriggio

b. `nel`
`<TIMEX3 tid=”01”> pomeriggio </TIMEX3>`

(171) a. per la settimana prossima

b. `per`
`<TIMEX3 tid=”01”> la settimana prossima </TIMEX3>`

All pre- and post-modifiers of a temporal expression must be included into the tag, with the exception of post-modifiers describing/denoting an event, adverbial premodifier like ”appena” and similar:

(172) a. quattro mesi fa

b. `<TIMEX3 tid=”01”> quattro mesi fa </TIMEX3>`

(173) a. non meno di un anno

b. `<TIMEX3 tid=’01’> non meno di un anno </TIMEX3>`

(174) a. circa sei anni di studio

b. `<TIMEX3 tid=”01”> circa sei anni </TIMEX3>`
`di <EVENT eid=’01’> studio </EVENT>`

(175) a. appena dodici anni fa

b. `appena`
`<TIMEX3 tid=’01’> dodici anni fa </TIMEX3>`

(176) a. il futuro dei nostri popoli

b. `<TIMEX3 tid=’01’> il futuro </TIMEX3>`
`dei nostri popoli`

The Italian word *dopo* can be either a temporal adverb, a temporal preposition or an adjective. Only in this last case it must be included into the markable:

(177) a. due ore dopo se n’era andato

b. <TIMEX3 tid="01"> due ore dopo </TIMEX> se n'era andato

Appositives are considered as post-modifiers, and thus are to be included into the tag span. However, if the appositives contain a lexical trigger (as explained in TIDES), then we have two separate expressions:

(178) a. gli anni '60, gli anni del libero amore.

b. <TIMEX3 tid="01"> gli anni '60 </TIMEX3>,
<TIMEX3 tid="02"> gli anni del libero amore </TIMEX3>.

Things are complicating when two temporal expressions are in a specification relation (e.g., *venerdì sera alle otto, martedì 26 giugno, dicembre del 1980...*). In the following examples, for clarity's sake, the TIMEX3 span is underlined.
In these cases:

- the temporal expressions will be marked up in a single tag if:
  - the two expressions belong to the same temporal unit, according to Annex A.2.2.1, Table 6, (e.g., *venerdì sera*, *alle 11 di mattina*, *martedì 26 giugno*, *giugno 1969*).
  - the second temporal expression is introduced by the prepositions *di* or *del* and it represents a definite time specification.
    (e.g., **la mattina del 20 giugno**, **ottobre del 1963**, **alle 11 di ieri mattina**, **alle 20.00 di giovedì**).

In all other cases, two tags must be created (e.g., *venerdì sera alle 20.00*, *ieri alle 11.00*).

NOTE It is important to stress the difference between temporal expressions of the form *NP+ PP*, where the head of the PP is realized by the prepositions *"di"* or *"del"*, and those cases where the head of the PP is realized by the prepositions *"a"* or its contracted variants. In the former case the expressions are viewed as belonging to the same syntactic constituent, while in the latter the temporal expression realized by the PP can attach either to the NP constituent or to a higher syntactic constituent like the IP or the VP.

Temporal expressions in an anchoring relation and in a conjunction relation will be marked up with two different tags (e.g., *due settimane da oggi*, *tre giorni prima di ieri*, *tra sei mesi o un anno*.) Those expressions denoting Times of Day, such as *le tre meno un quarto*, *dieci alle quattro*, *le tre e dieci*, are marked with a unique tag.

### D.9 The tag <SIGNAL>

Readers are referred to Annex A.2.3 for definitions and instructions on annotation and to A.2.3.2 for the BNF.

However, for Italian, it is necessary to consider the proper annotation of those SIGNALs which are realized by complex prepositions of the kind *"alle"*, *"dalle"*, *"dal"*, *"del"*, *"sul"*, *"al"*..., where a definite article is merged with a preposition (*"al=a+il"*). In such cases, the annotation must be conductucted as illustrated below:

Example (1): *dalle 3 di oggi*

```
<SIGNAL sid="s1">
dalle
</SIGNAL>
<TIMEX3 tid= "t1" type= "DATE" value= "2006-12-20T15:00">
3 di oggi
</TIMEX3>
```

Example (2): *l'incontro del 27 ottobre 2006*

```
l'
<EVENT eid="e1"  eiid="ei1" class="OCCURRENCE">
incontro
</EVENT>
<SIGNAL sid="s1">
del
</SIGNAL>
<TIMEX3 tid="t1" type="DATE" value="2006-10-26">
27 ottobre 2006
</TIMEX3>
```

### D.10 The link tags

Readers are referred to Annex A.3 for definitions, instructions on the annotation and the BNF for the three link tags.

### D.11 Examples of `tense` and `aspect` annotation in Italian

In this section we present some rules for annotating `tense` and `aspect` in Italian. Assignment of more than one value for `aspect` is due to the fact that the same tense can have more than one aspectual value according to co-textual and con-textual factors.

a) Events realized by finite verb forms:

- tense= "PRESENT"
  - *gioca* aspect= IMPERFECTIVE | PERFECTIVE | NONE
  - *sta giocando* aspect= PROGRESSIVE
  - *ha l'abitudine di giocare* aspect= IMPERFECTIVE
  - *ha mangiato* aspect= IMPERFECTIVE
  - *(che) mangi* aspect= IMPERFECTIVE | NONE
  - *mangerebbe* aspect= NONE
- tense= "PAST"
  - *giocò* aspect= PERFECTIVE
  - *ha giocato* aspect= PERFECTIVE | IMPERFECTIVE
  - *ebbe l'abitudine di giocare* aspect= PERFECTIVE
  - *fu mangiato* aspect= PERFECTIVE
  - *è stato mangiato* aspect= PERFECTIVE
  - *(che) abbia mangiato* aspect= PERFECTIVE
  - *aveva giocato* aspect= PERFECTIVE | IMPERFECTIVE
  - *ebbe giocato* aspect= PERFECTIVE
  - *era stata mangiata* aspect= PERFECTIVE
  - *(che) avesse mangiato* aspect= PERFECTIVE
  - *avrebbe mangiato* aspect= PERFECTIVE
- tense= "IMPERFECT"
  - *giocava* aspect= IMPERFECTIVE | PERFECTIVE
  - *stava giocando* aspect= PROGRESSIVE
  - *aveva l'abitudine di giocare* aspect= IMPERFECTIVE
  - *era mangiata* aspect= IMPERFECTIVE
  - *(che) mangiasse* aspect= NONE
- tense= "FUTURE"
  - *giocherà* aspect= IMPERFECTIVE | PERFECTIVE

- *avrà giocato* aspect= PERFECTIVE | IMPERFECTIVE
- *sarà mangiata* aspect= IMPERFECTIVE

b) Events realized by non-finite verb forms:

- tense= "PRESENT" vform="INFINITIVE"
  - *giocare* aspect= PERFECTIVE | IMPERFECTIVE | NONE

NOTE Non-finite verb forms do not have automonous temporal reference. For clarity's sake we keep the `PRESENT` value for simple forms and `PAST` for the compound ones. In addition to this, the aspectual values of simple forms are based on compelx semantic and pragmatic factors. As Bertinetto (2001) points out, aspectual values of simple infinitives introduced by verbs are influenced by: (a) lexical semantics of the main (finite) verb, (b) the semantic aspect of the main verb, and (c) the lexical aspect, or *Aktionsaart*, both of the main verb and of the infinitive form, together with pragmatic factors. Simple (or present) infinitives introduced by modals have `PERFECTIVE` aspect because modals force a prospective/futurate reading of the infinitive; if modals have an epistemic reading, the infinitives have `IMPERFECTIVE` aspect, because modals force a simultaneous/present reading.

a) tense="PAST" vform="INFINITIVE"

-*aver giocato* aspect= PERFECTIVE

b) tense= "PAST" vform="PASTPART"

- *giocato* aspect="PERFECTIVE"

c) tense= "PRESENT" vform="PRESPART"

- *giocante* aspect= NONE

d) tense= "PRESENT" vform="GERUNDIVE"

- *giocando aspect* = IMPERFECTIVE | PERFECTIVE | NONE

NOTE Like present (*simple*) infinitive, present (*simple*) gerunds in Italian receive aspectual values according to contextual interpretation.

a) tense= "PRESENT" vform="GERUNDIVE"
   - *avendo giocato* aspect= PERFECTIVE

b) Modal verbs:

- tense= "PRESENT | IMPERFECT" aspect= "IMPERFECTIVE"
  - *devo [andare a casa]*
  - *dovevo [andare a casa]*
- tense= "PAST" aspect= "PERFECTIVE"
  - *è dovuto [andare a casa]*
  - *dovette [andare a casa]*
  - *era dovuto [andare a casa]*

NOTE In epistemic reading modals have IMPERFECTIVE aspect.

a) Events realized by adjectives or nouns:

- tense= "NONE"
- aspect= "NONE"

- *La coppia, residente a Milano, era in vacanza.*
- *I lavoratori chiedono nuove trattative.*

b) Events realized by adjectives or nouns in predicative clauses. Events are in bold face, whereas the copular verb, from which temporal and aspectual information can be recovered, is underlined.

- tense= "PRESENT | IMPERFECT | PAST | FUTURE"
- aspect= "PERFECTIVE | IMPERFECTIVE | NONE"

- *Questo è un **accordo** per altri negoziati.*
  (PRESENT - NONE)
- *La coppia è **residente** a Milano.*
  (PRESENT- IMPERFECTIVE)
- *Questo era un **tentativo** per un accordo.*
  (IMPERFECT - NONE)
- *La coppia era **residente** a Milano.*
  (IMPERFECT - PERFECTIVE)
- *Questo è stato un **accordo** per altri negoziati.*
  (PAST - NONE)
- *Questo fu un **tentativo** di accordo.*
  (PAST - NONE)
- *La coppia è stata **residente** a Milano.*
  (PAST - PERFECTIVE | IMPERFECTIVE)
- *La coppia fu **residente** a Milano.*
  (PAST - PERFECTIVE)
- *Questo era stato l'**accordo** per i negoziati.*
  (PAST - NONE)
- *La coppia era stata **residente** a Milano.*
  (PAST - PERFECTIVE | IMPERFECTIVE)
- *Al Sayed sarà il nuovo **padrone** della Fermenta.*
  (FUTURE - NONE | )
- *Questo sarà stato un **tentativo** di aprire nuovi negoziati.*
  (FUTURE - NONE)
- *La coppia sarà **residente** a Milano.*
  (FUTURE - IMPERFECTIVE).

For events realized by prepositional phrases, the values for tense and aspect are the same as for those for adjectives and nouns, both if they occur alone or as a predicative complement of a copular phrase.

### D.12 Sample of Italian annotation

```
<ISO-TimeML>
  Repubblica
  <TIMEX3 functionInDocument="PUBLICATION_TIME"
  temporalFunction="false" tid="t2" type="DATE"
  value="1985-01-30">
  30/01/1985
  </TIMEX3>
  economia La Fiom contesta le scelte dell' Flm .
  I DELEGATI RESPINGONO L' ACCORDO CORNIGLIANO .

  GENOVA ??L'
  <EVENT class="OCCURRENCE" eid="e1" eiid="ei1" tense="NONE"
  aspect="NONE" pos="NOUN" mood="NONE" polarity="POS">
  assemblea
  </EVENT>
```

```
dei lavoratori Italsider di Cornigliano
( erano presenti duemila operai ) ha sostanzialmente
<EVENT class="I_ACTION" eid="e2  eiid="ei2" tense="PAST"
aspect="IMPERFECTIVE" pos="VERB" mood="NONE"
polarity="POS" >
contestato
</EVENT>
l'
<EVENT class="OCCURRENCE" eid="e3" eiid="ei3" tense="NONE"
aspect="NONE" pos="NOUN" mood="NONE" polarity="POS" >
accordo
</EVENT>
<EVENT class="STATE" eid="e4" eiid="ei4" tense="PASTPART"
aspect="NONE" pos="VERB" mood="NONE" polarity="POS" >
raggiunto
</EVENT>
<TIMEX3 anchorTimeID="t2" functionInDocument="NONE"
temporalFunction="false" tid="t3" type="DATE"
value="1985-01-25">
venerdi' scorso
</TIMEX3>
tra la Finsider e la Flm nazionale e regionale ,
in base al quale lo stabilimento genovese
<EVENT class="ASPECTUAL" eid="e5" eiid="ei5" tense="FUTURE"
aspect="NONE" pos="VERB" mood="NONE" polarity="POS" >
riprendera'
</EVENT>
a
<EVENT class="STATE" eid="e6" eiid="ei6" tense="INFINITIVE"
aspect="NONE" pos="VERB" mood="NONE" polarity="POS" >
produrre
</EVENT>
<SIGNAL sid="s1">
dal
</SIGNAL>
<TIMEX3 anchorTimeID="t2" functionInDocument="NONE"
temporalFunction="false" tid="t4" type="DATE"
value="1985-05-01">
primo maggio
</TIMEX3>
con 1600 addetti e sara'
<EVENT class="OCCURRENCE" eid="e7" eiid="ei7" tense="FUTURE"
aspect="PERFECTIVE" pos="VERB" mood="NONE" polarity="POS" >
gestito
</EVENT>
da una societa' pubblica
( Nuova Italsider , Dalmine e Acciaierie di Piombino ) ,
<EVENT class="STATE" eid="e29" eiid="ei29" tense="NONE"
aspect="NONE" pos="PREPOSITION" mood="NONE" polarity="POS" >
in attesa
</EVENT>
dei privati . I delegati della lega Fiom
di Cornigliano e dell' ' Oscar Senigallia ' ,
in particolare , come
<SIGNAL sid="s2">
gia'
</SIGNAL>
```

```
il consiglio di fabbrica , hanno
<EVENT class="OCCURRENCE" eid="e8"eiid="ei8" tense="PAST"
aspect="PERFECTIVE" pos="VERB" mood="NONE" polarity="POS" >
attaccato
</EVENT>
la Flm nazionale e regionale
<EVENT class="I_ACTION" eid="e9" eiid="ei9" tense="PRESPART"
aspect="NONE" pos="VERB" mood="NONE" polarity="POS" >
accusando
</EVENT>
le di averli
<EVENT class="I_ACTION" eid="e10" eiid="ei10"  tense="INFINITIVE"
aspect="PERFECTIVE" pos="VERB" mood="NONE" polarity="POS" >
esclusi
</EVENT>
dalle
<EVENT class="OCCURRENCE" eid="e11" eiid="ei11"  tense="NONE"
aspect="NONE" pos="NOUN" mood="NONE" polarity="POS" >
trattative
</EVENT>
, e non hanno
<EVENT class="OCCURRENCE" eid="e12" eiid="ei12"  tense="PAST"
aspect="PERFECTIVE" pos="VERB" mood="NONE" polarity="NEG" >
firmato
</EVENT>
la bozza di intesa . Essi
<EVENT class="I_ACTION" eid="e13" eiid="ei13"  tense="PRESENT"
aspect="NONE" pos="VERB" mood="NONE" polarity="POS" >
chiedono
</EVENT>
, e l’ assemblea e’ stata d’ accordo , alcune
fondamentali
<EVENT class="OCCURRENCE" eid="e20" eiid="ei20"  tense="NONE"
aspect="NONE" pos="NOUN" mood="NONE" polarity="POS" >
integrazioni
</EVENT>
, che
<EVENT class="I_STATE" eid="e30" eiid="ei30"  tense="PRESENT"
aspect="NONE" pos="VERB" mood="SUBJUNCTIVE" polarity="POS" >
dovrebbero
</EVENT>
<EVENT class="I_ACTION" eid="e16" eiid="ei16"  tense="INFINITIVE"
aspect="NONE" pos="VERB" mood="NONE" polarity="POS" >
scaturire
</EVENT>
da nuove
<EVENT class="OCCURENCE" eid="e17" eiid="ei17"  tense="NONE"
aspect="NONE" pos="NOUN" mood="NONE" polarity="POS" >
trattative
</EVENT>
che i delegati Fiom
<EVENT class="I_STATE" eid="e18" eiid="ei18"  tense="PRESENT"
aspect="NONE" pos="VERB" mood="NONE" polarity="POS" >
intendono
</EVENT>
<EVENT class="OCCURRENCE" eid="e19" eiid="ei19"  tense="INFINITIVE"
aspect="NONE" pos="VERB" mood="NONE" polarity="POS" >
```

```
aprire
</EVENT>
con la Nuova Italsider a livello locale .
I temi in discussione
<EVENT class="STATE" eid="e21" eiid="ei21"  tense="PRESENT"
aspect="NONE" pos="VERB" mood="NONE" polarity="POS" >
riguardano
</EVENT>
il numero fisso dei 1600 addetti ,
le condizioni di lavoro , la
<EVENT class="OCCURRENCE" eid="e22" eiid="ei22"  tense="NONE"
aspect="NONE" pos="NOUN" mood="NONE" polarity="POS" >
produzione
</EVENT>
iniziale troppo limitata
<SIGNAL sid="s3">
per
</SIGNAL>
<EVENT class="OCCURRENCE" eid="e24" eiid="ei24"  tense="INFINITIVE"
aspect="NONE" pos="VERB" mood="NONE" polarity="POS" >
rendere
</EVENT>
redditizio e competitivo lo stabilimento
( meno di un milione di tonnellate all' anno ) ,
e numerose altre condizioni , che
<EVENT class="I_STATE" eid="e31" eiid="ei31"  tense="PRESENT"
aspect="NONE" pos="VERB" mood="NONE" polarity="POS" >
devono
</EVENT>
' essere
<EVENT class="STATE" eid="e25" eiid="ei25"  tense="INFINITIVE"
aspect="PERFECTIVE" pos="VERB" mood="NONE" polarity="POS" >
rese
</EVENT>
piu' esplicite ' . Piu' positivo , invece ,
il giudizio di Fim-Cisl e Uilm-Uil , che hanno
<EVENT class="REPORTING" eid="e26" eiid="ei26"  tense="PAST"
aspect="PERFECTIVE" pos="VERB" mood="NONE" polarity="POS" >
annunciato
</EVENT>
per
<TIMEX3 anchorTimeID="t2" functionInDocument="NONE"
temporalFunction="false" tid="t5" type="DATE"
value="1985-01-30">
oggi
</TIMEX3>
una
<EVENT class="OCCURRENCE" eid="e27" eiid="ei27"  tense="NONE"
aspect="NONE" pos="NOUN" mood="NONE" polarity="POS" >
conferenza
</EVENT>
stampa e che sono
<EVENT class="I_STATE" eid="e32" eiid="ei32"  tense="PRESENT"
aspect="NONE" pos="ADJECTIVE" mood="NONE" polarity="POS" >
favorevoli
</EVENT>
 ad una
```

```
<EVENT class="OCCURRENCE" eid="e28" eiid="ei28"  tense="NONE"
aspect="NONE" pos="NOUN" mood="NONE" polarity="POS" >
votazione
</EVENT>
referendaria sulla bozza di accordo .

<TLINK lid="l11" origin="USER" relType="BEFORE"
relatedToTime="t2" timeID="t3" />
<TLINK lid="l12" origin="USER" relType="INCLUDES"
relatedToTime="t2" timeID="t5" />
<TLINK lid="l13" origin="USER" relType="BEFORE"
relatedToTime="t4" timeID="t5" />
<TLINK eventInstanceID="ei4" lid="l14" origin="USER"
relType="IS_INCLUDED" relatedToTime="t3" />
<TLINK eventInstanceID="ei2" lid="l15" origin="USER"
relType="AFTER" relatedToEventInstance="ei4" />
<TLINK eventInstanceID="ei6" lid="l19" origin="USER"
relType="BEGUN_BY" relatedToTime="t4" signalID="s1" />
<TLINK eventInstanceID="ei7" lid="l20" origin="USER"
relType="BEGUN_BY" relatedToTime="t4" signalID="s1" />
<TLINK eventInstanceID="ei29" lid="l21" origin="USER"
relType="DURING" relatedToEventInstance="ei7" />
<TLINK eventInstanceID="ei33" lid="l23" origin="USER"
relType="DURING" relatedToTime="t5" signalID="s3" />
<TLINK eventInstanceID="ei26" lid="l25" origin="USER"
relType="DURING" relatedToTime="t5" />
<TLINK eventInstanceID="ei11" lid="l27" origin="USER"
relType="BEFORE" relatedToTime="t3" />
<TLINK eventInstanceID="ei13" lid="l28" origin="USER"
relType="IS_INCLUDED" relatedToTime="t5" />
<TLINK eventInstanceID="ei2" lid="l30" origin="USER"
relType="BEFORE" relatedToTime="t2" />
<TLINK eventInstanceID="ei8" lid="l31" origin="USER"
relType="BEFORE" relatedToTime="t2" />
<TLINK eventInstanceID="ei18" lid="l32" origin="USER"
relType="BEGINS" relatedToTime="t5" />
<SLINK eventInstanceID="ei2" lid="l1" origin="USER"
relType="FACTIVE" subordinatedEventInstance="ei3" />
<SLINK eventInstanceID="ei9" lid="l3" origin="USER"
relType="FACTIVE" subordinatedEventInstance="ei10" />
<SLINK eventInstanceID="ei10" lid="l4" origin="USER"
relType="FACTIVE" subordinatedEventInstance="ei11" />
<SLINK eventInstanceID="ei13" lid="l5" origin="USER"
relType="FACTIVE" subordinatedEventInstance="ei20" />
<SLINK eventInstanceID="ei30" lid="l6" origin="USER"
relType="INTENSIONAL" subordinatedEventInstance="ei16" />
<SLINK eventInstanceID="ei16" lid="l7" origin="USER"
relType="FACTIVE" subordinatedEventInstance="ei17" />
<SLINK eventInstanceID="ei18" lid="l8" origin="USER"
relType="INTENSIONAL" subordinatedEventInstance="ei19" />
<SLINK eventInstanceID="ei31" lid="l9" origin="USER"
relType="INTENSIONAL" subordinatedEventInstance="ei25" />
<SLINK eventInstanceID="ei26" lid="l10" origin="USER"
relType="EVIDENTIAL" subordinatedEventInstance="ei27" />
<SLINK eventInstanceID="ei33" lid="l24" origin="USER"
relType="FACTIVE" subordinatedEventInstance="ei32" />
```

 

# Annex E
(informative)
# Temporal annotation of verbs in Korean

NOTE This annex is a slightly reformatted version of an article by Kiyong Lee, which is being reviewed for publication in a Korean journal.

