TIARA - a Lightweight Annotation Tool for Hierarchical-structured Discourse Relation

Jan Wira Gotama Putra
School of Computing
Tokyo Institute of Technology

October 1, 2019

This document describes the functions and actions you could perform in `TIARA`[1]. The tool was originally developed as an argument-relation annotation tool (Putra et al., 2019), which then extended for a more general-purpose use. The following are key points of the tool. It would also be appreciated if you send us a courtesy email, so we could survey what task is the tool used for.

1. **Free**. The tool is free and licensed under the MIT license[2] becase we want to encourage studies on discourse analysis. However, please kindly cite our paper (whenever possible) when using this tool for research purposes.

2. **Customizable**. This tool is customizable by changing the configuration file to suit your annotation scheme (Section 12).

3. **Visualization**. The tool visualizes annotated relations. We employ `JsPlumb` (community edition) and `Treant-js`.

4. **Lightweight**. The tool is (relatively) lightweight. By means, it harms neither your CPU nor memory (RAM).

5. **Works across platforms**. We wanted to provide a tool with easy installation so it can be used easily, even for a layman[3]. This tool is developed using `javascript`, and you do not need to install anything (except having a web browser). We recommend using Google Chrome.

6. **Refresh is necessary**. After you have finished working on a file, please refresh your web browser before working on another file.

Despite its appealing points, we need to admit that the tool has its drawbacks. Particularly, it is designed for the annotation of short-texts (approx. 300 words, $\leq$ 20 sentences).

[1] Originally, stands for **Ti**tech **A**rgument **A**nnotation tool
[2] `https://opensource.org/licenses/MIT`
[3] Of course, with some compromises

1 Directory

You are given a zip file containing our annotation tool and sample text(s) pre-formatted in“`.txt`” format. The extracted zip file contains files and folders as shown in Figure. 1. Several important files and folders are highlighted in Figure. 1.

- “`js`” folder contains “`annotation-globalsetting.js`” which you can edit to define the relations you can use in the tool. This configuration defines your annotation scheme.
- “`manual`” folder contains this document, i.e., the description of the tool’s functionalities.
- “`sample_annotated`” folder contains the sample(s) of annotated text(s) which you can try the tool with.
- “`sample_original`” folder contains the sample(s) of un-annotated text(s) which you can try the tool with.
- “`index.html`” is the main program of the tool itself. Open “`index.html`” in your web browser to use `TIARA`.

We recommend you to open “`index.html`” using **Google Chrome** (ver 74 or higher) web browser. The tool was not tested using other web browsers. You can use other web browsers, but we do not guarantee that the tool will work perfectly as expected. Figure. 2 shows the web interface when you open “`index.html`”.

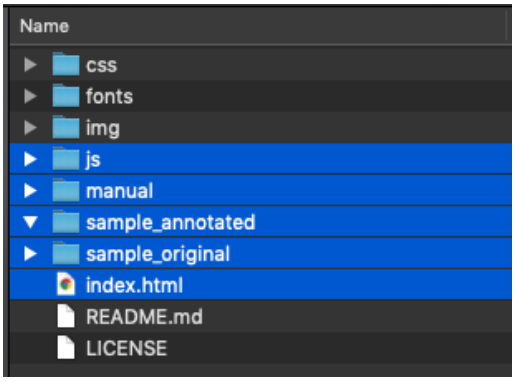


Figure 1. Directory

Figure 2. Annotation interface after you open "`index.html`" on web browser

2 Text Formatting

You should format the texts you would like to annotate in "`.txt`". Each line contains a discourse unit, e.g., sentence or clause. Figure. 3 shows an example.

