Guideline for Essay Revision Annotation

Jan Wira Gotama Putra
School of Computing
Tokyo Institute of Technology

October 15, 2018

Thank you for participating in our discourse annotation study. You will be given a set of argumentative essays on two different topics. The topic is given to the writers (students) in the form of a "prompt"; i.e., a sentence giving a statement to be discussed, for instance "*smoking should be banned at all restaurants in the country.*" Students are told to produce a stand-alone text that can be read without knowing the prompt. We would like you to perform the following three tasks.

1. Annotating relations or dropping sentences
   For each sentence in the text, determine another sentence that is most closely related to it and indicate what their relationship is. Alternatively, remove the sentence if it does not contribute to the entire argument.

2. Reordering sentences
   If necessary, reorder the sentences to improve the overall logical flow. The reordered text should be a more well-structured argument than the original one.

3. Repairing text
   If it is necessary for understanding the reordered text and only then, you may change the way how people and things are described or conjoined in the text.

1   Annotating Relations or Dropping Sentences

Sentences in a well-formed discourse are related to each other. Your task is the following. Read through the whole essay at least once to understand its content. Then, find the main claim (as explained in Section 1.1). Then, for each sentence (source) other than the main claim, determine another sentence (target) that is strongly related to it. Then, assign one of four relation labels to their relation. Table. 1 shows the relation labels you can choose from. They are explained in more detail in Section 1.2-1.5.

Label	Name	Description
sup	support	The source sentence asserts or justifies reasons and ideas for supporting the target sentence; it contains evidence or examples for the target sentence.
det	detail	The source sentence further explains, describes, elaborates or provides background for the concept(s) mentioned in the target sentence.
att	attack	The source sentence considers counter-arguments that argue for the opposite opinion.
=	restatement	Two sentences are connected with "=" label if they are restatements of each other.

Table 1. Relation labels

1.1   Finding Main Claim

The first step of the relation annotation is to find the statement that expresses the author's opinion at the highest level of abstraction, i.e., the main claim. It expresses the author's overall stance toward

a discussion topic. After determining the main claim, you can proceed to identifying all remaining relations existing in the text. Consider the following example.

(Prompt) Smoking should be banned at all the restaurants in the country.

(1) Supported by the utilitarian perspective, I believe smoking should be completely banned at all restaurants.
(2) This is because there is overall harm if smoking is not prohibited.

In this example, sentence (1) is the author's opinion at the highest level of abstraction denoting the author's stance in response to the discussion topic. Sentence (1) is the main claim, and other sentences will be connected according to their stance towards it. Some expressions like "*In my opinion*", "*I strongly believe*" or "*I feel that*" indicate the sentence in question might be the main claim.

1.2 Support

A statement in an argumentative essay can be supported by several reasons/ideas. They assert why readers of the essay should believe the statement, in general, by providing new argumentative material. Consider the following example.

(1) From my point of view, banning smoking in all restaurants is necessary.
(2) First, I think it is essential to protect the citizens' health.
(3) Second, banning smoking also allows all diners to eat in peace.

Sentence (1) is a statement about banning smoking in restaurants. It is further reinforced by sentence (2) and (3), which present different ideas from each other. Both are in favour of sentence (1), i.e., making the opinion in sentence (1) becomes more convincing. Therefore, both sentence (2) and (3) point at sentence (1) via the `sup` label.

The presence of enumeration, exemplification or reasoning expressions might indicate that the sentence in question is in a support relationship (either as a source or target). Example of the expressions we meant are "*first*", "*second*", "*for example*", "*it is because*", "*for this reason*". Please note, however, the presence of those expressions does not always mean the sentence in question is in a support relationship! You should always judge the proper relationship based on the context.

1.3 Detail

The `det` relation label is applied to those sentences which present additional details (further explanations, descriptions or elaborations) about a particular sentence in question, but without providing new argumentative material. Consider the following example.

(1) It is difficult to balance studying and working.
(2) Especially if the students cannot manage their time well, because it only breaks the focus of their studying.

