THYME Temporal Annotation Guidelines

Developed by Will Styler, Guergana Savova, Martha Palmer and James Pustejovsky, based on the ISO TimeML temporal relations specification

Last updated: March 14th, 2012

Contents

1 Introduction and Philosophical Notes

1.1 Recent Changes

- 3/14/12 - Explicitly stated that all TLINKs between tests and their results must be of type CONTAINS. Added an example to support this. Also discussed handling (in 6.1.2) of tests, their results, and their date of occurrence. Finally added a few more ALINK examples.

- 2/22/12 - Discussed the aspectuality (or not) of the word "recurrence".

- 2/16/12 - Explicitly banned cross-sentence ALINKs.

- 2/11/12 - Removed four now-unneeded references to Transition/State/Process from the guidelines. Added an example to HEDGED based on an annotator question. Specified that TIMEX3s of type SET should always use TLINK type OVERLAP.

- 1/30/12 - Clarified the usage of HYPOTHETICAL when discussing future EVENTs, cleaned up OVERLAP and CONTAINS examples to remove ambiguities

- 1/25/12 - Removed all EVENT types except "aspectual" (removing the difficult Transition/State/Process distinction), rewrote that portion of the guidelines, and added "N/A" as a type.

- 1/7/12 - Changed the guidelines title to reflect our new acronym, THYME

- 1/5/12 - Added a section discussing sections not to be annotated.

- 1/5/12 - Added "Narrative Containers" section, completely re-did the When-to-TLINK guidelines in hopes to better constrain the possible TLINK annotations.

- 12/07/11 - Added examples for History of, added "plan" to special cases. Also clarified the "What is not an EVENT" section, removing some typoes and adding "unusual" and "necessary" discussion.

- 9/6/11 - Added SET to TIMEX3 types, discussed handling of things like "twice daily", revised a typo in DURATION

- 8/17/11 - Fixed broken Figure numbers (Thanks Dave!)

- 8/08/11 - Discussed the Pre- and Post- expression types for TIMEX3

- 8/03/11 - Added an outline of the schema as needs to be implemented for annotation

- 6/28/11 - Revised TLINK guidelines as well as clarifying that EVENTs without DocTimeRel are not EVENTs at all.

- 6/15/11 - Revised the handling of Pre- and Post- expressions, clarified handling of missing body parts.

- 5/25/11 - Added reference to the newly-created "Accessing Clinical Data in Protege/Knowtator" document

- 5/17/11 - Revised typoes based on Dave Harris's feedback. Also modified all discussions of "DocTime" and "SectionTime" to reflect a small change in the schema.

- 5/12/11 - No substantive changes to the schema's functioning, just stylistic changes and massive improvements of the documentation. This included typo squishing, major rebuilding of the EVENT and "making annotations" sections for smoothness, rewriting of some text throughout the document for better readability, and the addition of an "Overview of the annotation process" section.

- 5/9/11 - Changed the title from "Draft temporal relations annotation guidelines" to "Clinical temporal relations annotation guidelines", completed the Annotating using Knowtator section, added DOCTIME as an annotatable attribute.

- 5/1/11 - Added additional examples of pre- and post- expressions, renamed the title of that section to be more generic

- 4/27/11 - Created a "Special Cases and Constructions" section, and discussed "postoperative", "family history", "patient history". Removed those items from "known issues"

- 4/21/11 - Section type labels added as Appendix I. Additional examples added.

- 4/13/11 - Span guidelines updated to differentiate verbal and non-verbal EVENTS, use of subsumed annotations removed.

- April 2011 - EVENT annotation section reorganized and expanded, with many more examples, Discussion of patient preference, and of the reasoning behind marking things like "we discussed the treatment plan..." based on Piet's feedback

1.2 Known Issues (schema and documentation)

- If Propbank is to be used to label eventive nouns, we'll need to describe how that will work for annotators.

- We will need discussion of UMLS event preannotations, once those are incorporated into the schema

1.3 Introduction

Thank you for becoming an annotator! The THYME project (short for "Temporal Histories of Your Medical Events") aims to develop a human-annotated corpus of medical records in which the temporal relations (or relations in time) between different events, occurrences, states, dates, and procedures are clearly annotated. This corpus will then be used to aid in machine learning, in hopes that this annotation can then be passed onto a machine such that the more than 40,000,000 records in the Mayo Database can be searched both by doctors and by researchers hoping to find long-span correlations and tracking patient outcomes.

Please note that we will only be working with medical records that have been de-identified. This means that no references to the patient's name or birthdate will be included, the names of different doctors will be obscured, and that all dates have been modified from their original form. However, because we're examining temporal relations, all relations between the dates have been preserved. So, to give an example, every date for a given patient may have had 16 days added to it. The number added (or sometimes subtracted) is a random number which is kept consistent across each individual patient and record, but is not revealed to anybody but the clinic personnel preparing the documents, and changed between patients. Even with this de-identification done, remember, this is still actual patient data, and your discretion and caution with the data is necessary.

Your annotation will consist of two main tasks. First, you will search the document and annotate events relevant to the patient's treatment and life, and provide additional temporal and modal information about these events. Then, you'll find discrete references to times and dates which we can use as pins on the grand timeline of the patient's history. Then, finally, you'll go through and mark the temporal relations of these events to one another. By doing this, you will be generating a useful timeline of the patient's medical care which is temporally rich enough for any number of uses.

The final goal is to produce annotations rich enough that a computer, using complex inferencing, co-reference, and domain-specific algorithms, would be able to construct an accurate timeline of when the events in a given medical record occur relative to any fixed dates present and relative to one another. This timeline could then be viewed directly by physicians reviewing a patient's case history, or queried by means of natural language questions ("Did the patient report any bleeding before her diagnosis of Colon Cancer?"). Such a tool could be very useful for increasing both the speed with which a doctor can "get current" with a patient's status, and potentially, increase the power of statistical queries for running analyses of large banks of medical data to find patterns too temporally distributed for humans to notice.

1.4   Acknowledgements

1.5   About the examples in this guide

In our examples, words in [brackets] are to be annotated with the type of annotation or attribute being discussed. The brackets represent the proper span. Brackets coupled with angle brackets (> <) indicate disjointed (or non-continuous) spans. So "[Urine Cytologies> were all <negative]" indicates that the span starts with "Urine cytologies", drops off at the angle brackets, then picks back up for "negative".

2   Overview of the annotation process

2.1   Finding an annotating events and timepoints

First, you'll need to go through the document and find any events or states relevant to the patient's clinical timeline. This could range from procedures to diseases to diagnoses to patient complaints and states, as well as nearly anything else relevant to the patient's case. Each of these clinically relevant events will be given an *EVENT* annotation

Then, for each EVENT annotation, you'll specify further attributes. These attributes specify the temporal relation of the event to the time of service (called "*DocTimeRel*", which specifies whether an EVENT happened before, after, during, or before and during the patient's visit), whether the event really happened or whether it was hypothetical, negation, its aspectual type, and more. This will be done for each event at the same time that you create the annotation.

Once you've found all the events in the note, you'll go back through and look for discrete references to time (called "*TIMEX3s*" in this schema). These can be dates (like "June 9th" or "May 2006"), times ("5:15pm"), relative times ("In six months" or "Next week"), or durations ("For 15 years", "Since last year"). These will be marked as TIMEX3 annotations, and then marked for their types.

2.2   Temporally linking events and timepoints together

At this point, you've found all the events and dates that make up the clinical timeline. Your next step is to mark the relations of these EVENTs to one another by creating a series of *temporal link* ("*TLINK*" for short) annotations. TLINKs connect two EVENTs, or an EVENT and a TIMEX3 together, specifying the temporal relationship between them (before, after, overlap, contains, begins-on and ends-on). So, if one event occurred before another, you would link them with a BEFORE TLINK. If one event started on a certain TIMEX3, you would link them with a

TLINK of the "BEGINS-ON" type. Armed with these types of TLINK, you'll go through the entire document and mark the temporal relations between EVENTs and TIMEX3s within each section.

This is a very important step, as every event needs to be anchored to the timeline somehow. Whether this anchoring is as specific as a TLINK to a TIMEEX3 or general as the DocTimeRel marking (which temporally links the EVENT to the time of the document's writing), there needs to be something. An EVENT that's just floating with no DocTimeRel or TLINK isn't very useful, and in general, we need to make sure that no event is left stranded and unable to attach itself to the timeline. When in doubt, make the annotation, whether it be an EVENT, a TLINK, or an attribute. Redundancy will never hurt.

2.3 *Finishing up*

Once you've finished TLINKing the document, you'll want to make sure that any aspectual EVENTs (like "continues" or "restarted") are linked to the EVENT they modify using *aspectual links* ("*ALINKs*"), and finally, you'll mark the overall document time (the time of patient service) and the time referred to by specific sections using the *DOCTIME* annotation.

Once you've completed all of these steps, the clinical note can be considered fully annotated. This guide is designed to walk you through each step of the process, first familiarizing you with all the annotations, their usage, and their default values, and any odd cases or exceptions, and will then show you how to create and modify these annotations using Protege and Knowtator, our annotation tools.

3 EVENT Annotation

3.1 *What is an EVENT?*

What is an EVENT? In our schema, an EVENT is anything that's relevant on the clinical timeline. Put differently, anything that would show up on a detailed timeline of the patient's care or life would be considered an EVENT. So, a diagnosis would certainly appear on such a timeline, as would a tumor, illness, or procedure, but span-less entities like people (the patient's mother-in-law or the doctor), organizations (the emergency room), or objects (the patient's car) would be quite unlikely to appear.

Another way to think about this is that if it wouldn't make any sense to relate a given span of text to the document time, chances are, the span doesn't represent an EVENT. So, it doesn't make sense to relate a patient's arm to the document's time, but it makes a great deal of sense to relate an MRI, for instance, to the document time ("The MRI occurred AFTER the document was written"). So, in this example, MRI would be an EVENT, but the patient's arm would not.