## E.1 Introduction

In Korean, verbal endings carry information on tense, aspect, modality, and mood. This annex focuses on ways of annotating such temporal and event-related information conveyed by verbal endings. As in other languages, addverbials and nouns also carry time and event-related information. But the annotation of these expressions is not treated here, for they can be annotated according to the general annotation guidelines as specified in ISO-TimeML.

ISO-TimeML is an XML-based formal language that provides the general guidelines for annotating temporal and event-related information in natural language texts. These guidelines are found in an ISO working draft for international standard, entitled *ISO 24617-1:2007 Language Resources Management – Semantic annotation framework – Part 1: Time and events.* The present annex is part of this document.

## E.2 Basic references

This annex heavily relies on the following references:

- Chang, Suk-Jin (1996), *Korean*, John Benjamins Publishing Co., Amsterdam.
- Lee, Iksop, and S. Robert Ramsey (2000), *The Korean Language*, State University of New York Press, Albany.
- Sohn, Ho-Min (1999), *The Korean Language*, Cambridge University Press, Cambridge.

These are general introductory books on the Korean language and its grammar for foreign readers, published most recently.

## E.3 Morpholoy of Korean predicates

In Korean, adjectives function as predicates like verbs. They have inflectional endings instead of a copula.

NOTE In this document, the term “verbal endings" refers to both adjectival and verb endings, unless a clear distinction is necessitated.

```
(1) 미아가 잔다 (verb)
    mia-ka ca-n-ta
    mia-NOM sleep-IND-DECL
    'Mia sleeps'
(2) 미아가 예쁘다 (adjective)
    mia-ka yey.ppu-ta
    mia-NOM pretty-DECL
    'Mia is pretty'
```

NOTES
1.There are several different ways of romanizing Hangul. Here, the Yale romanization of Hangul is adopted because it best reflects the morpho-syntactic properties of Korean.
2. In the romanization used here, the dot indicates a syllable break and the hyphen a morphological break. The verb 잔다 *can.ta*, for instance, consists of two syllables, while it consists of three morphemes, *ca-n-ta*, STEM-IND-DECL.
3. NOM stands for NOMINATIVE case.
4. IND stands for INDICATIVE mood marker.
5. DECL stands for sentence-final DECLARATIVE ending.

Being an agglutinative language, Korean concatenates nominal or verbal stems with a sequence of nominal particles or verbal endings, respectively, which can be of great length. Here is an example of a verbal concatenation:

(3) 뒷받쳐 주셨겠습니다
twis-patchy-e cwu-sy-ess-keyss-swup-ni-ta
PFX-STEM-BR AuxVerbSTEM-SH-PAST-CONJEC-AH-IND-DECL
'must have supported'

The example above is a complex verbal construction, consisting of two verbal parts : one is the main verbal part and the other its auxiliary verbal part. These two are connected by a bridge BR *-e*. The main verbal part *twis-patchy-e* 'support' consists of a prefix (PRX), a STEM, and a BRidge in this order. This sequence is then followed by an auxiliary verbal part *cwu-sy-ess-keyss-swup-ni-ta* that consists of an auxiliary verbal stem *cwu-* 'give' and a sequence of six verbal endings in the following order:

- SH: Subject Honorific suffix *-si*
- PAST: past tense ending *-ess*
- CONJEC: CONJECTURAL modal ending *-keyss*
- AH: Addressee Honorific ending *-swup*
- IND: INDICATIVE mood ending *-ni*
- DECL: declarative sentence-final ending *-ta*

Some of these verbal endings express tense, aspect, modality or mood. The following examples illustrate the ordering of endings that express such temporal information.

| | stem-cType | -aspect | -tense | -modality | -mood | -sType |
|---|---|---|---|---|---|---|
| (4) | 만나-고 | 있 | -었 | -겠 | -더 | -라 |
| | manna-ko | iss | -ess | -keyss | -te | -la |
| | meet-BR | PROG | PAST | CONJEC | RETRO | DECL |

'must/may have been meeting'

NOTES
1. cType stands for the type of sentence-non-final verbal ending, clause type.
2. sType stands for the type of sentence-final verbal ending, sentence type.
3. BR again stands for a bridge between a main verb and its auxiliary verb.
4. PROG stands for PROGRESSIVE aspect.
5. RETRO stands for RETROSPECTIVE mood.

Each of these temporal features will be discussed in the following clauses.

## E.4 Temporal structure: informative

Time and tense are entities of two different types. Time is part of the world with a certain structure, while tense is a linguistic feature related to time. But the use of terms like *present*, *future* and *past* can be ambiguous: they can either be used in an ontological sense or in a linguistic sense.

NOTE To make this distinction clear, the ontological present, future, and past are written in the lower case, while `PRESENT`, `FUTURE`, and `PAST`, written in the upper case, are understood to be grammatical entities that serve as values of the attribute `tense`.

The temporal structure assumed here is a quadruple $< T, I, n, R >$, where

- $T$ is a set of points of time,
- $I$ is a set of intervals of time,
- $n$ is the uniquely designated point of time in $T$, known as the present moment of time,
- $R$ is a set of temporal relations over $T$ or $I$ that include:
  - the partial (precedence) relation,
  - the overlap relation over intervals, and
  - the neighborhood $N$ of a point of time $t$ in $T$ such that $N(t)$ is an open interval that includes the point of time $t$.

On the basis of this structure, the following intervals are defined.

- The present (time) refers to an open interval $N(n)$, the neighborhood of the designated point of time $n$
- The past (time) refers to an interval preceding the designated point of time $n$.
- The future (time) refers to an interval which is not preceded by nor identical with the designated point of time $n$

NOTES
1. According to this definition, the past and the future time interval each may overlap with the present time interval, $N(n)$. The end point of the past may be included in $N(n)$ and the beginning point of the future may also be in $N(n)$.
2. By introducing the notion of neighborhood, as discussed in K. Lee (1998), the present moment of time can be understood as a temporal construct to define other temporal structures, whereas the linguistic expression like 지금 *cikum* 'now' in Korean can be understood as referring to its neighborhood, namely $N(n)$.
3. Such an interpretation allows the co-occurrence of 지금 *cikum* 'now' with the PAST tensed verb as in 지금 도착했다 *cikum tochakha-yss-ta* (now arrive-PAST-DECL) 'have arrived now'.

## E.5 Temporal annotation of non-Latin texts

The tense and other temporal features of an event are annotated according to ISO-TimeML. For texts written in characters other than Latin characters, as in English or French texts, however, two extra attributes may be introduced: `romanization` and `content`. The first attribute specifies how non-Latin character texts like Korean texts are romanized, whereas the second provides semantic content in some recognizable language, say English. Here is an example:

```
(5) 미아가 지금(now)
    <EVENT eid="e1" eiid="ei1"
     yaleRomanization="to.chak.ha-yss-ta" content="ARRIVE"
     tense="PAST" vForm="sFINAL">
     도착했다
    </EVENT>
```

NOTES
1. As noted earlier, this annex adopts the Yale romanization with the attribute `yaleRomanization`.
2. The value of `content` is `CDATA`, any non-empty sequence of charaters, representing the information content of a string romanized. If desirable for some other purposes, this content may also be represented by a logical form like `ARRIVE(mia)`.
3. ISO-TimeML also provides guidelines for annotating temporal adverbs and nominal expressions. But here to focus on the annotation of temporal features associated with events, other temporal expressions like temporal adverbs may sometimes be annotated simply with their English equivalents in parentheses in the text.
5. As will be discussed presently, the interpretation of tense in Korean depends on the type of a `vForm` which completes a verb or adjective stem as a word form.
6. The value of `vForm` -다 *-ta* is a `DECLARATIVE` verbal ending. But, for the purpose of interpreting tense, it is just sufficient to specify it with a more general value `sFinal`, standing for *sentence-final* verbal endings.

## E.6 Tense

### E.6.1 Tense markers

Despite its allomorphic variants, Korean has a single `tense` value, namely `PAST`, for a single tense morpheme -ㅆ *–ss*. This form has, however, four other variants: *-e-ss*, *-a-ss*, *-ye-ss*, and *-ay-ss*, where *-e*, *-a*, *-ye*, and *-ay* are considered by K. Lee (1999) as bridging vowels the choice of which depends on the syllable structure of a verbal stem to which the bridge is attached.

The PAST tense marker has a doubled form *-ess.ess*, being treated as PAST-PAST, PAST-PERFECTIVE, PLUPERFECT or REMOTE PAST. In this annex, it is annotated with `tense="PAST" aspect="PERFECTIVE"`.

Neither PRESENT nor FUTURE is morphologically marked in Korean verbs and adjectives. These tense values are, however, displayed as part of the adnominal forms of adjectives and verbs like 먹는 *mek-nun* (eat-adnomPRESENT), 먹을 *mek-ul* (eat-adnomFUTURE), and 먹은 *mek-un* (eat-adnomPAST). Here is the list of `PAST` tense endings:

#### E.6.1.1 List of `PAST` tense endings

- PAST: -ㅆ(-ss), -었(-ess), -았(-ass), 였(-yess), -ㅣㅆ(-yss)
- PAST-PAST: -ㅆ었(-ss-ess), -었었(-ess-ess), -았었(-ass-ess), -였었(-yess-ess), -ㅣㅆ었(-yss-ess)

NOTE The form -lㅆ *-yss* allows the abridged form of a *ha*-verb like 했 *ha-yss*".

#### E.6.1.2 Examples

```
(6)  어제        나는     미아를      만났다
     ecey       na-nun    mia-lul   manna-ss-ta
     yesterday  I-TOP     Mia-ACC   meet-PAST-DECL
         'Yesterday I met Mia'
(7)  전에  나는      미아를      만났었다
     ceney na-nun    mia-lul   manna-ss.ess-ta
     before I-TOP    Mia-ACC   meet-PASTperfective-DECL
         'I had met Mia before'
(8)  지금   나는    커피를      마신다
     cikum na-nun khephi-lul masi-n-ta
     now  I-TOP coffee-ACC drink-IND-DECL
        'Now I drink (am drinking) meal'
(9)  지금/내일      나는   미아를     만난다
     cikum/nayil  na-nun Mia-lul  manna-n-ta
     now/tomorrow I-TOP  Mia-ACC meet-IND-DECL
        'Now/tomorrow I meet (will be meeting) Mia'
```

NOTE TOP and ACC are case markers, respectively standing for TOPICALIZER and ACCUSATIVE. Both Subject and Object can be marked by TOP, while Object is marked by ACC.

### E.6.2 Annotation guidelines for tense

Annotation depends on the surface information of markables only. At least for implementation purposes, the task of annotation should be carried out purely on surface information through routine manner. Tense is thus annotated independently of its contextual information, although its interpretation varies contextually.

- PAST: If a verb or an adjective word form contains a tense marker -었 –ess or its morphological variants, then it is annotated with tense="PAST".
- PAST PERFECTIVE: If a verb or an adjective word form contains a tense marker -었었 –ess.ess or its morphological variants, then it is annotated with tense="PAST" with the additional specification of aspect="PERFECTIVE".
- NONE: If a verb or an adjective contains no tense marker, then it is annotated with tense="NONE".

NOTES
1. The so-called INDICATIVE mood markers -는 *-nun* or -ㄴ *n* occurs with verbs which are interpreseted as referring to an event in the present or future, as in 내일 떠난다 *nayil ttena-n-ta* tomorrow, leave-IND-DECL. But even in such a case, the value of tense is assigned NONE.
2. The verbal sentence-pre-final ending *-n/-nun/-nu* is often treated as PRESENT tense marker. See Chang (1996: 118). But Sohn (1999) treats it as the (NON-PAST) INDICATIVE mood marker, for it occurs only with the sentence-final DECLARATIVE ending *-ta* or the sentence-final APPERCEPTIVE DECLARATIVE sentence ending *–kuna*, when the stem is a type of genuine (action) VERB which is not an adjectival stem.
3. The ending *-keyss* is treated as a CONJECTURAL modal marker. It can be used with the PAST tense marker, as in *manna-ss-keyss-ta* 'must/might have met'.

### E.6.3 Contextual interpretation of tense

Tense is interpreted differently depending on its context. Consider the following EVENT-annotated complex example:

(10) 미아가 사온 사과를 먹고 잤다고 용이 말했다
'Yong said that he ate an apple which Mia had bought and slept'

```
미아가
<EVENT eid="e1" eiid="ei1"
yaleRomanization="sa-on" content="BUY"
class="OCCURRENCE" pos="VERB" vForm="ADNOMINAL"
tense="PAST" >
사온
</EVENT>
사과를
<EVENT eid="e2" eiid="ei2"
yaleRomanization="mek-ko" content="EAT"
class="OCCURRENCE" pos="VERB" vForm="CONJUNCTIVE"
tense="NONE" >
먹고
</EVENT>
<EVENT eid="e3" eiid="ei3"
yaleRomanization="ca-ss-ta-ko" content="SLEEP"
class="OCCURRENCE" pos="VERB" vForm="COMPLEMENTIZER"
tense="PAST" >
```

```
잤다고
</EVENT>
용이
<EVENT eid="e4" eiid="ei4"
yaleRomanization content="SAY"
class="OCCURRENCE" pos="VERB" vForm="sFINAL"
tense="PAST" >
말했다
</EVENT>
```

Each of the four `EVENT`s here is specfied with `tense` and its value. But each occurrence of `tense` values should be interpreted differently. There are, for instnace, three `EVENT` elements identified as `e1`, `e3` and `e4` that are all specified with the same type of tense specification `tense="PAST"`. But they are interpreted as having a different temporal relation with each other: the event `e1` is understood to have occurred `BEFORE e2` and `e2 BEFORE e3`.

To capture such a difference in temporal ordering, the interpretation of tense in Korean may be described by sub-featuring tense into:

- Absolute tense or simply tense
- Embedded tense
- Relative tense
- Inherited tense

#### E.6.3.1 Absolute interpretation of tense

Consider the following:

```
(11)  어제(yesterday) 나는 미아를
     <EVENT eid="e11" yaleRomanization="manna-ss-ta"
     content="MEET" class="OCCURRENCE" pos="VERB"
     tense="PAST" vForm="sFINAL">
     만났다
     </EVENT>
(12) 전에(before) 나는 미아를
    <EVENT eid="e12" yaleRomanization="manna-ss.ess-ta"
    content="MEET" class="OCCURRENCE" pos="VERB"
    tense="PAST" aspect="PERFECTIVE" vForm="sFINAL">
     만났었다
    </EVENT>
(13) 내일(tomorrow) 나는 미아를 또
    <EVENT eid="e13" yaleRomanization="manna-nta"
    content="MEET" class="OCCURRENCE" pos="VERB"
    tense="NONE" vForm="sFINAL">
    만난다
    </EVENT>
(14) 미아는
    <EVENT eid="e14" yaleRomanization="yeppu-ta"
    content="PRETTY" class="STATE" pos="ADJECTIVE"
    tense="NONE" vForm="sFINAL">
    예쁘다
    </EVENT>
```

Each occurrence of the tensed events referred to in the above examples is interpreted indepent of any other event, although some of them are temporally related to the time period referred to by the temporal expressions like 어제 *ecey* 'yesterday', 전에 *ceney* 'before' or 내일 *naeyil* 'tomorrow'. It should, however, be noted that the interpretation of a verb with the specificaiton `tense="NONE"` may depend on a temporal expression, for it can be interpreted as being ambiguous, either referring to an event in the present or to an event in the future. Consider the following:

```
(15) 내일(tomorrow) 나는 미아를
    <EVENT eid="e11"
    yaleRomanization="manna-nta" content="MEET"
    class="OCCURRENCE" pos="VERB" vForm="sFINAL"
    tense="NONE">
    만난다
    </EVENT>
```

**Interpretation conditions for absolute `tense`:**

Given an `EVENT` element specified with `vForm="sFINAL"`,

**PAST:** If the `tense` attribute in that `EVENT` element is specified with the value `PAST`, then that event is interpreted as having occurred in the past.

**PASTperfective:** If the `tense` attribute in that `EVENT` element is specified with the value `PAST`, but also if that `EVENT` element contains the specification `aspect="PERFECTIVE"`, then that event is interpreted as having occurred and also *completed* in the past.

**NONE:** If the value of `tense` attribute in that `EVENT` element is `NONE`, then that event is interpreted as occurring in the present.

But the present time interval can be extended to the future time, if the event is contextualized particularly by a temporal expression referrring to the future.

NOTES
1. In Korean, the present and the future interpretations of verbal expressions are differentiated not by a tense value, but can be specified by a temporal adverb like 내일 *nayil* 'tomorrow' or a modal operator like the CONJECUTARAL *-keyss*. This particular modal ending has thus been treated as FUTURE tense marker or FUTURITY modal marker (C. Lee (1987)).
2. Subordinate endings like the CONCESSIVE ending -지만 *-ciman* may also allow the absolute interpretation of tense, which will be discussed on another occasion.

#### E.6.3.2 Embedded Tense

Verbs of saying or asking have tensed sentences as complements. Here are examples:

```
(16) 미국에 언제 갔냐고 미아가 물었다
     mikuk-ey encey ka-ss-nya-ko      mia-ka mwul-ess-ta
     US-to    when go-PAST-INTERRO-COMP mia-NOM ask-PAST-DECL
     "Mia asked when Yong went to US"
(17) 곧 간다고 용이 대답했다
     kot ka-n-ta-ko        taytapha-yss-ta
     soon go-IND-DECL-COMP answer-PAST-DECL
     "Yong answered that (he) would go soon"
     그리곤  어제     떠났다
     kurikon ecey     ttena-ss-ta
     and   yesterday  leave-PAST-DECL
     "And then (he) left yesterday"
```

These sentences can be annotated in a straightforward manner as below:

```
(18) 미국에
     언제(when)
     <EVENT eid="e1" eiid="ei1"
     yaleRomanization="ka.ss,nya" content="GO"
     class="OCCURRENCE" pos="VERB" vForm="COMP"
     tense="PAST">
     갔냐
     </EVENT>
     <SIGNAL sid="s1"
     yaleRomanization="-ko">
     -고
     </SIGNAL>
     <EVENT eid="e2" eiid="ei2"
     yaleRomanization="mwul.ess.ta" content="ASK"
     class="OCCURRENCE" pos="VERB" vForm="sFINAL"
     tense="PAST">
     물었다
     </EVENT>
     <TLINK eventInstanceID="ei1" signalID="s1"
     relatedToEventInstance="ei2" relType="BEFORE"/>
(19) 곧(soon)
     <EVENT eid="e3" eiid="ei3"
     yaleRomanization="kan.ta" content="GO"
     class="OCCURRENCE" pos="VERB" vForm="COMP"
     tense="NONE">
     간다
     </EVENT>
     <SIGNAL sid="s2" yaleRomanization="-ko">
     -고
     </SIGNAL>
     용이
     <EVENT eid="e4" eiid="ei4"
     yaleRomanization="tay.tap.hayss.ta" content="ANSWER"
     class="OCCURRENCE" pos="VERB" vForm="sFINAL"
     tense="PAST" >
     대답했다
     </EVENT>
     <TLINK eventInstanceID="ei3" embeddedTense="PRESENT"
     signalID="s2" relatedToEventInstance="ei4" relType="AFTER"/>

     그리곤
     <TIMEX3 tid="t1"
     yRomanization="e.cey" content="YESTERDAY"
      type="DATE" value="2007-05-06>
     어제
     <EVENT eid="e5" eiid="ei5"
     yRomanization="tte.nass.ta" content="LEAVE"
     class="OCCURRENCE" pos="VERB" vForm="sFINAL"
     tense="PAST" >
     떠났다
     </EVENT>
     <TLINK eventInstanceID="ei5" relatedToTime="t1"
      relType="IS_INCLUDED"/>
```

#### E.6.3.2.1 Temporal linking of embedded tense

Given a sequence of three elements $\alpha, \sigma$, and $\beta$ such that

- $\alpha$ is an `EVENT` element specified with `eid="ej", eiid="eij"` for some integer `j` and also with "`vForm="COMP"`,
- $\sigma$ is a `SIGNAL` element specified with `sid="sk"` for some integer `k` for `vForm="COMP"` in $\alpha$, and
- $\beta$ is another `EVENT` element specified with `eid="em"` and `eiid="eim"` for some integer `m` that immediately follows both of the elements, $\alpha$ and $\sigma$,

  introduce:

  ```
  <TLINK eventInstanceID="eij" relatedToEventInstanceID="eim"/>
  ```

- specified with `relType="BEFORE"`, if $\alpha$ is specified with `tense="PAST"`,
- specified with `relType="AFTER"` otherwise.

#### E.6.3.2.2 Interpretation of embedded tense

The interpretation of embedded tense is governed by its related `TLINK` and is almost obvious by the `relType` between the given event instances.

**Interpretaiton conditions**

- If an event instance `ei1` is related to another event instance `ei2` and the type of their relation is `AFTER`, then `ei1` is understood to occur after `ei2`.
- But if this relation is `BEFORE`, `ei1` is understood to have occurred before `ei2`.