By reading sentence (1) on its own, we can infer that "*difficult to balance*" is talking about the time management between studying and working. When we read sentence (2) afterwards, we understand that sentence (2) elaborates the information which we inferred from sentence (1). Sentence (2) explains what the phrase "*difficult to balance*" means. Because no new argumentative material is introduced, and because sentence (2) provides additional detail (in this case, an explanation) to sentence (1), you should annotate sentence (2) pointing at sentence (1) via the `det` label.

There is another usage of the `det` label. If you find a statement that introduces the topic of the discussion in a neutral way by providing general background but without any argumentative material, please mark it as being in the `det` relation to the target sentence. Consider the following example.

(1) Today, more and more college students are taking part-time jobs.
(2) I think having a part-time job is a good thing for them.

Sentence (1) is an introduction to a discussion topic in a neutral way. It enables the readers to comprehend sentence (2) by giving some contextual information. In this case, the correct annota-

tion is to relate sentence (1) to sentence (2) via the `det` label; i.e., sentence (1) is now pointing at sentence (2).

1.4 Attack

The `att` relation label denotes sentences arguing for the opposite opinion. Consider the following example.

---

(1) From my point of view, banning smoking in all restaurants is necessary.
. . .
(2) On the other hand, I admit that some restaurants are popular because men are allowed to smoke.

Sentence (1) is an example of the main claim in a smoking-themed essay. It states that smoking should be banned. However, sentence (2) argues against banning smoking because smoking makes restaurants popular. In this example, the correct annotation is sentence (2) pointing at sentence (1) via the `att` label.

These are some useful expressions indicating that the sentence in question might be in an attack relationship: "*on the other hand*", "*but*", "*however*", "*in contrast*", "*contrary to*" and "*in another way*". The presence of those expressions, however, does not always indicate the presence of an attack relationship! Once again, you judge the proper relationship based on the context.

1.5 Restatement

Two sentences are connected with "=" label if they are restatements of each other. Consider the following example.

---

(1) I agree that college students should have a part-time job.
. . .
(4) Second, having a part-time job is a valuable way to pick up communication skills that will be needed in the workforce.
. . .
(8) Therefore, it is better for college students to have a part-time job to exercise communication skills.

Sentence (1) is an example of the author's opinion at the highest level of abstraction which is in favour of a part-time job. Sentence (4) states one of the reasons why students need a part-time job, i.e., to acquire communication skills. After further elaboration, the entire meaning of sentence (4) is restated as sentence (8). Notice that sentence (4) does not say anything about agreement or disagreement towards the question of part-time jobs while sentence (8) explicitly states it. However, we understand from reading sentence (1) that sentence (4) is implicitly in favour of a part-time job. This means that sentence (4) is basically restated as sentence (8). You should therefore connect sentence (4) and (8) with the "=" label. Such restatements often happen in a situation such as the one above, where large parts of an argument are summarised for the second time. The two restatement sentences are treated as an equivalence class with respect to all outgoing and incoming relations they participate in.

1.6 Relation Selection

A sentence can relate to many sentences at once. For example, a sentence might elaborate on two ideas at once. But in this task, you have to choose which sentence it relates to the most. There are several factors to consider.

1. Closeness in position.
   A sentence tends to more relate to closer sentences.

2. Directness of relation.
   A direct relation is preferred over an indirect relation. For example, consider three sentences (1), (2) and (3). Sentence (3) attacks sentence (2), and sentence (2) attacks sentence (1). In this case, sentence (3) also indirectly supports sentence (1) by attacking sentence (2). However,

since we prefer a direct relation, you should annotate the relation between sentence (3) and (2), not the one between sentence (3) and (1).

3. Preferential ordering.
   You should choose the `sup` relation over `det` when a sentence both explains another sentence further and contains a new idea. This is because the new idea is more informative.

Sometimes, a relation can hold between the parts of a long sentence. But relationships inside a sentence cannot be expressed in our system, so do not worry about them. In this case, please annotate only the function of the entire sentence as a whole.

1.7 Dropping Criteria

You may find it hard to connect some sentences. In this case, you can also consider dropping them. We list the criteria to judge whether a sentence should be dropped.