Our definition of "event" differs from the definition used in everyday life, and that may be counterintuitive, but remember the eventual goal of the annotations: to teach a computer how to turn a clinical note into a timeline of relevant occurrences and states in the patient's treatment and life. If in doubt, make it an EVENT.

In this schema[1], EVENT items don't necessarily have to be actual events in the sense in which the word is conventionally used. Here, states and conditions can just as easily be EVENTs as can surgeries or conditions. Also important is that EVENTs do not have to be verbs, and that for the sake of ease of annotation, adjectives and predications are marked instead of copulas ("the [eye is swollen]" rather than "the eye [is] swollen").

Don't worry about annotating two events in the text that refer to the same actual event. You can make the same "Liver tumor" an EVENT different ten times in the document, so long as there's temporal information relevant to each. There's a secondary co-reference annotation task which is going on in parallel with this annotation which will help us to unravel those links later on.

---

[1]Note that in the TimeML ISO standard (ISO 24617-1:2009(E):2009(E)), EVENTs such as those discussed here are more broadly referred to as "eventualities". "event" has a more specific meaning in ISO TimeML.

Once you've found an EVENT worth marking and the proper span (by referring to the section on determining what to mark for EVENTs), you'll make the annotation using Protege and Knowtator (as described in the "Annotating in Protege and Knowtator" section). Here are some example EVENTs from clinical data:

```
The patient [continues] to [do well] as an outpatient.

Comorbidities include [coronary artery disease]

There are again [noted] [postoperative changes] consistent with prior [right frontoparietal
craniotomy for resection of a right frontal brain tumor].
```

Once you've selected a span, you'll then go through and specify further information about each event. Filling in this additional information will allow you to express the relation of the event to the document time, whether or not it's negated, if it's hypothetical, and a variety of other properties. Please see the EVENT section below for a full description of these properties.

3.1.1   So, what *isn't* an EVENT?

Aside from the obvious, things like people, places, and organizations, there are other verbs and nouns which are not annotated as EVENTs in this schema.

Parts of the patient's natural anatomy are not usually marked as EVENTs. So, in "the patient has a puncture would next to her spleen", "spleen" itself would not be considered an EVENT. The only exception to this would be when the patient is missing a part of the anatomy, or when that part is being otherwise modified. So in "the patient has no spleen", [spleen] would be a negated EVENT, or "the patient's spleen is swollen", the EVENT would be "spleen ... swollen". Once again, you couldn't relate a spleen to the document's time of writing, but you could very easily relate spleen swelling or the patient no longer having a spleen to the document's time.

Note as well that markers of probability or necessity are not events either. So, in a sentence like:

```
It would be unusual for [colonic colon cancer] to [metastasize] to the distal esophagus
```

"Unusual" here is not a relevant state, as it's just the doctor's evaluation of the odds of the metastasis, and thus, it would not be an EVENT. Similarly, "prudent" would seldom be an EVENT ("An MRI is prudent in this case"), and "necessary" is unlikely to be an EVENT either ("If follow-up is necessary in the future, I would be happy to see her.").

Some verbs, when used just to introduce a noun-phrase into the sentence, don't need to be included. The most common example that you'll come across is, well, "is". Because we care more about the actual clinical and medical actions and occurrences that EVENTs represent, we want to mark what's going on, not the copular verb which introduces it. So, in:

```
Given these considerations, this is likely just a [dilated esophagus].

The patient is [stable] at this time.
```

Here, we would not want to mark "is" as the EVENT, instead, we would want to choose "dilated esophagus" or "stable", as that's what is clinically relevant. The same applies for verbs of having and experiencing, as in:

```
The patient has [nausea] and [peripheral edema]

She has been experiencing [neck pain] since July
```

Here, once again, we want to mark "nausea" and "peripheral edema" rather than "has". Note, however, that as previously discussed, verbs of reporting or perception should still be marked (so, "She reported nausea in June" would still have an EVENT for "reported".

There are also other verbs which serve little purpose in advancing the clinical narrative except to introduce nouns. These are light verbs (take in "Jane took a walk"), and they don't represent an occurrence or event of their own, but instead, are just used in discourse. For instance:

```
We will perform a [hemicolectomy]
```

Here, "perform", by definition, occurs simultaneously with the surgery, and is not terribly separate from the event itself. So here, the only EVENT marked would be 'hemicolectomy'.

```
    April 2005 - developed [constipation] for several days
```

Here, "developed" is not marked as an EVENT at all, as it just serves to introduce the constipation (which occurs over several days), rather than truly marking initiation.

```
    If there is evidence of an abnormality at the lower end of the esophagus, biopsies
    can be obtained
```

We could just as easily say "if there is evidence [...], we will biopsy it."

```
    She will try to come back in in approximately four months with to undergo a [CT of
    the chest, abdomen, and pelvis]
```

As before, here, 'undergo' doesn't indicate anything clinically relevant, so the real event, the CT scan, is marked instead. It's worth noting, though, that this is not a property of the verb itself, and in another situation, "perform" might be a real event ("She [reports] that she [performed] in a school play five days after undergoing [surgery]". (It's also worth noting that because a CT of the chest, abdomen and pelvis is a single procedure, it is considered here a single EVENT). Another example of these light verbs would be the verb "administered" in "we administered antibiotics daily".

Simply put, if a verb adds no meaning beyond the EVENT itself, it can safely be skipped. Saying that a patient "underwent surgery", "received antibiotics", or "experienced muscle spasms" is no different than simply making a note in the chart saying "surgery", "antibiotics" or "muscle spasms". Put differently, the generic test for a light verb is whether it can be replaced by simply using the action verb ("He's taking a walk" is not notably different than "He's walking").

Compare this to a sentence like "The patient reported chest pain", we get secondary information from the verb "reported", namely, that the patient provided the information, rather than the doctor witnessing the spasms firsthand. Similarly, "ordered" as in "the procedure should be ordered as a complex procedure" represents a temporally distinct EVENT from the surgery itself. Although this may seem like a very fine-grained distinction, the (un)reliability of patient reports can be a highly relevant to clinical practice, and it's important to capture the distinction where possible.

3.2 Determining what to mark for EVENTs

In Knowtator, every EVENT object is associated with a span (or section) of the text. This span can be a single word, can span several words, and even be non-continuous (where it has two sections separated by words not part of the span). Simply put, any verb, noun, or noun-phrase that is relevant to the understanding of a patient's clinical timeline should be marked as an EVENT, and if you're not sure whether a given span should be an EVENT, discuss it with your supervisor, or failing that, annotate it.

It's worth noting, however, that in this schema, you should never have one EVENT which is entirely contained within the span of another. So, rather than having something like "The patient [reported [nausea]]", we'll simply have "The patient [reported] [nausea]".

EVENTs in this schema, as discussed, can be nouns, verbs, adjectives, or otherwise. They are all marked with the same basic attribute, but the span of the EVENT must differ between verbs and other parts of notes.

3.2.1 Annotating verbal EVENTs

When the EVENT is verbal, as in the below examples, you simply mark the verb, and none of its arguments. So, in the sentence "We discussed the course of treatment and the possible side effects of FOLFOX", you would select only the word "discussed" as the first EVENT (rather than, say "discussed the course of treatment and the...."). Mind you, "treatment", "FOLFOX", and "side effects" (the last two marked as hypothetical) would still

be annotated as EVENTs, but you would not want or need to create a gigantic span to capture all that was discussed.

Verbs of showing, reporting or demonstrating are included as EVENTs and treated as "discussed" was above, as the source of information can be just as important as the information itself in a clinical context. The proper annotation spans are shown in the following examples (in which all EVENTS are marked with [brackets]:

    The [CT] [showed] a [small rectal abcess].

    The patient [reports] [nausea] and [vomiting]

    She does [note] [darkness of the stools]

Events of patient preference are just as relevant as the actual medical treatments, and you will often have to make larger spans including both the patient's preference and what it is that they would prefer:

    She is not [interested] in pursuing [chemotherapy] at this time but is [interested]
    in [continued] [surveillance]

    The patient [prefers] [sedated colonoscopy].

    She [feels] slightly weak but has [resumed] [most of her normal activities]

    I [reviewed] at length with Ms. LastName and her sister-in-law the results of the
    [pathology] from her [surgery] and the fact that there is a [significant risk for tumor
    recurrence.]

    I [reviewed] with her information from Adjuvant! Online suggesting that for someone
    her age there might be a 5% mortality benefit derived from [chemotherapy]

    We also [discussed] some of the [toxicities of fluoropyrimidine-based chemotherapy].

All of the EVENTs above are clinically relevant facts (although some are mostly relevant for documentation and establishing a record of proper care), and the spans are as small as they can be while still describing reality (she may not objectively be "slightly weak", but she feels like it). Note, of course, that within the larger sentence, there are other relevant EVENTs as well (pathology, surgery, risk for recurrence, tumor recurrence (hypothetical))

As you might suspect, the treatments themselves (chemotherapy, surveillance, sedated colonoscopy) would be marked as EVENTs as well. Don't worry, though, these annotations of verbs will be later combined with linguistic structural annotations to capture what was treated, shown, discussed or performed.

3.2.2   Annotating other EVENTs

In this project, EVENT spans are often composed of nouns or adjectives, and we treat them slightly differently than verbal EVENTS. So, with an example like:

    Patient [presents] with [nausea].

You would mark two spans. First, [presents], marks that the patient has come to the office, which is a medically relevant fact, and should be annotated according the verbal guidelines laid out above. The second span would be [nausea], which is clearly also medically relevant, but doesn't have any sort of verb introducing it. In many cases, though, the patient will present with something more complex than a one-word symptom. When determining the span of a given EVENT's annotation, you want to try and find the most specific noun-phrase span possible, which includes all the included information about the event. So, if you have:

    The patient has a [Grade 3 carcinoma with a tubovillious adenoma 10 cm inside Sigmoid
    Colon].