### E.6.3.3 Relative tense

#### E.6.3.3.1 Adnominal endings

Adnominal verb endings and adnominal adjective endings both carry temporal information. Verbs and adjectives, however, have different sets of adnominal endings. Here are the adnominal endings for verbs:

#### E.6.3.3.2 Adnominal endings for verbs

**Present** : -는 *-nun*

**Past** : -ㄴ *-n*, -은 *-un*

**Future** : -ㄹ *-l*, -을 *-ul*

**PASTdurative** : -던 (*-ten*)

**PASTperfective** : -었던 (*-ess.ten*)

Here are examples:

```
(20) 미아가 데이트하는 남자를 순이가 사랑했다
     mia-ka tey.i.tu.ha-nun nam.ca-lul
                              swuni-ka sa.lang.ha-yss-ta
     mia-NOM date-presADNOM male-ACC
                              swuni-NOM love-PAST-DECL
     "Swuni loved a man Mia dates/dated"
(21) 미아가 데이트할 남자를 순이가 사랑했다
     mia-ka tey.i.tu.ha-l nam.ca-lul
                            swuni-ka sa.lang.ha-yss-ta
     mia-NOM date-futADNOM male-ACC
                            swuni-NOM love-PAST-DECL
     "Swuni loved a man Mia will/would date"
(22) 미아가 데이트한 남자를 순이가 사랑했다
     mia-ka tey.i.tu.ha-n nam.ca-lul
                            swuni-ka sa.lang.ha-yss-ta
     mia-NOM date-pastADNOM male-ACC swuni-NOM love-PAST-DECL
     "Swuni loved a man Mia dated"
(23) 미아가 데이트하던 남자를 순이가 사랑했다
     mia-ka tey.i.tu.ha-ten nam.ca-lul
                              swuni-ka sa.lang.ha-yss-ta
     mia-NOM date-pastDurADNOM male-ACC
                              swuni-NOM love-PAST-DECL
     "Swuni loved a man Mia used to date"
(24) 미아가 데이트하였던 남자를 순이가 사랑했다
     mia-ka tey.i.tu.ha-yess-ten nam.ca-lul
                                swuni-ka sa.lang.ha-yss-ta
     mia-NOM date-pastPerfADNOM male-ACC
                                swuni-NOM love-PAST-DECL
     "Swuni loved a man Mia had dated"
```

Adjectives have a different set of adnominal endings:

#### E.6.3.3.3 Adnominal endings for adjectives

**Present** : -ㄴ *-n*, -은 *-un*

**Future** : -ㄹ *-l*, -을 *-ul*

**PASTdurative** : -던 (*-ten*)

**PASTperfective** : -었던 *-ess.ten*

Examples are as follows:

```
(25) 미아가 돈많은 남자를 사랑하였다
     mia-ka tonmanh-un namca-lul salangha-yess-ta
     mia-NOM money-much-presADNOM man-ACC love-PAST-DECL
     "Mia loved a man who had a lot of money""
(26) 미아가 돈많을 남자를 사랑하였다
     mia-ka tonmanh-ul namca-lul salangha-yess-ta
     mia-NOM money-much-futADNOM man-ACC love-PAST-DECL
     "Mia loved a man who will/would have a lot of money""
(27) 미아가 돈많던 남자를 사랑하였다
     mia-ka tonmanh-ten namca-lul salangha-yess-ta
```

```
     mia-NOM money-much-pastADNOM man-ACC love-PAST-DECL
     "Mia loved a man who used to have a lot of money""
(28) 미아가 돈이 많았던 남자를 사랑하였다
     mia-ka tonmanh-ass-ten namca-lul salangha-yess-ta
     mia-NOM money-much-pastPerfADNOM man-ACC love-PAST-DECL
     "Mia loved a man who had had a lot of money""
```

#### E.6.3.3.4 Annotation of relative tense

```
(29) 미아가
     <EVENT eid="e1" eiid="ei1"
     yaleRomanization="sa.lang.ha.nun" content="LOVE"
     class="STATE" pos="VERB" tense="PRESENT"
     vForm="ADNOMINAL">
     사랑하는
     </EVENT>
     남자를 순이가
     <EVENT eid="e2" eiid="ei2"
     yaleRomanization="sa.lang.ha.yess.ta" content="LOVE"
     class="STATE" pos="VERB" tense="PAST" vForm="sFINAL">
     사랑하였다
     </EVENT>
     <TLINK eventInstanceID="ei1" relatedToEventInstance="ei2"
     relType="SIMULTANEOUS"/>
(30) 미아가
     <EVENT eid="e1" eiid="ei1"
     yaleRomanization="sa.lang.han" content="LOVE"
     class="STATE" pos="VERB" tense="FUTURE" vForm="ADNOMINAL">
     사랑할
     </EVENT>
     남자를 순이가
     <EVENT eid="e2" eiid="ei2"
     yaleRomanization="sa.lang.ha.yess.ta" content="LOVE"
     class="STATE" pos="VERB" tense="PAST" vForm="sFINAL">
     사랑하였다
     </EVENT>
     <TLINK eventInstanceID="ei1" relatedToEventInstance="ei2"
     relType="AFTER"/>
(31) 미아가
     <EVENT eid="e1" eiid="ei1"
     yaleRomanization="sa.lang.han" content="LOVE"
     class="STATE" pos="VERB" tense="PAST" vForm="ADNOMINAL">
     사랑한
     </EVENT>
     남자를 순이가
     <EVENT eid="e2" eiid="ei2"
     yaleRomanization="sa.lang.ha.yess.ta" content="LOVE"
     class="STATE" pos="VERB" tense="PAST" vForm="sFINAL">
     사랑하였다
     </EVENT>
     <TLINK eventInstanceID="ei1" relatedToEventInstance="ei2"
     relType="BEFORE"/>
(32) 미아가
     <EVENT eid="e1" eiid="ei1"
     yaleRomanization="sa.lang.ha.ten" content="LOVE(mia,x)"
     class="STATE" pos="VERB" tense="PAST" aspect="DURATIVE"
     vForm="ADNOMINAL">
```

```
     사랑하던
     </EVENT>
     남자를 순이가
     <EVENT eid="e2" eiid="ei2"
     yaleRomanization="sa.lang.ha.yess.ta" content="LOVE(swuni,x)"
     class="STATE" pos="VERB"
     tense="PAST" vForm="sFINAL">
     사랑하였다
     </EVENT>
     <TLINK eventInstanceID="ei1" relatedToEventInstance="ei2"
     relType="BEFORE"/>
```

NOTE As was stated earlier, the attribute `content` may have a predicate-logic-type logical form like "LOVE(mia,yong)" as in (32).

Sentences with adnominal adjectives can also be annotated in a similar manner.

```
(33) 미아가
     <EVENT eid="e1" eiid="ei1"
     yaleRomanization="ton.manh.un" content="RICH"
     class="STATE" pos="ADJECTIVE"
     tense="PRESENT" vForm="ADNOMINAL">
     돈많은
     </EVENT>
     남자를
     <EVENT eid="e2" eiid="ei2"
     yaleRomanization="sa.lang.ha.yess.ta" content="LOVE"
     class="STATE" pos="VERB" tense="PAST"
     vForm="DECLARATIVE">
     사랑하였다
     </EVENT>
     <TLINK eventInstanceID="ei1" relatedToEventInstance="ei2"
     relType="SIMULTANEOUS"/>
(34) 미아가
     <EVENT eid="e1" eiid="ei1"
     yaleRomanization="ton.manh.ul" content="RICH"
     class="STATE" pos="ADJECTIVE" tense="FUTURE"
     vForm="ADNOMINAL">
     돈많을
     </EVENT>
     남자를
     <EVENT eid="e2" eiid="ei2"
     yaleRomanization="sa.lang.ha.yess.ta" content="LOVE"
     class="STATE" pos="VERB" tense="PAST"
     vForm="sFINAL">
     사랑하였다
     </EVENT>
     <TLINK eventInstanceID="ei1" relatedToEventInstance="ei2"
     relType="SIMULTANEOUS"/>
(35) 미아가
     <EVENT eid="e1" eiid="ei1"
     yaleRomanization="ton.manh.ten" content="RICH"
     class="STATE" pos="ADJECTIVE" tense="PAST"
     vForm="ADNOMINAL">
     돈많던
     </EVENT>
```

```
     남자를
     <EVENT eid="e2" eiid="ei2"
     yaleRomanization="sa.lang.ha.yess.ta" content="LOVE"
     class="STATE" pos="VERB" tense="PAST"
     vForm="sFINAL">
     사랑하였다
     </EVENT>
     <TLINK eventInstanceID="ei1" relatedToEventInstance="ei2"
     relType="BEOFRE"/>
(36) 미아가
     <EVENT eid="e1" eiid="ei1"
     yaleRomanization="ton.manh.ass.ten" content="RICH"
     class="STATE" pos="ADJECTIVE" tense="PAST"
     aspect="PERFECTIVE" vForm="ADNOMINAL">
     돈많았던
     </EVENT>
     남자를
     <EVENT eid="e2" eiid="ei2"
     yaleRomanization="sa.lang.ha.yess.ta" content="LOVE"
     class="STATE" pos="VERB" tense="PAST"
     vForm="sFINAL">
     사랑하였다
     </EVENT>
     <TLINK eventInstanceID="ei1" relatedToEventInstance="ei2"
     relType="BEFORE"/>
```

If a predicate is adnominal, then its tense may also be relativized with respect to the tense of the main clause or the clause in which the adnominoal or relativized section is embedded. For this case, a relevant `TLINK` is introduced. Otherwise, the tense of an adnominal clause carries its given value, being interpreted in an absolute sense, such that it is anchored to the designated time of origin or utterance.

Consider sentence (20), which is repeated here:

```
(37=20) 미아가 데이트하는 남자를 순이가 사랑했다
        mia-ka tey.i.tu.ha-nun nam.ca-lul
                                swuni-ka sa.lang.ha-yss-ta
        mia-NOM date-presADNOM male-ACC
                                swuni-NOM love-PAST-DECL
        "Swuni loved a man Mia dates/dated"
```

Here the time of Mia's dating can be anchored to the present time, instead of taking place in the past. This becomes clearer, if the temporal adverb 지금 *cikum* 'now' is added to it or make the main verb carry `PAST PERFECTIVE`.

```
(38) 미아가 지금 데이트하는 남자를 순이가 사랑했다
     mia-ka cikum tey.i.tu.ha-nun nam.ca-lul
                                swuni-ka sa.lang.ha-yss-ta
     mia-NOM now date-presADNOM male-ACC
                                swuni-NOM love-PAST-DECL
     "Swuni loved a man Mia dates now/is dating now"
(39) 미아가 데이트하는 남자를 순이가 사랑했었다
     mia-ka tey.i.tu.ha-nun nam.ca-lul
                                swuni-ka sa.lang.ha-yss.ess-ta
     mia-NOM date-presADNOM male-ACC
                                swuni-NOM love-pastPerfective-DECL
     "Swuni loved a man Mia dates/is dating"
```

Because of such an ambiguity of adnominal tense, TLINK is optional.

#### E.6.3.3.5 Annotation guidelines

Given two events, e1 and e2, such that their respective event instances are identified with eiid=ei1 and eiid=ei2:
if an EVENT element associated with e1 and ei1 contains vForm="ADNOMINAL",
then ei1 may be related to ei2 by TLINK such that the tense of ei1 is relativized with respect to that of ei2 as below:

- if the tense of ei1 is PRESENT, then relType="SIMULATANEOUS",
- if the tense of ei1 is FUTURE, then relType="AFTER", and
- if the tense of ei1 is PAST, then relType="BEFORE".

#### E.6.3.3.6 Interpretation conditions

The conditions for interpreting the above annotations need not be stated here, for they are provided by the general guidelines.

### E.6.3.4 Inherited tense

Verbal expressions with a CONJUNCTIVE ending like *-ko* often lack a tense marker, but inherit the tense value from the main verbal expression at the end of a sentence. Here is an example:

```
(40) 어제 밤 사과를 먹고 차를 마시고 잤다
     ecey pam sakwa-lul mek-ko  cha-ul     masi-ko   ca-ss-ta
    apple-ACC eat-CONJ   tea-ACC drink-CONJ sleep-PAST-DECL
    "ate an apple, drank tea, and slept"
```

Here, neither the verb 먹고 *mek-ko* 'eat-CONJ' nor the verb 마시고 *masi-ko* 'drink-CONJ' is marked with a tense, but inherits the tense information from the sentence final verb 잤다 *ca-ss-ta*, marked with PAST tense. Through this inheritance process, these CONJUNCTIVE tenseless verbs can refer to the PAST events of eating and drinking that occurred before the event of sleeping in the past.

Consider anothe case of tense inheritance. The sentence-non-final verbal ending 다가 *-taka* may occurr with the PASTperfective marker -었었 *-ess.ess*. Here are examples:

```
(41) 미아가 미국에     가다가  서울로     왔다
     mia-ka mikwuk-ey ka-taka seowul-lo wa-ss-ta
     mia-NOM US-GOAL   go-CONJ Seoul-DIR come-PAST-DECL
     'While going to the States Mia came back to Seoul'
(42) 미아가  미국에     갔다가     서울로     왔다
     mia-NOM mikwuk-ey ka-ss-taka seowul-lo wa-ss-ta
     US-GOAL   go-PERFECTIVE-CONJ
                         Seoul-DIR come-PAST-DECL
     'Having been to the States Mia came back to Seoul'
```

Both of the events of one's going to the States took place before his coming back to Seoul. But in (41) his going to the States was incomplete, while it was completed in (42).

#### E.6.3.4.1 Annotation of inherited tense information

```
(43) <TIMEX3 tid="t1"
     yaleRomanization="ecey pam" content="LAST NIGHT"
     type="TIME" value="2007-03-31TNT"
     temporalFunction="TRUE" anchorTime="t0"
     comment="TNI stands for 'night time'">
     어제 밤
     </TIMEX3>
     사과를
     <EVENT eid="e1" eiid="ei1"
     yaleRomanization="mek-ko" content="EAT"
     class="OCCURRENCE" pos="VERB"
     tense="NONE" vForm="CONJ">
     먹고
     </EVENT>
     차를
     <EVENT eid="e2" eiid="ei2"
     yaleRomanizaiton="masi-ko" content="DRINK"
     class="OCCURRENCE" pos="VERB"
     tense="NONE" vForm="CONJ">
     마시고
     </EVENT>
     <EVENT eid="e3" eii="ei3"
     yaleRomanization="ca-ss-ta" content="SLEEP"
     class="OCCURRENCE" pos="VERB"
     tense="PAST" vForm="sFINAL">
     잤다
     </EVENT>.
     <TLINK eventInstandID="ei1" relatedToEventInstance="ei3"
     relType="IBEFORE"/>
     <TLINK eventInstandID="ei2" relatedToEventInstance="ei3"
     relType="IBEFORE"/>
     <TLINK eventInstandID="ei3" relatedToTime="t1"
     relType="INCLUDED"/>
(44) 미아가
     <TIMEX3 tid="t1"
     yaleRomanization="ecey" content="YESTERDAY"
     type="TIME" value="2007-05-05"
     temporalFunction="TRUE" anchorTime="t0">
     어제
     </TIMEX3>
     미국에
     <EVENT eid="e1" eiid="ei1"
     yaleRomanizaiton="ka-taka" content="GO"
     class="OCCURRENCE" pos="VERB"
     tense="NONE" vForm="CONJ">
     가다가
     </EVENT>
     서울로
     <EVENT eid="e2" eiid="ei2"
     yaleRomanizaiton="oass-ta" content="COME"
     class="OCCURRENCE" pos="VERB"
     tense="PAST" vForm="sFINAL">
     왔다
     </EVENT>
     <TLINK eventInstandID="ei1" relatedToEventInstance="ei2"
```

```
     relType="IBEFORE"/>
     <TLINK eventInstandID="ei2" relatedToTime="t1"
     relType="INCLUDED"/>
(45) 미아가
     <TIMEX3 tid="t1"
     yaleRomanization="ecey" content="YESTERDAY"
     type="TIME" value="2007-05-05"
     temporalFunction="TRUE" anchorTime="t0">
     어제
     </TIMEX3>
     미국에
     <EVENT eid="e1" eiid="ei1"
     yaleRomanizaiton="ka-taka" content="GO"
     class="OCCURRENCE" pos="VERB"
     tense="PAST" vForm="-taka,WHILE">
     갔다가
     </EVENT>
     서울로
     <EVENT eid="e2" eiid="ei2"
     yaleRomanizaiton="oass-ta" content="COME"
     class="OCCURRENCE" pos="VERB"
     tense="PAST" vForm="sFINAL">
     왔다
     </EVENT>
     <TLINK eventInstandID="ei1" relatedToEventInstance="ei2"
     relType="IBEFORE"/>
     <TLINK eventInstandID="ei2" relatedToTime="t1"
     relType="INCLUDED"/>
```

#### E.6.3.4.2 Interpretation of `PAST` in `vForm="-taka, WHILE"`

- If an EVENT element contains `vForm="-taka, WHILE"` and also `PAST`, then it is interpreted as referring to an event that has been *completed* or *accomplished*.

### E.6.3.5 Annotation of tense in conditonal sections

Tense `PAST` in a conditional section may not refer to an event in the past. Consider:

```
(46) 내일 미아가 오면 좋겠다
     nayil mia-ka o-meyn coh-keyss-ta
     tomorrow Mia-NOM come-COND nice-CONJEC-DECL
     'It would be nice if Mia comes tomorrow'
(47) 내일 미아가 왔으면 좋겠다
     nayil mia-ka o-ass-umeyn coh-keyss-ta
     tomorrow Mia-NOM come-PAST-COND nice-CONJEC-DECL
     'It would be nice if Mia comes/would come tomorrow'
(48) 어제 미아가 왔으면 좋았겠다
     ecey mia-ka o-ass-umeyn coh-keyss-ta
     yesterday Mia-NOM come-PAST-COND nice-PAST-CONJEC-DECL
     'It would have been nice if Mia had come yesterday'
(49) 어제 미아가 왔더라면 좋았겠다
     ecey mia-ka o-ass-te-ra-meyn coh-keyss-ta
     yesterday Mia-NOM come-PAST-RETRO-IND-COND nice-PAST-CONJEC-DECL
     'It would have been nice if Mia had come yesterday'
```

These sentences can be annotated in a straightforward way as below:

```
(50) 내일(tomorrow) 미아가
    <EVENT eid="e1" eiid="ei1"
    yaleRomanization="o.myen" content="COME"
    class="OCCURRENCE" pos="VERB" vForm="CONDITIONAL"
    tense="NONE">
    오면
    </EVENT>
    <EVENT eid="e2" eiid="ei2"
    yaleRomanization="coh.keyss.ta" content="NICE"
    class="I_STATE" pos="ADJECtIVE" vForm="sFIANAL"
    tense="NONE" modality="CONJECTURAL">
    좋겠다
    </EVENT>
    <TLINK eventInstandID="ei1" relatedToEventInstance="ei2"
    relType="SIMULATANEOUS">
    </TLINK>
(51) 내일(tomorrow) 미아가
    <EVENT eid="e1" eiid="ei1"
    yaleRomanization="oass.u.myen" content="COME"
    class="OCCURRENCE" pos="VERB" vForm="CONDITIONAL"
    tense="PAST">
    왔으면
    </EVENT>
    <EVENT eid="e2" eiid="ei2"
    yaleRomanization="coh.keyss.ta" content="NICE"
    class="I_STATE" pos="ADJECtIVE" vForm="sFIANAL"
    tense="PAST" modality="CONJECTURAL">
    좋았겠다
    </EVENT>
    <TLINK eventInstandID="ei1" relatedToEventInstance="ei2"
    relType="SIMULATANEOUS">
(52) 어제(yesterday) 미아가
    <EVENT eid="e1" eiid="ei1"
    yaleRomanization="oass.te.ra.myen" content="COME"
    class="OCCURRENCE" pos="VERB" vForm="CONDITIONAL"
    tense="PAST" mood="RETROSPECTIVE">
    왔더라면
    </EVENT>
    <EVENT eid="e2" eiid="ei2"
    yaleRomanization="coh.keyss.ta" content="NICE"
    class="I_STATE" pos="ADJECtIVE" vForm="sFIANAL"
    tense="PAST" modality="CONJECTURAL">
    좋았겠다
    </EVENT>
    <TLINK eventInstandID="ei1" relatedToEventInstance="ei2"
    relType="SIMULATANEOUS"/>
```

## E.7 Aspect

### E.7.1 Aspect markers

In Korean, aspect PROGRESSIVE is expressed by complex verbal structures such as follows:

- -고 있(STEM-ko iss) mostly for OCCURRENCE class verbs
- -어 있(STEM-e iss) for STATE class verbs

Here are examples:

```
(53) 사과를 먹고 있었다
    sakwa-lul mek-ko iss-ess-ta
    apple-ACC eat-BR auxVerbPROG-PAST-DECL
    'was/were eating an apple'
(54) 하루 종일  앉아 있었다
    haru cong-il sit-BR auxVerbDURA-PAST-DECL
    'was/were sitting all day'
```

NOTES
1. Unlike English, stative verbs like 가지- *kaci-* 'have/own' and 알- *al-* 'know' may form the PROGRESSIVE construction: 가지고 있다 *kaci-ko iss-ta* and 알고 있다 *al-ko iss-ta*.
2. Unlike the *-ko iss-ta* construction, the *-e iss-ta* construction is very restricted. Only few intransitive verbs like 눕/누우- *nup-/nuw-* 'lie', 앉- *anc-* 'sit', and 서- *se-* 'stand' or passivized verbs with the ending *-ci* like 알려지- *allyeci-* 'be known' take up this form.

### E.7.2 Annotation of aspect markers

The annotation of aspect features is straighforward. The complex aspectual constructions are treated as single chunks without separating the main STEM part from the auxiliary part. Here are examples:

Examples

```
(55) 사과를
     <EVENT eid="e46" eiid="ei46"
     yaleRomanization="mek-ko iss-ess-ta" content="EAT"
     class="OCCURRENCE" pos="VERB" tense="PAST"
     aspect="PROGRESSIVE" vForm="sFINAL">
     먹고 있었다
     </EVENT>
(56) 어제(yesterday) 하루 종일(all day)
     <EVENT eid="e47" eiid="ei47"
     yaleRomanization="anc-a iss-ess-ta" content="SIT"
     class="STATE" pos="VERB" tense="PAST"
     aspect="PROGRESSIVE" vForm="sFINAL">
     앉아 있었다
     </EVENT>
(57) 미아는 돈을 많이
     <EVENT eid="e48" eiid="ei48"
     yaleRomanization="kaci-ko iss-ta" content="HAVE/OWN"
     class="STATE" pos="VERB" tense="NONE"
     aspect="PROGRESSIVE" vForm="sFINAL">
     가지고 있다
     </EVENT>
(58) 지금(now) 미아가 미국에
     <EVENT eid="e14"
     yaleRomanization="ka iss-ta" content="GO"
     class="STATE" pos="VERB" tense="NONE"
     aspect="PROGRESSIVE">
     가 있다
     </EVENT>
```

Sentence (57) means 'Mia has a lot of money' and sentence (58) means 'Mia has gone to US, and she is there now'.

### E.7.3 Interpretation of aspectual features

The concept of aspect involves the event ontology, the types and structure of events. ISO-TimeML introduces seven classes or types of events:

```
class ::='OCCURRENCE' | 'PERCEPTION' | 'REPORTING' |
         'ASPECTUAL' | 'STATE' | 'I_STATE' | 'I_ACTION'
```

The structure of an event may be segmented into the initial, medial, and terminal part. Depending on the perspective of a language user or interpreter, one's focus shifts from the initial to the medial (on-going) or to the terminal (completed) part. Suppose someone says 'Mia is eating an apple'. This statement is understood as focusing on the medial or on-going part of the event of Mia's eating an apple. On the other hand, if one says 'Mia is wearing a beautiful hat', its focus is on the terminal or resulting part of the event of Mia's putting on a hat. This statement is interpreted as restulting STATE rather than as an OCCURRENCE of activity.