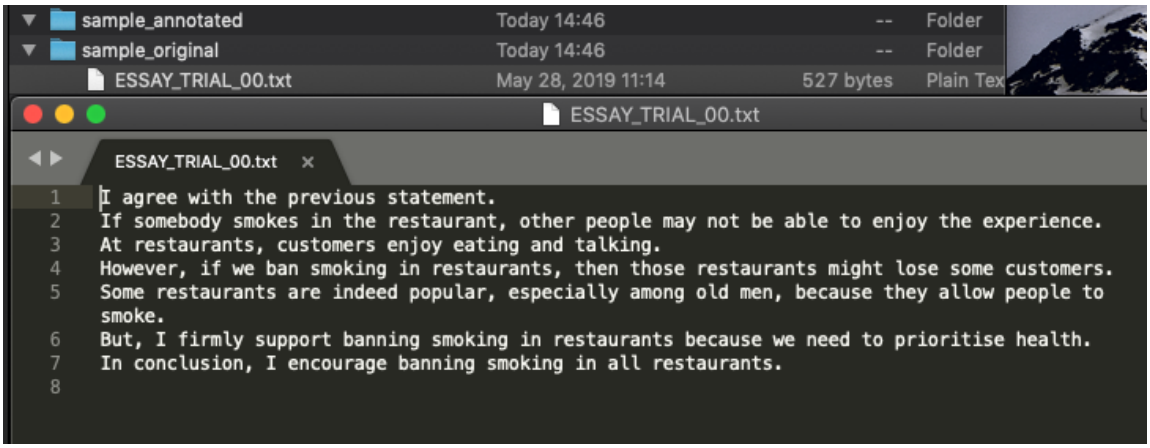


Figure 3. Un-annotated text example

3 Loading File

To load a text into the interface, click the "`Load`" button on top, and browse a text you need to annotate (Figure. 4). Figure. 5 shows an example of a loaded text. The tool presents you the text you want to annotate (in boxes) and the available relations as configured. You can change the relation labels that would be used in the tool and their colors (explained further in Section 12). You can also load an annotated file ("`.html`" file, described in Section 9).

