

ISO TC37/SC4 N471 rev01

Notes to ISO Meeting

Compiled by James Pustejovsky

January 11, 2008 Hong Kong

Revised by Kiyong Lee (marked in red)

May 25, 2008 Marrakech

Revised by Harry Bunt (marked in blue)

June 1, 2008

Summarized by Kiyong Lee

June 4, 2008

ACTION ITEMS

Foreword.

Actions: include remarks on the expected future parts of SemAF.

Delete Note regarding TimeML WG.

Include mention that the current effort is based on the TimeML work conducted by the TimeML WG, etc. see timeml.org. Also explain that TimeML in the document is used to refer to the language from that effort, and ISO-TimeML refers to the annotation language developed in this document.

1. Scope.

Actions: Note will be moved to the forward. Edit and correct content of note.

2. Normative References

Actions: Add LAF, ISO WD 24612, Language Resource Management-Linguistic Annotation Framework

Delete last two regarding feature structures.

3. Terms and Definitions

Actions:

3.1: replaced by eventuality: jdp

3.2, 3.3, deleted

3.11: Rename: utterance-time. Time-span during which utterance occurs.

3.12: Rename: event-time. Time-span at which the event mentioned in a given utterance occurs

3.13: Rename: reference-time.

3.14: delete it. Deleted also from Meta-model. Some possibility that we keep it, rename it "temporal perspective point" (Kamp/Reyle), and define it in the syntax and semantics.

3.18, 3.19: delete

3.20-3.22: modified according to our notes. jdp

Add: temporal entity: instance, intervals

Add: time span: sequence of temporal entities which is not necessarily contiguous

Add: temporal quantity: equivalence class defined by a pair consisting of a numerical element and a temporal unit: e.g 5 milliseconds; 2.5 hours

4. (old 5) Motivation and requirement

Actions: delete names

5. (old 4) Overview

Actions: make reference to events and temporal objects clearer.

6. Basic concepts and metamodel

Actions: delete name

7. Abstract Syntax for ISO-TimeML

Actions: drafted by Harry Bunt and James Pustejovsky, reviewed by Thierry Declerck, Nancy Ide. Meta-review by Kiyong Lee.

Merge: 6. Basic concepts, meta-model, and abstract syntax for ISO-TimeML

Note: Abstract syntax may be dropped if time prohibits.

New Outline of Annexes

A. Concrete Specification of ISO-TimeML: Move to the main part again(klee)

B. Core Annotation Guideline

C. Completely Annotated Examples

D. Illustrative Concrete Semantics of ISO-TimeML: move to the main part again (klee)

E. rest of annexes get pushed down

Keep the original table of contents.

7. Concrete specification of ISO-TimeML

Stand-off annotation:

OLD: John

```
<EVENT eid="e1"> left </EVENT>
<TIMEX3 tid="t1"> yesterday</TIMEX3>
<TLINK .../>
```

Proposed: John left yesterday

```
<EVENT eid="e1" markable="left"/>
<TIME tid="t1" markable="yesterday"/>
<TLINK .../>
```

For <EVENT>

OLD class ::= "STATE"|"OCCURRENCE"|

Proposed: type eventClass ::= "STATE"|"PROCESS"|"TRANSITION"

OLD attributes: pos, vform

Proposed: delete purely syntactic attributes that are not relevant for doing semantics

OLD: <TIMEX3>

type ::= "TIME"|"DATE"|"DURATION"|"SET"

Proposed: <TIME>

timeClass Type ::= "'DATE'"|"INSTANT"|"INTERVAL"|"TIME_SPAN"

<TIME_QUANTITY>

type ::= "DURATION"|"FREQUENCY" |

Examples: The plan crashed in the hillside at 10 am yesterday

<EVENT eid="e1" eiid="ei1" markable="crashed" .../>

<TIME tid="t1" type="INSTANT" clockTimevalue="T10:00" />

<TIME tid="t2" type="INTERVAL" datevalue="2008-05-24"/>

<TLINK .../> <TLINK .../>

John swims twice a day in the summer

<EVENT markable="swims" type="SET"/>

<TIMEQ markable="twice"/>

<TIMEQ markable="a day" timeClass="INTERVAL"/>

<TIME markable="the summer" timeClass="INTERVAL">

Notes by HB:

The attribute 'type' for event tags is proposed to be replaced by 'eventClass' not just for the name's sake, but because the list of possible values: DATE|TIME|DURATION|SET is semantically really not acceptable. Dates are not temporal objects (see below); times and durations are incomparable types of entities (although sometimes confused), and "SET" is again an entirely different thing. It is both semantically better and conceptually clearer for annotators to separate these things. The attribute 'type' by default has the value 'ELEMENT', and has the value "SET" when the text is referring to a set of events rather than to a single event.

"DATE" has been omitted as a possible value of eventClass ('type'), because dates are better regarded not as temporal entities but as names identifying a certain period. Hence date information is stored in the values of the attribute 'date', replacing 'value' as an attribute of temporal entities of the class INTERVAL.

Example: "May 25, 2008" is annotated as

<TIME id="t1" timeClass="INTERVAL" date="2008-05-25">

"every Friday" is annotated as <TIME timeClass="INTERVAL" date="XXXX-YY-05">

Similarly, clocktimes are a way of identifying an instant. (At least, that's how they work in natural language.) So they are stored similarly in the 'clockTime' attribute (replacing 'value') of INSTANTs.

Example: "twelve o'clock" is annotated as

<TIME id="t2" timeClass="INSTANT" clockTime="T12:00:>

(To be formally complete we should also have time zones, but from an annotator's point of view that seems pretty silly...)

The length of an interval or a time-span is altogether a different kind of entity than the object whose length it measures, so it has different kinds of properties (attributes) and is

conceptually better to be represented by a separate tag, different from TIME. Omit <SIG markable="in"/>, for it is reflected in relType in TLINK

Frequencies are properties of sets of events, and consist of two parts: a number of occurrences and a period during which the occurrences fall. They are to be distinguished from sheer number of occurrences of events of a certain kind; it is a peculiarity of English (and French), not shared with e.g. German, Italian, Spanish, Danish, Dutch, ... that the word for time is also used for number of occurrences. A sentence like "John called Mary twice" simply refers to a set of two events, and seems best annotated as:

<EVENT ... type="SET" card="2"> using an attribute for indicating the number of elements in a set.

Frequencies, by contrast, require two attributes for being expressed properly:

"Harry teaches twice a week" should be annotated as something like:

<EVENT ... type="SET" freq="2" period="P1W">

Again different are cases like "Harry teaches four hours a week", which do not describe a frequency but the total length of a certain set of events. Temporal annotation should perhaps not have the ambition to this quantification precisely. We propose to annotate a set of teaching events, a temporal quantity ('four hours') and a period ('a week') and relate them through TLINKS, one with relType=DURATION and one with relType=Is_INCLUDED.

Summary

By Kiyong Lee, 2008-06-04

For <EVENT>

attributes ::= eid eiid **markable eventClass type** tense aspect
polarity mood [modality] [comment]

eid ::= ID

{eid ::= eventID

eventID ::= e<integer>}

eiid ::= ID

{eiid ::= eventInstanceID

eventInstanceID ::= ei<integer>}

markable ::= CDATA

eventClass ::= 'STATE'|'PROCESS'|'TRANSITION'

type ::= 'SET'|'ELEMENT'

{default, if absent, is 'ELEMENT'}

card ::= <integer>

freq ::= <integer>

period ::= 'P<positiveRealNumber>timeUnit'

Question (KL): for two hours and a half => period='P2.30H'???