Sentence (59) is ambiguous:

```
(59) 미아가  옷을         입고      있다 (OCCURRENDCE or STATE)
     mia-ka  os-ul        ip-ko     iss-ta
     Mia-NOM clothes-ACC wear-COMP auxVerbPROG-DECL
     'Mia is wearing clothes'
(60) 미아가 지금 샤워를 하고 옷을 입고 있다 (OCCURRENCE)
     mia-ka cikum syawe-lul ha-ko os-ul ip-ko iss-ta
     Mia-NOM now  shower-ACC do-AND clothes wear-COMP auxVerbPROG-DECL
     'Mia took a shower and is now putting on clothes'
(61) 미아가 오늘은 빨간 옷을 입고 있다 (STATE)
     mia-ka onul-un ppalkan os-ul ip-ko iss-ta
     Mia-NOM today-TOPIC red clothes-ACC wear-COMP auxVerbPROG-DECL
     'Today Mia is dressed in red'
```

The PROGRESSIVE form of the verbal expression in sentence (59) can be interpreted in two ways: either in the on-going OCCURRENCE sense as in (60) or in the resulting STATE sense as in (61). Sentence (59) is thus annotated differently as shown below:

```
(62) 미아가 옷을
     <EVENT eid="e1" eiid="ei1"
     yaleRomanization="ip-ko iss-ta" content="WEAR"
     class="OCCURRENCE" pos="VERB" tense="NONE"
     aspect="PROGRESSIVE" vForm="sFINAL">
     입고 있다
     </EVENT>
(63) 미아가 옷을
     <EVENT eid="e2" eiid="ei2"
     yaleRomanization="ip-ko iss-ta" content="WEAR"
     class="STATE" pos="VERB" tense="NONE"
     aspect="PROGRESSIVE" vForm="sFINAL">
     입고 있다
     </EVENT>
```

The basic difference between these two is that (62) refers to an event instance of class OCCURRENCE, while (63) refers to an event instance of class STATE. In order to make this distinction clear, PROGRESSIVE STATE may be annotated as class="STATE" aspect="DURATIVE".

The *verbalSTEM-e iss-ta* construction constrasts with a particular set of verbs with the PAST verbal form like `ip-` 'wear' and `cuk-` 'die'. Consider:

```
(64) 미아가 예쁜 옷을 입었다
     mia-ka yeyppun ot-ul ip-ess-ta
     Mia-NOM pretty clothes-ACC wear-PAST-DECL
     'Mia wore/is wearing pretty clothes'
(65) 새가 죽었다
     say-ka cuwk-ess-ta
     bird-NOM die-PAST-DECL
     '(the) bird died/is dead'
(66) 미아는 돈을 많이 가졌다
     mia-nun ton-ul manhi kacy-ess-ta
     Mia-TOP money-ACC much have-PAST-DECL
     'Mia had/has much money'
(67) 봄이 왔다
     pom-i wa-ss-ta
     spring-NOM come-PAST-DECL
     'Spring came/has come'
```

Each of the verbs in these examples is marked with the PAST tense ending. But the examples above each refer to a STATE at the present, the terminal phase of some event that may have occurred or started in the past . Sentence (64), for instance, means that Mia has put on pretty clothes and is still wearing them. Sentence (65) means that a bird died and is dead now, sentence (66) that Mia (has earned money and) owns a lot of money now, and sentence (67) that spring came and that it is spring now.

#### E.7.3.0.1 Annotation of `RESULTATIVE`

By introducing a new aspect feature `RESULTATIVE` these examples with tense `PAST` can be properly annotated, as illustrated below:

```
(68) 새가 죽었다
     `(The) bird died'

    새가
    <EVENT eid="e1" eiid="ei1"
    yaleRomanization="cuk-ess-ta" content="DIE"
    class="STATE" pos="VERB" aspect="RESULTATIVE">
    죽었다
    </EVENT>
```

### E.7.3.1 Interpretation conditions of aspect

`PROGRESSIVE` **activity** If an `EVENT` element is specified with `aspect="PROGRESSIVE"` and `class ="OCCURRENCE"`, then that event is interpreted as an activity being continued at some interval of time.

`PROGRESSIVE/DURATIVE` **state** If an `EVENT` element is specified with `aspect="PROGRESSIVE"` and `class= "STATE"`, then that event is interpreted as a state being uniformly retained at some interval of time. Note that this state may be annotated as `aspect="DURATIVE"`.

`PRESENT PROGRESSIVE` If an `EVENT` element is specified with `aspect="PROGRESSIVE"` and `tense="NONE"`, then that event is interpreted as taking place in the neighborhood of the present moment of time, namely*N(n)*.

`PAST PROGRESSIVE` If an `EVENT` element is specified with `aspect="PROGRESSIVE"` and `tense="PAST"`, then that event is interpreted as taking place in the neighborhood of some moment of time *t* in the past interval of time, *N(t)* which is a subinterval of the past time interval.

PAST RESULTATIVE If an EVENT element is specified with aspect="RESULTATIVE" and tense="PAST", then an event associated with that element is interpreted as an event the initial OCCURRENCE of which might have completed in the past, but with the resulting state retained in the neighborhood of the present moment of time, namely*N(n)*.

NOTE PAST RESULTATIVE may be understood as equivalent to PRESENT DURATIVE.

## E.8 Modality

### E.8.1 Conjectural modal markers

The CONJECTURAL modality is expressed by the verbal ending 겠 *-keyss* or -ㄹ/을 것 *-l/ul kes*. It is claimed by K. Lee (1998) that these two differ in the degree of certainty expressed by each. But here this difference is ignored and only the ending 겠 *-keyss* is illustrated. Here are examples:

(69) 지금 부산에 비가 오고 있겠다
cikum pusan-ey rain-ka o-ko iss-keyss-ta
now Busan-LOC rain-NOM come-COMP PROG-CONJEC-DECL
'It must be raining in Busan now'
(70) 미아는 어제 떠났겠다
mia-nun ecey ttena-ss-keyss-ta
Mia-TOP yesterdat leave-PAST-CONJEC-DECL
'Mia must have left yesterday'
(71) 어릴 때에 미아는 예뻤겠다
eri-l ttay-ey mia-ka yeyppe-ss-keyss-ta
young-ADNOM time-LOC Mia-NOM pretty-PAST-CONJEC-DECL
'When she was young, Mia must have been pretty'

When used with the first person Subject, the CONJECTURAL 겠 *-keyss* often expresses her or his intention, as illustrated by:

(72) 나는 내일 떠나겠다
na-nun nayil ttena-keyss-ta
I-TOP tomorrow leave-CONJEC-DECL
'I will leave tomorrow'

But note that the following examples do not express the speaker's intention:

(73) 나는 지금 아파 죽겠다
na-nun cikum aph-a cwuk-keyss-ta
I-TOP now sick-CONJUNCTIVE die-CONJEC-DECL
'I'm now sick and so must die'
(74) 이번에는 내가 복권에 당첨되겠다
ipen-ey-nun nay-ka pokkwon-ey tangchemtoy-keyss-ta
this time-LOC-TOP I-NOM lottery win-CONJEC-DECL
'This time I must win/be winning a lottery'

### E.8.2 Annotation of modality **CONJECTURAL**

The annotation of the CONJECTURAL modality is again very simple.

(75) 미아는 어제(yesterday)

```
     <EVENT eid="e65" eid="ei65"
     yaleRomanization="ttena-ss-keyss-ta" content="LEAVE"
     class="OCCURRENCE" pos="VERB" tense="PAST"
     modality="CONJECTURAL" vForm="sFINAL">
     떠났겠다
     </EVENT>
(76) 부산에는 지금 비가
     <EVENT eid="e66" eid="ei66"
     yaleRomanization="o-ko iss-ul-kes-ita" content="COME"
     class="OCCURRENCE" pos="VERB" tense="NONE"
     modality="CONJECTURAL" vForm="sFINAL">
     오고 있을 것이다
     </EVENT>
```

#### E.8.3 Interpretation of modality **CONJECTURAL**

- If the modality of an EVENT element is specified with the value CONJECTURAL, then it is understood as expressing the speaker's conjectural certainty of the occurrence of that event referred to.
- Furthermore, if the event referred to is controllable by the speaker himself, then it is understood as expressing the speaker's intention to make that event realized.

### E.9 Mood

#### E.9.1 Mood markers

The verbal ending 더 *-te* is treated as the RETROSPECTIVE mood marker. Here are examples:

```
(77) 베트남은 참 덥더라
     peytunam-un cham tep-te-la
     Vietnam-TOP really hot-RETRO-DECL
     'Vietnam was really hot, I recall'
(78) 전쟁 때 베트남 사람들의 삶이 비참하였겠더라
     cencayng ttay peytunam saramtul-uy salm-i pichamha-yess-keyss-te-la
     war time Vienam peopple-GEN life-NOM terrible-PAST-CONJEC-RETOR-DECL
     'In the war Vietnamese people's life must have been terrible, I think'
(79) 십년 후엔 베트남이 잘 살겠더라
     sipnyen hwu-ey-n peytunam-i cal sal-keyss-te-la
     10 years after Vietnam-NOM well live-CONJEC-RETRO-DECL
     'Ten years later Vietnam must live well, I think'
(80) 점을 쳤더니 미아가 내년에 결혼을 하더라
     cem-ul chy-ess-te-ni mia-ka nayneyn-ey kyelhon-ul ha-te-la
     fortune tell-RETRO-CONN Mia-NOM next year marriage do-RETRO-DECL
     'As I saw the fortune telling, Mia will be marrying next year'
(81) 일기 예보를 들으니 내일 비가 오겠던데
     ilki yeypo-lum tul-u-ni nayil pi-ka o-keyss-te-nte
     weather forecast-ACC hear tomorrow rain-NOM come-CONJEC-RETRO-DECL
     'As I listened to the weather forecast, it must be raining tomorrow, I think'
```

#### E.9.2 Annotation of mood **RETROSPECTIVE**

Verbal expressions with the verbal ending 더 *-te* can be automatically marked with mood RETROSPECTIVE.

```
(82) 베트남은 참
```

```
     <EVENT eid="e72" eid="ei72"
     yaleRomanization="tep-te-la" content="HOT"
     class="STATE" pos="ADJECTIVE" tense="NONE"
     mood="RETROSPECTIVE">
     덥더라
     </EVENT>
(83) 전쟁 때(at the time of the war) 베트남 사람들의 삶이
     <EVENT eid="e73" eiid="ei73"
     yaleRomanization="pichamha-yess-keyss-te-la" content="TERRIBLE"
     class="STATE" pos="ADJECTIVE" tense="PAST" modality="CONJECTURAL"
     mood="RETROSPECTIVE">
     비참하였겠더라
     </EVENT>
(84) 10년 후엔(10 years later) 베트남이 잘
     <EVENT eid="e74" eid="ei74"
     yaleRomanization="sal-keyss-te-la" content="LIVE"
     class="STATE" pos="VERB" tense="NONE" modality="CONJECTURAL"
     mood="RETROSPECTIVE">
     살겠더라
     </EVENT>
```

NOTE Chang (1996) treats the verbal ending 겠 *-keyss* as one of the mood markers, namely the VOLITIONAL mood. But, since it can co-occur with the RETROSPECTIVE marker 더 *-te* as shown above, thus illicitly duplicating the value of the attribute *mood*, it is treated here in this annex as a modality marker.

### E.9.3 Interpretation of mood **RETROSPECTIVE**

Chang (1996: 131) states: “The retrospective mood denotes the speaker’s experience or observation in retrospect." Sohn (1999: 359) also expresses a similar view: "Retrospective mood denotes a situation in which someone recalls a fact or an event he witnessed or experienced, and thus has meanings such as ‘I saw, observed, experienced’ in declaratives and ‘did you see, oberserve, experience’ in interrogatives." To formalize this notion of experience or observation in situation-theoretic terms, K. Lee (1993) introduces the notion of observational accessibility. The retrospective mood is used when the event referred to is observationally accessible to the speaker-observer.

- If the mood of an EVENT element is specified with the value RETROSPECTIVE, then the event referred to by that EVENT element is interpreted as implying that the event is/was within the (spatio-temporal) location observationally accessible to the speaker.

## E.10 BNF values for <EVENT> in Korean

```
attributes ::= eid eiid yaleRomanization content class pos tense [aspect]
[polarity] [modality] [mood] [vForm]
eid ::= e<integer>
eiid ::=ei<integer>
yaleRomanization ::= CDATA
content ::= CDATA
class ::= 'REPORTING'|'PERCEPTION'|'ASPECTUAL'|'I_ACTION'|'I_STATE' |
      'STATE'|'OCCURRENCE'
pos ::= 'ADJECTIVE'|'NOUN'|'VERB'|'NONE'
tense ::= 'FUTURE'|'PAST'|'PRESENT'|'NONE' {default, if absent, is 'NONE'}
aspect ::= ('DURATIVE')|'PROGRESSIVE'|'PERFECTIVE'|'RESULTATIVE'|'NONE'
             {default, if absent, is 'NONE'}
polarity ::= 'NEG'|'POS' {default, if absent, is 'POS'}
mood ::=  'RETROSPECTIVE'|'NONE' {default, if absent, is 'NONE'}
```

```
modality ::= 'CONJECTURAL'|'INDICATIVE' {default, if absent, is 'INDICATIVE'}
vForm ::= 'COMP'|'CONJUNCTIVE'|'CONDITIONAL'|'sFINAL'|CDATA
                {default, if absent, is 'sFINAL'}
```

### E.11 Summary

The temporal annotation of verbal endings in Korean can be summarized as in the following table:

Table: Verbal temporal endings

| **ending** | **class** | **tense** | **aspect** | **modality** | **mood** | **example** |
|---|---|---|---|---|---|---|
| nun | OCCUR | | | | IND | mek-nun-ta (eats) |
| ess-supni | OCCUR | PAST | | | IND | mek-ess-sup-ni-ta (ate, polite) |
| ess | OCCUR | PAST | | | | mek-ess-ta (ate) |
| ess | STATE | PAST | RESULT | | | kaci-ess-ta (has) |
| ess.ess | OCCUR | PAST | PERF | | | mek-ess-ess-ta (had eaten) |
| keyss | OCCUR | | | CONJEC | | mek-keyss-ta (may eat) |
| ess-keyss | OCCUR | PAST | | CONJEC | | mek-ess-keyss-ta (might have eaten) |
| te | OCCUR | | | | RETRO | mek-te-la (I recall, ate) |
| ess-te | OCCUR | PAST | | | RETRO | (I recall, had eaten) |
| keyss-te | OCCUR | | | CONJEC | RETRO | (I recall, may have eaten) |
| ko iss | OCCUR | | PROG | | | mek-ko iss-ta (is eating) |
| ko iss | STATE | | PROG | | | al-ko iss-ta (know) |
| ko iss-ess-keyss | OCCUR | PAST | PROG | CONJEC | | mek-ko iss-ess-keyss-ta (must have been eating) |
| e iss | STATE | | PROG(DURA) | | | nwu-e iss-ta (is lying) |
| e iss-keyss-te | STATE | | PROG(DURA) | CONJEC | RETRO | nwu-e iss-keyss-te-la (must be lying, I recall) |

Table: Verbal temporal endings

# Annex F
(Arabic Sample For Testing)
# Arabic Sample For Testing

جُحَا وَ حِمَارُهُ *ǧuḥā wa-ḥimāruhu*

*ʾatā ṣadīqun ʾilā ǧuḥā yaṭlubu minhu ḥimārahu li-yarkabahu fī safratin qaṣīratin wa-qāla lahu: sawfa ʾuʿīduhu ʾilayka fī ʾl-masāʾi, wa-ʾadfaʿu laka ʾuǧratan. fa-qāla ǧuḥā:*

أَتَى صَدِيقٌ إِلَى جُحَا يَطلُبُ مِنهُ حِمَارَهُ لِيَركَبَهُ فِي سَفرَةٍ قَصِيرَةٍ وَقَالَ لَهُ: سَوفَ أُعِيدُهُ إِلَيكَ فِي المَسَاءِ، وَأَدفَعُ لَكَ أُجرَةً. فَقَالَ جُحَا:

*ʾanā ʾāsifun ǧiddan ʾannī lā ʾastaṭīʿu ʾan ʾuḥaqqiqa laka raġbataka, fa-ʾlḥimāru laysa hunā ʾl-yawma. wa-qabla ʾan yutimmu ǧuḥā kalāmahu badaʾa ʾl-ḥimāru yanhaqu fī ʾṣṭablihi.*

أَنَا آسِفٌ جِدًّا أَنِّي لَا أَستَطِيعُ أَن أُحَقِّقَ لَكَ رَغبَتَكَ، فَالحِمَارُ لَيسَ هُنَا اليَومَ. وَقَبلَ أَن يُتِمَّ جُحَا كَلَامَهُ بَدَأَ الحِمَارُ يَنهَقُ فِي اصطَبلِهِ.

*fa-qāla lahu ṣadīquhu: ʾinnī ʾasmaʿu ḥimāraka yā ǧuḥā yanhaqu. fa-qāla lahu ǧuḥā: ġarībun ʾamruka yā ṣadīqī! ʾa-tuṣaddiqu ʾl-ḥimāra wa-tukaḏḏibunī?*

فَقَالَ لَهُ صَدِيقُهُ: إِنِّي أَسمَعُ حِمَارَكَ يَا جُحَا يَنهَقُ. فَقَالَ لَهُ جُحَا: غَرِيبٌ أَمرُكَ يَا صَدِيقِي! أَتُصَدِّقُ الحِمَارَ وَتُكَذِّبُنِي؟

# Annex G
(informative)
# ISO-TimeML DTD

NOTES
1. This annex is a revised version of *timeml_1.2.1.dtd*.
2. New changes were incorporated in it by James Pustejovsky after the January 2007 Tilburg meeting (2007-02-24).

```
<!ELEMENT TimeML ( #PCDATA | ALINK | CONFIDENCE | EVENT |
    | SIGNAL | SLINK | TIMEX3 | TLINK )* >
<!ATTLIST TimeML xsi:noNamespaceSchemaLocation CDATA #IMPLIED >
<!ATTLIST TimeML xmlns:xsi CDATA #IMPLIED >

<!ATTLIST TimeML comment CDATA #IMPLIED >

<!ELEMENT EVENT ( #PCDATA ) >
<!ATTLIST EVENT eid ID #REQUIRED >
<!ATTLIST EVENT eiid ID #REQUIRED >
<!ATTLIST EVENT type ( STATE | PROCESS | TRANSITION) #REQUIRED >
<!ATTLIST EVENT class ( ASPECTUAL | I_ACTION | I_STATE |
  OCCURRENCE | PERCEPTION | REPORTING | STATE ) #REQUIRED >
<!ATTLIST EVENT stem CDATA #IMPLIED >
<!ATTLIST EVENT pos ( ADJECTIVE | NOUN | VERB | PREPOSITION
   | NONE ) #REQUIRED >
<!ATTLIST EVENT tense ( FUTURE  | NONE | PAST |
    PRESENT | IMPERFECT) #REQUIRED >
<!ATTLIST EVENT aspect ( NONE | PERFECTIVE | IMPERFECTIVE |
   PERFECTIVE_PROGRESSIVE | PROGRESSIVE |
   IMPERFECTIVE_PROGRESSIVE ) #REQUIRED >
<!ATTLIST EVENT vform ( NONE | INFINITIVE | GERUNDIVE |
   PRESPART | PASTPART) #REQUIRED >
<!ATTLIST EVENT polarity ( POS | NEG ) #REQUIRED >
<!ATTLIST EVENT mood ( SUBJUNCTIVE | NONE ) #REQUIRED >
<!ATTLIST EVENT modality CDATA #IMPLIED >
<!ATTLIST EVENT comment CDATA #IMPLIED >

<!ELEMENT TIMEX3 ( #PCDATA ) >
<!ATTLIST TIMEX3 tid ID #REQUIRED >
<!ATTLIST TIMEX3 type ( DATE | DURATION | SET | TIME ) #REQUIRED >
<!ATTLIST TIMEX3 value NMTOKEN #REQUIRED >

<!ATTLIST TIMEX3 anchorTimeID IDREF #IMPLIED >

<!ATTLIST TIMEX3 beginPoint IDREF #IMPLIED >

<!ATTLIST TIMEX3 endPoint IDREF #IMPLIED >

<!ATTLIST TIMEX3 freq NMTOKEN #IMPLIED >

<!ATTLIST TIMEX3 functionInDocument ( CREATION_TIME |
EXPIRATION_TIME | MODIFICATION_TIME | PUBLICATION_TIME |
RELEASE_TIME | RECEPTION_TIME | NONE ) #IMPLIED>
```

```
<!ATTLIST TIMEX3 mod ( BEFORE | AFTER | ON_OR_BEFORE | ON_OR_AFTER
   | LESS_THAN | MORE_THAN | EQUAL_OR_LESS | EQUAL_OR_MORE | START |
   MID | END | APPROX )  #IMPLIED >

<!ATTLIST TIMEX3 quant CDATA #IMPLIED >

<!ATTLIST TIMEX3 temporalFunction ( false | true ) #IMPLIED >
<!ATTLIST TIMEX3 valueFromFunction IDREF #IMPLIED >

<!ATTLIST TIMEX3 comment CDATA #IMPLIED >

<!ELEMENT SIGNAL ( #PCDATA ) >
<!ATTLIST SIGNAL sid ID #REQUIRED >
<!ATTLIST SIGNAL comment CDATA #IMPLIED >

<!ELEMENT ALINK EMPTY >
<!ATTLIST ALINK lid ID #REQUIRED >
<!ATTLIST ALINK relType ( CONTINUES | CULMINATES | INITIATES |
   REINITIATES | TERMINATES ) #REQUIRED >

<!ATTLIST ALINK eventInstanceID IDREF #REQUIRED >

<!ATTLIST ALINK relatedToEventInstance IDREF #REQUIRED >

<!ATTLIST ALINK signalID IDREF #IMPLIED >

<!ATTLIST ALINK syntax CDATA #IMPLIED >

<!ATTLIST ALINK comment CDATA #IMPLIED >

<!ELEMENT SLINK EMPTY >

<!ATTLIST SLINK lid ID #REQUIRED >

<!ATTLIST SLINK relType ( CONDITIONAL | COUNTER_FACTIVE |
   EVIDENTIAL | FACTIVE | INTENSIONAL | NEG_EVIDENTIAL ) #REQUIRED >
<!ATTLIST SLINK eventInstanceID NMTOKEN #REQUIRED >

<!ATTLIST SLINK subordinatedEventInstance NMTOKEN #REQUIRED >
<!ATTLIST SLINK signalID NMTOKEN #IMPLIED >

<!ATTLIST SLINK syntax CDATA #IMPLIED >

<!ATTLIST SLINK comment CDATA #IMPLIED >

<!ELEMENT TLINK EMPTY >
<!ATTLIST TLINK lid ID #REQUIRED >
<!ATTLIST TLINK relType ( BEFORE | AFTER | INCLUDES | IS_INCLUDED
   | DURING | DURING_INV | SIMULTANEOUS | IAFTER | IBEFORE | IDENTITY
   | BEGINS | ENDS | BEGUN_BY | ENDED_BY ) #REQUIRED >
<!ATTLIST TLINK eventInstanceID IDREF #IMPLIED >

<!ATTLIST TLINK timeID IDREF #IMPLIED >

<!ATTLIST TLINK relatedToEventInstance IDREF #IMPLIED >
```

```
<!ATTLIST TLINK relatedToTime IDREF #IMPLIED >

<!ATTLIST TLINK signalID IDREF #IMPLIED >

<!ATTLIST TLINK origin CDATA #IMPLIED >

<!ATTLIST TLINK syntax CDATA #IMPLIED >

<!ATTLIST TLINK comment CDATA #IMPLIED >

<!ELEMENT CONFIDENCE EMPTY >

<!ATTLIST CONFIDENCE tagType ( EVENT | MAKEINSTANCE | TIMEX3 |
   SIGNAL | ALINK | SLINK | TLINK ) #REQUIRED >

<!ATTLIST CONFIDENCE tagID IDREF #REQUIRED >

<!ATTLIST CONFIDENCE attributeName CDATA #IMPLIED >

<!ATTLIST CONFIDENCE confidenceValue CDATA #REQUIRED >

<!ATTLIST CONFIDENCE comment CDATA #IMPLIED >
```