1. Metainformation. You should drop sentences which only make statements about other sentences, without adding any real material. For example, "*I have two reasons for supporting this opinion.*" Unlike details, this kind of sentences contributes nothing substantial toward the argument.

2. Redundant material. For example, a student may state twice that smoking is dangerous as it causes lung cancer. Please note the difference to restatements, which contain the same information, but typically at a higher level of argumentation (claims, not facts) and with a real function in the overall discourse. Unlike redundant material, dropping a restatement might affect the discourse and flow of argumentation. The material considered as redundant here typically consists of mere facts, rather than real argumentative material such as claims or conclusions.

3. Truly disconnected sentences with no proper connection to the argument. Sometimes a sentence is logically isolated; i.e., it does not really relate to any other sentence. In this case, you should drop it.

The remaining sentences after dropping should be connected to each other.

2 Reordering Sentences

A good text usually places semantically related sentences close to each other, forming semantically consistent segments. If the sentences in the current text are not already in the best order they could be, please arrange them into a logically well-structured argument; i.e., the best arrangement of sentences that you can think of. Remember that you can also drop sentences. Please make sure that you keep the original meaning of the text intact while doing so. Consider the following example.

> (1) If people smoke in the restaurant, other people may think the food isn't delicious.
> (2) At restaurants, people enjoy eating and talking.
> (3) They might have a sore throat and be unable to enjoy talking.

This text talks about the effect of smoking in restaurants, then talks about how dining should be an enjoyable experience while moving back again to the effect of smoking in restaurants. A better order is to place sentence (2) before sentence (1) as below.

> (2) At restaurants, people enjoy eating and talking.
> (1) If people smoke in the restaurant, other people may think the food isn't delicious.
> (3) They might have a sore throat and be unable to enjoy talking.

3 Repairing Text

After reordering the sentences in the previous step, you might have made the text itself harder to understand in certain superficial ways. In order to revert these negative changes, you are allowed to perform the following operations.

1. Change the text material used to connect two sentences or sentence parts.
   Examples of what we mean are "*however*", "*therefore*", "*but*" or "*eventhough*".

2. Change the text material used to identify people or things.
   Examples of what we mean are "*she*", "*the woman*", "*Maria*" or "*Sister of Kim*".

Please only make minimal repairs necessary for keeping the meaning the same. You should edit, i.e., add, delete, substitute, parts of the text (cf. material 1 and 2 above) only if it is needed for understanding the reordered text correctly.

Example 1

> (1) I think it is okay when poor students have part-time jobs.
> (2) Generally speaking, there are challenges in part-time jobs.
> (3) For instance, my girlfriend cannot focus on her studies.
> (4) I don't think she needs part-time jobs as she is not in a dire state for money.

After reordering the sentences in a more natural way, you might have the following text.

> (2) Generally speaking, there are challenges in part-time jobs.
> (3) For instance, my girlfriend cannot focus on her studies.
> (4) I don't think she needs part-time jobs as she is not in a dire state for money.
> (1) However, I think it is okay when poor students have part-time jobs.

After moving sentence (1) at the end of the text, it becomes more coherent. In this case, it might be useful to use the expression "*however*" at the beginning of sentence (1) as well.

Example 2

> (1) I don't like when my girlfriend is smoking.
> (2) She doesn't look cute while doing so.
> (3) But I think my grandmother has lived long enough.
> (4) It is okay if she smokes.

When you reorder the text, you might have the following text.

> (1) I don't like when my girlfriend is smoking.
> (2) She doesn't look cute while doing so.
> (4) It is okay if she smokes.
> (3) But I think my grandmother has lived long enough.

In this example, sentences (3) and (4) are swapped in position to make a better text; the sentences are now arranged in the form of opinions followed by reasons. But after reordering, the expression "*she*" in sentence (4) wrongly refers to "*my girlfriend*" instead of "*my grandmother*". To preserve the meaning of the statement, it is therefore necessary to replace "*she*" with "*my grandmother*" in sentence (4), as in the following text.

> (1) I don't like when my girlfriend is smoking.
> (2) She doesn't look cute while doing so.
> (4) It is okay if my grandmother smokes.
> (3) But I think my grandmother has lived long enough.