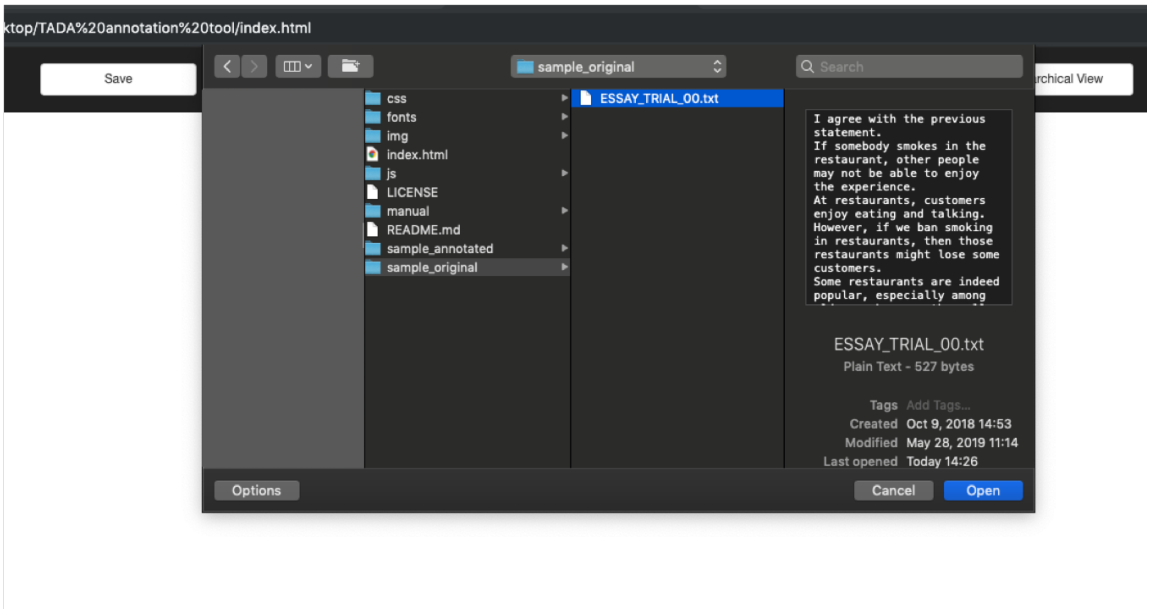


Figure 4. Select the text you need to annotate

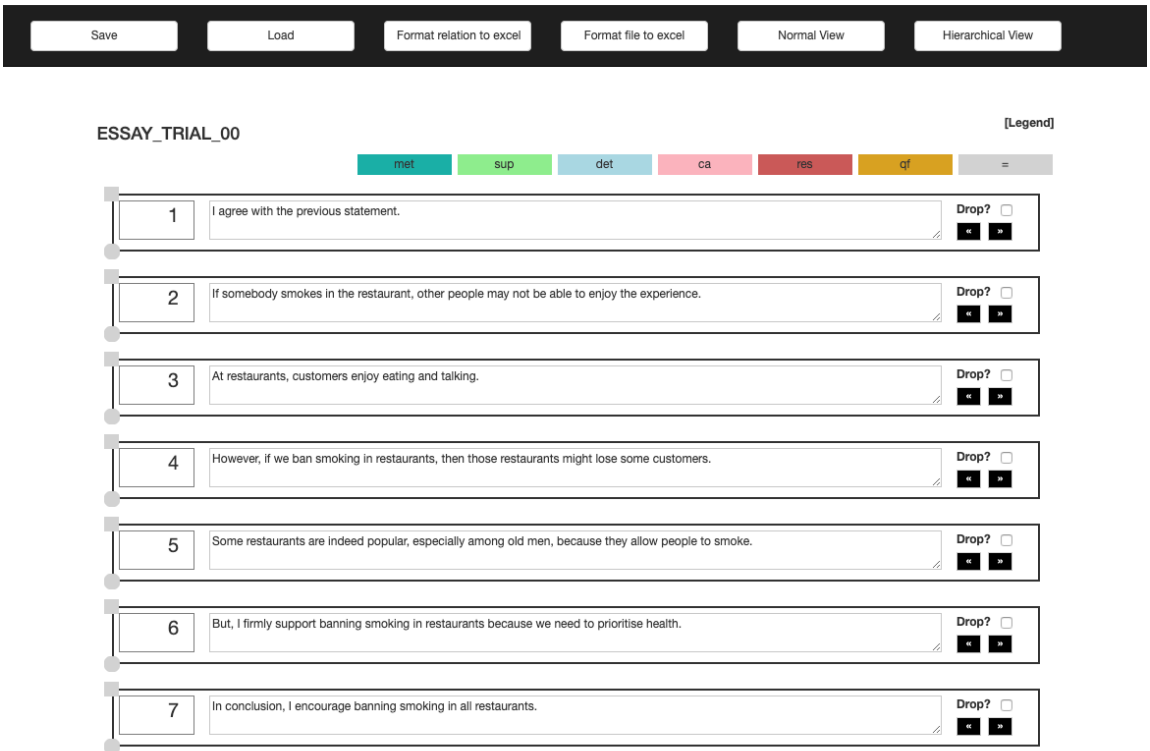


Figure 5. Loaded text example

4 Establishing a Relation

You establish a connection between two units by dragging an arrow from the source to the target unit. An arrow is established by dragging your mouse from the rectangular endpoint of the source unit to the circular endpoint of the target unit. Figure. 6 shows an example. You are also shown a dialog box to choose the label of the link or to drop the link you have just established. Figure. 7 shows the example of an established link.

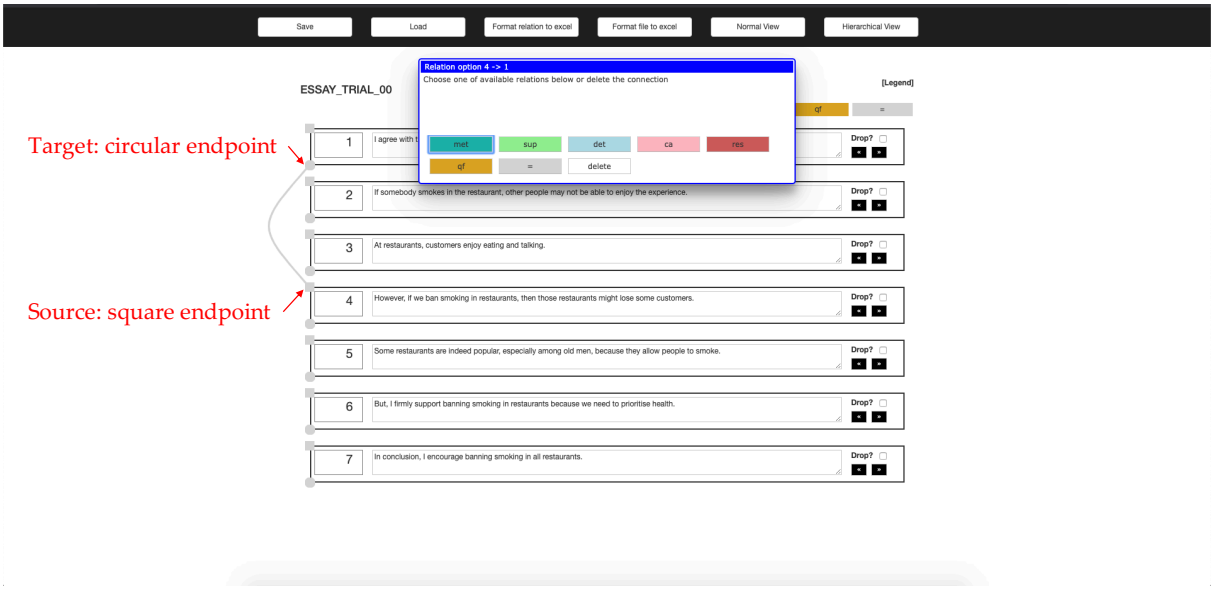


Figure 6. Drag an arrow from the source unit endpoint (square) to the target unit endpoint (circle). A dialog box appears when you try to establish a link