 

# Annex H
(informative)
# ISO-TimeML schema

NOTES
1. This annex is a revised version of *TimeML Schema*.
2. New changes were incorporated in it by James Pustejovsky after the January 2007 Tilburg meeting (2007-02-24).

```
<?xml version="1.0"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
 xml:lang="en" version="1.2">

  <xs:annotation>
    <xs:documentation xml:lang="en-us">
      SemAFTimeML schema
      <a href="http://www.timeml.org/">TANGO</a>
    </xs:documentation>
  </xs:annotation>

    <xs:simpleType name="ISO-ISO-TimeMLID">
        <xs:annotation>
          <xs:documentation xml:lang="en-us">
        A <b>ISO-TimeMLID</b> is a sequence
        of lowercase letters
        followed by a number (decimal digits).
          </xs:documentation>
        </xs:annotation>
        <xs:restriction base="xs:string">
            <xs:pattern value="[a-z]+\p{Nd}+" />
        </xs:restriction>
    </xs:simpleType>

    <xs:simpleType name="EventID">
    <xs:annotation>
      <xs:documentation xml:lang="en-us">
        A <b>EventID</b> is an <b>e</b> followed
        by a number (decimal digits).
      </xs:documentation>
    </xs:annotation>
        <xs:restriction base="ISO-TimeMLID">
            <xs:pattern value="e\p{Nd}+" />
        </xs:restriction>
    </xs:simpleType>

    <xs:simpleType name="EventInstanceID">
    <xs:annotation>
      <xs:documentation xml:lang="en-us">
        A <b>EventInstanceID</b> is an <b>ei</b>
        followed by a number (decimal digits).
      </xs:documentation>
    </xs:annotation>
        <xs:restriction base="ISO-TimeMLID">
```

```
        <xs:pattern value="ei\p{Nd}+" />
    </xs:restriction>
</xs:simpleType>

<xs:simpleType name="TimeID">
<xs:annotation>
  <xs:documentation xml:lang="en-us">
    A <b>TimeID</b> is an <b>t</b>
    followed by a number (decimal digits).
  </xs:documentation>
</xs:annotation>
    <xs:restriction base="ISO-TimeMLID">
        <xs:pattern value="t\p{Nd}+" />
    </xs:restriction>
</xs:simpleType>

<xs:simpleType name="SignalID">
<xs:annotation>
  <xs:documentation xml:lang="en-us">
    A <b>SignalID</b> is an <b>s</b>
    followed by a number (decimal digits).
  </xs:documentation>
</xs:annotation>
    <xs:restriction base="ISO-TimeMLID">
        <xs:pattern value="s\p{Nd}+" />
    </xs:restriction>
</xs:simpleType>

    <xs:simpleType name="LinkID">
    <xs:annotation>
      <xs:documentation xml:lang="en-us">
        A <b>LinkID</b> is an <b>l</b>
        followed by a number (decimal digits).
      </xs:documentation>
    </xs:annotation>
        <xs:restriction base="ISO-TimeMLID">
            <xs:pattern value="l\p{Nd}+" />
        </xs:restriction>
</xs:simpleType>


<xs:simpleType name="Probability">
    <xs:annotation>
  <xs:documentation xml:lang="en-us">
    The <b>Probability</b> of an annotation
    is a number between 0.0 and 1.0.
  </xs:documentation>
</xs:annotation>
    <xs:restriction base="xs:decimal">
        <xs:minInclusive value="0" />
        <xs:maxInclusive value="1" />
    </xs:restriction>
</xs:simpleType>


<xs:simpleType name="Date">
    <xs:restriction base="xs:string">
```

```
        <xs:pattern value="[0-9X]{1,4}(-[0-9X]
        {1,2}(-[0-9X]{1,2})?)?" />
    </xs:restriction>
</xs:simpleType>

<xs:simpleType name="Time">
    <xs:restriction base="xs:string">
        <xs:pattern value="([0-9X]{1,4}-[0-9X]
        {1,2}-[0-9X]{1,2})?T(([0-9]{2}
        (:[0-9]{2}(:[0-9]{2})?)?)|
        (MO|MI|AF|EV|NI|DT))" />
    </xs:restriction>
</xs:simpleType>

<xs:simpleType name="WeekDate">
    <xs:restriction base="xs:string">
        <xs:pattern value="[0-9X]{1,4}-W[0-9X]
        {1,2}(-([1-7X]|WE))?"/>
    </xs:restriction>
</xs:simpleType>

<xs:simpleType name="WeekTime">
    <xs:restriction base="xs:string">
        <xs:pattern value="[0-9X]{1,4}-W[0-9X]
        {1,2}-[1-7X]T(([0-9]{2}(:[0-9]{2}
        (:[0-9]{2})?)?)|(MO|MI|AF|EV|NI|DT))"/>
    </xs:restriction>
</xs:simpleType>


<xs:simpleType name="Season">
    <xs:restriction base="xs:string">
        <xs:pattern value="[0-9X]{1,4}-(SP|SU|WI|FA)"/>
    </xs:restriction>
</xs:simpleType>

<xs:simpleType name="PartOfYear">
    <xs:restriction base="xs:string">
        <xs:pattern value="[0-9X]{1,4}-(H[1-2X]|Q[1-4X])"/>
    </xs:restriction>
</xs:simpleType>

<xs:simpleType name="PaPrFu">
    <xs:restriction base="xs:string">
        <xs:enumeration value="PAST_REF"/>
        <xs:enumeration value="PRESENT_REF"/>
        <xs:enumeration value="FUTURE_REF"/>
    </xs:restriction>
</xs:simpleType>


<xs:simpleType name="Duration">
    <xs:restriction base="xs:string">
        <xs:pattern value="P((((\p{Nd}+|X{1,2})Y)
        ?((\p{Nd}+|X{1,2})M)?((\p{Nd}+|X{1,2})D)
        ?(T((\p{Nd}+|X{1,2})H)?((\p{Nd}+|X{1,2})M)
        ?((\p{Nd}+|X{1,2})S)?)?)|
```

```
            ((\p{Nd}+|X{1,2}))(W|L|E|C|Q))"/>
            <xs:minLength value="3"/>
        </xs:restriction>
    </xs:simpleType>

    <xs:element name="ISO-TimeML">

        <xs:unique name="link_id">
            <xs:selector xpath="TLINK|SLINK|ALINK" />
            <xs:field xpath="@lid" />
        </xs:unique>

        <xs:key name="event_id">
            <xs:selector xpath="EVENT" />
            <xs:field xpath="@eid" />
        </xs:key>

        <xs:key name="time_id">
            <xs:selector xpath="TIMEX3" />
            <xs:field xpath="@tid" />
        </xs:key>

        <xs:key name="signal_id">
            <xs:selector xpath="SIGNAL" />
            <xs:field xpath="@sid" />
        </xs:key>


<!--    these keyrefs assume that no other (non-ISO-TimeML)
        elements occur within the ISO-TimeML root tag.
        Only use them if non-ISO-TimeML tags have been removed.

        <xs:keyref name="event_id_ref_mi" refer="event_id">
            <xs:selector xpath="EVENT" />
            <xs:field xpath="@eventID" />
        </xs:keyref>

        <xs:keyref name="time_id_ref_tx_ati" refer="time_id">
            <xs:selector xpath="TIMEX3" />
            <xs:field xpath="@anchorTimeID" />
        </xs:keyref>

        <xs:keyref name="signal_id_ref_tl" refer="signal_id">
            <xs:selector xpath="TLINK" />
            <xs:field xpath="@signalID" />
        </xs:keyref>

        <xs:keyref name="time_id_ref_tl_ti" refer="time_id">
            <xs:selector xpath="TLINK" />
            <xs:field xpath="@timeID" />
        </xs:keyref>

        <xs:keyref name="time_id_ref_tl_rtt" refer="time_id">
            <xs:selector xpath="TLINK" />
            <xs:field xpath="@relatedToTime" />
        </xs:keyref>
```

```
<xs:keyref name="time_id_ref_tx_bp" refer="time_id">
    <xs:selector xpath="TIMEX3" />
    <xs:field xpath="@beginPoint" />
</xs:keyref>

<xs:keyref name="time_id_ref_tx_ep" refer="time_id">
    <xs:selector xpath="TIMEX3" />
    <xs:field xpath="@endPoint" />
</xs:keyref>

<xs:keyref name="event_instance_id_ref_tl-eid"
refer="event_instance_id">
    <xs:selector xpath="TLINK" />
    <xs:field xpath="@eventInstanceID" />
</xs:keyref>

<xs:keyref name="event_instance_id_ref_tl-rte"
refer="event_instance_id">
    <xs:selector xpath="TLINK" />
    <xs:field xpath="@relatedToEventInstance" />
</xs:keyref>

<xs:keyref name="signal_id_ref_sl"
refer="signal_id">
    <xs:selector xpath="SLINK" />
    <xs:field xpath="@signalID" />
</xs:keyref>

<xs:keyref name="event_instance_id_ref_sl-eid"
refer="event_instance_id">
    <xs:selector xpath="SLINK" />
    <xs:field xpath="@eventInstanceID" />
</xs:keyref>

<xs:keyref name="event_instance_id_ref_sl-sei"
refer="event_instance_id">
    <xs:selector xpath="SLINK" />
    <xs:field xpath="@subordinatedEventInstance" />
</xs:keyref>


<xs:keyref name="signal_id_ref_al"
refer="signal_id">
    <xs:selector xpath="ALINK" />
    <xs:field xpath="@signalID" />
</xs:keyref>

<xs:keyref name="event_instance_id_ref_al-eid"
refer="event_instance_id">
    <xs:selector xpath="ALINK" />
    <xs:field xpath="@eventInstanceID" />
</xs:keyref>

<xs:keyref name="event_id_ref_al-rte"
refer="event_instance_id">
    <xs:selector xpath="ALINK" />
    <xs:field xpath="@relatedToEventInstance" />
```

```
    </xs:keyref>
 -->
</xs:element>


<xs:element name="EVENT">
    <xs:complexType mixed="true">
        <xs:sequence minOccurs="1" maxOccurs="1">
            <xs:any namespace="##any" minOccurs="0"
            maxOccurs="unbounded" processContents="lax"/>
        </xs:sequence>
        <xs:attribute name="eid" type="EventID" use="required" />
 <xs:attribute name="type" use="required">
            <xs:simpleType>
                <xs:restriction base="xs:string">
                    <xs:enumeration value="STATE" />
                    <xs:enumeration value="PROCESS" />
                    <xs:enumeration value="TRANSITION" />
                </xs:restriction>
            </xs:simpleType>
        </xs:attribute>
        <xs:attribute name="class" use="required">
            <xs:simpleType>
                <xs:restriction base="xs:string">
                    <xs:enumeration value="OCCURRENCE" />
                    <xs:enumeration value="PERCEPTION" />
                    <xs:enumeration value="REPORTING" />
                    <xs:enumeration value="ASPECTUAL" />
                    <xs:enumeration value="STATE" />
                    <xs:enumeration value="I_STATE" />
                    <xs:enumeration value="I_ACTION" />
                    <xs:enumeration value="INTENSIONAL" />
                </xs:restriction>
            </xs:simpleType>
        </xs:attribute>
        <xs:attribute name="eiid" type="EventInstanceID" use="required" />
        <xs:attribute name="eventID" type="EventID" use="required" />
        <xs:attribute name="mood" use="required">
          <xs:simpleType>
            <xs:restriction base="xs:string">
              <xs:enumeration value="SUBJUNCTIVE" />
              <xs:enumeration value="NONE" />
                </xs:restriction>
            </xs:simpleType>
        </xs:attribute>
        <xs:attribute name="modality" type="xs:string" />
        <xs:attribute name="polarity" default="POS">
            <xs:simpleType>
                <xs:restriction base="xs:string">
                    <xs:enumeration value="NEG" />
                    <xs:enumeration value="POS" />
                </xs:restriction>
            </xs:simpleType>
        </xs:attribute>
        <xs:attribute name="tense" use="required">
            <xs:simpleType>
                <xs:restriction base="xs:string">
```

```
                    <xs:enumeration value="PAST" />
                    <xs:enumeration value="PRESENT" />
                    <xs:enumeration value="FUTURE" />
                    <xs:enumeration value="IMPERFECT" />
                    <xs:enumeration value="NONE" />
                </xs:restriction>
            </xs:simpleType>
        </xs:attribute>
        <xs:attribute name="aspect" use="required">
            <xs:simpleType>
                <xs:restriction base="xs:string">
                    <xs:enumeration value="PROGRESSIVE" />
                    <xs:enumeration value="PERFECTIVE" />
                    <xs:enumeration value="PERFECTIVE_PROGRESSIVE" />
                    <xs:enumeration value="IMPERFECTIVE_PROGRESSIVE" />
                    <xs:enumeration value="IMPERFECTIVE" />
                    <xs:enumeration value="NONE" />
                </xs:restriction>
            </xs:simpleType>
        <xs:attribute name="vform" use="required">
            <xs:simpleType>
                <xs:restriction base="xs:string">
                    <xs:enumeration value="PASTPART" />
                    <xs:enumeration value="PRESPART" />
                    <xs:enumeration value="INFINITIVE" />
                    <xs:enumeration value="GERUNDIVE" />
                    <xs:enumeration value="NONE" />
                </xs:restriction>
            </xs:simpleType>
        </xs:attribute>
        </xs:attribute>
        <xs:attribute name="nf_morph" use="required">
            <xs:simpleType>
                <xs:restriction base="xs:string">
                    <xs:enumeration value="ADJECTIVE" />
                    <xs:enumeration value="NOUN" />
                    <xs:enumeration value="INFINITIVE" />
                    <xs:enumeration value="PRESPART" />
                    <xs:enumeration value="PASTPART" />
                    <xs:enumeration value="NONE" />
                </xs:restriction>
            </xs:simpleType>
        </xs:attribute>

    </xs:complexType>
</xs:element>

<xs:element name="TIMEX3">
    <xs:complexType mixed="true">
        <xs:sequence minOccurs="1" maxOccurs="1">
            <xs:any namespace="##any" minOccurs="0"
            maxOccurs="unbounded" processContents="lax"/>
        </xs:sequence>

        <xs:attribute name="tid" type="TimeID" use="required" />

        <xs:attribute name="type" use="required">
```

```
    <xs:simpleType>
        <xs:restriction base="xs:string">
            <xs:enumeration value="DATE" />
            <xs:enumeration value="TIME" />
            <xs:enumeration value="DURATION" />
            <xs:enumeration value="SET" />
        </xs:restriction>
    </xs:simpleType>
</xs:attribute>

<xs:attribute name="functionInDocument" default="NONE">
    <xs:simpleType>
        <xs:restriction base="xs:string">
            <xs:enumeration value="CREATION_TIME" />
            <xs:enumeration value="EXPIRATION_TIME" />
            <xs:enumeration value="MODIFICATION_TIME" />
            <xs:enumeration value="PUBLICATION_TIME" />
            <xs:enumeration value="RELEASE_TIME" />
            <xs:enumeration value="RECEPTION_TIME" />
            <xs:enumeration value="NONE" />
        </xs:restriction>
    </xs:simpleType>
</xs:attribute>

<xs:attribute name="temporalFunction"
type="xs:boolean" default="false" />

<xs:attribute name="value" use="required">
    <xs:simpleType>
        <xs:union memberTypes="Date Time WeekDate
        WeekTime Season PartOfYear PaPrFu Duration"/>
    </xs:simpleType>
</xs:attribute>

<xs:attribute name="valueFromFunction" type="xs:IDREF" />

<xs:attribute name="mod">
    <xs:simpleType>
        <xs:restriction base="xs:string">
            <xs:enumeration value="BEFORE" />
            <xs:enumeration value="AFTER" />
            <xs:enumeration value="ON_OR_BEFORE" />
            <xs:enumeration value="ON_OR_AFTER" />
            <xs:enumeration value="LESS_THAN" />
            <xs:enumeration value="MORE_THAN" />
            <xs:enumeration value="EQUAL_OR_LESS" />
            <xs:enumeration value="EQUAL_OR_MORE" />
            <xs:enumeration value="START" />
            <xs:enumeration value="MID" />
            <xs:enumeration value="END" />
            <xs:enumeration value="APPROX" />
        </xs:restriction>
    </xs:simpleType>
</xs:attribute>

<xs:attribute name="anchorTimeID" type="TimeID" />
```

```
        <xs:attribute name="beginPoint" type="TimeID"/>

        <xs:attribute name="endPoint" type="TimeID"/>

        <xs:attribute name="quant" type="xs:string"/>

        <xs:attribute name="freq" type="Duration"/>
    </xs:complexType>
</xs:element>

<xs:element name="SIGNAL">
    <xs:complexType mixed="true">
        <xs:sequence minOccurs="1" maxOccurs="1">
            <xs:any namespace="##any" minOccurs="0"
            maxOccurs="unbounded" processContents="lax"/>
        </xs:sequence>
        <xs:attribute name="sid" type="SignalID" use="required" />
    </xs:complexType>
</xs:element>

<xs:element name="TLINK">
    <xs:complexType>
        <xs:attribute name="lid" type="LinkID" />
        <xs:attribute name="origin" type="xs:string"/>
        <xs:attribute name="eventInstanceID"
            type="EventInstanceID" />
        <xs:attribute name="timeID" type="TimeID" />
        <xs:attribute name="signalID" type="SignalID" />
        <xs:attribute name="relatedToTime" type="TimeID" />
        <xs:attribute name="relatedToEventInstance"
            type="EventInstanceID" />
        <xs:attribute name="relType" use="required">
            <xs:simpleType>
                <xs:restriction base="xs:string">
                    <xs:enumeration value="BEFORE" />
                    <xs:enumeration value="AFTER" />
                    <xs:enumeration value="INCLUDES" />
                    <xs:enumeration value="IS_INCLUDED" />
                    <xs:enumeration value="DURING" />
                    <xs:enumeration value="DURING_INV" />
                    <xs:enumeration value="SIMULTANEOUS" />
                    <xs:enumeration value="IAFTER" />
                    <xs:enumeration value="IBEFORE" />
                    <xs:enumeration value="IDENTITY" />
                    <xs:enumeration value="BEGINS" />
                    <xs:enumeration value="ENDS" />
                    <xs:enumeration value="BEGUN_BY" />
                    <xs:enumeration value="ENDED_BY" />
                </xs:restriction>
            </xs:simpleType>
        </xs:attribute>
        <xs:attribute name="magnitude" type="TimeID" />
    </xs:complexType>
</xs:element>
<xs:element name="SLINK">
    <xs:complexType>
        <xs:attribute name="lid" type="LinkID" />
```

```
        <xs:attribute name="origin" type="xs:string"/>
        <xs:attribute name="eventInstanceID"
           type="EventInstanceID" use="required"/>
        <xs:attribute name="subordinatedEventInstance"
           type="EventInstanceID" use="required" />
        <xs:attribute name="signalID" type="SignalID" />
        <xs:attribute name="relType" use="required">
            <xs:simpleType>
                <xs:restriction base="xs:string">
                    <xs:enumeration value="INTENSIONAL" />
                    <xs:enumeration value="EVIDENTIAL" />
                    <xs:enumeration value="NEG_EVIDENTIAL" />
                    <xs:enumeration value="FACTIVE" />
                    <xs:enumeration value="COUNTER_FACTIVE" />
                    <xs:enumeration value="CONDITIONAL" />
                </xs:restriction>
            </xs:simpleType>
        </xs:attribute>
    </xs:complexType>
</xs:element>

<xs:element name="ALINK">
    <xs:complexType>
        <xs:attribute name="lid" type="LinkID" />
        <xs:attribute name="origin" type="xs:string"/>
        <xs:attribute name="eventInstanceID"
           type="EventInstanceID" use="required" />
        <xs:attribute name="signalID" type="SignalID" />
        <xs:attribute name="relatedToEventInstance"
           type="EventInstanceID" use="required" />
        <xs:attribute name="relType" use="required">
            <xs:simpleType>
                <xs:restriction base="xs:string">
                    <xs:enumeration value="INITIATES" />
                    <xs:enumeration value="CULMINATES" />
                    <xs:enumeration value="TERMINATES" />
                    <xs:enumeration value="CONTINUES" />
                    <xs:enumeration value="REINITIATES" />
                </xs:restriction>
            </xs:simpleType>
        </xs:attribute>
    </xs:complexType>
</xs:element>

<xs:element name="CONFIDENCE">
    <xs:complexType>
        <xs:attribute name="tagType" use="required">
            <xs:simpleType>
                <xs:restriction base="xs:string">
                    <xs:enumeration value="EVENT" />
                    <xs:enumeration value="TIMEX3" />
                    <xs:enumeration value="SIGNAL" />
                    <xs:enumeration value="TLINK" />
                    <xs:enumeration value="SLINK" />
                    <xs:enumeration value="ALINK" />
                </xs:restriction>
            </xs:simpleType>
```

```
            </xs:attribute>
            <xs:attribute name="tagID" type="ISO-TimeMLID" use="required" />
            <xs:attribute name="attributeName">
                <xs:simpleType>
                    <xs:restriction base="xs:string">
                        <xs:enumeration value="anchorTimeID"/>
                        <xs:enumeration value="aspect"/>
                        <xs:enumeration value="beginPoint"/>
                        <xs:enumeration value="cardinality"/>
                        <xs:enumeration value="class"/>
                        <xs:enumeration value="endPoint"/>
                        <xs:enumeration value="eventID"/>
                        <xs:enumeration value="eventInstanceID"/>
                        <xs:enumeration value="freq"/>
                        <xs:enumeration value="functionInDocument"/>
                        <xs:enumeration value="lid"/>
                        <xs:enumeration value="mod"/>
                        <xs:enumeration value="modality"/>
                        <xs:enumeration value="mood"/>
                        <xs:enumeration value="polarity"/>
                        <xs:enumeration value="origin"/>
                        <xs:enumeration value="quant"/>
                        <xs:enumeration value="relType"/>
                        <xs:enumeration value="relatedToEventInstance"/>
                        <xs:enumeration value="signalID"/>
                        <xs:enumeration value="subordinatedEventInstance"/>
                        <xs:enumeration value="temporalFunction"/>
                        <xs:enumeration value="tense"/>
                        <xs:enumeration value="timeID"/>
                        <xs:enumeration value="type"/>
                        <xs:enumeration value="value"/>
                        <xs:enumeration value="valueFromFunction"/>
                    </xs:restriction>
                </xs:simpleType>
            </xs:attribute>
            <xs:attribute name="confidenceValue"
               type="Probability" use="required" />
        </xs:complexType>
    </xs:element>
</xs:schema>
```

# Annex I
(informative)
# Past and current activities on temporal and event annotation

NOTES
Much of the discussion here is adapted from an introduction to Part IV Temporal Annotation, in Mani, Pustejovsky, and Gaizauska (2005).

