Annotations Instructions & Guidelines

The goal of this annotation task is to differentiate absolute from relative or incomplete time expressions, and to link the latter to anchor dates.

Absolute time expression : an expression which contains all the information needed to normalize it to a standard date. Eg : “12/05/2020”

Relative time expression : an expression whose temporal meanings is stated as a relative value against another time expression. Eg “ two days before the admission”

Incomplete time expression : an expression which holds only partial information : the context is needed to determine the calendar date. Eg : “in December”

Anchor date : The reference point which can be used to infer the normalized value of a relative or incomplete temporal expression.

# Instructions

1. **Download** the annotations folder from the python\_timex OneDrive folder
2. **Launch the MAE tool** by clicking on the **“Mae\_v0.9.3”** file
3. Once the window is open, click on **“Load DTD”** from the File menu
4. Choose the **RITempAnnotation.dtd** file

*Once loaded, the following tabs should appear :*

*- RTIMEX3 : these are the relative time expressions, the main focus of the annotation process*

*- ATIMEX3 : the absolute time expressions, here to provide anchorage for the RTIMEX3*

*- SECTIME : These are special annotations for the admission and discharge date*

*- ANCHORLINK : These links will be created by the annotator to join a RTIMEX3 to an anchor date (an ATIMEX3 or SECTIME).*

1. Load a file : from the File menu, select the “**Load File**” option.

*The text should appear along with the annotations outlined in different colors. The RTIMEX3 are in blue, the ATIMEX3 in red, the discharge and admission date are usually double annotated as both absolute time expressions and SECTIME so they are underlinded. The ANCHORLINKs are empty as they will be created during the annotation process*

6. The process is as follows : The focus should be on each RTIMEX3 annotations until they are all annotated

* First, the “**relative**” column must be filled :

with “TRUE” if the expression is indeed a relative time expression, “FALSE” if it is an absolute timex3 which was not correctly filtered.

*Common example would be : “May 1997”, “On Christmas of 2002”, “April 2nd 2015”*

* If “relative” is TRUE, an **ANCHORLINK** has to be created

This is done by holding down the ctrl key (or the command key, if you are on a Mac) and left click each of the entities that will be included in the link, with the RTIMEX3 first and the Anchor Date second.

*For precise instructions on how to select the appropriate anchor date, see the “Guidelines” section.*

A link window will pop up and ask you to confirm the two dates and the link type.

Special case - if the anchor date is the admission or discharge date : because these are double annotated as ATIMEX3 and SECTIME, the program will let you choose between the two instances. They have the same value so it does not matter too much but the SECTIME should be preferred.

* Once the ANCHORLINK is created, the **“relation”** attribute has to be filled with either BEFORE, EQUAL or AFTER

7. **Check the** **ATIMEXEs** : sometimes, an expression marked as an absolute time expression is in fact a relative one. For this, the “absolute” attribute of the ATIMExes as to be filled with True or False. If the expression is in fact a relative one, it has to be anchored.

8. **Output the file** Once all the RTIMEXES are filtered and anchored, choose the “Export as XML” option in the File menu, and save the file with its original name in a separate folder.

9. **Upload** the annotated files in the python\_timex directory

# Guidelines for Ambiguous Cases :

* SELECTING THE ANCHOR DATE

When selecting the anchor date, the first potential anchor dates to study are : the previous absolute timex, the previous timex, the admission date and the discharge date. One should prioritize absolute anchor dates over relative ones, and if there is still an ambiguity, “EQUAL” relations over “BEFORE” and “AFTER”.

These four possibilities are to be prioritised, but other anchor dates are valid as well.

* POSTOPERATIVE DAYS

As a general rule, expressions relating to the “post-operation” concept should be anchored to the day of the operation. The exception to that rule is if there is another time expression which can serve as anchor for the POD with the relation “EQUAL”

* AGE RELATED EXPRESSIONS :

Some time expressions annotated as dates age in fact age expressions. If this case arises, one has to change the type of the expression to “AGE\_RELATED”.

* INCOMPLETE EXPRESSIONS

Some expressions are not relative but rather incomplete: their normalized value depends on one or more missing information, such as the year. ex “Labor Day”

In this case, they should still be annotated as RTIMEXEs, and if they cannot be anchored, it is possible to change the “mod” attribute of the expressions to “EXT”, to signify that there is a need for external information.

* NON-ANNOTATED EXPRESSIONS

Sometimes, expression which should be annotated as either R-TIMEXEs or A-TIMEXEs are not annotated at all : this is likely an error coming from i2b2’s gold standard, and we should let them as is. No annotation should be added to the documents.

* INCOMPLETE TIMES

Eg “2.30 pm”

They are to be anchored to the day they belong to. Usually they are wrongly annotated as absolute time expressions.

* SECTION TIMES

Usually, imprecise expressions found at the beginning of a document relate to the Admission date, and those found at the end to the Discharge date.