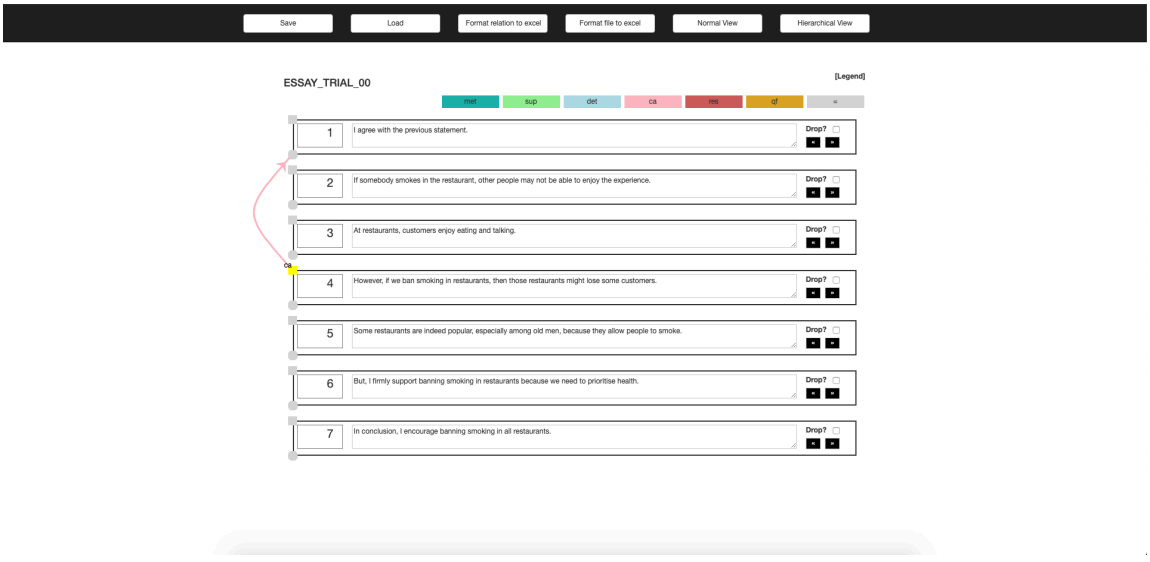


Figure 7. Example of an established relation

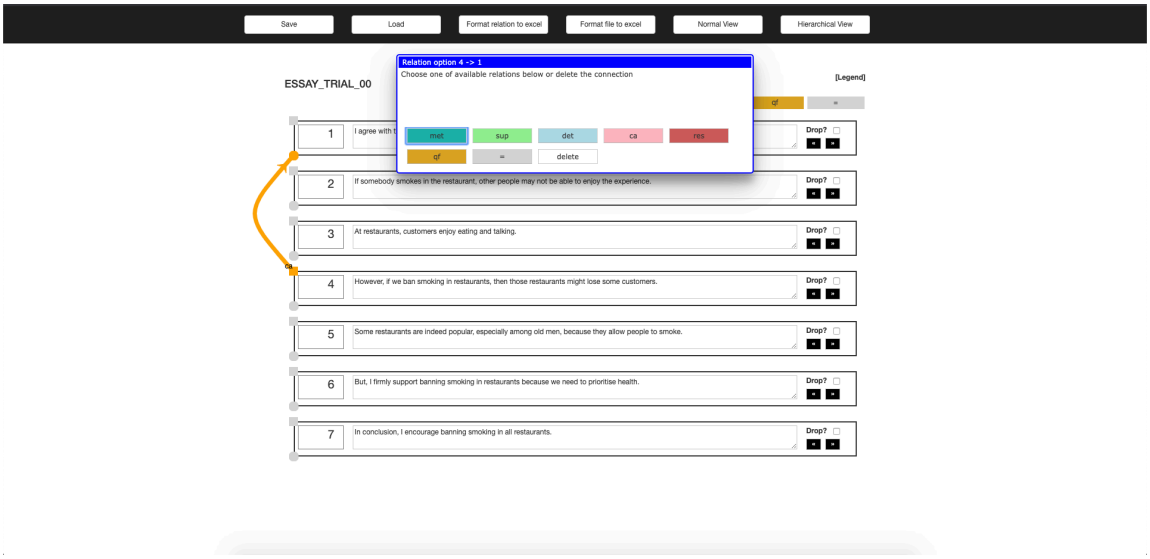


Figure 8. To modify an existing relation, click the link indicating the relation. The link is highlighted, and a dialog box appears

To modify an existing relation, e.g., changing the relation label or removing the connection, click the link indicating the relation. The corresponding arrow is highlighted, and a dialog box appears. Figure. 8 shows an example. A circular connection is not allowed and the tool checks for it in real-time. Whenever you try to establish a circular connection, the tool outputs an error message (Figure. 9) and the new relation will be automatically deleted, i.e., Figure. 9 will be reverted back to Figure. 7. This means, **your final annotation should form a hierarchical structure**.