Note that in sentence (3), the repetition of "*my grandmother*" now sounds a bit unnatural while the meaning of the text is not affected. As we ask you to make only minimal changes, please leave "*my grandmother*" in sentence (3) as it is. However, the way of connecting sentences is unnatural in a different way too. To make the structure of "opinions followed by reasons" apparent, we can modify the text as follows.

(1) I don't like when my girlfriend is smoking.
(2) She doesn't look cute while doing so.
(4) But, It is okay if my grandmother smokes.
(3) ~~But~~ I think my grandmother has lived long enough.

By introducing "*but*" at the beginning of sentence (4) and deleting "*but*" from sentence (3), the text now expresses the contrast relationship better.

Special Case: Repairing Main Claim

You may also have to fix the main claim if the author makes the error of assuming that the prompt is read alongside the text. Remember that according to our treatment of the prompt described on page 1, the prompt is not considered as part of the text. For example, he/she may write the following main claims.

1. I think *so.*

2. I *agree with the prompt.*

3. *But,* I do not think $\cdots$

4. $\cdots$ is bad *indeed.*

The example sentences above shows the case when the main claim appears in the beginning of the text. As the writers (students) are supposed to produce stand-alone texts, we should assume that readers do not read the prompt. You should repair the examples above by including some information from the prompt and/or editing phrases indicating discourse connection to the prompt. Some possible repairs for sentences above are as follows.

1. I think *smoking should be banned at all restaurants in the country.*

2. I *strongly believe smoking should be banned.*

3. I do not think $\cdots$

4. $\cdots$ is bad, *I think.*

Formatting

Editing should be done by placing the edited part inside a bracket "[*before* | *after*]". The "*before*" part denotes the expression before edit while the "*after*" part denotes the expression after edit. We will now give a formatting example of each operation.

1. **Addition**. Suppose you want to add an expression "*therefore,*" before the phrase "*the old man*". You rewrite this phrase as "[| *therefore,*] *the old man*", leaving the "*before*" part as blank (space).

2. **Deletion**. Suppose you want to delete the word "*old*" from the phrase "*the old man*". You rewrite this phrase as "*the* [*old* |] *man*", leaving the "*after*" part as blank (space).

3. **Substitution**. Suppose you want to substitute the word "*instead*" with "*but*" in the phrase "*I don't have a pen. Instead, I have a pencil.*" You rewrite this phrase as "*I don't have a pen.* [*Instead* | *But*]*, I have a pencil.*" You put the original phrase in the "*before*" part and the new phrase in the "*after*" part.

4 Annotation Procedure

Before going into the actual annotation procedure, it is better for you to understand the general characteristics of an argumentative essay. Roughly, an argumentative essay can be divided into three main parts: `introduction`, `body` and `conclusion`. Figure. 1 shows an illustration.

   `Introduction` typically presents a general background about the discussion topic. It also contains the main claim that begins the argumentation. Since an argumentative essay aims to persuade the readers to adhere to the main claim, a deeper level of argumentation usually follows in the `body`. The `body` contains one or several ideas that support or attack the main claim. For example, arguing why students should (or not) have a part-time job from the economics and education viewpoint. The essay's author may also describe a viewpoint on a deeper level of argumentation. For example, arguing about economics from the viewpoint of practising financial management and lessening family's burden. Finally, the `conclusion` part sums up the entire argument, most often, by restating the main claim as has been explained in Section 1.5. Please note that the `conclusion` part does not strictly consist of only one sentence. It might be composed of several sentences.
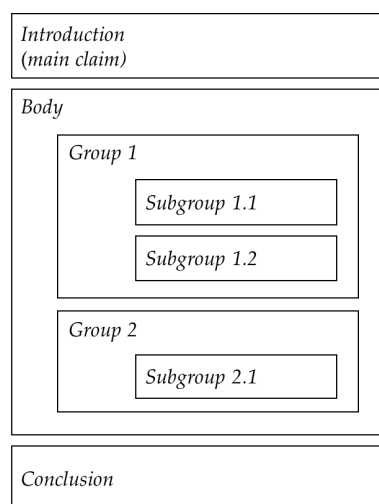

Figure 1. General structure of an argumentative essay

The followings are **the sequential steps you should perform** when annotating an essay.