The automatic recognition of temporal and event expressions in natural language text has recently become an active area of research in computational linguistics and semantics. In this annex, we review the work done on temporal and event annotation.

## I.1 Annotating temporal expressions

The most obvious temporal feature to annotate in texts, and the one which historically was addressed first, is temporal referring expressions (as found in temporal adverbials, for example); that is, expressions which refer to times (July 1, 1867), durations (three months), or frequencies (weekly).

Being able to identify and distinguish these types of expression is crucial to being able to situate the events described in text either absolutely in terms of some conventional calendrical time frame or relatively with respect to other events. The examples just given perhaps understate the complexity of the phenomena to be addressed. When devising an annotation scheme to capture temporal referring expressions o ne must deal with a variety of complications:

(1) indexicals: expressions like now, yesterday –and other contextually dependent expressions such as partially specified calendrical times (e.g. Wednesday– which Wednesday? ) or relatives such as next week, three weeks ago, all of which depend for their interpretation on knowledge of a deictic centre;

(2) relational expressions: expressions which explicitly specify times in relation to other times (two weeks after Christmas) or to events (5 seconds after the first explosion); and

(3) vagueness: expressions referring to times whose boundaries are inherently vague (spring, evening) or which contain modifiers which blur the time reference (several days ago, sometime after 7 p.m.).

Work to devise annotation schemes for temporal referring expressions appears to have begun as part of the Named Entity (NE) tagging subtask within the DARPA Message Understanding Conference (MUC) series of evaluations, specifically in MUC-6 (MUC 1995). In this task participants' systems were to tag (by inserting SGML tags into running text) expressions which named persons, organizations, locations, dates, times, monetary amounts, and percentages. A key part of this exercise was that a set of texts was manually tagged by human annotators to provide a 'gold standard' measure of correctness. Metrics, principally the recall and precision metrics adapted from information retrieval research, were used to compare system-supplied annotations (or responses) against human-supplied annotations (or answer keys). Recall, the proportion of the answer keys for which a correct response is supplied, is a measure of coverage or completeness of a system; precision, the proportion of responses which are correct, i.e. match the answer key, is a measure of correctness or soundness of a system.

In MUC-6 date and time (of day) expressions were labeled using a TIMEX tag. Only absolute time expressions were to be annotated, i.e. expressions which indicated a specific minute, hour, day, month, season, year, etc. Relative time expressions (e.g. last July) were excluded, though subexpressions within them (e.g. July in this example) were to be tagged. A set of thirty manually annotated newswire texts were used for a blind evaluation. The top scoring automated system scored .97 recall and .96 precision on the TIMEX tagging task.

In MUC-7 (MUC 1998) the principal change was to capture relative as well as absolute date and time expressions, though the two did not need to be distinguished in the tagging. Thus indexicals, such as yesterday, last July, were to be marked, as were so-called 'time-relative-to-event' phrases such as the morning after the July 17 disaster. For the final blind evaluation a set of 100 tagged texts was used and the highest scoring system scored .89/.99 recall/precision on the date tagging task and .81/.97 recall/ precision on the time tagging task.

One of the principal limitations of the date and time NE task in both MUC-6 and MUC-7 is that while identifying temporal referring expressions in text is useful, what is really needed is the ability to interpret or evaluate or dereference these expressions to obtain the time they denote. Thus, according to the MUC-7 TIMEX tagging guidelines, an expression such as yesterday in an article datelined "June 12, 1998" would be tagged as a TIMEX of type DATE. However, what applications really need is the knowledge that in this context yesterday refers to June 11, 1998. This requirement is addressed by the TIMEX2 tagging guidelines (Wilson et al., 2002). Interpretation is handled by adding the full calendrical time value for every temporal referring expression as an attribute of the tagged element, using an ISO standard time format as the attribute's value. Wilson et al. (2002) also describe an implemented tagger which annotates newswire text (in English and Spanish) with TIMEX2 tags with impressively high scores, achieving 96.2 f-measure (a combined measure of recall and precision) for tagging surface expressions and 83.2 f-measure in interpreting them.

The ability to evaluate a relational or indexical time expression, returning a calendrical time value, is clearly needed as part of the temporal interpretation process. However, there is utility in separating the evaluation process into two stages, first mapping the time expression into a semantic representation in the form of a functional expression, and second evaluating the functional expression. So, for example last Thursday might in the first stage be mapped into the expression thursday (predecessor (week DCT)), where DCT is the document-creation time of the article and in the second stage an absolute calendrical time is computed from this latter representation given the DCT. This separation of semantic interpretation from full evaluation has number of advantages. It fosters discussion of the correct semantic interpretation of complex temporal referring expressions, it permits separate evaluation of the two stages (an algorithm could be good at working out the semantics of last expressions, but bad at finding their anchors), it allows unevaluated semantic representations to be made available to other interpretation components which may require them rather than their values, and it permits taggers to defer the evaluation of temporal functions until their values are actually required. Pustejovsky et al. (2003) propose an extension of the TIMEX2 standard to include temporal functional representations, and call the extended standard TIMEX3 (TIMEX3 includes a number of other refinements to the TIMEX2 standard, but this is the most significant).

Most of the work described above has been driven by the English-speaking research community, though as noted TIMEX2 has been applied to English and Spanish, and recently to Korean, French, Chinese, and Hindi. However, Schilder and Habel (2001) independently propose an approach for annotating German newswire texts which aims to capture the same sort of temporal referring expressions as the TIMEX2 and 3 standards. Their tagger outputs a semantic representation of relative time expressions which are evaluated in a subsequent stage, making its handling of these expressions similar to that proposed in TIMEX3.

### I.2 Annotating events

To interpret a text temporally means not just identifying the times, durations, and frequencies mentioned in a text; it means positioning the events and states described in the text with respect to these times and to each other. However, before it is possible to discuss how to annotate relations between events, states, and times, agreement must be reached on how to annotate events and states themselves. To do this in turn requires making decisions about (a) what we are trying to annotate – just events? events and states? and what do we take the difference to be? (b) how events/states are realized in text; (c) what textual representative of the event/state will be annotated; (d) what attributes should be associated with annotated events/states.

At the most general level, temporal annotation can be taken as the task of correctly annotating the temporal position of all temporal entities in a text, i.e. of all things that happen or are situated in time. If, for purposes of the following discussion, we assume a top-level ontological class of eventualities or situations which is divided into events and states, this would mean annotating all events and states.

Such a task is daunting, and since practical applications are primarily concerned with events, it might appear reasonable to start out with the more modest aim of annotating events, but excluding states. However, drawing

a firm conceptual distinction between events and states is not straightforward, as the discussion in Part I, Mani, Pustejovsky, and Gaizauskas (2005), has shown. One common distinguishing test is the so-called subinterval property (Dowty, 1979): for any state p that holds over an interval t, p must hold for every sub-interval of t. However, this is not a particularly easy test to apply and not one to expect annotators of texts to be able to carry out efficiently or effectively.

A second way to distinguish events and states is via linguistic tests. States tend to be expressed via constructions with the copula, or via certain verbs such as have, know, believe. This is perhaps a more practical approach in the context of producing realistic guidelines for annotation. If the point of making the distinction is to capture genuine semantic differences between events and states, however, then this approach depends on determining an accurate and complete set of linguistic correlates for states.

Most approaches to event annotation reported in this part, however, do not attempt to make a distinction between events and states. In general, the approach is to treat all verbs as expressing temporal entities suitable for tagging. This 'lumping' together assumes that the distinction is not important, or is too difficult, for purposes of annotation. While dismissing the problem in the short term, this ignores the fact that there are genuine semantic differences between events and states, and that these have consequences in terms of the inferences that can be drawn and the likely questions that can be asked concerning each. For example, states typically invite questions about when they began, ended, and how long they lasted; events invite questions about when they happened, but not so typically about their duration. Furthermore, the process of positioning states in time may differ from that of positioning events, so that an algorithm that attempts do this positioning automatically would need to know which it was dealing with.

The only work in this part which does propose to distinguish events and states and to annotate both is that of Pustejovsky et al. (2003). Note, however, that they treat states as a subtype of events – effectively identifying events with what we have here termed eventualities. In fact they go further than simply distinguishing events and states, and propose distinguishing seven types of events in their annotation scheme, two of which are stative and all of which are held to have distinctive temporal significance. Their distinguishing criteria, as presented, are primarily linguistic, though concerning states they do appeal to something like the subinterval property cited above. Further, they do not propose to annotate all states: they propose to annotate only those states which "are directly related to a temporal expression including those states that identifiably change over the course of a document".

To date then, the work on temporal annotation of 'events' in text has not worried overly about the semantic distinction between events and states and has assumed that the "things which are situated in time" which need to be annotated can be identified via a set of syntactic or lexical linguistic criteria. Katz and Arioso (2001), for example, define their task in a (deliberately) restrictive way: "The temporal interpretation of a sentence, for our purposes, can simply be taken to be the set of temporal relations that a speaker naturally takes to hold among the states and events described by the verbs of the sentence". Thus, for example, event nominals such destruction, election, war are excluded, as are, presumably, stative adjectives such as sunken. However, their investigation is exclusively concerned with sentence-internal temporal relations and they are not aiming to position every event or state reference in time, or in relation to another event or state.

Filatova and Hoy (2001) take the locus of events to be syntactic clauses which contain a subject (one or more noun phrases) and predicate (verb phrase with one or more verbs), as output by a specific parser. Their concern is to time-stamp these clauses, that is, to associate a calendrical time reference with each clause. They too, ignore, event nominals and stative adjectives. However, again, they are not aiming at complete temporal interpretation, but at a more limited task.

Schilder and Habel (2001) have a broader target. They identify two types of event-denoting expressions: sentences and event-denoting nouns, especially nominalizations. The most inclusive treatment is that of Pustejovsky et al. (2003), who consider events expressed by tensed or untensed verbs, nominals, adjectives, predicative sections, or prepositional sections.

Once a set of linguistic signals for events has been decided there is still the issue of deciding precisely what text spans will be annotated, i.e. what will count as the textual representative of the event. For the most part this follows straightforwardly from decisions made about the linguistic realizations of events and states. However, those decisions do not entirely specify the annotation.

Concerning events conveyed by clauses containing verbs, one could decide that the entire section is the appropriate span to be annotated. This is the position taken by Filatova and Hovy (2001). Or, one could decide to annotate just verb groups or just the heads of verb groups. This latter approach has been adopted by the other authors in this part, perhaps because it simplifies matters when dealing with embedded clauses or clauses with multiple verbs (Filatova and Hovy (2001) acknowledge problems with their approach for cases of co-ordinated verb phrases where the verbs have different tenses).

As well as tagging a text span as event representative, some approaches chose to associate attributes with the event. In Schilder and Habel's (2003) approach, for example, each event has a `sem` attribute that holds a predicate-argument representation of the event. It also has a `temp` attribute whose value is triple consisting of a binary temporal relation, the time id of the event itself, and the id of a time related to the event time by the temporal relation. This attribute gets its value computed as part of the interpretation process.

These event attributes are effectively part of Schilder and Habel's (2003) implementation of a computational mechanism to assign times to events. Another sort of information that can be associated with events is descriptive linguistic information which may be of use during the interpretation process. So, for example, Filatova and Hovy (2001) make use of tense information associated with event clauses by their parser. Pustejovsky et al. (2003) associate tense, aspect, and subtype information with events. The event subtypes they propose are: occurrence (crash, merge), state (on board, love), reporting (say, report), i-action (attempt, offer), i-state (believe, want), aspectual (begin, stop), and perception (see, hear). These classes are distinguished because of the distinctive sorts of temporal inferences that may be drawn for events within them.

In the foregoing we have discussed what is to be annotated when annotating events or states. Now we briefly discuss the state of play with implemented systems that do event tagging. These include: Filatova and Hovy (2001), Schilder and Habel (2003) and Li et al. (2004). However, for none of these researchers is event tagging itself a goal – rather they are aiming to anchor events in time and possibly also to relate events to each other temporally (Li et al. (2004). Only Filatova and Hovy (2001) provide separate evaluation results for their system's ability to recognize events – in their case the ability to recognize clauses, since for them clauses are the textual representatives of events. They report figures of around 61 per cent recall and 56 per cent precision, errors being due in part to the parser they use and in part to their shallow algorithm for extracting clauses from the parse tree. As noted the others do not evaluate event recognition separately from temporal relation annotation.

Given an approach to annotating temporal referring expressions and event/state denoting expressions, the next challenge for a programme of temporal annotation is to establish conventions for annotating the relations between times and events or between events and events.

NOTE from now on we will use the term 'event' loosely to refer to events and possibly to states as well, making clear if necessary where remarks may only pertain to states or to nonstative eventualities.

### I.3 Annotating relations between times and events

Time-event relational information may be conveyed in a variety of ways. The most explicit route is via a prepositional phrase in which a preposition signals a relation between a temporal referring expression (the complement of the phrase) and an event denoting expression (typically a verb or an event nominal modified by the phrase); for example, John flew to Boston on Friday. Sometimes the explicit prepositional marker is omitted and temporal referring expressions are used in adverbial (Friday John flew to Boston), nominal modifier (John's Friday flight to Boston) or elliptical/reduced relative section (John's flight, Friday at 5, will be crowded ) contexts. We refer to these cases as instances of syntactically implicit time-event relations. However, in many cases the relational information may be implicit in a much less direct way, to be derived by the reader using world or lexical semantic knowledge, or narrative convention and discourse interpretation. In many of these cases relations between times and events are established indirectly by first establishing relations between events and then inferring relations between times and events.

One position to take is that relations between time and events should be marked only in cases where explicitly signalled by prepositions or where they are syntactically implicit. This position is adopted by Schilder and Habel (2001), who assume a default semantic relation of inclusion for all syntactically implicit relations. Time-event relations for events which do not occur in such syntactic contexts are simply not supplied. Another possible position is to assign a calendrical time point or interval to all events in a text – so-called time-stamping of events.

Filatova and Hovy (2001) pursue this line, developing a heuristic algorithm for news texts which assigns to each event a calendrical date, date range, or open-ended date interval (i.e. the interval before or after a given date). They use one set of rules which apply to cases of explicit time reference (e.g. temporal PPs), and another set that apply when no implicit information is available.

A further position to take is that time-event relations should only be marked in cases where they are explicitly signaled or are syntactically implicit (as with Schilder and Habel), but that event-event temporal relations (to be discussed later) should also be marked, so that calendrical time-points for some events can be recovered by inference from combinations of time-event and event-event relations (so, for example, if e1 occurs at t and e2 occurs after e1 then we know e2 occurs after t). The approaches of both Li et al. (2004) and Pustejovsky et al. (2003) admit event-event relations to be tagged as well as time-event relations and hence support this sort of indirect positioning of events in time.

Before discussing the annotation of event-event relations in detail, it is worth con- sidering the time-stamping project in more detail. Time-stamping –by which we mean the assignment of a calendrical time reference (point or interval) to every event in running text– is an appealing aim. Motivating it is the intuition or wish, which is especially strong as concerns narrative texts such as newswires, that all events should be placeable on a time-line. This goal suggests that the target representation for a temporal annotator should be a mapping or anchoring of all events in a text on a calendrical time-line. Despite its intuitive appeal, time-stamping all events has serious drawbacks which stem ultimately from the fact that natural language narratives underspecify event positions in time in a way that makes a time-line representation problematic. Put another way, narratives may only specify a partial ordering between events; a time-line representation commits one to assigning a total ordering, information which simply may not be present in the text. This position is elaborated by Setzer and Gaizauskas (2002) who prefer a time-event graph, in which the nodes are times or events and the arcs are temporal relations, to a time-line as a target representation for temporal relation annotation.

As with time-event relations, event-event temporal relations may be conveyed explicitly or implicitly. The primary mechanism for explicit relation is the temporal conjunction, typically used to relate the event expressed in a subordinated section to one in a main section; for example: While chopping vegetables, John cut his finger or After the game John called Bob. As with time-event relations, event-event temporal relations are frequently expressed implicitly, relying on world or lexical semantic knowledge, or narrative convention and discourse interpretation. Katz and Arioso (2001) are interested in the temporal relations between events, as signaled by verbs, within single sentences. Their primary concern is the study of how temporal information is conveyed within sentences such as John kissed the girl he met at the party where there are no explicit temporal relational markers. Is, for example, our knowledge that the kissing took place after the meeting dependent on lexical semantic knowledge of these two verbs? or on the recognition of the syntactic structure of matrix and subordinate sections both with past tense verbs?

To answer this question they propose adding to a large corpus of syntactically annotated sentences further annotations which capture temporal relational information. This resource could then be used for the induction of the sort of knowledge needed to resolve questions of temporal ordering in implicit contexts. In their annotation scheme a human annotator adds labeled, directed edges between nodes in a graph which are the verbs in a syntactically annotated sentence. In addition to verb nodes, each sentence also has associated with it a node corresponding to its speech time. The edges represent temporal relations and the edge labels and direction specify the relation (their set of relations contains just the two relations of precedence and inclusion, though their duals are also available by reversing the directionality of an edge). As noted above, they do not consider event nominals.

While Katz and Arioso (2001) are concerned only with intrasentential temporal relations between verbs, the TimeML scheme proposed by Pustejovsky et al. (2003) aims to capture event-event temporal relations as completely as possible and in a way that will facilitate the development of time, event, and temporal relational tagging systems for use in applications such as question answering and summarization. To that end they propose an approach to relational tagging that allows event?vent relations be marked between any two event-denoting expressions. The approach relies on implementing a relational graph by using XML elements which consume no text but link, via pointers, XML elements surrounding event representatives and associate a relation type with the link. The set of relation types they employ are the thirteen proposed by Allen (1984). Note that these links, called `TLINK`s, can be asserted between any two event-denoting expressions (or between event and temporal referring expressions), regardless of whether or not they occur in the same sentence. In their model, single event

statements are related to times (i.e. placed in a temporal relation to a calendrical time- point), while in multiple event statements the events are related to each other, using one of Allen's (1984) thirteen temporal relations. Thus, like Katz and Arioso (2001), event-event relations are only marked within sentences. However, presumably event-event temporal relational information for events in separate sentences is available indirectly via the temporal relation of these single events to times on a time-line.

### I.4 Subordinating and aspectual relations

If one considers verbs as event signals and examines sentences with multiple verbal elements with a view to labeling their temporal relations, several problem cases soon emerge. Consider, for example, John might have kissed the girl he met at the party or John hoped to kiss the girl he met at the party (and did/did not). In neither case can we mark a temporal relation between kiss and met, because we do not know whether or not it occurred. These cases reveal that in contexts where verbs are modally subordinated, or occur as arguments in intensional constructions, they cannot straightforwardly be taken as denoting real events. However, there are some such contexts where the events the subordinated verbs denote are guaranteed to have occurred, such as John forgot that he had already paid the bill or John knew Bill had gone.

A further class of problem cases are those involving aspectual verbs, such as start, keep, which may signal the beginning, culmination, termination, or continuation of an activity, as in John started chopping vegetables or Sue kept talking. These verbs do not signal events distinct from the ones denoted by their verbal arguments, but rather draw attention to an aspect of these events. Attempting to assert a temporal relation between them, therefore, is problematic. These cases demonstrate that proposing to annotate temporal relations between all verbs within a sentence is not sensible. There are two other possibilities. One is to ignore them; the other is to annotate these verb-verb relations in some other way. Ignoring these contexts might have no impact on certain uses of temporal annotation, for example on Katz and Arioso's (2001) project of building an annotated corpus from which to induce the temporal profile of lexical items. For other applications, such as question answering or summarization, however, the ability to distinguish these contexts is certainly needed. Either to learn to ignore them, or to handle them appropriately, an annotation scheme for these contexts is desirable. This has been proposed in the ISO-TimeML specification, via the addition of two further sorts of relational links. Subordination links (`SLINK`s) are introduced to deal with cases of subordinate relations, and aspectual links (`ALINK`s) are introduced to deal with the cases of relations introduced by aspectual verbs.

# Annex J
(informative)
# Tools and templates

NOTE Written by Bran Boguraev

NOTE List of acronyms

- ACE: Automatic Content Extraction
- GATE: General Architecture for Text Engineering
- MUC: Message Understanding Conference
- TANGO: TimeML ANnotation Graphical Organizer
- TARSQI: Temporal Awareness and Reasoning Systems for Question Interpretation
- TERN: Time Exrepssion Rexcognition and Normalizaiton
- TERQAS: Temporal and Event Recognition for Question Answersing Systems
- TIDES: Translingual Information Detectiona and Extraction
- UIMA: Unstructured Information Management Architecture

Broadly speaking, there are two categories of tools in the domain of computational analysis of time. On the one hand, there are the *analytic tools*, developed primarily for the purpose of identifying and extracting time-related data from text. On the other hand, there are *annotation tools*, whose purpose is to assist with the preparation of annotated corpus data.