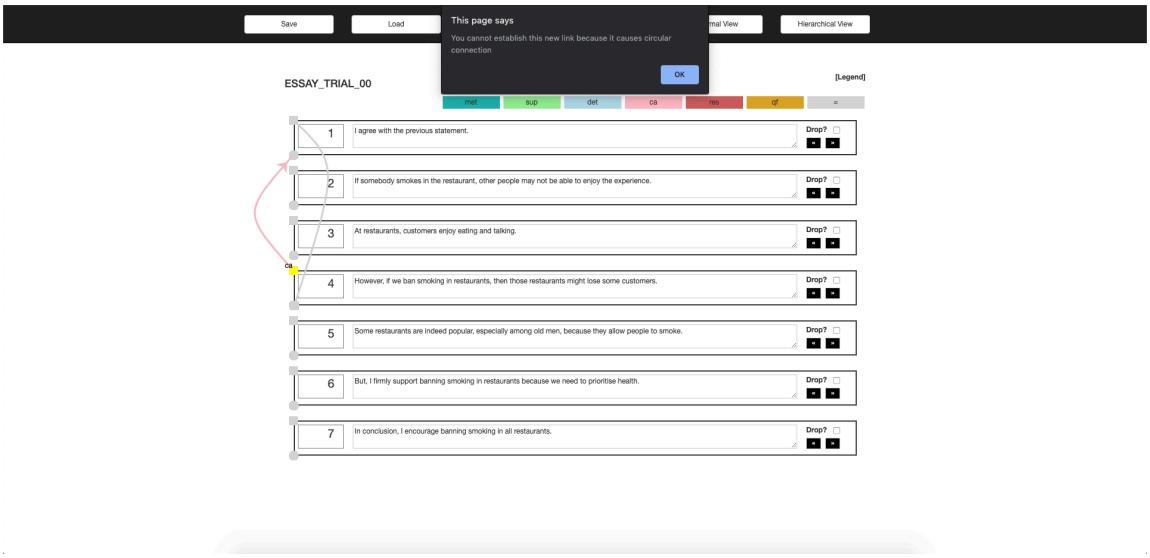


Figure 9. An error message indicating a circular connection is prohibited

5 Dropping a Unit

Sometimes, you might feel there are units (sentences, clauses) that cannot conect to the rest of the discourse. If your annotation scheme disregards those units, you can drop (delete) those textual units. Please note that this functionality is optional (you can turn it on and off, explained in Section 12).

To drop a unit, check the "`Drop?`" checkbox on the right-hand side of the unit as shown in Figure. 10. All associated relations to that unit (both incoming and outgoing) are all deleted. For example, if you click the "`Drop?`" checkbox of unit (4) in Figure. 7, its connection to unit (1) will be deleted. When you drop a unit, the box's background is changed to black. You cannot establish a link from or to a dropped unit as shown in Figure. 11.

