COSI 140B
Prof. Nianwen Xue
3.28.2020
Group member: Zhuoran Huang, Xiaoyu Lu, Yonglin Wang

Annotation Guidelines for Temporal Relations

Table of contents

Table of contents

Part I. Annotation Schema
- 1.1 Definition of an Event
- 1.2 Definition of the four Temporal Relation types

Part II. Useful Diagnostics
- 2.1 Diagnostic for Generics and Non-generics
- 2.2 Possible world analysis
- 2.3 Causality, condition, and corresponding relation

Part III. Annotation Process
- 3.1 Phase I. (More prepping) Cleaning the bad verb tags
- 3.2 Phase II. (the real deal!) Deciding on the TempRel Type

Part IV. Annotation Work Assignment
- 4.1 Phase I work assignment
- 4.2 Phase II work assignment
  - 4.2.1 Work Assignment
  - 4.2.2 Annotation Task Walkthrough

Part I. Annotation Schema

1.1 Definition of an Event

We use a simplified representation of an event: each event is represented by a non-auxiliary verb in the sentence.

E.g. In "as soon as the environment becomes stable again, the habit starts to reassert itself", the events are the environment becoming stable, the habits starting to do so, and the habit reasserting itself, represented by *becomes*, *starts*, and *reassert*.

Examples of auxiliary verbs that should NOT be tagged for links (bold & italicized):

- *They'**ve** been more successful.* (*been* is a main verb in this case, since it is not followed by a progressive or passive main verb)
- *The president **has** never needed a new message to voters as he **does** now.*
- ***Having** reached the top of a major corporation and amassed a fortune, chief executives often wonder: What's next?*
- *He **doesn't** talk about it very convincingly.*

(basically, auxiliary verbs are certain forms of **have**, **be**, and **do** that are not the main verb in a clause, as defined in Syntactic Theory--if you feel unsure about this, feel free to reach out to us)

1.2 Definition of the four Temporal Relation types

For an event pair (A, B), choose, based on the startpoints of event A and event B, one of **the four TempRel types**:

a. BEFORE: A takes place before B takes place (startpoint of A is before startpoint of B), regardless of the relation between their endpoints.
*E.g.The first thing I **ask** is that they **help** writing this column.*

b. AFTER: A takes place after B takes place.
*E.g.The first thing I **dislike** is that they **stop** writing this column.*

c. SIMULTANEOUS: A and B take place at the same time.
*E.g. As soon as the environment **becomes** stable again, the habit **starts** to reassert itself.*

d. VAGUE: A might take place before, after, or at the same time as B takes place.
*E.g. Perhaps Mr. Sanders will **win** a party-run primary or caucus — in Alaska, Hawaii or Wyoming over the coming weeks — as he did in North Dakota last week. Otherwise Mr. Biden is the favorite in every remaining primary state. And without victories, Mr. Sanders will be **deprived** of opportunities to claim the momentum and favorable news coverage to change the trajectory of the race.*

Part II. Useful Diagnostics

2.1 Diagnostic for Generics and Non-generics

First, here are the necessary, together sufficient criteria for a non-generic event:

1. Non-generic events are **perceptible**
2. Non-generic events can be **located in space and time**
3. Non-generic events **have a unique manner of realization**

According to this definition, an event is generic if it fails (i.e. produces a semantically bizarre meaning) one of the Linguistic diagnostics for events:

i. **Event expressions can serve as infinitival complements of perception verbs.**
   E.g. see + VP[bare], watch + VP[bare]
   Pass: I saw Jaco eat a bone. → *eats* in *"Jaco eats a bone"* may be non-generic
   Fail: ?# I saw Jaco be a dog. → *is* in *"Jaco is a dog"* is generic

ii. **Event expressions combine with locative and temporal modifiers.**
   E.g. at Brandeis, tomorrow, yesterday
   Pass: Jaco eats a bone at Brandeis. → *eats* in *"Jaco eats a bone"* may be non-generic
   Fail: ?# Jaco is a dog at Brandeis → *is* in *"Jaco is a dog"* is generic

iii. **Event expressions combine with manner adverbials, comitatives, etc.**
   Manner adverbial: slowly, quickly
   Pass: Jaco eats a bone slowly. → *eats* in *"Jaco eats a bone"* may be non-generic
   Fail: # Jaco is a dog slowly. → *is* in *"Jaco is a dog"* is generic

   From the tests above, we can see that *"Jaco eats a bone"* fails none of the tests and therefore is non-generic. In contrast, *"Jaco is a dog"* fails at least one of the tests above, and therefore is a generic event.