The span will be [Grade 3 carcinoma with a tubovillious adenoma 10 cm inside Sigmoid Colon], not just *[carcinoma] or *[has]. If you're not sure what a term means but think it might be attached to the EVENT, include it.
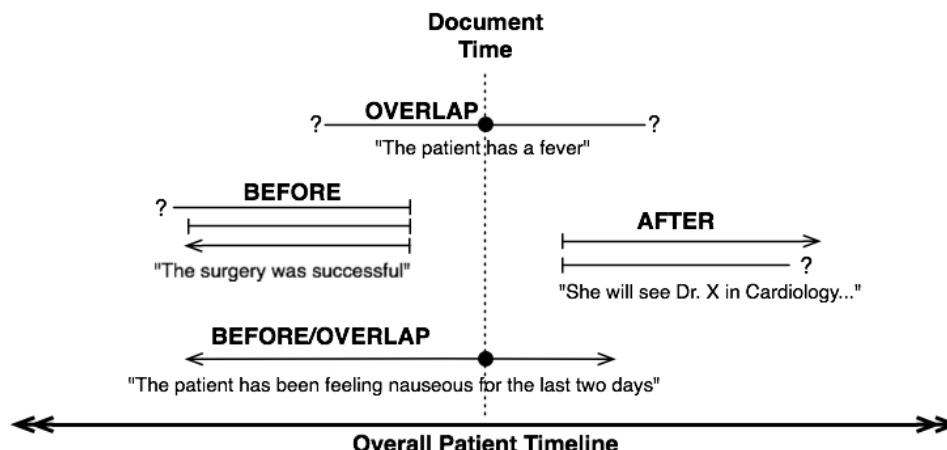
Figure 1: Schematic view of all the DocTimeRel possibilities relative to DOCTIME

Do your best to make sure that each EVENT span includes only one event (so "[nausea] and [vomiting]" rather than "[nausea and vomiting]". Also, because these spans will be used on the timeline, feel free to cut out unnecessary grammatical words using non-continuous spans, as below:

```
[Urine Cytologies> were all <negative]
```

Here, there is just one span, which has both "urine cytologies" and "negative", but not "were all". Once you have a span for your event and have created it, it's time to further specify the attributes of that event.

So, remember, if you're marking a noun as an EVENT, grab the most specific span you can, but if you're annotating verbs, you only need to mark the verb itself. We can figure the rest out from other annotations.

3.3   Annotating DocTimeRel of EVENTs

DocTimeRel is short for "Document Creation Time Relation", and represents the temporal relation between the EVENT in question and the time when the medical record in question was created (the "document time"). For the purposes of this schema, we are assuming that writing of the record itself is functionally equivalent to the time of the patient's visit to the physician. So, anything considered true during the visit will be considered true when the visit was documented by the physician.

DocTimeRel allows us to avoid the linguistic ambiguities inherent in explicitly marking the grammatical tense of the verb (like "past", "present", or "past perfect"), instead marking the actual temporal relations of the event to the time when the document was created (marked with DOCTIME in this schema). We've chosen to use the same temporal relations present in TLINK annotations.

When annotating DocTimeRel on EVENTs, remember that this is the relation of the EVENT in question relative to the moment when the record was written. So an event which occurs before the time of writing would be given the BEFORE value, but an EVENT which will occur after DOCTIME will be given the AFTER value for DocTimeRel. Thus we see that marking DocTimeRel on the EVENT can be thought of as a faster, easier way to temporally link the EVENT to DOCTIME (rather than making a TLINK for every event).

Unlike the other EVENT attributes, DocTimeRel has no default value, as it should be filled in individually for every event. Currently, our schema includes four relations between the event and DOCTIME: BEFORE, AFTER, OVERLAP, and the combined relation BEFORE/OVERLAP.

3.3.1 BEFORE

BEFORE is used where the event occurred before the patient was seen (and thus, before the document itself was written). The bracketed events below would be marked as "BEFORE":

> He is taking in adequate nutrition and adequate fluids; [consumed 3500 calories] and [drank 2-3 liters of fluid]

> This is unchanged and may be related to [treatment changes].

> Today's study demonstrates a marked improvement compared to the prior 9-16-03 [study].

> Until last week, the patient had had no [nausea].

> She had experienced no [dizziness] until the start of chemotherapy.

> The patient had had no [fever] before the start of her surgery.

3.3.2 OVERLAP

OVERLAP is used for events or states which are happening at the time that the patient was seen and thus, when the document was written:

> The patient [continues] to [do well] as an outpatient.

> The patient is [alert], [cooperative], and appears to be in no [acute distress].

> [Moderate sized retention cyst or polyp in the right maxillary antrum] again [noted]

> She is not [interested] in pursuing chemotherapy at this time but is interested in continued [surveillance]

3.3.3 AFTER

AFTER is used where the event is scheduled or planned to occur following the document time:

> [Levaquin 750 mg p.o. q. day] (will restart today)

> The patient will [return> tomorrow <for labs and exam].

> She is not interested in pursuing chemotherapy at this time but is interested in [continued] surveillance

> *(This example (and its twin in AFTER, above) shows the interaction between DocTimeRel and ALINK to cover the idea of "already happening, and will now continue")*

3.3.4 BEFORE_OVERLAP

BEFORE_OVERLAP is used where an event has been happening for an indeterminate amount of time (BEFORE the DOCTIME) and continues into and through the DOCTIME (OVERLAP with DOCTIME). Simply put, this is used where an event started before the exam or patient visit and continues through to the present, and often (but not always) corresponds with the use of the English present perfect tense:

> The [patient> has <felt quite well] and his [appetite> has been <good].

> She has not [seen a cardiologist].

> She has had no [fever].

3.4 Annotating type of EVENTs

Some EVENTs don't actually represent clinical events, but instead, provide aspectual information (starting, stopping, continuing about other EVENTs. To differentiate these EVENTs from the traditional clinical EVENTs which occur on a timeline, we use the "type" marker. It has two values: "N/A" and "Aspectual".

3.4.1 N/A

"N/A" is the default value, and represents the vast majority of EVENTs in the schema, and unless explicitly mentioned otherwise (below or in the ALINK section), all EVENTs used in examples in the Guidelines. Unless the EVENT is of the specific, relatively closed class, listed below under "aspectual", it will be marked as "N/A" (or, as it is the default value, left blank).

One other note on the word "recurrence", which is often troublesome for annotators working with the N/A/ASPECTUAL distinction. In an example like:

    She has significant risk for [tumor recurrence].

"recurrence" does actually carry some aspectual information (the tumor would have restarted). However, because "recurrence" would not be aspectual in "she has risk of recurrence", we have chosen never to mark the work "recurrence" as an ASPECTUAL EVENT. Instead, this will be an EVENT of type N/A, with a span of "tumor recurrence", as shown above. No ALINK annotations will be made here.

3.4.2 ASPECTUAL

The other EVENT type is ASPECTUAL, which is used to indicate an event whose function is to emphasize or code the aspect of a later event, like "continues" or "restart". Every EVENT of type "aspectual" must later participate in an ALINK.

    The rash has not [reappeared] and we will monitor closely

    His anterior chest rash has not [reoccurred].

    The patient will [continue] treatment.

    She is not interested in pursuing chemotherapy at this time but is interested in [continued]
    surveillance

These represent a relatively closed class, and you will find yourself marking the same words as aspectual EVENTs over and over again. This is expected, and should not be cause for concern.

3.5 Annotating polarity of EVENTs

In order to express the polarity of an EVENT, the "polarity" attribute of an event is specified. Polarity in this schema is relatively straightforward, and there are only two possible types: POS and NEG

3.5.1 POS

The first and the most commonly polarity value used is POS. This is the marker of positive polarity. This is used for an EVENT that did, in fact, occur. Most events annotated are of this polarity, and this is the default degree. POS need not be specified in annotation.

    The patient has [hepatosplenomegaly].

    PO [changes] right pterional craniotomy

3.5.2 NEG

The opposite of POS, as you might guess, is NEG, which is used to indicate when the event didn't take place, or has an otherwise negative polarity:

```
No evidence for new [suprasellar mass].

Otherwise, he has not had any [nausea], [vomiting], [diarrhea], [chest pain], [shortness
of breath], or [fever]

She is not [interested] in pursuing [chemotherapy] at this time.
```

3.6 Annotating degree of EVENTs

In order to express an incomplete degree of an EVENT, the "degree" of an event is specified. We have three different values for degree: N/A, MOST, and LITTLE.

3.6.1 N/A, MOST and LITTLE

Our three different degrees are N/A, MOST and LITTLE. N/A is used where there is no need to mark either of the other two degrees on the EVENT, and is the default value for degree. These are used when there has been "a little" of an event, or a large (but not complete) change:

```
There is a small amount of [bright T1 signal].   (LITTLE)

[Abdominal tenderness> has nearly <disappeared].   (MOST)
```

3.7 Annotating contextual modality of EVENTs

Our current schema has three contextual modalities, ACTUAL, HYPOTHETICAL, and HEDGED. Please note that this is unrelated to grammatical mood or modality, and these modalities give information about the modality expressed in the document, not about the grammatical forms used to express them.

3.7.1 ACTUAL and HYPOTHETICAL

The first is ACTUAL, which is used most of the time, and is the default option (that need not be specifically marked). The second modality is HYPOTHETICAL. This is useful when annotating diagnoses, theories, or other medically relevant but hypothetical events. Hypothetical EVENTs will often follow "if" statements ("If X happens, then we'll use Y to treat Z") or other sorts of conditionals ("Depending on the patient's response, we might treat A with B or with C").

```
I've warned the patient that this new medication may cause [peripheral edema]
```

It's worth noting that having a DocTimeRel of AFTER does not imply that the EVENT is HYPOTHETICAL (although most hypothetical EVENTs will be AFTER DocTime). Although it's true that there's always a degree of uncertainty with anything happening in the future, HYPOTHETICAL marks explicit uncertainty in the text, and should not be used just to indicate this future-uncertainty.