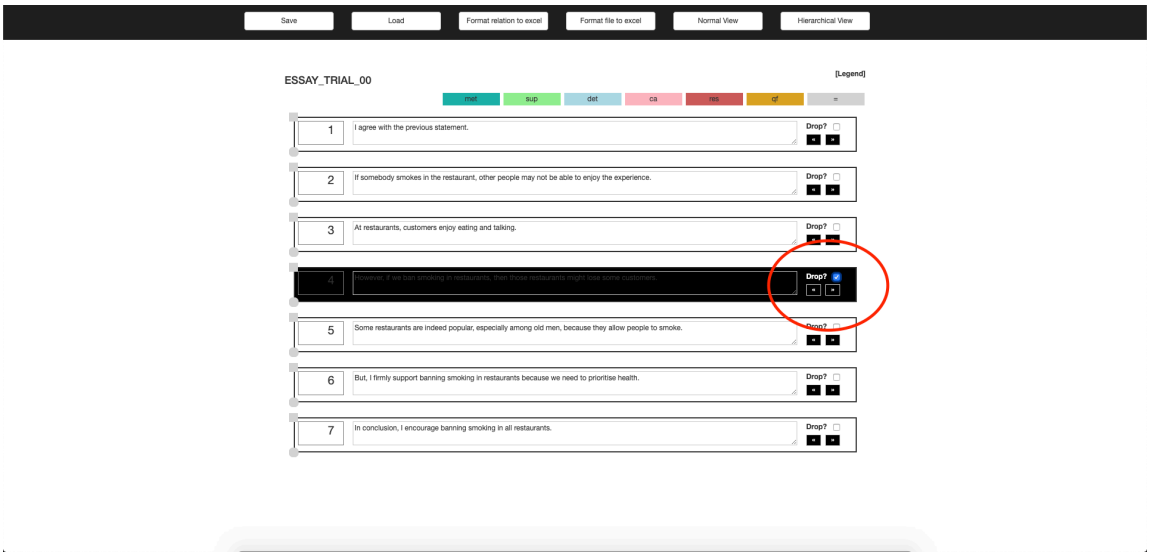


Figure 10. Example of dropping a unit

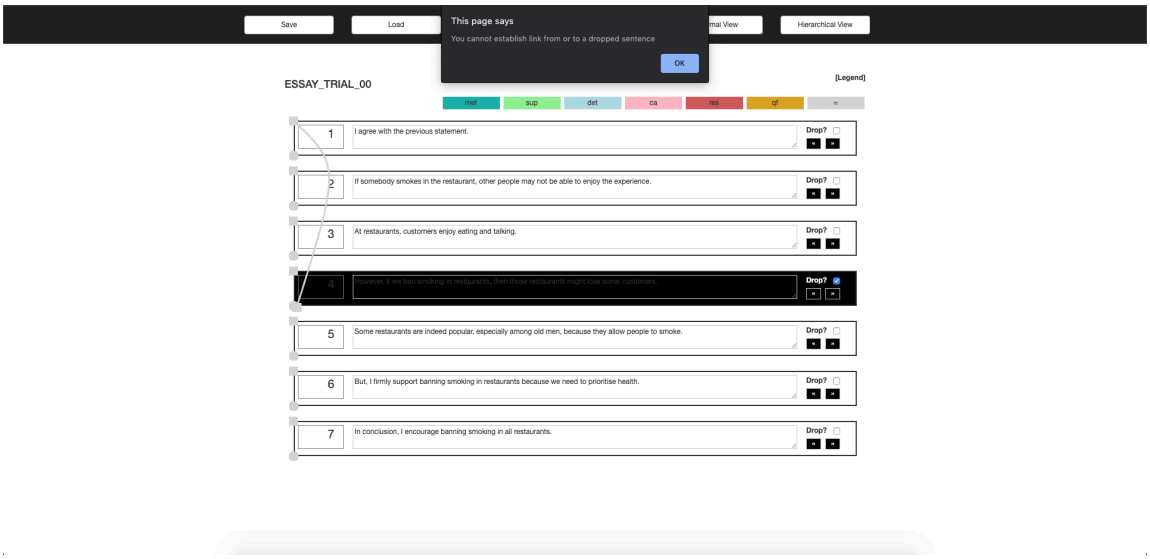


Figure 11. An error message appears when you try to establish a link from or to a dropped unit

6 Moving a Unit's Position

To move a unit (for reordering), click the box containing the unit number of your left-hand side. The whole box is highlighted as illustrated in Figure. 12. Please hold your click while moving the box around. If you have found the place where you want to position the unit, simply release your click. Please note that this functionality is optional (you can turn it on and off, explained in Section 12).

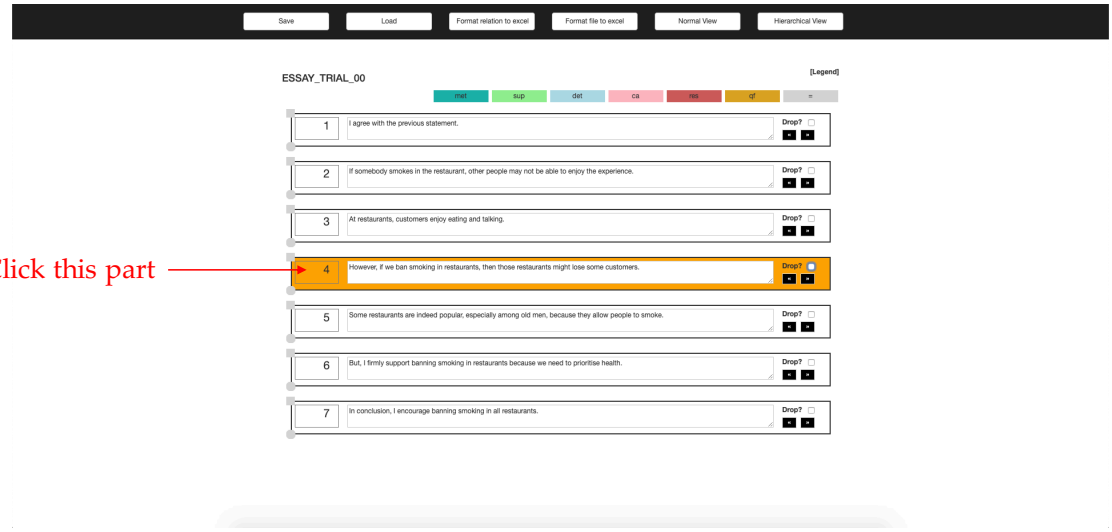


Figure 12. Select the unit you want to move by grabbing the unit-number box on your left-hand side. The whole unit box is highlighted while you do so

7 Unit Indentation

As has been described before, this tool assumes that the relations in text form a hierarchical structure. When you annotate a text, you also need to recognise units' *grouping*. To enable you to simulate the hierarchical structure and grouping, you can change the indentation for the units. A deeper level argument is indented to the right. Figure. 13 shows an example.