Recently—and especially where temporal analysis is concerned—the line between the two is becoming somewhat blurred. Automating the task of hitherto manual corpus creation is a growing enterprise. Analytic tools are being incorporated into annotation frameworks; conversely, some of the tool components built into an annotation framework can be used in the development and deployment of analysis engines. This kind of re-purposing of toolkit components is made possible by developments in text processing methodology where separation of a language model from the engine that interprets it—be it a statistical model for *e.g.* a core classification engine [?], or a symbolic grammar for a finite state device [?]—is strictly maintained, and further facilitated by emerging notions of pipelined, composable, and re-configurable text processing architectures such as GATE or UIMA [?, ?].

For expository purposes largely, we will keep the two categories separate in the remainder of this annex.

## J.1 Annotation tools and templates

At some level of generalization, there have been a few community-wide activities for creating annotated corpora with temporal markup. Initially, there was some temporal annotation within the context of the Message Understanding Conferences [?, ?]. More recently, the Translingual Information Detection and Extraction (TIDES) effort focused on what eventually emerged as TIMEX2: temporal expressions in documents [?, ?]. The TERQAS initiative (Temporal and Event Recognition for Question Answering Systems), in addition to developing the more general purpose markup language for time TimeML, produced a corpus annotated according to the TimeML annotation guidelines [?].

NOTE See `http://www.timeml.org/terqa/index.html`.

The TimeBank corpus is, at the time of writing, on its second revision cycle (TimeBank Version 1.2 is now available via the Linguistic Data Consortium).

NOTE See `http://www.ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2006T08`.

Meanwhile, some of the community was engaged in temporal information extraction in the context of the TERN initiative (Time Expression Recognition and Normalization, [?]), with the TERN corpus focusing on TIMEX2 annotation (cf. [?]) for the purposes of identifying and normalizing temporal expressions. Most recently, the ACE program (Automatic Content Extraction), in some of the latest tasks, defines focus on temporal expression analysis, coupled with some event identification and temporal linking.

NOTE See `http://www.nist.gov/speech/tests/ace/index.html`.

Specifically for TimeML-style annotation, three annotation toolkits have been either adapted (with suitable templates developed for TimeML-related tags), or developed specifically.

### J.1.1 The ALEMBIC workbench

The ALEMBIC workbench is a general purpose annotation tool developed at MITRE. It has been used for a number of annotation projects, including some of the annotation tasks under the TIDES initiatives.

NOTE See `http://www.mitre.org/tech/alembic-workbench/`.

For TimeML-style annotations, in the process of preparing TimeBank 1.1, *task definitions* (together with accompanying *tagset definitions*) were developed especially for the purpose. EVENTs, TIMEX3's and SIGNALs can be annotated within the base ALEMBIC paradigm, by manipulating text extents. For temporal links, ALEMBIC's table-based annotation was used; tables are the organizational device which makes it possible to relate tags (*e.g.* for EVENTs and/or TIMEX3's) to each other.

NOTE Originally, table-based relation specification was introduced specifically for the purposes of annotating co-reference chains; over time, the mechanism was developed sufficiently to generalize enough so it could be used for the annotation of temporal links as well; this is indicative of some of the development history of ALEMBIC, as it was driven by specific annotation project requirements.

### J.1.2 The CALLISTO toolkit

The CALLISTO annotation toolkit was developed as a rational generalization of the annotation methodology and design ideas behind the ALEMBIC workbench. In particular, it aims to support a broad range of linguistically motivated annotations, it rationalizes the interface design for supporting basic annotation tasks (both for entity markup and relational linking), and it remains open-ended—by means of adopting a plugin-based architecture and stand-off annotations—with respect to task definitions. Importantly (at least from the point of view of a multi-lingual project like TimeML definition), it supports annotation for *any* Unicode-supported language.

CALLISTO has been 'template'-ized for TimeML annotation by means of defining a TimeML task. Thus the general annotator interface offering tag editing capabilities through a highlighted text display, tag attribute tables, and relation-argument association tables, has been adapted to the extent that TimeML-style annotation could be applied to a corpus like TimeBank; indeed, the re-annotation of TimeBank for Version 1.2 was carried out within a CALLISTO environment.

### J.1.3 The TANGO temporal relation editor

Annotating ISO-TimeML is an example of a 'dense' annotation task: there is a very large number of (temporal) relations which need specifying, among arguments which range from local to very long-term spans. There are some overheads in the cognitive mapping between, say a TLINK with its TIMEX3 and/or EVENT arguments, and their tabular representations (in *e.g.* CALLISTO), which may be many rows apart, and requiring the manipulation of indexed objects not perspicuously linked to their textual counterparts. For temporal relation markup, in particular, a dedicated effort following the release of the first version of TimeBank focused on developing a custom annotation tool.

TANGO, a TimeML ANnotation Graphical Organizer [?], alleviates these problems by means of a special-purpose interface which lays out a visual 'map' of all temporal objects in the document and provides direct manipulation mechanisms for linking any selected such objects.

NOTE See also http://timeml.org/site/tango/index.html.

The TANGO functionality can be seamlessly integrated on top of CALLISTO's base annotation functions (thanks to CALLISTO's modular and extensible design)—thus the combination of the two is a particularly powerful annotation tool for TimeML.

TANGO is only an initial exploration into the space of visual support for the display and manipulation of temporal information. Recent work looks at alternative ways of presenting such complex and interconnecting relationships, both for the purposes of developing felicitous cognitive models of the task of temporal analysis, and for facilitating corpus annotation with temporal markup.

## J.2 Analytic tools

Prior to TimeML, a particular focus of computational analysis of time was that of the identification and normalization of temporal expressions. This was, in fact, the focus of the TERN evaluation already mentioned [?]; prior to that, a pivotal effort in that area was the work by Mani & Wilson [?], which was offered to the community (via a Web-based interface) as one of the earliest broadly available analytic tools.

The range of research efforts looking at temporal analysis is too broad for this annex, and by no means all of them can be viewed as 'tools'. We will focus here on a representative sample of relatively self-contained analytic components.

### J.2.1 The TARSQI toolkit

The TARSQI project has created a series of tools for temporal information extraction [?].

NOTE See http://www.timeml.org/tarsqi.

All of them can be used as stand-alone programs that automatically identify TimeML tags in a document; their design, however, makes it possible for them to be composed in a sequence. Thus, the project defines, in effect, a 'reconfigurable' TimeML analyzer, whose components can be individually, and incrementally, developed and enhanced.

Verhagen *et al.* [?] describe component analyzers:

- GUTIME extracts normalizes TIMEX2 expressions and instantiates normalized values; more recently, it has been extended to handle expressions also based on TimeML's TIMEX3 definition;
- EVITA is a robust event recognizer, which recognizes TimeML EVENTs, and adds CLASS, TENSE, and ASPECT attribute tags;
- GUTENLINK (recently renamed to Blinker) is a TLINK parser: it uses 187 syntactic and lexical rules to infer and label TLINKs between tagged events and other tagged events or times.
- SLINKET is a partial modal parser for for SLINKs.

More recently, this initial set of tools has been augmented by

- an S2T module, which creates new TLINKs from the SLINKs provided by SLINKET, using several simple rules,
- an event duration analyzer [?], and
- a TLINK classifier which automatically identifies temporal relations (TLINKs) between already tagged EVENTs in text [?].

### J.2.2 The IBM TimeML annotator

Somewhat in contrast to the modular approach of TARSQI, Boguraev and Ando [?] cast the entire task of TimeML analysis as an information extraction task, with, broadly speaking, TIMEX3's and EVENTs being considered as named entities, and LINKs as relations among them. That work targets the full temporal mark-up language—seeking to extract both temporal expressions and events, and further looking for temporal relations (TLINKs).

The design is that of a hybrid TimeML annotator (realized as a UIMA text analysis engine [?]), which is trained on TimeBank, and deploys a hybrid analytical strategy of mixing aggressive finite-state processing over linguistic annotations with a state-of-the-art machine learning technique capable of leveraging large amounts of unannotated data.

### J.2.3 The Amsterdam temporal component extractor

While not directly targeting TimeML, Ahn *et al.* [?] develop a framework for machine learning of temporal expression recognition in ways which directly enhances the normalization problem. More recently, the insights from such a methodology are applied to the complementary task of (ACE) EVENT detection and recognition [?].

### J.2.4 The Time Calculus analyzer

Han *et al.* go one step further than developing and evaluating a tool for temporal analysis: in addition to a complete implementation of a temporal expression recognizer and normalizer, they define a constraint-based representation of time within a self-contained *Time Calculus* framework [?]. This is used in an application (time-based anchoring of e-mails) which demonstrates the representational and functional completeness of a versatile temporal analytic tool.

As mentioned earlier, some of the analytical tools developed for time analysis are beginning to find their way in corpus annotation frameworks. A particularly good example of such a shift can be found in the context of the TARSQI project, where tagging functionalities for TimeML components are used within a purpose-designed annotation environment for an evaluation task of time extraction (TempEval).

NOTE See `http://www.timeml.org/tempeval`.

This also illustrates the more general state of affairs with respect to infrastructure base of an emerging representational framework for analysis of time: ISO-TimeML has become more mature, more stable, more tractable, and more broadly utilized.

# Annex K
(informative)
# Editors, contributors and meetings

## K.1 Editors

As part of ISO/TC 37/SC 4/WG 2 activities, this project, entitled ISO 24617-1 Semantic annotation framework (SemAF) - Part 1: Time and events, was carried out by the following group of editors:


- Kiyong Lee (KATS)
  WG 2 convenor and project leader
  Korea University
  klee@korea.ac.kr
- James Pustejovsky (ANSI/TAG)
  head of the TimeML Group
  ANSI/TAG, Brandeis University
  jamesp@cs.brandeis.edu
- Harry Bunt (NEN)
  head of the LIRICS group
  Tilburg University
  Harry.Bunt@uvt.nl
- Branimir Boguraev (ANSI/TAG)
  The IBM Thomas J. Watson Research Center
  bran@us.ibm.com
- Nancy Ide (ANSI/TAG)
  Vassar College
  ide@cs.vassar.edu


## K.2 Contributors

The following lists main contributors excluding the editors.

- Tommaso Caselli, ILC-CNR, Pisa, Italy
- Yuzhen Cui, City University of Hong Kong
- Alex Chengyu Fang, City University of Hong Kong
- Roland Hausser, Erlangen-Nuernberg University, Germany
- Manigo Kit, City University of Hong Kong
- Rainer Osswald, University of Hagen, Germany
- Haihua Pan, City University of Hong Kong
- Ian Pratt-Harman, Universiity of Manchester, UK
- Amanda Schiffrin, Tilburg University, The Netherlands

### K.3 Meetings

(1) 2006-04-20/22
At the TDG 3 and LIRICS Working Group Meeting, USC/ISI, Marina del Rey, CA, U.S.A., the ISO/TC 37/SC 4/WG 2 (Convenor: Kiyong Lee) and the SemAFTimeML Working Group (headed by James Pustejovsky) agreed on joint activities on establishing an ISO standard on temporal annotation, entitled ISO NP 2461x-1 Semantic annotation framework - Part 1: Time and event (SemAF/Time), proposing an editorial group consisting of James Pustejovsky, Harry Bunt, Branimir Boguaraev, Nancy Ide and Kiyong Lee.

(2) 2006-08-24
The ISO/TC 37/SC 4 plenary meeting in Beijing approved the New Project SemAF/Time and nominated the proposed editorial group with Kiyong Lee acting as project leader.

(3) 2006-10-26/28
The first working group meeting was held at Brandeis University, Waltham, MA, U.S.A., with 21 experts participating, and the drafting task was divided and distributed among the experts with its submission due 2006-12-24.

(4) 2007-01-08/10 The second working group meet was helt at Tilburg University in conjunction with IWCS-7. Here, the general structure of the document was decided on with the addition of Clause 8 Semantics of ISO-TimeML. Ian Pratt-Harman was asked to write the first draft with the help of James Pustejovsky, Harry Bunt, Rainer Osswald and Kiyong Lee.

(5) 2007-05-07/09
The third working group meeting was held at AFNOR, Paris, in conjunction with the GL-2007 conference (2007-05-10/12). The first draft was reviewed at the meeting. It was agreed that the finalized document be submitted for the first CD ballot by the end of June, 2007.

(6) 2007-08-17
The ISO/TC 37/SC 4 resolved that the document be submitted for the first CD ballot by 2007-09-15. Most recent edits are by James Pustejovsky and Kiyong Lee.

NOTE The bibliography need be rechecked.

# Bibliography


[1] ADVANCED RESEARCH PROJECTS AGENCY. *Proceedings of the Sixth Message Understanding Conference (*MUC-6*)* (1995), Software and Intelligent Systems Technology Office.

[2] ADVANCED RESEARCH PROJECTS AGENCY. *Proceedings of the Seventh Message Understanding Conference (*MUC-7*)* (1998), Software and Intelligent Systems Technology Office.

[3] AHN, D. The stages of event extraction. In Boguraev et al. [?], pp. 1–8. Proceedings of a Coling-ACL Workshop.

[4] AHN, D., ADAFRE, S. F., AND DE RIJKE, M. Towards task-based temporal extraction and recognition. In Katz et al. [?]. <http:// drops.dagstuhl.de/ opus/volltexte/2005/ 318> [date of citation: 2006-12-01].

[5] ALLEN, J. Maintaining knowledge about temporal intervals. *Communications of the ACM 26*, 11 (1983).

[6] ALLEN, J. A general model of action and time. *Artificial Intelligence 23*, 2 (1984).

[7] ALLEN, J. Time and time again: The many ways to represent time. *International Journal of Intelligent Systems 6*, 4 (July 1991), 341–356.

[8] ALLEN, J., AND FERGUSON, G. Actions and events in interval temporal logic,. *Journal of Logic and Computation 4*, 5 (1994).

[9] ALLEN, J., AND HAYES, P. Moments and points in an interval-based temporal logic. *Computational Inteligence* (Jan. 1990).

[10] AMERICAN ASSOCIATION FOR ARTIFICIAL INTELLIGENCE. *New Directions in Question-Answering (Working Papers)* (Stanford, CA, March 2003), Spring Symposium Series.

[11] ANDO, R. K. Exploiting unannotated corpora for tagging and chunking. In *Proceedings of ACL-04* (2004).

[12] ANDO, R. K. Semantic lexicon construction: Learning from unlabeled data via spectral analysis. In *Proceedings of CoNLL-04* (2004).

[13] ANDROUTSOPOULOS, I. Temporal meaning representations in a natural language front-end. *CoRR cs.CL/9906020* (1999).

[14] ANDROUTSOPOULOS, I., RITCHIE, G. D., AND THANISCH, P. Time, tense and aspect in natural language database interfaces. *CoRR cmp-lg/9803002* (1998).

[15] ASHER, N., AND LASCARIDES, A. *Logics of Conversation*. Cambridge University Press, Cambridge, England, 2003.

[16] BACH, E. The algebra of events. *Linguistics and Philosophy*, 9, 5–16. (1986).

[17] BERTNETTO, P.M. Tempo, Aspetto e Azione nel verbo Italiano Accademia della Crusca, Firenze, 1986.

[18] BERTNETTO, P.M. Il Verbo In *Grande Grammatica Italiana di Consultazione* L. Renzi-G. Salvi (a cura di), Il muliono, 13-163, 1991.

[19] BERTNETTO, P.M. Suelle proprieta tempo-aspettuali dell'Infinito italiano available at http://alphalinguistica.sns.it/QLL/QLL01/PMB.Infinito.pdf 2001.

[20] BOGURAEV, B. Towards finite-state analysis of lexical cohesion. In *Proceedings of the 3rd International Conference on Finite-State Methods for NLP,* INTEX-3 (Liege, Belgium, June 2000)..

[21] BOGURAEV, B., AND ANDO, R. K. TimeBank-driven TimeML analysis. In Katz et al. [?]. <http:// drops.dagstuhl.de/ opus/volltexte/2005/ 318> [date of citation: 2006-12-01].

[22] BOGURAEV, B., AND ANDO, R. K. TimeML-compliant text analysis for temporal reasoning. In *Nineteenth International Joint Conference on Artificial Intelligence (IJCAI-05)* (Edinburgh, Scotland, 2005).

[23] BOGURAEV, B., AND ANDO, R. K. Analysis of TimeBank as a resource for TimeML parsing. In *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC-2006)* (Genoa, Italy, May 2006).

[24] BOGURAEV, B., AND NEFF, M. An annotation-based finite state system for UIMA: User documentation and grammar writing manual. Tech. rep., IBM T.J. Watson Research Center, Yorktown Heights, New York, Dec. 2006.

[25] BOGURAEV, B., PUSTEJOVSKY, J., ANDO, R., AND VERHAGEN, M. Evolution of TimeBank as a community resource for TimeML parsing. *Language Resources and Evaluation* (2007). Forthcoming.

[26] BOGURAEV, B., PUSTEJOVSKY, J., AND MUÑOZ, R., Eds. *ARTE: Annotating and Reasoning about Time and Events* (Sydney, Australia, July 2006), The Association for Computational Linguistics. Proceedings of a Coling-ACL Workshop.

[27] BRISCOE, T., AND CARROLL, J. Robust accurate statistical annotation of general text. *Proceedings of the Third International Conference on Language Resources and Evaluation (LREC 2002), Las Palmas, Canary Islands, May 2002* (2002), 1499–1504.

[28] BUNT, H. *Mass Terms and Model-theoretic Semantics*, Cambridge University Press. 1985.

[29] BUNT, H. AND C. OVERBEEKE An Extensible Compositional Semantics for Temporal Annotation. In: N. & et al., *Proceedings of LAW II, the Second Workshop on Linguistic Annotation*, Marrakech, May 2008.

[30] CARRERAS, X., AND MÀRQUEZ, L. Phrase recognition by filtering and ranking with perceptrons. In *Recent Advances in Natural Language Processing III*, N. N. et al., Ed., vol. 260 of *Current Issues in Linguistic Theory (CILT)*. John Benjamins, Amsterdam/Philadelphia, 2004, pp. 205–216..

[31] CHANG, SUK-JIN Korean, John Benjamins, Amsterdam/Philadelphia, 1996.

[32] CHKLOVSKI, T., AND PANTEL, P. VerbOcean: Mining the web for find-grained semantic verb relations. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing* (Barcelona, Spain, 2004).

[33] COMRIE 1986.

[34] CROUCH, D., SAURÍ, R., AND FOWLER, A. Aquaint pilot knowledge-based evaluation: Annotation guidelines. Tech. rep., Palo Alto Research Center, May 2005.

[35] CUNNINGHAM, H., Information exxtraction, automatic *Encyclopedia of Language and Linguistics* 2nd Edition, Elsevier, 2005.

[36] DALE, R., AND MAZUR, P. Local semantics in the interpretation of temporal expressions. In Boguraev et al. [?], pp. 9–16. Proceedings of a Coling-ACL Workshop.

[37] DARPA TIDES (TRANSLINGUAL INFORMATION DETECTION, EXTRACTION AND SUMMARIZATION). The TERN evaluation plan; time expression recognition and normalization. Working papers, TERN Evaluation Workshop, 2004. <http://timex2.mitre.org/tern.html> [date of citation: 2005-07-12].

[38] DAY, D., ABERDEEN, J., HIRSCHMAN, L., KOZIEROK, R., ROBINSON, P., AND VILAIN, M. Mixed-Initiative Development of Language Processing Systems. In *Fifth Conference on Applied Natural Language Processing Systems* (Washington D.C., U.S.A., 1997), pp. 88–95.

[39] DECHTER, R., MEIRI, I., AND PEARL, J. Temporal constraint networks. *Artificial Intelligence 49* (1991), 61–95.

[40] DERSHOWITZ, N., AND REINGOLD, E. *Calendrical Calculations*. Cambridge University Press, Cambridge, UK, 1997.

[41] DOWTY, D. 1979.

[42] FERRO, L. TIDES: Instruction manual for the annotation of temporal expressions. Tech. Rep. MTR 01W0000046V01, The MITRE Corporation, 2001.

[43] FERRO, L. ET AL.

[44] FERRO, L., GERBER, L., MANI, I., AND WILSON, G. TIDES 2003 standard for the annotation of temporal expressions. Tech. rep., The MITRE Corporation, 2004.

[45] FERRUCCI, D., AND LALLY, A. UIMA: an architectural approach to unstructured information processing in the corporate research environment. *Natural Language Engineering 10*, 4 (2004). Special Issue on Software Architectures for Language Engineering.

[46] FIKES, R., JENKINS, J., AND FRANK, G. JTP: A system architecture and component library for hybrid reasoning. Tech. Rep. KSL-03-01, Knowledge Systems Laboratory, Stanford University, 2003.

[47] FIKES, R., JENKINS, J., AND ZHOU, Q. Including domain-specific reasoners with reusable ontologies. Tech. Rep. KSL-03-05, Knowledge Systems Laboratory, Stanford University, Stanford, CA, 2003.

[48] FILATOVA, E., AND HOVY, E. Assigning time-stamps to event-sections. In *Proceedings of the 10th Conference of the European Chapter of the ACL* (Toulouse, France, 2001).

[49] FLORIAN, R., ITTYCHERIAH, A., JING, H., AND ZHANG, T. Named entity recognition through classifier combination. In *Proceedings of CoNLL-2003* (2003).

[50] FORĂSCU, C., PISTOL, I. C., AND CRISTEA, D. Temporality in relation with discourse structure. In *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC-2006)* (Genoa, Italy, May 2006).

[51] FORASCU, C., AND SOLOMON, D. Towards a Time Tagger for Romanian. In *Proceedings of the 9th ESSLLI Student Session* (2004), P. Egre and L. Alonso, Eds.

[52] FORBES, K., MILTSAKAKI, E., PRASAD, R., SAKAR, A., JOSHI, A., AND WEBBER, B. D-LTAG system: Discourse parsing with a lexicalized tree adjoining grammar. In *Proceedings of the ESSLLI 2001: Workshop on Information Structure, Discourse Structure and Discourse Semantics* (2001).

[53] FREKSA, C. Temporal reasoning based on semi-intervals. *Artificial Intelligence 54* (1992), 199–227.

[54] GAIZAUSKAS, R., HARKEMA, H., HEPPLE, M., AND SETZER, A. Task-oriented extraction of temporal information: The case of clinical narratives. In Montanari et al. [?].

[55] GAIZAUSKAS, R., AND SETZER, A., Eds. *Annotation Standards for Temporal Information in NL* (Las Palmas, Spain, 2002).