For instance:

```
The patient's [myringotomy] will take place on Friday.
```

Here, Myringotomy has a DocTimeRel of AFTER, but is ACTUAL. Compare that to:

```
If she has [additional bleeding] next week, she should [come back in]
```

Here, both "additional bleeding" and "come back in" would be HYPOTHETICAL, in addition to being AFTER the DocTime.

3.7.2 HEDGED

EVENTS are marked as hedged when the doctor mentions a given EVENT with any sort of hedging. This hedging can be lexical ("seems", "likely", "suspicious", "possible", "consistent with"), or phrasal ("I suspect that...", "It would seem likely that"). These EVENTs are strongly implied, but, for safety, liability, or due to lack of comprehensive evidence, are not stated as fact by the doctor. As such, it's very important that these hedged diagnoses and findings be included in the timeline, but be marked so that they can be easily differentiated from hard and fast diagnoses.

```
Ultrasound findings were felt to be consistent with a [T3, N1 rectal tumor]

An approximately 3cm nodular region of intermediate T2 signal involving the body of
the corpus callosum is suspicious for residual or recurrent [tumor] but appears unchanged
from the patient's prior examination.

She has a rash not inconsistent with [measles].

The patient may have undergone a [mild stroke].
```

It is also possible for HEDGED to be used for EVENTs which occur in the future, but in which the doctor explicitly hedges whether or not they will occur.

```
Recommendations were given to [resume] the [Cipro] and [Flagyl] and obtain a [CT of
the chest, abdomen and pelvis].
```

Although this is likely functionally equivalent to "We will resume...", the extra statement is worth marking with hedged, just to be safe.

3.8 Annotating contextual aspect of EVENTs

Aspect is used to express aspectual ideas about the events which are not coded explicitly with aspectual EVENTs and ALINKs. We have three values for contextual aspect in the schema, N/A, INTERMITTENT and NOVEL. Please note that this is unrelated to grammatical aspect, and these two aspects give information about the temporal relations in the document, not about the grammatical forms used to express them.

3.8.1 N/A

N/A is the most common value (and our default value) for contextual aspect, and simply represents that a given EVENT is neither NOVEL nor INTERMITTENT. If neither of the other contextual aspects seems to fit for a given EVENT, leave the contextual aspect field blank (which will then be auto-filled with N/A).

3.8.2 NOVEL

The NOVEL indicates, well, novelty, and is associated with predicate adjectives like "new".

```
The patient's new [tumor] is 3.5cm from the epiglottis.

The newest [MRI] revealed a previously undiscovered mass.
```

3.8.3 INTERMITTENT

INTERMITTENT is used in situations where there may be a series of smaller events within a single EVENT, rather than a single, constant event. Those events are usually marked with words like "intermittently" or "occasionally", and when such phrasing is used, the EVENT is marked as INTERMITTENT. These indicate that, for instance, the patient has had vomiting since a certain time, but he/she has not been vomiting 24/7 since that point.

Please note that INTERMITTENT will only be used for irregular, unpredictable periods (like the span between seizures or vomiting) and not for things like medications or dialysis (which occur on a set schedule). An event which occurs at an explicit interval can either be treated as a constant ("the patient is taking montelukast for asthma") or the interval can be marked as a TIMEX3 ("She undergoes dialysis [every three days]"), which is then TLINKed to the original event.

It's important to note that we are only marking INTERMITTENT when there is an explicit mention of intermittency in the sentence. Even if you happen to know that a given disorder or symptom often manifests intermittently, if it's not stated explicitly as doing so, you should not mark it as such. As with all of these annotations, we are marking the relations mentioned in the document, not those you can infer from your own background knowledge.

```
He reports occasional [bright red bleeding] from the rectum.

Patient complains of intermittent [chest pain]
```

If you are unsure about the consistency of a given EVENT, mark it as UNDETERMINED.

3.9 Annotating permanence of EVENTs

Because we'll be creating a timeline from this data, sometimes, it's valuable to know whether a given condition or disease is FINITE (it might go away with treatment), or is PERMANENT (and the patient will have it for the rest of their lives).

3.9.1 FINITE, PERMANENT and UNDETERMINED

This is a property of the event and disease itself, not of the patient's specific disease. This means that diseases like pneumonia or different cancers are always treated as FINITE even if the individual patient is likely to die before the disease can be resolved. This is as opposed to a condition like Multiple Sclerosis or Diabetes, which is unlikely to be completely cured (given the state of medicine at the time of writing) and is thus PERMANENT. A disease that the patient will have for the rest of their lives, despite any medical intervention, is classified as PERMANENT.

This is medical-domain specific knowledge, and as such, the default value is UNDETERMINED. If you are unsure as to the Permanence of an event or condition, leave it unmarked and it will be referred to a domain expert.

4 TIMEX3 Annotation

Because we're looking specifically at temporal relations, the next step of the annotation process is to find and annotate TIMEX3 objects. These are definitive references to time, and will provide concrete temporal references throughout the document or section. Examples of these might be words like "today", "tomorrow", "24 hours ago", and "early March". In addition, specific dates are annotated as TIMEX3 objects as well.

4.1 Annotating TIMEX3 class

4.1.1 DATE

The majority of TIMEX3 annotations you make will be of the class DATE or TIME. DATE represents dates. These can be calendar dates (such as [January 4]) and relative dates (such as [Last week]).

    MRI of the brain without and with gadolinium contrast utilizing tumor followup protocol
    compared with prior studies of [29, February 2005] and [28, January 2005].

    His anterior chest rash has not reoccurred since the PCN VK was discontinued [24-hours
    ago].

    The last cyclosporine level was 373 in [January]. His dose was adjusted downward to
    300-mg twice-daily. A cyclosporine level will be repeated on [Friday morning].

4.1.2 TIME

TIME is used for specific time points within a day, for instance, [3:00PM] or [23:45], and once again can be relative ([20 minutes from now])

    The patient's MRI is scheduled for [5:30pm]

    Following the patient's latest seizure, [20 minutes ago], we are re-evaluating her
    medications.

    Surgery will need to be completed by [2:45] to have the biopsy to the lab sooner.

4.1.3 DURATION

Sometimes, an entire span will be specified in a TIMEX3. These are things like "For 24 hours" or "All of February", and these are marked with the class DURATION.

    The patient continuously experienced nausea for [nearly two weeks].

    [For the next 12 hours], we will lower the patient's morphine drip and then we will
    re-evaluate his pain.

    [Since August], she has not had any episodes.

4.1.4 QUANTIFIER

Although it may seem odd at first, expressions like "Twice", "four times", and "18 times in the month of May" are all TIMEX3s. These are annotated with the QUANTIFIER class.

    The patient vomited [twice] before the surgery.

    We have seen Mr. Lastname [three times] for his ulcerative colitis.

    [On two to three incidents] she has had blood in the stools

4.1.5 PREPOSTEXP

Similarly odd, Pre- and Post- expressions ("preoperative", "post-exposure", "post-surgery", "prenatal") all actually designate specific temporal spans ("The time before the surgery", "The time after exposure"), and thus, are TIMEX3s, marked with the class PREPOSTEXP. Usage of this TIMEX3 is discussed in more detail in the Special Cases and Considerations section of the handout.

```
Patient underwent a partial hemicolectomy in July 2009.  [Postoperative] scarring noted
during exam.
```

```
The patient exhibits [post-exposure] changes.
```

4.1.6 SET

SET is used exclusively in our schema for covering expressions which give both a quantifier and an interval (like "Three times weekly", "monthly" or "1/day"). This is different from QUANTIFIER ("twice") which only gives a quantifier, and different from DURATION ("all week") which only gives a span.

```
Will administer Lariam [twice daily].
```

```
Patient has checked into the ER [roughly three times a month].
```

```
We will proceed with [weekly] consultations to monitor the patient's condition.
```

```
Mirtazapine [REMERON] 7.5-mg tablet 1 tablet by mouth [every bedtime].  Simvastatin
[ZOCOR] 20-mg tablet 1 tablet by mouth [one-time daily].
```

TIMEX3s of type SET should always be TLINKed to EVENTs using the TLINKs of the type OVERLAP.

5   Narrative Containers

One very important concept for the marking of temporal relations is that of the narrative container, discussed extensively in Pustejovsky and Stubbs 2011 [1]. A narrative container can be thought of as a temporal bucket into which an EVENT or series of EVENTs may fall. These are often dates or other temporal expressions (within which a variety of different EVENTs occur), although they can also be more abstract concepts ("recovery" which might involve a variety of EVENTs) or even EVENTs themselves (many other EVENTs can occur during a surgery).

It's worth explicitly stating, though, that these narrative containers are not anything which need to be marked explicitly by annotators, as they're a property of the timeline itself. Instead, the idea of these containers are being presented as an aid to visualizing the temporal relations within a document. If the annotator does his or her job correctly, properly implementing DocTimeRel and creating proper TLINKs, a good understanding of the narrative containers present in a note will naturally emerge from the annotated text.

To give an example of this somewhat abstract concept, let's look at some of these containers in the wild.

5.1   *Narrative Containers from DocTimeRel*

All documents have at least three narrative containers, marked by DocTimeRel. One bucket contains all EVENTs which occur BEFORE the document time, one contains all EVENTs which occur AFTER the document time, and one contains all EVENTs which overlap the document time. Take, for example, this very simplified note section:

```
The patient recovered well after her first surgery to remove the adenocarcinoma, and
today we reviewed the pathology results and the results of her follow-up CT.  Based
on these results, pending her consult with Dr. Hart, I feel she is fit to undergo her
second surgery in three weeks.
```

Here, the red-underlined portion of the note all occurs in a narrative container which ends at the day of the visit, the blue portion all occurs during the course of the visit, and the green section all occurs after the visit, in a narrative container which is bounded by the visit. In this case, all of this information is captured explicitly through DocTimeRel, and the boundaries of these narrative containers are all set by the boundaries of the patient's visit which is being described in the present note.

