

# **An Extended Specification for Annotating Time Anchor of Events v1.0**

## **Content**

- 1. General Description of This Task**
- 2. Several Time Anchor Categories of Events**
  - 2.1 Single-Day Events (Certain/Uncertain)
    - 2.1.1 Certain Date
    - 2.1.2 Uncertain Date
  - 2.2 Multiple-Day Events
  - 2.3 Frequency Occurrence
- 3. Other Expressions**
  - 3.1 The n-th Day of a Period (week/month/season/quarter/year, etc.)
  - 3.2 Early/middle/late of a Time Period
  - 3.3 Past / future with Uncertain Distance
  - 3.4 AND / OR
- 4. Narrow Down the Date range as Precise as Possible**
- 5. Verb Tense**
  - 5.1 Future
  - 5.2 Past
  - 5.3 Present
  - 5.4 Verb Tense in Subordinate Clauses
- 6. Event Attributes**
  - 6.1 Verbs Preceded by Modal Auxiliary
  - 6.2 Intensional Events
- 7. Preferred Annotation Cases**
  - 7.1 Time information in the same sentence is preferred
  - 7.2 Ambiguous Single-day or Multiple-day Events
  - 7.3 Accurate Time Period Calculation
  - 7.4 Uncertain Time Period

# 1 General Description of This Task

**Time anchor of event** defines a period of time in a time axis that an **event** *happen* or *occur*. Events can occur in one day, for a period of time, frequently, etc.

## Time anchor example:

*He was **sent** into space on May 26, 1980. He **spent** six days aboard the Salyut 6 spacecraft.*

For instance in this text, the time anchor of ‘**sent**’ can be annotated with the normalized date 1980-05-26. The time anchor of ‘**spend**’ can be annotated as a period with multiple days from 1980-05-26 to 1980-06-01.

In this work, we intend to apply this annotation guidelines for the TimeBank corpus (Pustejovsky et al.,2003). The TimeBank corpus contains 183 new articles that have been annotated with add events, time expression and temporal links following the TimeML specification ([www.timeml.org](http://www.timeml.org)). The annotators of this work are required to infer the time anchor of each individual event (total 7,935) which was annotated in the TimeBank corpus.

## An example of the TimeBank corpus:

*<DCT>03/08/1998 06:26:00<DCT>  
Some 1,500 ethnic Albanians  
<EVENT eid="e2" class="OCCURRENCE"> **marched**</EVENT>  
<TIMEX3 tid="t68" type="DATE" value="1998-03-08"> Sunday </TIMEX3>  
in downtown Istanbul.*

In this example, the annotator is required to read the whole text first. Based on those temporal information i.e. document create time (<DCT>) and Sunday (<TIMEX3>), the annotator can infer the time anchor of **marched** is 1998-03-08, and annotate it with an new property of this event. For instance, <EVENT eid="e2" class="OCCURRENCE" tanchor="1998-03-08"> **marched**</EVENT>.

\* We used days as the minimal granularity for the annotation.

## 2 Several Time Anchor Categories of Events

### 2.1 Single-Day Events (Certain/ Uncertain)

#### 2.1.1 Certain Date

If the exact time of a **event** can be written in the format of *YYYY-MM-DD*.

- ◇ *He was **sent** into space on May 26, 1980.*
- ➔ ***sent** : Tanchor="1998-05-26"*

#### 2.1.2 Uncertain Date

In case the exact event time is not stated in the document, the annotators are asked to narrow down the possible event time to a **tuple** (*after YYYY-MM-DD, before YYYY-MM-DD*) as precisely as possible.

- ◇ *In 1996 he was **appointed** military attache at the Hungarian embassy in Washington.*
- ➔ ***appointed** : Tanchor="(after 1996-01-01, before 1996-12-31)"*

The event appointed is annotated (*after 1996-01-01, before 1996-12-31*) as the event must have happened some day in 1996. “*after 1996-01-01*” is **inclusive** of the day “*1996-01-01*” in our definition.

### 2.2 Multiple-Day Events

Except the single-day time anchors describe in Section 2.1, we will introduce several other types of time anchors.

A multiple-day event contains a span over multiple days and is annotated as a **tuple** (*begin:single-day time anchor, end:single-day time anchor*).

- ◇ *He was **sent** into space on May 26, 1980. He **spent** six days aboard the Salyut 6 spacecraft.*
- ➔ ***spend** : Tanchor="(begin:1980-05-26, end:1980-06-01)"*

**spent** is a multiple-day Event with the event time beginning from precise *1980-05-26* to the precise end *1980-06-01*.