In our annotation task, we only want to compare the start points of two generic events or two non-generic ones. If one word in a pair is generic while the other is non-generic, we should immediately label the pair as VAGUE.

2.2 Possible world analysis

Each event should be treated as having occurred, whether or not the text implied that it had occurred. Negated events and hypotheticals are treated similarly. One assumes the event does occur, and all other events are ordered accordingly. Use context (word meaning or tense) to infer when the event would happen, if possible.

- Hypothetical example (with context for inferring time of event available):
  *E.g. We should have been **talking** about this last week. We **cut** off people's utilities this week...*
  (talking, cut): before; if the talking last week were to have happened, it must have happened before the cutting off of utilities this week.
- Negation example (with tense for inferring time of the event available):
  *E.g. the governor **claims** that the crew was not **anticipating** such drastic measures.*

(claimed, anticipating): after;

1) claims is present tense, while anticipating is part of a past tense construction.
2) According to the tense cue, if the anticipating were to have happened, then it must have happened before the claiming

- Infinitival example:
  *E.g. Investors **expected** to **see** a surge in the stock market after the reinforcement of the new policies.*
  (expected, see): before; if seeing a surge in the stock were to happen, it would happen after the expecting, according to the context.

2.3 Causality, condition, and corresponding relation

a. Causality: If event A causes/enables/prevents B, then we say A happens before B
*E.g. Chancellor Angela Merkel later said she herself was **going** into isolation because her doctor had **tested** positive for the coronavirus.*
(going, test): after; testing positive caused going into isolation.

b. Condition: If A is the necessary condition for B, then we say A happens before B
*E.g. The president wants to **delay** candidate status so that he can **travel** as president making speeches.*
(delay, travel): before; delaying the status is the necessary condition for traveling as so.

Part III. Annotation Process

3.1 Phase I. (More prepping) Cleaning the bad verb tags

Before we start deciding on the temporal relations, we need to first clean the bad labels created by the preprocessing script. For each verb in each paragraph file assigned to you, please do the following:

1. **Is the verb a main verb?**
   If yes, move to step two.
   If not, delete the tag on this verb and move on to the next verb. Then, **record the verb you have deleted** in the correction.txt text file with your name in the project root directory!
2. **If the verb is indeed a main verb, is the category (form) label of the verb correct?**
   If yes, we're done with this verb, you can move on to the next one now.
   If not, please change the category to the correct one before moving on to the next.
   Here's list of all possible verb categories and their explanations:

Verb form key

Category Label	Explanation	Examples
VB	Base/bare form	Be, take
VBD	Past tense	was/were, took
VBG	Gerund or present participle	Being, taking
VBN	Past participle	Been, taken
VBP	Present, non-3rd person form	am/are, take
VBZ	3rd person present	Is, takes

3.2 Phase II. (the real deal!) Deciding on the TempRel Type

This is where you will choose a TempRel type for each of our previously generated, ready-for-choosing links. The following is our general decision process for your reference:

1. **Based on the definitions in 2.1, is it true that one event is generic while the other is not?**
   a. If yes, choose VAGUE.
   b. Else, the events are both generic or both non-generic. Good, go to step 2.
2. **Based on the instructions in 2.2, imagine a possible world where both events have actually happened as described, in order to pin-pointing their temporal position.**
   a. Pay attention to the contextual cues of when the events happened or would happen!
3. **Now, determine the TempRel of the two events:**

a. Using causality and condition (as described in 2.3), **your knowledge of the world** and **your understanding of the English language**, determine the TempRel of the start points of the two events.
b. If still not clear, choose VAGUE.

Below are some sample decisions made based on the process above (verbs are bolded, duplicate verbs are marked, and contextual cues underlined for your reference):

- *E.g. He has **lost** every county **bordering** Wisconsin, even if it **shapes** up as a relatively good state for him.*

  (lost, bordering): vague;

  1) *lost* is a non-generic event while *bordering* is not.

  (bordering, shapes): before;

  1) Both are generic events.
  2) We imagine a possible world where *bordering* and *shaping up* both have happened.
  3) The beginning of *bordering* must have started (long) before *shaping up*.

- E.g. He would have **faced** an even greater deficit if Ohio had **voted_1** on Tuesday as originally **scheduled**, based on results so far. Georgia and Louisiana, originally **scheduled** to **vote** over the next few weeks, might have **voted_2** for Mr. Biden by 50 points or more. And Mr. Biden might have **won** an outright majority of delegates and **wrapped** up the nomination by late April. Now it **is** hard to **say** when he will **clinch** it.