5.2   Annotating additional narrative containers

However, there can be many more narrative containers in a given note than just "Before DocTime", "Overlapping DocTime", and "After DocTime". Take, for example, this note:

> The patient recovered well after her first surgery on December 16th to remove the ade
> nocarcinoma, although on the evening of January 3rd she was admitted with a fever and
> treated with antibiotics. Today we reviewed the pathology results and the results of
> her follow-up CT. She will have her second surgery in three weeks, followed by a third
> surgery on April 19th which will involve both additional resection as well as the tak
> ing of additional biopsies.

Here, there are actually six containers, the contents of each marked again with differently colored underlines. The first surgery occurred entirely within the December 16th container, the fever, admission and antibiotics occurred on January 3rd (and within that container), the followup CT and her recovery occurred at an ambiguous point between the surgery and the DocTime, the during-visit container, the second surgery occurs three weeks from the DocTime, and the final container involves everything occurring on April 19th (the surgery, resection, and taking of biopsies).

In this, more complex example, DocTimeRel alone is not sufficient to mark the temporal relations between all of the EVENTs and to clearly associated all of the EVENTs the narrative containers which they are a part of. Additional detail is needed here to capture that in addition to being BEFORE the Document Time, the first surgery happened on December 16th, whereas the fever occurred on January 3rd. We could also state that the follow-up CT occurred after December 16th. We also need to know that, for instance, the April 19th surgery will happen AFTER the surgery in three weeks. This clarification of narrative containers, showing which EVENTs happened in which containers, as well as the temporal relations between them, is the role of TLINKs.

So far, we've been discussing temporal expressions (TIMEX3s) as providing the boundaries of narrative containers, but EVENTs themselves can provide boundaries too. Take this example:

> December 19th: The patient underwent an EKG as well as emergency surgery. During the
> surgery, the patient experienced another MI, and repeated bouts of tachycardia

Here, we can see that in addition to the overall container of December 19th (containing both surgery and an EKG), the surgery itself is a container, containing another MI (myocardial infarction or "heart attack") as well as tachycardia. In this case, for the temporal structure of the note to really be captured, both EKG and surgery need to be linked to December 19th, and both the MI and Tachycardia need to be linked to the surgery, all done using TLINK annotations.

5.3   Ordering within narrative containers

Because of the difficulty of capturing detail within a given narrative container, not all relations between EVENTs will be captured. For instance, in the above sentence, we do not explicitly give the relationships between EKG and Surgery, nor between the MI and Tachycardia. In many cases, those relationships within a narrative container will follow a set progression (first colonoscopy, then biopsy (which occurs during the colonoscopy), then pathological analysis, then histology), but you as an annotator are not asked (or allowed) to make explicit those domain-specific progressions.

So, although it may seem like some of the narrative containers that you define may be a loose bucket of temporal EVENTs, that's not actually a problem, as the greater timeline of the patient's care is more important than the fine structure within a given narrative container, and as you'll find out, there are still plenty of TLINK annotations to be made.

6   TLINK Annotation

TLINKs, as previously mentioned, are relations you can mark between EVENTs, between TIMEX3s, or across the two categories to show the temporal relations present within the document and to clearly define the bounds of the narrative containers at work beyond what DocTimeRel will naturally give us. Although these are annotated in a more interactive fashion using Knowtator, they have the basic format:

```
EVENT1 [temporal relation] EVENT2/TIMEEX3
```

To give a realistic example, imagine the following sentence:

```
The patient was to follow-up with oncology this month
```

The EVENT [follow-up] is clearly related to the TIMEX3 [this month], because [follow-up] will occur during [this month], giving you more specificity than the followup just occurring after document time. So here, you would create a TLINK object, insert [follow-up] into the "event" slot, [this month] into the "related to" slot, and select CONTAINS as the relation. By doing so, "This month" is established as a narrative container as well, which may contain additional EVENTs later on in the note.

6.1   When to TLINK

TLINKs themselves are relatively straightforward, and in fact, the more difficult part of annotating TLINKs is to know when to stop. Without any constraint, one could see making TLINKs between every EVENT in the document, which leads to exponential growth of TLINKs and a tangle of relations which nobody, let alone a machine, would like to unpack.

So, to constrain this process a bit, we have developed four rules to govern TLINKing:

6.1.1   TLINK only when it captures more information than just marking DocTimeRel

Because the 'DocTimeRel' attribute of EVENTs expresses the relation of the event to the time the document or section was written, you will never need to TLINK to the DOCTIME annotations. Marking an EVENT as 'after' in the DocTimeRel field gives us the same information as making an 'after' TLINK between the EVENT and DOCTIME, so you need not explicitly mark that. Similarly, if one EVENT has a DocTimeRel of OVERLAP and another has a DocTimeRel of AFTER, there's no need to make a TLINK between those two EVENTs.

6.1.2   TLINK all EVENTs to their narrative container, if possible

As previously discussed, most EVENTs will fall into a narrative container of some kind. If a given EVENT is in a narrative container (like "August 22nd" or "during her recovery"), you should always TLINK that EVENT to the TIMEX3 or EVENT which represents that narrative container, using the appropriate link. Once again, though, this should only be done if the result will be more informative than just analyzing the DocTimeRels.

So, in:

```
December 19th:  The patient underwent an [EKG] as well as [emergency surgery].  During
the [surgery], the patient experienced another [MI], and [repeated bouts of tachycardia]
```

December 19th should be TLINKed (Type:CONTAINS) to both EKG and Emergency Surgery. Then, Surgery should be TLINKed (again, type: CONTAINS) to both the MI and the Tachycardia.

In most cases where you have a test and then a series of results, the test itself will form the narrative container for the results, and the day on which the test occurred will itself CONTAIN the test:

```
[Colonoscopy] ([January 7, 2010)]:
1.  [Diminutive polyps of the rectosigmoid], [removed].
```

Here, the colonoscopy will be TLINKed (type:CONTAINS) to the diminutive polyps and their removal, and then January 7th, 2010 will be TLINKed (CONTAINS) to the colonoscopy itself. Note that, as described below, all test results, regardless of whether the conditions they observe may have existed previously, are TLINKed to the test using CONTAINS.

Not every EVENT will be a part of a detailed narrative container (more specific than before, after, or during DocTime). However, it's vital that you, as an annotator, stop to ask yourself whether each EVENT you examine is a part of a narrative container, and whether the TLINKs you created are sufficient to mark that membership.

6.1.3 TLINK all explicitly stated temporal relations

If the doctor goes out of his or her way to make a temporal statement, a TLINK should be made to reflect that statement. So, if the sentence reads:

```
The patient [developed] a [rash] after [treatment]
```

You should explicitly mark both "developed" and "rash" as being AFTER treatment, using a TLINK. Note, though, that if no explicit temporal language is used, no TLINK should be created, and annotator knowledge should not be used to fill these TLINKs in.

6.1.4 TLINKs should only link EVENTs and TIMEX3s within the same sentence

Nearly all of the TLINKs you make should occur across two EVENTs or TIMEX3s in the same sentence. If, for some reason, you need to link to an EVENT or TIMEX3 in a previous sentence to put an EVENT in the proper narrative container, you may do so, but you should double-check to ensure that there's no other way of going about it. Because of the nature of the notes, **TLINKs should never link items in different sections.**

6.2 TLINK sub-types

There are six different temporal relations often used in this schema, BEFORE, AFTER and OVERLAP, BEGINS-ON, ENDS-ON and CONTAINS. There are other types of TLINK relations in the current version of the schema, but these five are the most commonly used. There is no default relation type for TLINKs

6.2.1 BEFORE and AFTER

BEFORE and AFTER are both fairly straightforward, and they simply order two events in time. When possible, BEFORE should be used for consistency, but in some cases when TIMEEX3 annotations are used, AFTER is necessary.

```
He does have a history of peri-rectal abscess with his last round of chemotherapy.
```

(chemotherapy BEFORE abscess)

```
His anterior chest rash has not [reoccurred] since the PCN VK was [discontinued] 24-hours
ago.
```

(discontinued BEFORE reoccurred, discontinued OVERLAP 24-hours ago)

```
She was seen by Dr.  XXX in cardiology following the stent placement
```

(seen by... AFTER stent placement)

6.2.2 CONTAINS

CONTAINS signals that the EVENT is completely contained within the span of the EVENT or TIMEX3 it's related to. In other words, the contained event occurs entirely within the temporal bounds of the event it's contained within. This relation is most often used to mark when an EVENT is contained entirely by a narrative container.

CONTAINS is a very specific relation implying complete containment within a narrative container, and if you annotate that X CONTAINS Y, it's assumed that there's also an OVERLAP relation between the two. You should only use CONTAINS when you're quite sure that the nature of the overlap is one of complete containment.

> March 2005 - Patient underwent appendectomy

(March 2005 CONTAINS appendectomy)

> Levaquin 750 mg p.o. q. day (will [restart] [today])

(restart OVERLAP today)

> Comparison is made with prior [MRI head examination without and with gadolinium] from [10-23-03].

(MRI head examination... OVERLAP 10-23-03)

> An ENT performed the myringotomy during Friday's surgery.

(surgery CONTAINS myringotomy)

> Gengraf 300-mg p.o. b.i.d. ([decreased] in [early June])

(early June CONTAINS decreased)

> The patient's first MI occurred while she was undergoing chemotherapy.

(chemotherapy CONTAINS MI)

In addition, we have made one specific regulation involving the use of CONTAINS: All test results or observations are to be linked to the test which generated them (their narrative container) using a CONTIANS relation. This is not always intuitive, because as humans with some inferencing ability, we realize that the tumor likely existed before the CT scan which revealed it, but from a machine-learning perspective, it's important to have that consistency. So, in a section like:

> [Colonoscopy] ([January 7, 2010)]:

> Postop Diagnosis:

> 1.   [Fair/adequate prep.]

> 2.   [Limited colonoscopy] to the distal sigmoid due to an [obstructive lesion].

> 3.   [Diminutive polyps of the rectosigmoid], [removed].