The realistic scenario could be much more complicated than precise time anchors of *begin/end*. The following is an example of both *begin* and *end* are the imprecise date.

- ◇ *McBride was **part** of a seven-member crew aboard the Orbiter Challenger in October 1984.*
- ➔ ***part** : Tanchor="(begin:(after 1984-10-01, before 1984-10-31), end:(after 1984-10-01, before 1984-10-31))"*

Consider another case, if we simply annotate **attend** with the previous imprecise multiple day time anchor. The *Tanchor* will be "(begin:(after 2018-01-1, before 2018-01-31), end:(after 2018-01-1, before 2018-01-31)". The temporal information of 'two-day', which is the duration of **attend** is obvious lost. In order to address this issue, we introduce 'dur:' into the annotation.

◇ I will **attend** a two-day conference in January, 2018.

➔ **attend** : *Tanchor*="(begin:(after 2018-01-1, before 2018-01-31), end:(after 2018-01-1, before 2018-01-31),dur:2D"

'dur:' could be:

*xD* (several days),

*2M* (two months),

*3Y*(three years), *1XY*(over a decade), *xXY*(several decades)

*4W* (four weeks),

*1S* (one season), etc.

## 2.3 Frequent Occurrence

Single day and multiple day time anchors are not capable to represent some types of events such as: a event frequently occur during a period of time.

◇ *DCT=2018-02-09: John **taught** a class every Monday this year.*

➔ **taught** : *Tanchor*="( begin:2018-01-01, end:2018-12-31, freq:1T/Mon)

◇ *DCT=2018-02-09: John **taught** twice on Monday and once on Tuesday this year.*

➔ **taught** : *Tanchor*="( begin:2018-01-01, end:2018-12-31, freq:2T/Mon;1T/Tue

◇ *DCT=2018-02-09: John **taught** twice a week this year.*

➔ **taught** : *Tanchor*="( begin:2018-01-01, end:2018-12-31, freq:2T/1W

Other 'freq:' could be:

*1T/1D* (one time per day),

*1T/2M* (one time every two months),

*2T/3Y* (two times every three years),

*1T/4W* (one time every four weeks),

*xT/1S* (several times per season), etc.

Instead of a definite occurrence of events, some words can also express the meaning of probability, e.g. usually, often, sometimes, rarely. These can be used to replace the frequency number, e.g. *freq:often/Mon*

### 3 Other Expressions

#### 3.1 The *n*-th Day of a Period (week/month/season/quarter/year, etc.)

Use 'FIRST, LAST, 2nd, 3rd, 4th, etc.' to express the *n*-th day of a duration:

W<sub>x</sub>: the *x*-th week of a year

Season: {SP, SU, AU, WI}

Quarter: {Q1, Q2, Q3, Q4}

◇ *The first day of summer 2017*

➔ *Tanchor="2017-SU-1st"*

◇ *GM Hughes Electronics, said net income **fell** 22% in the third quarter.*

➔ *Tanchor="begin: 1989-Q3-FIRST, end: 1989-Q3-LAST"*

◇ *In the nine months, GM Hughes **fell** 6.6% to \$486.6 million*

➔ *Tanchor="begin: 1989-Q1-FIRST, end: 1989-Q3-LAST"*

#### 3.2 Early/middle/late of a Time Period

We use "EAR", "MID", "LAT" to label early/middle/late of a time period (week/month/season/quarter/year).

◇ *Vanguard's \$1 billion High Yield Bond Portfolio has **seen** \$161 million flow out since **early September**; \$14 million of that seeped out **Friday Oct. 13** alone.*

➔ *Tanchor="begin:after1989-09-EAR-FIRSTbefore1989-09-EAR-LAST,end:1989-10-27"*

#### 3.3 Past / Future events with Uncertain Distance

The past/future events occur with a date distance to another date.

◇ *DCT=2018-02-09: I **studied** Latin several years ago.*

➔ ***studied** : Tanchor="begin:before2018-02-09, end:before2018-02-09, dis=PxY"*

◇ *DCT=2018-02-09: I will **enter** a competition a few week later.*

➔ ***enter** : Tanchor="DCT, dis=FxW"*

#### 3.4 AND / OR

AND/OR can be used as a conjunction of multiple time anchors for the target event.

◇ *DCT=2018-02-08: I will **go** to school tomorrow and for the whole of the next week.*

➔ ***go** : Tanchor="2018-02-09 AND begin=2018-02-12,end=2018-02-16"*

◇ *DCT=2018-02-08: I will **go** to school tomorrow or on Monday of the next week.*

➔ ***go** : Tanchor="2018-02-09 OR 2018-02-12"*

## 4 Narrow Down the Date range as Precise as Possible

In a news article, there could be several time expressions distributed throughout the text. Some time expressions with long distance to a event could be import to make annotation more precise.