1. Read through the whole text at least once to understand its content.

2. Find the statement that expresses the author's opinion at the highest level of abstraction, i.e., the main claim.

3. Iteratively, determine the parts or *groups* exist in the text, i.e., `introduction`, `body` and `conclusion`. A part (especially `body`) might be recursively divided into several subgroups denoting deeper level of argumentation. In a logical representation, relations in argumentative texts form hierarchical structures.

4. Determine relations existing in the text and drop irrelevant sentences. We recommend you to determine the relations inside small groups first.

5. If necessary, please reorder sentences in such a way that a logically better-structured text results. From now on, ignore the dropped sentences (if present).

6. As has been explained in Section 3, reordering may cause changes in how people and things are described or connected. Edit the expressions describing people or things strictly if it is needed.

7. Read through the whole text, again, at least once to assess whether the current annotation is already the most proper annotation you can think. If it is not, repeat the process from Step 3.

After the annotation process, the resulting relations in the text should form a hierarchical structure in which the main claim (conclusion, in the absence of the main claim) is on top of the hierarchy. The main claim is then supported or attacked at a deeper level of argumentation. The hierarchical structure is arguably the most natural form of relationships between arguments. To illustrate the whole annotation process, read the text below and follow the step-by-step illustration of its annotation with full attention. The text below is used through the rest of this section.

`(Prompt) Smoking should be banned at all restaurants in the country.`

(1) I agree with the previous statement.
(2) If somebody smokes in the restaurant, other people may not be able to enjoy the experience.
(3) At restaurants, customers enjoy eating and talking.
(4) However, if we ban smoking in restaurants, then those restaurants might lose some customers.
(5) Some restaurants are indeed popular, especially among old men, because they allow people to smoke.
(6) But, I firmly support banning smoking in restaurants because we need to prioritise health.
(7) In conclusion, I encourage banning smoking in all restaurants.

Step 1

Read through the whole text at least once to understand its content.

Step 2

The main claim is sentence (1).

Step 3

The `body` part can be divided into three groups. The first group, consists of sentence (2)–(3), argues from the "*enjoyment of eating and talking*" viewpoint. The second group, composed of sentence (4)–(5), approaches from the "*smoking and the number of customers*" viewpoint. Lastly, sentence (6) argues from the "*health*" viewpoint forming the third group. The grouping is illustrated in Figure. 2.
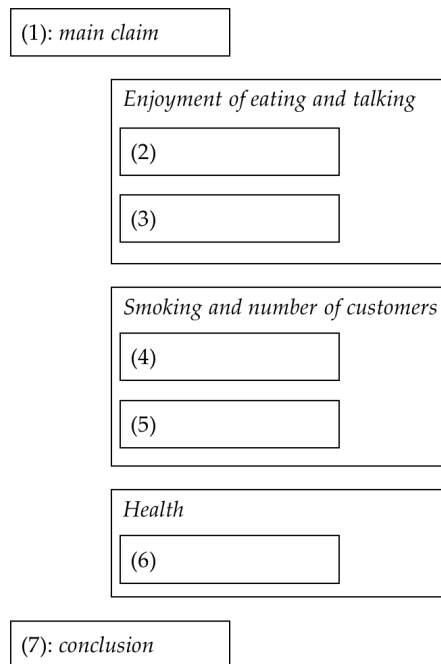

Figure 2. Illustration of recognising *groups* in text

Step 4

We consider relations existing inside smaller groups first. In the "*enjoyment of eating and talking*" group, there are two sentences. Sentence (2) acts as a *root* (representative) of the group since it is the main statement of the group. Sentence (3) provides additional information for the readers to comprehend the word "*experience*" in sentence (2). Therefore, sentence (3) points at sentence (2) via the `det` label. The group representative, sentence (2), supports sentence (1) by arguing for it.