(January 7, 2010 CONTAINS Colonoscopy)

(Colonoscopy CONTAINS Fair/adequate prep)

(Colonoscopy CONTAINS Limited Colonoscopy)

(Colonoscopy CONTAINS Obstructive Lesion)

(Colonoscopy CONTAINS Diminutive polyps of the rectosigmoid)

(Colonoscopy CONTAINS removed)

6.2.3 OVERLAP

OVERLAP is a single temporal relation that encompasses all the different notions of two things happening at the same time, but is less specific than CONTAINS . This can refer to two simultaneous events, an EVENT that occurs during another, larger EVENT or time reference (but where containment is not entirely sure), or any other sense in which two events are occurring in the same timeframe:

> The patient had some [rectal itching] and [mild pain] [this morning].

> *(rectal itching OVERLAP this morning, mild pain OVERLAP this morning)*

> She is not [interested in pursuing chemotherapy] at [this time]

> *(interested in pursuing chemotherapy (negated) OVERLAP this time)*

In short, OVERLAP is meant for situations where two events overlap in some way, but where you're not sure (or don't have enough information to tell) whether there is containment.

OVERLAP is also used for linking TIMEX3s of type SET with other EVENTs.

> We'll keep her on [rate-control medications] 100 mg [twice daily]

Here, [twice daily] OVERLAPs [rate-control medications]

6.2.4 BEGINS-ON

BEGINS-ON signals that the EVENT begins on the EVENT or TIMEX3 it's related to. This type of TLINK will only occur with EVENTs which have a non-trivial temporal span.

> She has had [Abdominal Cramping] since [January].

> *(abdominal cramping BEGINS-ON January)*

> He reports intermittent [chest pain] since his prior [MI]

> *(chest pain BEGINS-ON MI)*

6.2.5 ENDS-ON

ENDS-ON signals that the EVENT ends on the EVENT or TIMEX3 it's related to. As with BEGINS-ON, this type of TLINK will only occur with EVENTs which have a non-trivial temporal span.

> His [anterior chest rash] has not [reoccurred] since the [PCN VK> was <discontinued] 24-hours ago.

> *(anterior chest rash ENDS-ON PCN VK...discontinued)*

7 ALINK Annotation

The final sort of annotation performed on the data is the ALINK (or "aspectual link") annotation. ALINKs are created between an aspectual EVENT and a non-aspectual EVENT. Any EVENT previously marked with the class ASPECTUAL will be ALINKed to another, non-aspectual event, and you will never make an ALINK which includes a TIMEX3 or two non-aspectual EVENTS. ALINKs only occur when you have an aspectual EVENT, and are much less common in the text than TLINK annotations. Relatively few EVENTs will have an associated ALINK.

Unlike TLINKs, no matter the circumstance, an ALINK should never cross a sentence boundary.

As with TLINKs, they are created interactively within Knowtator, but have the basic form:

```
    EVENT1/PRD-ADJ [aspectual relation] EVENT2
```

7.1 ALINK sub-types

There are four different aspectual relations used in the schema, CONTINUES, INITIATES, REINITIATES, and TERMINATES.

7.1.1 CONTINUES

CONTINUES is used when an aspectual event shows the continuation of another event:

```
 We will [continue] to [monitor LFTs carefully] along with his weight.
```

(continue CONTINUES monitor)

```
 The patient will [remain] on [dialysis] until her condition changes.
```

(remain CONTINUES dialysis)

```
 She is not interested in pursuing chemotherapy at this time but is interested in [continued]
[surveillance]
```

(continued CONTINUES surveillance)

```
 We will [continue] to [monitor] heart rate and rhythm along with serial cardiac markers
and electrocardiograms to rule her out for any cardiac involvement
```

(continue CONTINUES monitor)

7.1.2 INITIATES

INITIATES is used when an aspectual event indicates the start or initiation of another event:

```
 Patient will [begin] a [high-fiber diet] upon release.
```

(begin INITIATES diet)

```
 We will [start] Ms.  Miller on a [normal saline infusion at 75 an hour for a total
of 1 L]
```

(start INITIATES normal saline infusion at 75 an hour for a total of 1 L)

7.1.3 REINITIATES

REINITIATES is used when an aspectual event indicates that another event will be restarted or reinitiated:

```
 [Levaquin 750 mg p.o.  q.  day] (will [restart] today)
```

(restart REINITIATES Levaquin 750 mg)

```
 His anterior chest [rash] has not [reoccurred] since the PCN VK was discontinued 24-hours
ago.
```

(reoccurred REINITIATES rash)

7.1.4 TERMINATES

TERMINATES is used when an aspectual event indicates the ending of another event:

> Because of this reaction, [Allegra] will be [discontinued]
>
> *(discontinued TERMINATES Allegra)*
>
> Patient [nausea] was successfully [stopped] by 1-mg Ativan p.r.n.
>
> *(stopped TERMINATES nausea)*

8 DocTime and SectionTime Annotation

Because all of these annotations are linking the EVENTs and TIMEX3s to the time of patient service, it's important that we specify what that time actually is. Each note should include a line at the very start or end of the document which states the service time, and we select that time as the DocTime. DocTime can be considered a special sort of TIMEX3, and is annotated similarly.

It's worth noting though, that because of the complexities of medical record extraction, some sections (separated by [start section id="xxxxx"] blocks) might have a separate, local time. These are often vital signs and medication sections, and they will always include a specific date, which will be marked not with DocTime, but with SectionTime. In these cases, any DocTimeRel annotations will link to this, section-specific time. *If no separate date is explicitly given for a section, assume that it shares the overall DocTime.*

So, at the top of the document, you'll find a line like the below:

> [head start date="[12/13/2010]" rev date="12/13/2010" rev="0002"]

Here, the [12/13/2010] would be marked as DocTime

Sometimes, as discussed before, a specific section has an explicit, different timestamp than the rest of the document.

> [start section id="20104"] Alendronate FOSAMAX 70 mg tablet 1 tablet by mouth every
> week. Colace 100-mg capsule 1 capsule by mouth two times a day. These are the patient's
> medications as of [Saturday, February 20, 2010 at 3:34 PM]. [end section id="20104]

Here, [Saturday, February 20, 2010 at 3:34 PM] would be considered SectionTime, and would be annotated as such.

9 Schema Description

In an effort to clarify the schema at a glance, here's a quick summary of the Clinical Temporal Relations Annotation Schema, using the structure described below:

- **Entity/Relation Type**
  - Property of that entity/relation type
    * Choice 1 for that property
    * Choice 2 for that property
    * ...
- **EVENT**
  - DocTimeRel
    * BEFORE
    * OVERLAP

           * AFTER
           * BEFORE/OVERLAP
     – Type
           * N/A
           * ASPECTUAL
     – Polarity
           * POS
           * NEG
     – Degree
           * N/A
           * MOST
           * LITTLE
     – Contextual Modality
           * ACTUAL
           * HYPOTHETICAL
           * HEDGED
     – Contextual Aspect
           * N/A
           * NOVEL
           * INTERMITTENT
     – Permanence
           * FINITE
           * PERMANENT
           * UNDETERMINED

- **TIMEX3**

     – Class
           * DATE
           * TIME
           * DURATION
           * QUANTIFIER
           * PREPOSTEXP
           * SET

- **TLINK**

     – Event
           * (This is a slot, designed to be filled with the EVENT being TLINKed)
     – Related to
           * (This is a slot, designed to be filled another EVENT or TIMEX3 to which the first EVENT is being temporally linked)
     – Type
           * BEFORE
           * AFTER
           * OVERLAP
           * CONTAINS
           * BEGINS-ON
           * ENDS-ON

- **ALINK**

     – Event
           * (This is a slot, designed to be filled with the EVENT (of type "ASPECTUAL") being ALINKed)
     – Related to
           * (This is a slot, designed to be filled another EVENT to which the first EVENT is aspectually linked)
     – Type

      ∗ CONTINUES
      ∗ INITIATES
      ∗ REINITIATES
      ∗ TERMINATES

- **DOCTIME**

- **SECTIONTIME**

9.1 Default Values

Some of the attributes of EVENTs, TLINKs, TIMEX3s and ALINKs have default values, which, if the attribute is not marked, will be assumed. This is a list of those default values:

Default Values for EVENT

EVENT.type	(N/A)
EVENT.DocTimeRel	Must be specified
EVENT.polarity	POS
EVENT.degree	(N/A)
EVENT.contextualaspect	(N/A)
EVENT.contextualmodality	ACTUAL
EVENT.permanence	UNDETERMINED

Default Values for TIMEEX3

All values for TIMEX3 annotations must be specified for each instance

Default Values for TLINK

All values for TLINK annotations must be specified for each instance

Default Values for ALINK

All values for ALINK annotations must be specified for each instance

10 Special Cases and Constructions

10.1 Family History/History of...

Patient history and family history are treated as EVENTs in our schema. So, in the sentence:

    Past [medical history]: [Common variable immunodeficiency]

    No [history] of [cough] or [aspiration pneumonias]

    (Common variable immunodeficiency CONTAINS medical history)

We have two (or more) EVENTs. [medical history] is an EVENT whose DocTimeRel is BEFORE_OVERLAP (because the patient's medical history started before the document time, but continues through the note). [Common variable immunodeficiency] is also an EVENT, with a DocTimeRel of BEFORE (as the patient is not experiencing this immunodeficiency at the time of writing). The TLINK between Common variable immunodeficiency and medical history simply indicates that the disorder occurs as part of the medical history, and because we're treating history as a persistent state, during the patient's medical history. As such, if the sentence read "Common variable immunodeficiency and diabetes", Diabetes would also be an EVENT, and would also be TLINKed to "medical history".

Family history works similarly:

> [Family history] Mother and father with [allergic rhinitis], siblings with [allergic rhinitis] and [atopic eczema]

> *(Family history CONTAINS allergic rhinitis), (Family history CONTAINS atopic eczema)*

Once again, "family history" is an EVENT, DocTimeRel BEFORE_OVERLAP, and each of the disorders described are EVENTs with DocTimeRel BEFORE. The TLINKs mark co-occurrence of the various disorders and the abstract "family history".