For instance in “AQA022\_APW19991008.0265.tml”:

*Another man has been **arrested** in connection with the bombings of U.S. embassies in Kenya and Tanzania.*

One annotator assigned **arrested** as *before DCT*. However, he missed an import time expression at the end of this news.

*The bombs went off almost simultaneously on Aug. 7 , 1998 at U.S. embassies in Nairobi , Kenya and Dar es Salaam , Tanzania .*

A more precise **annotation** could be infer as *(after 1998-08-07, before DCT)*.

Annotator should carefully review the relations between a target event to all the time expressions in this article. Because one missing time expression could possibly effect all the event annotation.

## 5 Verb Tense

Time anchor annotation could be highly related to verb tense. Different verb tense could provide us clear evidence of determine a temporal order ('before', 'after' or exact on) between a event and a time expression (yyyy-mm-dd). For instance, 'infinitive' and 'future tense' are often used when a event will occur in the future with 'after' yyyy-mm-dd.

One the other hand, verb tense can be helpful use to recognize a event as **single-day** or **multi-day**. For instance, 'continuous' and 'perfect continuous' possibly means the event occur over **multi-day**.

### 5.1 Future Tense

- Simple Future

◇ *The United States will **consider** helping set up a disaster response center in Kenya . (DCT)*

➔ *Tanchor=after DCT*

- Future Perfect
- Future Continuous
- Future perfect Continuous

### 5.2 Past Tense

- Simple Past

◇ *BBC forecasters **said** the snow would cause travel... (DCT)*

➔ *Tanchor="before DCT"*

◇ *President Fidel Castro **said** Tuesday(2000-03-28) his government told...*

➔ *Tanchor="2000-03-28", because a certain date Tuesday is in this sentence.*

- Past Perfect
- Past Continuous
  - ◇ *Castro said a group of Cuban-Americans in Miami was **studying** the possibility of taking Elian and his Miami relatives to a third nation. (DCT)*
  - ➔ *Tanchor="begin:before DCT, end:before DCT",*
- Past perfect Continuous

### 5.3 Present Tense

- **Simple Present**

It's often used to describe **habits**, **unchanging situations**, **general truths**, and **fixed arrangements**.

◇ *He lives in Tokyo. (DCT)*

➔ *Tanchor=" begin:before DCT, end:after DCT" (unchanging situations)*

◇ *We **want** to build on the long-term partnership that has existed between our countries to help innocent Kenyans, Satcher said at a news conference. (DCT)*  
 ➔ *Tanchor= "before DCT"* , In this case, **want** refer to the same time as *said*.

But sometimes, in some **weather**, **stock** news simple present or present continuous might only describe the situation until today. In these cases, we annotate **end** with certain today, because the situation of tomorrow is unknown.

- **Present Perfect**

- **Present Continuous:**

It's often about something **unfinished or incomplete** (end in the future).

◇ *Heavy snow is **causing** disruption to transport across the UK. (DCT)*

➔ *Tanchor= " begin:before DCT, end:after DCT"*

◇ *He is **living** in Tokyo (DCT)*

➔ *Tanchor= " begin:before DCT, end:after DCT"*

## 5.4 Verb Tense in Subordinate Clauses

News articles very often quote the speech of an official. The verb tense in such quoted speech or a subordinate clause describe a **relative** temporal order to the verb tense in the main clause. Words like said, reported, denied are often used. Such events will be reference point to infer Tanchor of events in subordinate clauses.

For instance:

*"We **want** to build on the long-term partnership that has existed between our countries to help innocent Kenyans, Satcher **said** at a news conference. "*

Simple present of **want** means that it occur on the same day as **said**.

We should anchor the time of **said** first (*before DCT*, in this case). Then the verb tense of **want** can indicate a relative same day to **said**. So we can finally infer **want** as *before DCT*.



## 6 Event Attributes

### 6.1 Verbs Preceeded by Modal Auxiliars

Modal auxiliars in this group include: *may, might, can, could, would* and etc. Such words often express that a event does not occur yet but will probabaly occur in the future.

For instance:

*Lazaro Gonzalez , denied that he would **send** Elian out of the country.*

We can infer **send** doesn't occur yet and is *after denied*. In this case, **denied** occur on DCT. So we know **send** is *after DCT*.