[56] GIMÉNEZ, J., AND MÀRQUEZ, L. Fast and accurate part-of-speech tagging; The SVM approach revisited. In *Recent Advances in Natural Language Processing III*, N. N. et al., Ed., vol. 260 of *Current Issues in Linguistic Theory (CILT)*. John Benjamins, Amsterdam/Philadelphia, 2004, pp. 153–162..

[57] GROVER, C., HITZEMAN, J., AND MOENS, M. Algorithms for analysing the temporal structure of discourse. In *Sixth International Conference of the European Chapter of the ACL* (Dublin, Ireland, 1995).

[58] HAN, B., GATES, D., AND LEVIN, L. Anchoring temporal expressions in scheduling-related Emails. In Katz et al. [?]. <http:// drops.dagstuhl.de/ opus/volltexte/2005/ 318> [date of citation: 2006-12-01].

[59] HAN, B., GATES, D., AND LEVIN, L. From language to time: a temporal expression anchorer. In Montanari et al. [?].

[60] HAN, B., AND KOHLHASE, M. A time calculus for natural language. In *The 4th Workshop on Inference in Computational Semantics* (Nancy, France, September 2003).

[61] HAN, B., AND LAVIE, A. A framework for resolution of time in natural language. *TALIP Special Issue on Spatial and Temporal Information Processing 3*, 1 (2004), 11–35.

[62] HAN, ET AL. 2005.

[63] HARPER, L., MANI, I., AND SUNDHEIM, B., Eds. *Workshop for Temporal and Spatial Information Processing, ACL-EACL'2001* (Toulouse, France, July 2001).

[64] HINDELY AND SELDIN 1986.

[65] HITZEMAN, J. Text type and the position of a temporal adverbial within the sentence. In Katz et al. [?]. <http:// drops.dagstuhl.de/ opus/volltexte/2005/ 318> [date of citation: 2006-12-01].

[66] HOBBS, J. On the coherence and structure of discourse. In *CSLI Technical Report 85-37* (Stanford, CA, USA, 1985), Center for the Study of Language and Information.

[67] HOBBS, J., FERGUSON, G., ALLEN, J., HAYES, P., AND PEASE, A. A DAML ontology of time, 2002.

[68] HOBBS, J., AND PAN, F. An ontology of time for the semantic web. *TALIP Special Issue on Spatial and Temporal Information Processing 3*, 1 (2004), 66–85.

[69] HOBBS, J., AND PUSTEJOVSKY, J. Annotating and reasoning about time and events. In *AAAI Spring Symposium on Logical Formalizations of Commonsense Reasoning* (Stanford, CA, March 2004).

[70] INGRIA, R., AND PUSTEJOVSKY, J. TimeML Specification 1.0. available at هتّپ://تِم۲۰۰۲اورگ *http://time2002org*, 2002.

[71] HOOPER, J. B. On assertive predicates. In *Syntax and semantics, IV*, J. Kimball, Ed. Academic Press, New York, 1975, pp. 91–124.

[72] ISO COMMITTEE. Information interchange: representation of dates and times. Tech. Rep. ISO 8601, International Organization for Standardization, 1997.

[73] JANG, S.B., BALDWIN, J., AND MANI, I. Automatic TIMEX2 Tagging of Korean News. *ACM Transactions on Asian Language Information Processing*, 3.1, (2004) pp. 51–65

[74] JESPERSON, J., AND FITZ-RANDOLPH, J. *From sundials to atomic clocks: understanding time and frequency*. Dower Publications, Inc., New York, 1977.

[75] KARTTUNEN, L. Implicative verbs. In *Language* (1970), vol. 47, pp. 340–358.

[76] KARTTUNEN, L. Some observations on factivity. In *Papers in Linguistics* (1971), vol. 4, pp. 55–69.

[77] KARTTUNEN, L., CHANOD, J.-P., GREFENSTETTE, G., AND SCHILLER, A. Regular expressions for language engineering. *Natural Language Engineering 4*, 1 (1996), 305–328.

[78] KATZ AND ARIOSO 2001.

[79] KATZ, G., PUSTEJOVSKY, J., AND SCHILDER, F., Eds. *Annotating, Extracting and Reasoning about Time and Events* (Dagstuhl, Germany, 2005), no. 05151 in Dagstuhl Seminar Proceedings, Internationales Begegnungs- und Forschungszentrum fuer Informatik (IBFI), Schloss Dagstuhl, Germany. <http:// drops.dagstuhl.de/ opus/volltexte/2005/ 318> [date of citation: 2006-12-01].

[80] KILLGARRIF, A., RYCHLY, P., SMRZ, P., AND TUGWELL, D. The sketch engine. In *Proceedings of Euralex, Lorient, France* (2004), pp. 105–116.

[81] KIPARSKY, P., AND KIPARSKY, C. Fact. In *Progress in Linguistics. A Collection of Papers*, M. Bierwisch and K. E. Heidolph, Eds. Mouton, The Hague, Paris, 1970, pp. 143–173.

[82] KOEN, D., AND BENDER, W. Time Frames: Temporal augmentation of the news. *IBM Systems Journal 39*, 3/4 (2000), 597–616.

[83] MANI, I. AND WILSON, G. Robust temporal processing of news *Proceedings of the aCL '2000 Conference*(October 2000, Hong Kong), 69-76.

[84] KUDO, T., AND MATSUMOTO, Y. Chunking with Support Vector Machines. In *Proceedings of the Second Meeting of North American Chapter of Association for Computational Linguistics (NAACL-2001)* (Pittsburgh, June 2001), pp. 192–199.

[85] LAPATA, M., AND LASCARIDES, A. Inferring sentence-internal temporal relations. In *HLT-NAACL* (2003).

[86] LAVELLI, A., MAGNINI B., NEGRI, M., PAINATA, E., SPERANZA, M., AND SPRUGNOLI, R. *Italian Content Annotation Bank (I-CAB): Temporal Expressions (V.1.0)* ITC First Technical Report. 2005.

[87] LEE, CHUNGMIN Temporal expressions in Korean. In Verschueren and Bertucceoo8-Papi (eds.), *The Pragmatic Perspective: Selected Papers from 1985 International Conference on Pragmatics* (John Benjamins, Amsterdam/Philadephia), 1987, pp.435-453.

[88] LEE, IKSOP, AND RAMSEY S. *The Korean Language*. State University of New York Press, Albany, NY, 2000.

[89] LEE, KIYONG A situation-theoretic analysis of the so-called retrospective aspect in Korean In Proceedings of the First International Conference on Linguistics at Chosun University, 1993.

[90] LEE, KIYONG *Tense and Modality: Possible Worlds Semantics*(written in Korean) Korea University Press, Seoul, 1998.

[91] LEE, KIYONG *Korean Morphology* (written in Korean) Korea University Press, Seoul, 1999.

[92] LEE, KIYONG Multilinguality in Temporal Annotation: A Case of Korean. In *Proceedings of PACLIC 20*(Wuhan, China, November 2006.)

[93] LEE, K., PUSTEJOVSKY, J., AND BOGURAEV, B. Towards an international standard for annotating temporal information. In *Third International Conference on Terminology, Standardization and Technology Transfer* (Beijing, China, August 2006).

[94] LEVIN, B. *English verb classes and alternations*. The University of Chicago Press, Chicago and London, 1993.

[95] LI, W., CAO, G., WONG, K.-F., AND YUAN, C. Applying machine learning to chinese temporal relation resolution. In *42nd Meeting of the Association for Computational Linguistics* (Barcelona, Spain, 2004).

[96] LIN, D. Automatic retrieval and clustering of similar words. In *Proceedings of COLING-ACL* (Montreal, Canada, 1998).

[97] LIN, D. 2003.

[98] MANI, I. Recent developments in temporal information extraction. In *Recent Advances in Natural Language Processing*, G. Angelova, K. B. N. Nicolov, and R. Mitkov, Eds. John Benjamins, 2004.

[99] MANI, I., PUSTEJOVSKY, J., AND SUNDHEIM, B. Introduction: special issue on temporal information processing. *ACM Transactions Asian Language Information Processing 3*, 1 (March 2004), 1–10.

[100] MANI, I., PUSTEJOVSKY, J., AND GAIZAUSKAS, R. (eds.) *The Language of Time*, Oxford University Press, 2005.

[101] MANI, I., SCHIFFMAN, B., AND ZHANG, J. Inferring temporal ordering of events in news. In *Human Language Technology/North American Chapter of the ACL (HLT-NAACL* (Edmonton, Canada, 2003).

[102] MANI, I., AND WELLNER, B. A pilot study on acquiring metric temporal constraints for events. In Boguraev et al. [?], pp. 23–29. Proceedings of a Coling-ACL Workshop.

[103] MANI, I., WELLNER, B., VERHAGEN, M., LEE, C. M., AND PUSTEJOVSKY, J. Machine learning of temporal relations. In *Proceedings of the 44th Annual Meeting of the Association for Computational Linguistics* (Sydney, 2006), ACL.

[104] MANI, I., AND WILSON, G. Robust temporal processing of news. In *Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics* (Hong Kong, 2000), ACL, pp. 69–76.

[105] MAYS, E., LANKA, S., JOSHI, A. K., AND WEBBER, B. L. Natural language interaction with dynamic knowledge bases: monitoring as response. In *European Conference on Artificial Intelligence* (1982), pp. 255–256.

[106] MILTSAKAKI, E., DINESH, N., PRASAD, R., JOSHI, A., AND WEBBER, B. Experiments on sense annotation and sense disambiguation of discourse connectives. In *Proceedings of the Fourth Workshop on Treebanks and Linguistic Theories (TLT 2005)* (Barcelona, Catalonia, 2005).

[107] MONTANARI, A., PUSTEJOVSKY, J., AND REVESZ, P., Eds. *TIME 2006: International Symposium on Temporal Representation and Reasoning* (Budapest, Hungary, June 2006).

[108] NEGRI, M., SAQUETE, E., MARTINEZ-BARCO, P., AND MUÑOZ, R. Evaluating knowledge-based approaches to the multilingual extension of a temporal expression normalizer. In Boguraev et al. [?], pp. 30–37. Proceedings of a Coling-ACL Workshop.

[109] NGAI, G., AND YAROWSKY, D. Rule Writing or Annotation: Cost-efficient Resource Usage for Base Noun Phrase Chunking. In *Proceedings of the Second Meeting of North American Chapter of Association for Computational Linguistics (ACL-2000)* (Hong Kong, 1-8 October 2000).

[110] PAN, F., AND HOBBS, J. R. Temporal arithmetic mixing months and days. In Montanari et al. [?].

[111] PAN, F., MULKAR, R., AND HOBBS, J. R. An annotated corpus of typical durations of events. In *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC-2006)* (Genoa, Italy, May 2006), pp. 77–83.

[112] PAN, F., MULKAR, R., AND HOBBS, J. R. Extending timeml with typical durations of events. In Boguraev et al. [?], pp. 38–45. Proceedings of a Coling-ACL Workshop.

[113] PAN, F., MULKAR, R., AND HOBBS, J. R. Learning event durations from event descriptions. In *Proceedings of the 44th Annual Meeting of the Association for Computational Linguistics (COLING-ACL)* (Sydney, Australia, 2006), pp. 393–400.

[114] PARSONS, T. *Events in the Semantics of English*. MIT Press, Cambridge, MA, 1990.

[115] PRAGER, J., CHU-CARROLL, J., BROWN, E., AND CZUBA, C. Question answering using predictive annotation. In *Advances in Question Answering*, 2003. To appear.

[116] PRIOR, A.N. (1967) *Past, present and Future*, The Clarendon Press, Oxford, 1967.

[117] PRATT-HARTMANN, I. From TimeML to TPL. In Katz et al. [?]. <http:// drops.dagstuhl.de/ opus/volltexte/2005/ 318> [date of citation: 2006-12-01].

[118] PUSTEJOVSKY, J. Events and the semantics of opposition. In *Events as Grammatical Objects*, C. Tenny and J. Pustejovsky, Eds. CSLI Publications, 2000.

[119] PUSTEJOVSKY, J., BELANGER, L., CASTAÑO, J., GAIZAUSKAS, R., HANKS, P., INGRIA, B., KATZ, G., RADEV, D., RUMSHISHKY, A., SANFILIPPO, A., SAURÍ, R., SETZER, A., SUNDHEIM, B., AND VERHAGEN, M. TERQAS Final Report. Tech. rep., The MITRE Corporation, Bedford, Massachusetts, 2002.

[120] PUSTEJOVSKY, J., CASTAÑO, J., INGRIA, R., SAURÍ, R., GAIZAUSKAS, R., SETZER, A., KATZ, G., AND RADEV, D. TimeML: Robust specification of event and temporal expressions in text. In *AAAI Spring Symposium on New Directions in Question-Answering (Working Papers)* (Stanford, CA, 2003), pp. 28–34.

[121] PUSTEJOVSKY, J., HANKS, P., SAURÍ, R., SEE, A., GAIZAUSKAS, R., SETZER, A., RADEV, D., SUNDHEIM, B., DAY, D., FERRO, L., AND LAZO, M. The TIMEBANK corpus. In *Corpus Linguistics* (Lancaster, 2003), T. McEnery, Ed., pp. 647–656.

[122] PUSTEJOVSKY, J. ET AL. 2004.

[123] PUSTEJOVSKY, J., HAVASI, C., SAURÍ, R., HANKS, P., RUMSHISKY, A., LITMAN, J., CASTAÑO, J., AND VERHAGEN, M. Towards a Generative Lexical resource: The Brandeis Semantic Ontology. In *Language Resources and Evaluation Conference, LREC 2006* (Genoa, Italy, 2006).

[124] PUSTEJOVSKY, J., KNIPPEN, R., LITTMAN, J., AND SAURÍ, R. Temporal and event information in natural language text. *Language Resources and Evaluation 39*, 2–3 (Mar. 2005), 123–164.

[125] PUSTEJOVSKY, J., LITTMAN, J., AND SAURÍ, R. Argument structure in TimeML. In Katz et al. [?]. <http:// drops.dagstuhl.de/ opus/volltexte/2005/ 318> [date of citation: 2006-12-01].

[126] PUSTEJOVSKY, J., KNIPPEN, R., LITTMAN, J., AND SAURÍ, R. Temporal and event information in natural language text. In H. Bunt and R. Muskens (eds.) *Computing Meaning, Volume 3*.. Springer, Dordrecht, 301–346, 2007.

[127] PUSTEJOVSKY, J., MANI, I., BÉLANGER, L., BOGURAEV, B., KNIPPEN, B., LITTMAN, J., RUMSHISKY, A., SEE, A., SYMONENKO, S., GUILDER, J. V., GUILDER, L. V., VERHAGEN, M., AND INGRIA, R. Graphical annotation kit for TIMEML. Tech. rep., TANGO (TIMEML Annotation Graphical Organizer) Workshop, 2003. Version 1.4, <http://www.timeml.org/tango> [date of citation: 2005-06-20].

[128] RABINER, L. R. A tutorial on Hidden Markov Models and selected applications in speech recognition. In *Proceedings of the IEEE* (1989), vol. 77, pp. 256–286.

[129] RAMSHAW, L., AND MARCUS, M. Text Chunking Using Transformation-Based Learning. In *Proceedings of the Third Workshop on Very Large Corpora* (Somerset, New Jersey, 1995), D. Yarovsky and K. Church, Eds., Association for Computational Linguistics, pp. 82–94.

[130] REICHENBACH, H. *Elements of Symbolic Logic*. Macmillan, New York, 1947.

[131] SAMUELSSON, C. Extending N-gram tagging to word graphs. In *Recent Advances in Natural Language Processing II*, N. Nicolov and R. Mitkov, Eds., vol. 189 of *Current Issues in Linguistic Theory (CILT)*. John Benjamins, Amsterdam/Philadelphia, 2000, pp. 3–20.

[132] SANG, E. F. T. K., AND VEENSTRA, J. Representing Text Chunks. In *Proceedings of European Chapter of the Association for Computational Linguistics* (1999), Bergen, Norway, pp. 173–179.

[133] SAQUETE, E. *Temporal information resolution applied to Question Answering*. PhD thesis, Universidad de Alicante, Alicante, Spain, Feb. 2005.

[134] SAQUETE, E., MARTINEZ-BARCO, P., MUÑOZ, R., NEGRI, M., SPERANZA, M., AND SPRUGNOLI, R. Automatic resolution rule assignment to multilingual temporal expressions using annotated corpora. In Montanari et al. [?].

[135] SAQUETE, E., MART?EZ-BARCO, P., AND MUÑOZ, R. Recognizing and tagging temporal expressions in Spanish. In *Workshop on Annotation Standards for Temporal Information in Natural Language, LREC-2002* (2002), pp. 44–51.

[136] SAQUETE, E., MART?EZ-BARCO, P., MUÑOZ, R., AND VICEDO, J. Splitting complex temporal questions for question answering systems. In *42nd Meeting of the Association for Computational Linguistics* (Barcelona, Spain, 2004).

[137] SAURÍ, R., KNIPPEN, R., VERHAGEN, M., AND PUSTEJOVSKY, J. Evita: A robust event recognizer for qa systems. In *Proceedings of the HLT/EMNLP 2005* (2005), pp. 700–707.

[138] SAURÍ, R., KNIPPEN, R., VERHAGEN, M., AND PUSTEJOVSKY, J. Temporal information in intensional contexts. In *Proceedings of the Sixth International Workshop on Computation Semantics (IWCS-6)* (Tilburg, Netherlands, 2005).

[139] SAURÍ, R., LITTMAN, J., KNIPPEN, B., GAIZAUSKAS, R., SETZER, A., AND PUSTEJOVSKY, J. TimeML annotation guidelines. Tech. rep., TERQAS Workshop, 2005. Version 1.4, <http://timeml.org/site/publications/timeMLdocs/AnnGuide_1.2.1.pdf> [date of citation: 2006-02-16].

[140] SAURÍ, R., VERHAGEN, M., AND PUSTEJOVSKY, J. Annotating and recognizing event modality in text. In *The 19th International FLAIRS Conference, FLAIRS 2006* (Melbourne Beach, Florida, USA, 2006).

[141] SAURÍ, R., VERHAGEN, M., AND PUSTEJOVSKY, J. SlinkET: A partial modal parser for events. In *Language Resources and Evaluation Conference, LREC 2006* (Genoa, Italy, 2006).

[142] SCHILDER, F. Extracting meaning from temporal nouns and temporal prepositions. *ACM Transactions on Asian Language and Information Processing* (2004).

[143] SCHILDER AND HABEL 2001.

[144] SCHILDER, F., AND HABEL, C. Temporal information extraction for temporal QA. In *AAAI Spring Symposium on New Directions in Question-Answering (Working Papers)* (Stanford, CA, 2003), pp. 35–44.

[145] SCHILDER, F., AND MCCULLOH, A. Temporal information extraction from legal documents. In Katz et al. [?]. <http:// drops.dagstuhl.de/ opus/volltexte/2005/ 318> [date of citation: 2006-12-01].

[146] SETZER, A. *Temporal Information in Newswire Articles; An Annotation Scheme and Corpus Study*. PhD thesis, University of Sheffield, 2001.

[147] SETZER, A., AND GAIZAUSKAS 2002.

[148] SETZER, A., GAIZAUSKAS, R., AND HEPPLE, M. The role of inference in the temporal annotation and analysis of text. *Language Resources and Evaluation 39*, 2–3 (Mar. 2005).

[149] SHIN, J. H., HAN, Y. S., AND CHOI, K.-S. A hmm part-of-speech tagger for korean with wordphrasal relations. In *Recent Advances in Natural Language Processing*, N. Nicolov and R. Mitkov, Eds., vol. 136 of *Current Issues in Linguistic Theory (CILT)*. John Benjamins, Amsterdam/Philadelphia, 1997, pp. 439–450..

[150] Time expression recognition and normalization (TERN) task, 2004.

[151] VENDLER, Z. (1967), *Linguistics in Philosophy*. Cornell University Press, Ithaca, NY, 1967.

[152] VERHAGEN, M. *Time between the Lines: Embedding a Temporal Closure Component in a mixed-Initiative Temporal Annotation Framework*. PhD thesis, University of Brandeis, 2004.

[153] VERHAGEN, M. Drawing TimeML relations with T-BOX. In Katz et al. [?]. <http:// drops.dagstuhl.de/ opus/volltexte/2005/ 318> [date of citation: 2006-12-01].

[154] VERHAGEN, M. Temporal closure in an annotation environment. *Language Resources and Evaluation 39* (May 2005), 211–241.

[155] VERHAGEN, M., KNIPPEN, R., MANI, I., AND PUSTEJOVSKY, J. Annotation of temporal relations with Tango. In *Language Resources and Evaluation Conference, LREC 2006* (Genoa, Italy, 2006).

[156] VERHAGEN, M., MANI, I., SAURI, R., LITTMAN, J., KNIPPEN, R., JANG, S. B.., RUMSHISKY, A., PHILLIPS, J., AND PUSTEJOVSKY, J. Automating temporal annotation with TARSQI. In *43rd Annual Meeting of the Association for Computational Linguistics (ACL-05)* (Ann Arbor, Michigan, 2005). Poster/Demo.

[157] VILAIN, M., AND KAUTZ, H. Constraint propagation algorithms for temporal reasoning. In *Fifth National Conference on Artificial Intelligence* (Philadelphia, PA, 1986), pp. 377–382.

[158] WEINER AND SIMPSON 1996.

[159] WELLNER, B., AND VILAIN, M. Leveraging machine readable dictionaries in discriminative sequence models. In *Language Resources and Evaluation Conference, LREC 2006* (Genoa, Italy, 2006).

[160] WILSON, G., MANI, I., SUNDHEIM, B., AND FERRO, L. A multilingual approach to annotating and extracting temporal information. In *Workshop for Temporal and Spatial Information Processing (ACL-EACL'2001)* (Toulouse, France, 2001).

[161] WILSON ET AL 2002.

[162] WOLF, F., AND GIBSON, E. Representing discourse coherence: A corpus-based analysis. *Computational Linguistics 31*, 2 (2005), 249–287.

[163] WOOD, F. T. *English Prepositional Idioms*. The Macmillan Press Ltd, London and Basingstoke, 1967.

[164] ZHANG, T., DAMERAU, F., AND JOHNSON, D. E. Text chunking based on a generalization of Winnow. *Journal of Machine Learning Research 2* (2002), 615–637.

[165] ZHANG, T., AND JOHNSON, D. E. A robust risk minimization based named entity recognition system. In *Proceedings of CoNLL-2003* (2003), pp. 204–207.