10.2   Pre- and Post- expressions (preoperative, post-treatment, etc)

The adjectives "preoperative" and "postoperative" (as well as related terms like "post-surgical", "post-injury", "pre-treatment", "post-partum") and present a particular difficulty for our schema when used without explicit mention of their referents, as they simultaneously express two inseparable temporal meanings: first, they introduce an EVENT (like an operation or an injury), and second, they tell us that the noun they modify occurred either before or after this operation.

Because "preoperative" and "post-treatment" (and their ilk) actually denote temporal spans ("everything until the surgery", "any time after the treatment"), they're actually a sort of TIMEX3. So, all of these pre- and post-expressions will be marked as being TIMEX3 of the type "prepostexp" ("Pre- Post- Expression").

> Patient underwent a [partial hemicolectomy] in [July 2009]. [Postoperative] [scarring] [noted] during [exam].

> *(partial hemicolectomy OVERLAP July 2009), (partial hemicolectomy BEFORE scarring), (partial hemicolectomy BEFORE noted), (partial hemicolectomy BEFORE exam), (exam CONTAINS noted), (scarring OVERLAP postoperative), (postoperative BEGINS-ON partial hemicolectomy)*

Here, you can see that there's a large amount of temporal information expressed here, but most important is the [postoperative] TIMEX3 and the two associated TLINKs. The scarring is TLINKed to postoperative with OVERLAP because the scarring occurs during the span of time which "postoperative" designates. Then, we mark that postoperative BEGINS-ON the surgery, because the surgery's end represents the start of the postoperative period.

This is relatively straightforward, but there's not always an explicit reference to the surgery in the note:

> The patient exhibits postoperative changes.

Here, we have no EVENT for the operation which we can TLINK the "changes" to. In this situation, "postoperative" would still be a TIMEX3, and the changes would OVERLAP

> The patient exhibits [postoperative] [changes].

> *[postoperative] is a TIMEX3 whose Type is PREPOSTEXP, (changes OVERLAP postoperative)*

A second example of this would be:

> Patient is [in recovery], no [post-operative] [nausea].

> *[postoperative] is a TIMEX3 whose Type is PREPOSTEXP, [nausea] is an EVENT whose DocTimeRel is OVERLAP and whose Polarity is NEG, [in recovery] is an EVENT whose DocTimeRel is OVERLAP, (nausea OVERLAP postoperative)*

Once again, post-operative here stands in for the surgery, allowing us to capture that the lack of nausea occurred following the unmentioned procedure.

    The patient's [preoperative] [health is good].

> *[preoperative] is an EVENT whose DocTimeRel is AFTER, [health is good] is an EVENT whose Doc-TimeRel is OVERLAP, (preoperative AFTER health is good)*

This is a bit more opaque, but preoperative in this case stands in for the surgery it introduces, which presumably will occur in the future (given the tenses in the sentence). So, "preoperative" will be an EVENT, which will be AFTER the document time. The "health is good" EVENT is occurring at the document time, so it has a DocTimeRel as OVERLAP, and then finally the TLINK shows us that the surgery will be occurring after the good health.

If the sentence read "The patient's preoperative health is good, and next week's tonsillectomy should proceed without difficulty", we would TLINK "health is good" to "tonsillectomy" (BEFORE) in addition to the link to the PrePostExp. The majority of cases involving pre- and post- expressions will include referents and thus should be relatively transparent to annotate, but in the event of a missing referent, the above guidelines should allow you to capture the temporal information implied.

10.3 Annotating the word "plan" or "plans"

Plan is a temporally difficult word. In some cases, it's clearly temporally relevant:

> *I'd discussed that she'll need to come in for the EGD with biopsies. followed by a CT scan. We're hoping to implement this plan in the next few weeks.*

In this case, it's clearly an EVENT as it needs to be implemented at a specific time and TLINKed with 'next few weeks'. But in some cases, its status as an EVENT is more ambiguous.

> *I have explained this plan to both Miss Mullins and the primary team.*

Here, there's no strong TLINK, and the plan itself isn't necessarily temporally bound.

To avoid this ambiguity and improve annotator agreement, we've decided that **in all cases, "plan" will be considered an EVENT**. Do your best to annotate DocTimeRel and all other EVENT properties, making TLINKs which may be necessary, but no matter the usage, we should always be marking 'plan' (or 'plans') as an EVENT. This rule applies to both verbs and nouns, where relevant.

11 Non-annotated sections of clinical notes

Not all portions of the clinical note should be temporally annotated. The following sections, given both with a Mayo section descriptor and a qualitative description, should not be annotated at all (for a variety of reasons). These sections should not be annotated even if not labeled with the section id given below, or if they're included in another section (as "Patient Education" and "Advance Directives" often are). That said, if you're not sure whether a given section of a note should be included, contact your annotation supervisor.

- 20104 - Current Medications
- 20105 - Allergies
- 20116 - Advance Directives
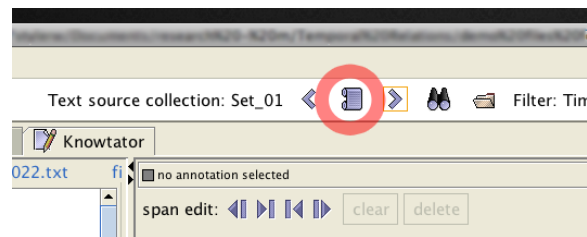- 20138 - Patient Education

Figure 2: The icon for selecting text sources. Note the "Next text source" and "Previous text source" buttons to either side

12 Annotating in Protege and Knowtator

Annotation for this project will be done in Protege and Knowtator, a set of open-source programs designed for creating, implementing and storing ontologies and annotations. These guidelines are written assuming that you're using Protege 3.3.1 and Knowtator 1.9 beta 2, and that you've already accessed the data referring to the *Accessing Clinical Data in Protege/Knowtator* document. The following instructions walk you through the steps required to annotate a document using the schema described above using Knowtator and Protege.

12.1 Saving your progress in Knowtator

Protege and Knowtator do not automatically save your work, so you'll need to do so manually. In order to save, simply click the Floppy Disc icon in the toolbar (directly under "Project").

Please remember, Protege and Knowtator are not perfect, nor is the connection to the server where the data resides. As such, a program bug or internet dropout could disrupt your annotation and cause you to lose all your changes at any moment. As such, **save your data often!**. It's not at all unreasonable to save following each completed EVENT, TIMEX3, ALINK, or TLINK.

12.2 Opening text sources in Knowtator

The first step for any annotation in this project is to open Protege/Knowtator and then open your assigned project (following the guidelines given to you elsewhere).

Once Protege is opened, click on the "Knowtator" tab under the menu bars to bring up Knowtator. At this point, a clinical note should automatically pop up, but if you're presented with a blank window, find the text source folder associated with this set by clicking the "Select Text Source from List" button (shown in Figure 2) and then selecting the folder with the medical records:

Once a text source is open, you'll be presented with a window which should look similar to the window below in Figure 3:

12.3 Annotating EVENTs in Knowtator

You will likely want to annotate EVENTs before annotating anything else. Because the first step in this schema is to create EVENTs, we'll start by doing that. First, click on "EVENT" in the schema pane, and then click "Fast annotate with EVENT" from the pop-up window. Once you're in fast annotate mode (you'll see a new bar show up beneath the toolbar), any span of text that you select with your cursor will be automatically turned into an EVENT. To exit fast annotate mode, click "quit" in the fast annotation toolbar.

So, select an EVENT in the text, and release the mouse button. You'll notice that the span you selected is now blue, and that the annotation editor pane has now changed to resemble figure 4:
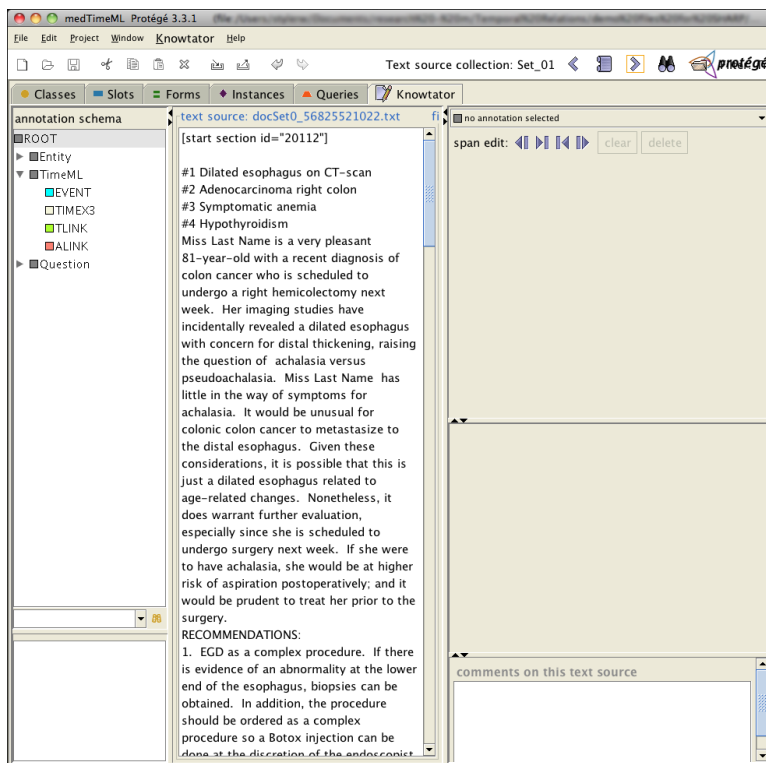
Figure 3: The Knowtator window used for annotation. Note the three panes, from left to right, the schema pane, the text source pane, and the annotation editor pane.
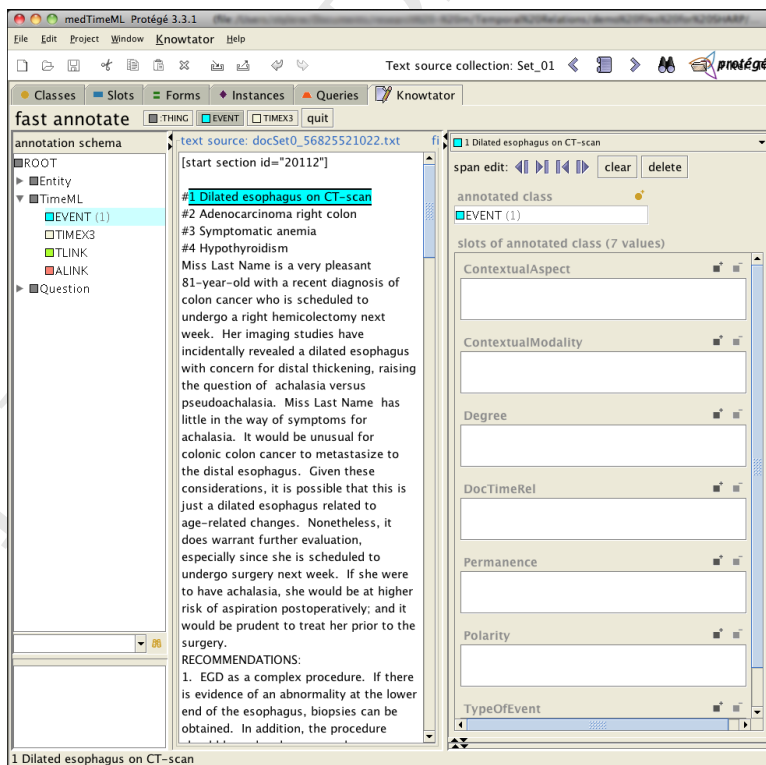