  (faced, scheduled): after;

  3) world knowledge: scheduling the vote should happen before the deficit being faced

  (faced, vote_1): before;

  2) propose a possible world where facing a deficit happened "on Tuesday" and voting happened "over the next few weeks"

  3) world knowledge: the thing happening on Tuesday must happen before the thing over the next few weeks.

  (vote, voted_2): simultaneously;

  3) They are basically the same event. Therefore, they happened simultaneously

  (won, wrapped): simultaneously

  3) world knowledge: winning and wrapping up should happen at the same time

  (faced, say): before

  2) facing would happen on Tuesday; saying would happen now

  3) world knowledge: the Tuesday thing happens before current/now thing

  (say, clinch): before

  3) world knowledge: *saying* in this case is similar to making a prediction. Therefore, the prediction happens before the actual clinching being predicted happens

  (is, say): vague

  1) *is* is generic while say is not.

- E.g. Mr. Loeser told me he **has** clients who **are** business executives considering a run for high public office.

(has, are): after

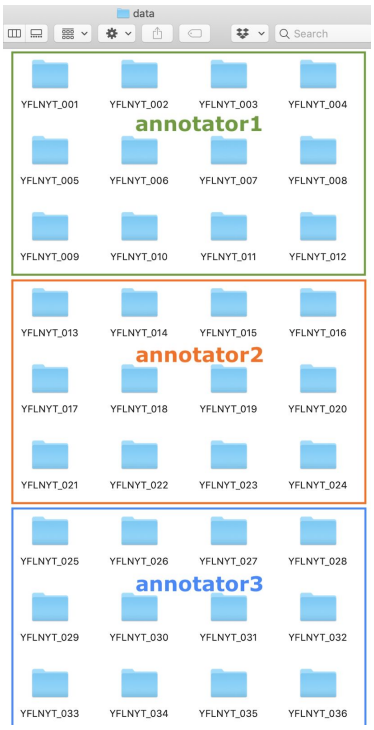
1) this *has* indicates possession and fails the tests; so does *are*.
2) the two events are happening in this world already.
3) condition: according to the context, it sounds like being an executive is the necessary condition for considering being Mr. Loeser's client; therefore, possessing the clients would happen after they are executives.

Part IV. Annotation Work Assignment

GitHub [This link will take you to the project Github](#).

4.1 Phase I work assignment

Since Phase I is only pre-processing and essentially correcting some sparse errors in the data to be annotated, we really hope to get these done in 1 week. But still, let us know if you run into any issues.
Estimated deadline: **23:59, Monday, April 6**



Annotator #	annotator name	annotation work allotted
1	Carmi Rothberg	YFLNYT_001~012
2	Jonne Saleva	YFLNYT_013~024
3	Jeff Zhong	YFLNYT_025~036

For instructions on Phase I, please see 3.1.

4.2 Phase II work assignment

Minimal requirement: (please do these or we probably won't have enough data to do anything) Finish these 9 articles: 1, 2, 3, 13, 14, 15, 25, 26, 27

Tentative deadline: <u>Friday, May 1st</u>. Please upload everything you have annotated by this time so that we can begin with some initial analysis.

4.2.1 Work Assignment

Annotator #	annotator name	annotation work allotted	Files saved in
1	Carmi Rothberg	13, 14, 15, 26, 27, 28 (in this order)	/Carmi
2	Jonne Saleva	26, 27, 28, 1, 2, 3 (in this order)	/Jonne
3	Jeff Zhong	1, 2, 3, 13, 14, 15 (in this order)	/Jeff

Please annotate the articles in the same order as given in the table! (e.g. Annotator 2 will start from article 25/folder YFLNYT_25, then YFLNYT_26, YFLNYT_27, YFLNYT_1… etc.)

4.2.2 Annotation Task Walkthrough

Here's a quick walkthrough for everything you'll need to start on Phase II:

1. Look at Section 1.2 to understand how we compare two events and the 4 temporal relation categories
2. Go through each section in Section 2, make sure you understand how each example works
3. Go through the 3-step decision process in Section 3.2. Quickly go through the example to make sure you agree with the annotation decision.
4. Now, open your assigned files and annotate~ (feel free to reach out to us about ***<u>any</u>*** questions you have regarding the annotation process, as this is also a crucial part for the project discussion!)
5. **<u>On Friday, May 1st</u>,** please upload everything you have annotated (and let us know which file you leave off from in the comment) so that we can begin with some initial analysis.