REPLACE

class ::= 'OCCURRENCE' | 'PERCEPTION' | 'REPORTING' |

```

    'ASPECTUAL' | 'STATE' | 'I_STATE' | 'I_ACTION'
DELETE
pos ::= 'ADJECTIVE' | 'NOUN' | 'VERB' | 'PREPOSITION' | 'OTHER'

tense ::= 'FUTURE' | 'PAST' | 'PRESENT' | 'IMPERFECT' | 'NONE'
aspect ::= 'PROGRESSIVE' | 'PERFECTIVE' | 'IMPERFECTIVE'
         | 'PERFECTIVE_PROGRESSIVE' | 'IMPERFECTIVE_PROGRESSIVE' |
'NONE'

DELETE
vform ::= 'INFINITIVE' | 'GERUNDIVE' | 'PASTPART' | 'PRESPART' |
'NONE'

polarity ::= 'NEG' | 'POS'
{default, if absent, is 'POS'}
modality ::= CDATA
mood ::= 'SUBJUNCTIVE' | 'NONE'
{default, if absent, is 'NONE'}
KL: what if we just have mood::=CDATA so that Korean can admit mood values like
evidential or retrospective? or
mood ::= 'SUBJUNCTIVE'|'EVIDENTIAL'|NONE
comment ::= CDATA

```

For <TIMEX3>

REPLACE <TIMEX3> with <TIME> and <TIME_QUNATITY>

BNF for <TIME>

```

attributes ::= tid timeClass clockTime date [functionInDocument] [beginPoint]
            [endPoint] [temporalFunction]
            (value | valueFromFunction)[mod] [anchorTimeID]
            [comment]
tid ::= ID
{tid ::= TimeID TimeID ::= t<integer>}

timeClass ::= 'INSTANT'|'INTERVAL'|'TIME_SPAN'

clockTime ::= 'TXX:XX' where X is an integer
date ::= "YYYY-MM-DD" where Y, M, and D are integers, representing Year, Month,
and Day, respectively.

```

DELETE 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}

DELETE quant ::= CDATA
DELETE freq ::= Duration

temporalFunction ::= 'true' | 'false' {default, if absent,
 is 'false'}
{temporalFunction ::= boolean}

DELETE 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

(3) For <TIME_QUANTITY>
Question(KL): Do we really need <TIME_QUANTITY>?

attributes ::= qid quant [comment]
qid ::= ID
{qid ::= timeQunatityID timeQuantityID ::= q<integer>}

quant ::= 'EVERY' | 'SOME' | or CDATA

comment ::= CDATA

Illustrations:

(1) Stand-off annotation with the attribute markable ::= CDATA
Question (KL): just head or chunk?

John has already left

- a. <EVENT eid="e1" eiid="ei1" markable="left" tense="PAST" aspect="PERFECTIVE" polarity="POS"/>
- b. <EVENT eid="e1" eiid="ei1" markable="has ... left" tense="PAST" ... />

John didn't leave

- a. <EVENT eid="e1" eiid="ei1" markable="leave" tense="PAST" aspect="NONE" polarity="NEG"/>
- b. <EVENT eid="e1" eiid="ei1" markable="didn't leave" tense="PAST" aspect="NONE" polarity="NEG"/>

2. <EVENT>

Harry teaches twice a week

<EVENT ... eventClass="PROCESS" type="SET" freq="2" period="P1W"/>

3. Deleting SIGNAL

The plane crashed into the hildside at 10 am yesterday

<EVENT eid="e1" eiid="ei1" markable="crashed" .../>
<TIME tid="t1" timeClass="INSTANT" clockTime="T10:00" />
<TIME tid="t2" timeClass="INTERVAL" date="2008-05-24"/>
<TLINK eventInstanceID="ei1" relatedToTime="t1" relType="SIMULTANEOUS"/>
<TLINK timeID="t1" relatedToTime="t2" relType="IS_INCLUDED"/>

4. For <TIME>

Date: May 25, 2008

<TIME tid="t1" timeClass="INTERVAL" date="2008-05-25"/>

Clock time: twelve o'clock

<TIME tid="t1" timeClass="INSTANT" clockTime="T12:00"/>

5. Temporal Quantity

every Friday

HB: <TIME timeClass="INTERVAL" date="XXXX-YY-05">

KL: On Friday

<TIME timeClass="INTERVAL" date="XXXX-YY-05">

but every Friday

<TIME_QUANTITY quant="EVERY" timeClass="INTERVAL" date="XXXX-YY-05">