Figure 4: The window once an EVENT annotation is made. Note the changes in the annotation editor pane, and the fast annotate mode toolbar
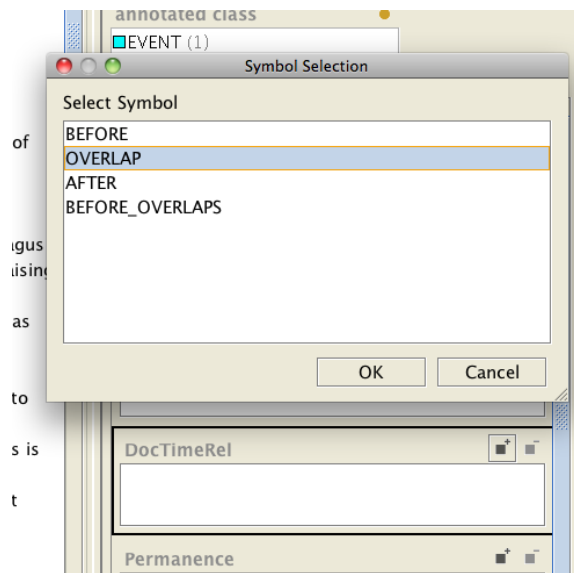
Figure 5: Selecting an attribute

Now, before making another EVENT, you'll need to mark the attributes of the EVENT. To do this, use the annotation editor window (feel free to resize the window or the sub-panes within it) and click on the ■⁺ button for each attribute where needed. This will pop up a secondary window, and you can simply select the attribute you'd like.

So, to mark an EVENT as having a DocTimeRel of OVERLAP, you would click the ■⁺ button next to DocTimeRel in the annotation editor pane, then simply select OVERLAP from the list, as shown in figure 5.

You'll then go through, select all necessary attributes (those where the desired value differs from the default value), and then move on to the next EVENT by selecting a different span of text. Once you've annotated all the EVENTs in the document, move on to TIMEX3s.

Note that EVENTs (as well as all other annotations) can be deleted using the "delete" button at the top of the annotation editor pane.

12.4   Annotating TIMEX3s in Knowtator

Annotating TIMEX3s in Knowtator is actually identical in practice to annotating EVENTs. Use fast annotate mode, and for every TIMEX3 in the document, select the span, set the TYPE attribute, and then move on to the next one. Once you've finished, exit fast annotate mode and move on to making TLINK annotations.

12.5   Annotating TLINKs in Knowtator

Annotating TLINKs is a bit different than EVENTs or TIMEX3s. To create a TLINK, make sure that no text is selected in the text source pane and then click "TLINK" in the schema pane, then click "create TLINK annotation" (Note that fast annotate mode will NOT work here). Once you've created the TLINK, the annotation editor pane will change to look as it does in figure 6, below.

Once you've created the annotation, find the first EVENT which you want to link, and click it in the text source pane. From the menu, select "fill Event slot of selected annotation with ...". Then, find the EVENT or TIMEX3 it's related to, click it, and select "fill RelatedTo slot of selected annotation with ...".
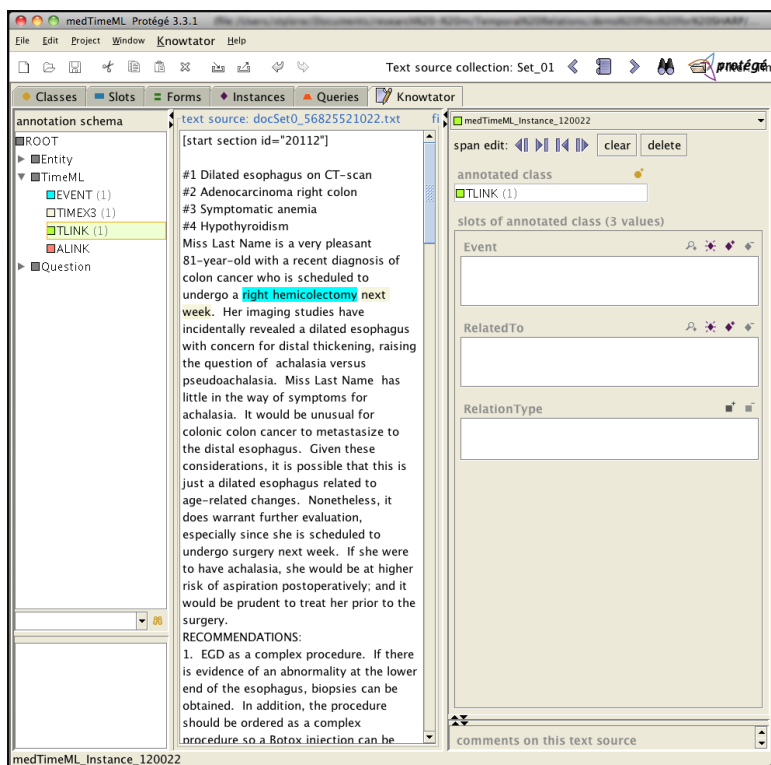
Figure 6: Knowtator during a TLINK annotation. Note that normally, far more EVENTs and TIMEX3s would be marked.

At this point, the Event and RelatedTo slots in the annotation editor pane should be filled. Now, all that remains is to select the TLINK RelationType, in the same way you chose attributes for EVENTs and TIMEX3s.

It's worth noting that due to the way that Knowtator works, once created, a TLINK will not show up at all in the text source pane, so once you select another annotation (or create a new one), the link will not be obvious, but you will still be able to find the annotation by clicking on TLINK in the schema pane and going through the "select:" options.

12.6   Annotating ALINKs in Knowtator

ALINK annotations are made identically to TLINKs in Knowtator. Just link the aspectual EVENT to the regular EVENT and specify the type.

12.7   Annotating DocTime and SectionTime in Knowtator

DocTime and SectionTime are easy. Simply select a span and mark the annotation.

13 Glossary

Default Value

All EVENT, TLINK, ALINK and TIMEX3 annotations have additional characteristics. If no other option is selected for these characteristics, the default values will be assumed and added.

Document Time/DOCTIME

The date of the note being annotated, usually at the top of the note. This is marked with DocTime.

Knowtator

Knowtator is the annotation tool to be used in this project

Span

An amount of text which is associated with an event. Spans do not have to be continuous, and in the sentence "Mr. X has experienced debilitating seizures", the span would be "experienced ... seizures"

Temporal Relations

Temporal relations is the relation between different events, actions, states and fixed timepoints in terms of time. If something happened before, after, during or around something else (or a fixed point in time), then those two things are temporally related, and that relationship must, in this schema, be annotated.

TimeML

This annotation schema is based on a system for marking the temporal relations of different words in a text called TimeML, developed by the TimeML working group. Please see http://timeml.org/site/index.html for more information.

14 References

References

[1] James Pustejovsky and Amber Stubbs, *Increasing Informativeness in Temporal Annotation*, Linguistic Annotation Workshop, 2011, pp. 152–160.

[2] GK Savova, S Bethard, Will Styler, J Martin, and M Palmer, *Towards temporal relation discovery from the clinical narrative*, (2009), 445 (English).

15   Appendix I: Clinical Note Section Labels

- 20100 - Revision History
- 20101 - Referral Source
- 20102 - Chief Complaint/Reason for Visit
- 20103 - History of Present Illness
- 20104 - Current Medications
- 20105 - Allergies
- 20106 - System Review
- 20107 - Past Medical/Surgical History
- 20108 - Social History
- 20109 - Family History
- 20110 - Vital Signs
- 20111 - Physical Examination
- 20112 - Impression and Report and Plan
- 20113 - Diagnosis
- 20114 - Administrative
- 20115 - Special Instructions
- 20116 - Advance Directives
- 20117 - Service Actors
- 20118 - Immunizations
- 20119 - Admission Findings and Test Results
- 20120 - Problem Oriented Hosp. Course
- 20121 - Final Physical Examination
- 20122 - Adverse Reactions
- 20123 - Diet / Nutrition
- 20124 - Discharge Condition
- 20125 - Condition at Discharge
- 20126 - Ongoing Care Orders
- 20127 - Admission Physical Exam
- 20128 - Ongoing Care
- 20129 - Follow Up Agreements
- 20130 - PHF and CVI Dates
- 20133 - Admission Medications
- 20135 - Anticipated Problems and Interventions
- 20136 - Post-op Services

- 20137 - EMTALA Statement
- 20138 - Patient Education
- 20147 - Dismissal Medications