### 6.2 Intensional Events

Some verbs (try, attempt) show intensional actions. The events introduced as arguments of such intensional actions have not occurred when the actions takes place, but occur in the future.

intentional action      argument  
↓                      ↓  
*The investigators **try** to **track** down the gunman . (DCT)*

In this case, we assign **try** as “DCT” and **track** as “after DCT”. Because **try** is a current intension and **track** doesn't occur yet.

*The investigators will **try** to **track** down the gunman . (DCT)*

In this case, we assign **try** as “after DCT” and **track** as “after DCT”. Because **try** is a intension in the future and **track** will occur *after try*.

Some intentional actions:

- attempt, try, scramble:

*Companies such as Microsoft or a combined worldcom MCI are **trying** to **monopolize** Internet access.*

*Israel has been **scrambling** to **buy** more masks abroad.*

- avoid, prevent, cancel:

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

- ask, order, persuade, request, beg, command, urge, authorize:

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

They were **asked** to **take** along important papers.

- promise, offer, assure, propose, agree, decide:

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.

## 7 Preferred Annotation Cases

Here, I summarized several error types. I provide our annotation preference for annotators to follow, unless the annotations have strongly confidence.

### 7.1 Time information in the same sentence is preferred

In the following example, the event time of “**wrote**”, “**intends**” could be inferred based on “**yesterday**” or *DCT*. In many cases, the time expression in the same sentence would be preferred to be anchored.

- ◇ The No. 1 U.S. auto maker then **wrote** Jaguar that it **intends** "to go to that 15%" level once it wins the U.S. clearance to go beyond \$15 million, a Jaguar spokesman said **yesterday**. (*DCT*: 1989-10-27)
  - ➔ Event: **wrote** Tanchor="before1989-10-26" (**Preferred**)  
Tanchor="before1989-10-27" (Not Preferred)
  - ➔ Event: **intends** Tanchor="after1989-10-26" (**Preferred**)  
Tanchor="after1989-10-27" (Not Preferred)

### 7.2 Ambiguous Single-day or Multiple-day Events

Sometimes, the context information is ambiguous for annotators to distinguish whether the target event occurred (once, in multiple times or for several days) during a time period. In this case, we prefer to have a multiple-day annotation.

- ◇ Dresdner Bank said it will also **buy** all shares tendered by shareholders on the Paris Stock Exchange at the same price from today through Nov. 17.
  - ➔ Tanchor="begin:1989-10-27,end:1989-11-17" (**Preferred**)  
Tanchor="after1989-10-27,before1989-11-17" (Not Preferred)
- ◇ GM **sought** U.S. antitrust clearance last week to purchase more than \$15 million worth of Jaguar shares but doesn't own any yet, according to GM officials here and at the company's Detroit headquarters.
  - ➔ Tanchor="begin:1989-10-27,end:1989-11-17" (**Preferred**)  
Tanchor="after1989-10-15,before1989-10-21" (Not Preferred)

### 7.3 Accurate Time Period Calculation

Accurate calculation of begin/end dates is preferred for “past, last, next, upcoming, etc.” expressions.

- ◇ *In the past two months* the nation's 88 junk funds have **lost** a total of about \$6 billion -- more than 15% of assets -- through sales or transfers of junk-fund shares, Dalbar says. (DCT: 1989-10-27)  
→ Tanchor="Begin:1989-08-27,end:1989-10-27" (**Preferred**)  
Tanchor="Begin:1989-09-01,end:1989-10-27" (Not Preferred)
- ◇ He adds: "Our fund has had net **sales** every month for **the last three years** -- until this month." (DCT: 1989-10-27)  
→ Tanchor="begin:1986-10-27,end:1989-10-27" (**Preferred**)  
Tanchor="begin:1987-01-01,end:1989-10-27" (Not Preferred)

### 7.4 Uncertain Time Period

Don't specify any additional time period in the expression of “the weeks ahead”, “recent years”, etc.

- ◇ If the heavy outflows continue, fund managers will face increasing **pressure** to sell off some of their junk to pay departing investors in **the weeks ahead**. (DCT: 1989-10-27)  
→ Tanchor=" begin:after1989-10-27,end:after1989-10-27, dur:xW" (**Preferred**)  
Tanchor=" begin:after1989-10-27 end:before1990-01-05, dur:xW "(Not Preferred)
- ◇ But tremors are likely in the junk-bond market, which has helped to **finance the takeover boom of recent years**. . (DCT: 1989-10-27)  
→ Tanchor=" begin:before1989-10-27,end:1989-10-27, dur:xY"(**Preferred**)  
Tanchor=" begin: after1984-10-27before1989-10-27,end:1989-10-27, dur:xY"(Not Preferred. The annotator specified recent 5 years.)