In "*smoking and the number of customers*" group, sentence (5) supports sentence (4) by presenting an opinion to increase readers' belief on it. Therefore, sentence (5) points at sentence (4) via the `sup` label. Sentence (4), the group representative, points at sentence (1) via the `att` label.

Sentence (6) presents an opposing opinion of sentence (4) by saying we should prioritise health. In this sense, sentence (6) supports sentence (1) by attacking sentence (4). However, as we prefer a more direct relation, you should annotate sentence (6) pointing at sentence (4) (cf. Section 1.6).

Finally, sentence (7) sums up the whole argument by basically restating the author's main claim. Even though sentence (7) is not the same as sentence (1), we understand that both expressions are equal. Therefore, we annotate sentence (7) pointing at sentence (1) via the "`=`" label. In this text, all sentences participate in the discourse, and thus no sentence should be dropped. The relations we have established so far are illustrated in Figure. 3. As you can see, the relations form a hierarchical structure in which the main claim is placed on top.
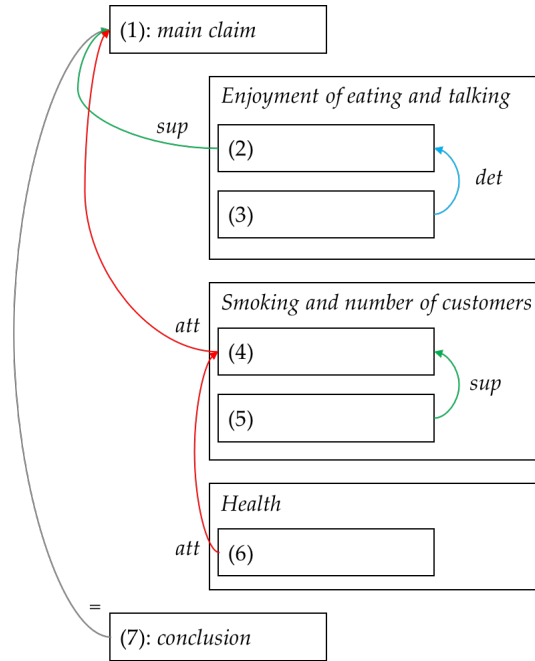

Figure 3. Illustration of annotating relations

Step 5

We can improve the arrangement of sentences by swapping sentence (2) and (3). It is because sentence (2) contains the expression "*the experience*". It implies some background information must have been said before. A more natural ordering is to place sentence (3) before sentence (2) to enable readers to comprehend the expression "*the experience*" easier.

> (1) I agree with the previous statement.
> (3) At restaurants, customers enjoy eating and talking.
> (2) If somebody smokes in the restaurant, other people may not be able to enjoy the experience.
> (4) However, if we ban smoking in restaurants, then those restaurants might lose some customers.
> (5) Some restaurants are indeed popular, especially among old men, because they allow people to smoke.
> (6) But, I firmly support banning smoking in restaurants because we need to prioritise health.
> (7) In conclusion, I encourage banning smoking in all restaurants.

Step 6

The author of the example text has made an error assuming that the prompt is read alongside the text (cf. Section 3). It is indicated by the expression "*with the previous statement*" in sentence (1). Furthermore, it is necessary to improve the transition from sentence (4) and (5). For example, we can append "*It is because*" at the beginning of sentence (5). The result of this final step is given as follows.

---

(1) I agree [with the previous statement | that smoking should be banned at all restaurants in the country].
(3) At restaurants, customers enjoy eating and talking.
(2) If somebody smokes in the restaurant, other people may not be able to enjoy the experience.
(4) However, if we ban smoking in restaurants, then those restaurants might lose some customers.
(5) [| It is because] some restaurants are indeed popular, especially among old men, because they allow people to smoke.
(6) But, I firmly support banning smoking in restaurants because we need to prioritise health.
(7) In conclusion, I encourage banning smoking in all restaurants.

Step 7

Read through the whole text, again, at least once to assess whether the current annotation is already the most proper annotation you can think. Relatively, the annotation example is already the most proper one.

5   General Comment

We appreciate your work and patience. After you annotate the essays, we would like to hear about your experience with and observations about the texts, the tool and the task.