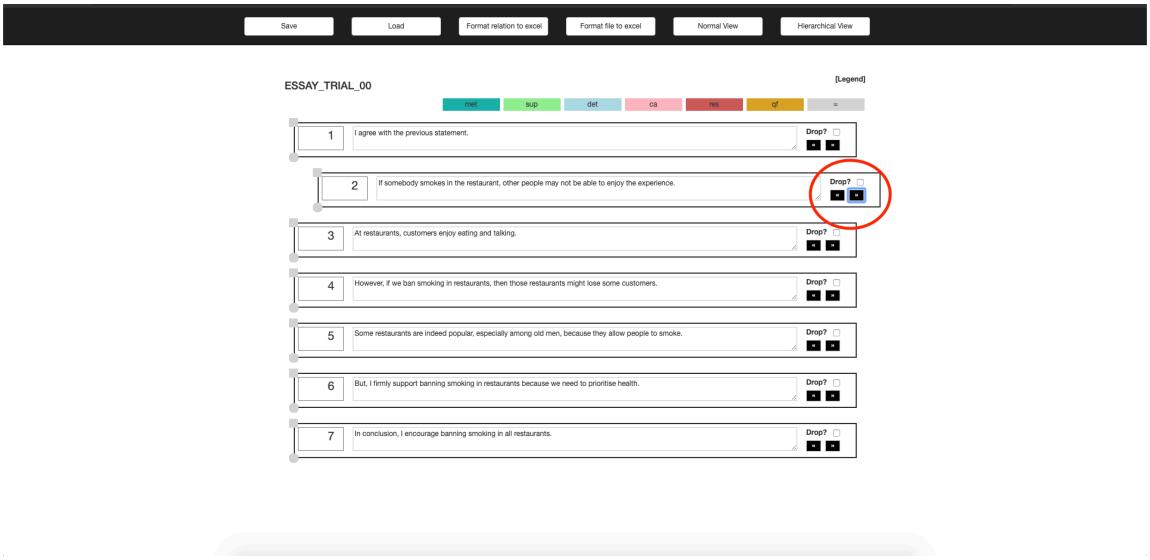


Figure 13. You change the unit indentation by clicking the left or right arrow located on the right-hand-side of the unit.

8 Editing Text

The tool enables you to edit the text. For example, you want to edit an expression describing people or things to make a better anaphora or simply correcting grammatical mistakes and typos (depending on your annotation scheme). 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.

Figure. 14 shows an example. In this case, you read unit (1) as "*I agree that smoking should be completely banned in our country*". The tool checks (shallowly) whether you satisfy this editing format. Figure. 15 shows an example when you do not follow the formatting standard.

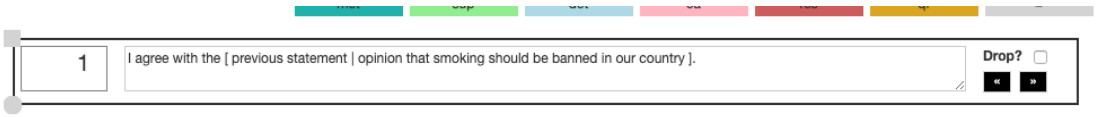


Figure 14. Example of editing the way people and things are described or connected

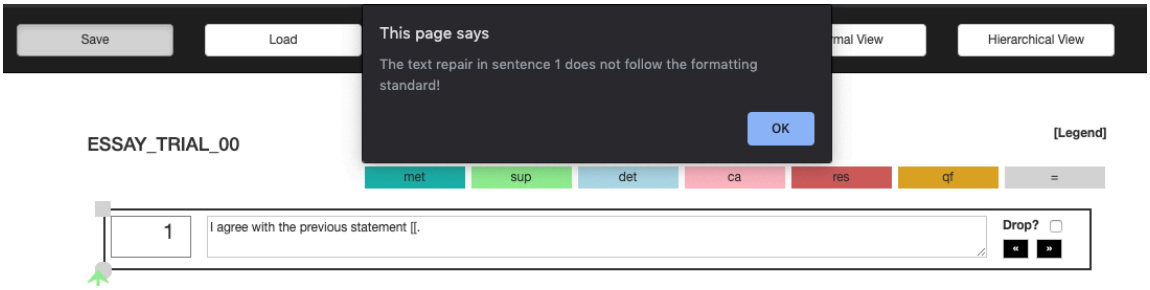


Figure 15. The tool checks whether you edit the text following our formatting standard

9 Saving an Annotated File

After performing the annotation steps, click the "`Save`" button on top of the interface. An annotated file will be automatically downloaded (Figure. 16). Find it in your "`download`" folder (to be precise, the default folder where your web browser stores downloaded files) and move it to the folder where you want to manage them. The saved file is stored in "`.html`" format. It means that you might not be able to understand the file (unless you are experienced in programming). The saved file contains a logging information which is useful for studying annotation behavior. Figure. 17 shows an example. Before working on another file, **refresh your web browser**.

If you want to edit an annotated file, you can load it into the interface (Section 3). On default, we do not allow midway saving. It means you must complete annotating a file before you can save it. However, you can turn this on and off (Section 12) at your own risk. An error message alongside with potential improvements appear when you try to save an incomplete annotation. Figure. 18 shows an example.

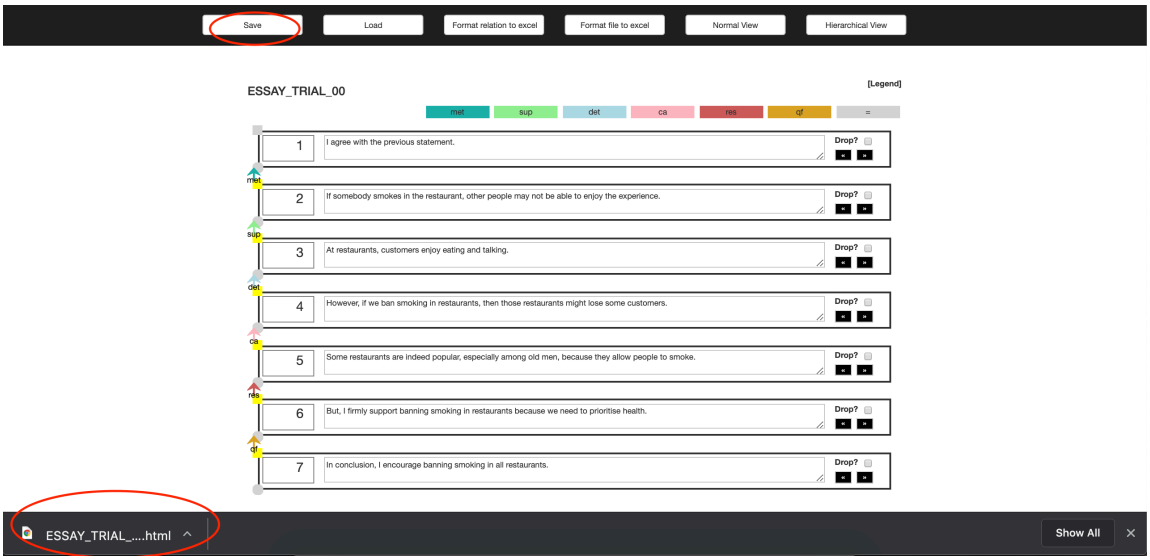


Figure 16. Saving file

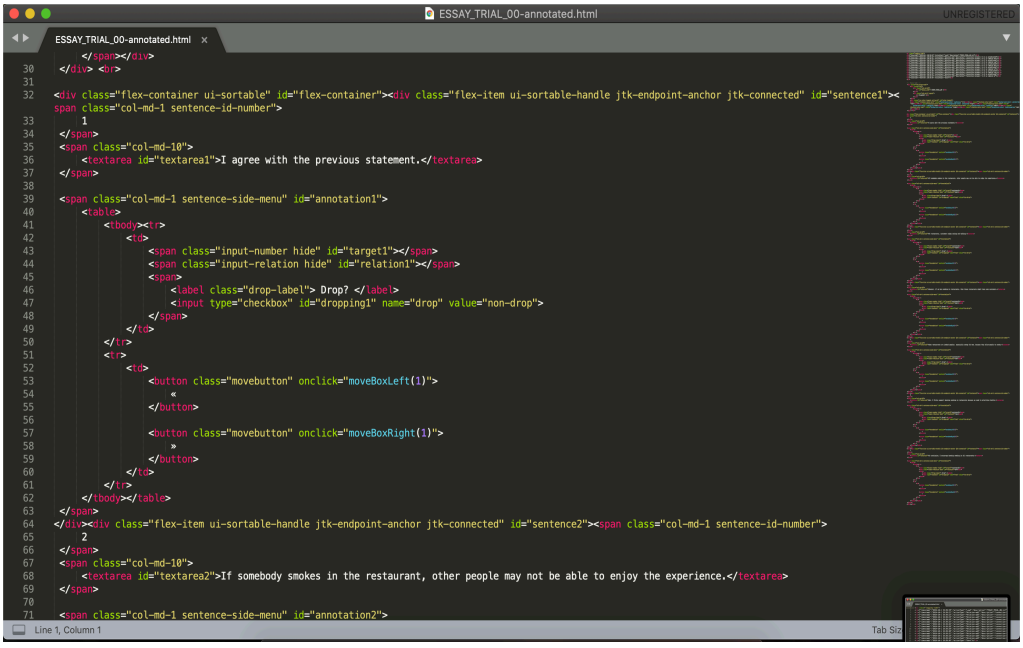


Figure 17. Saved annotation file in "`.html`" format

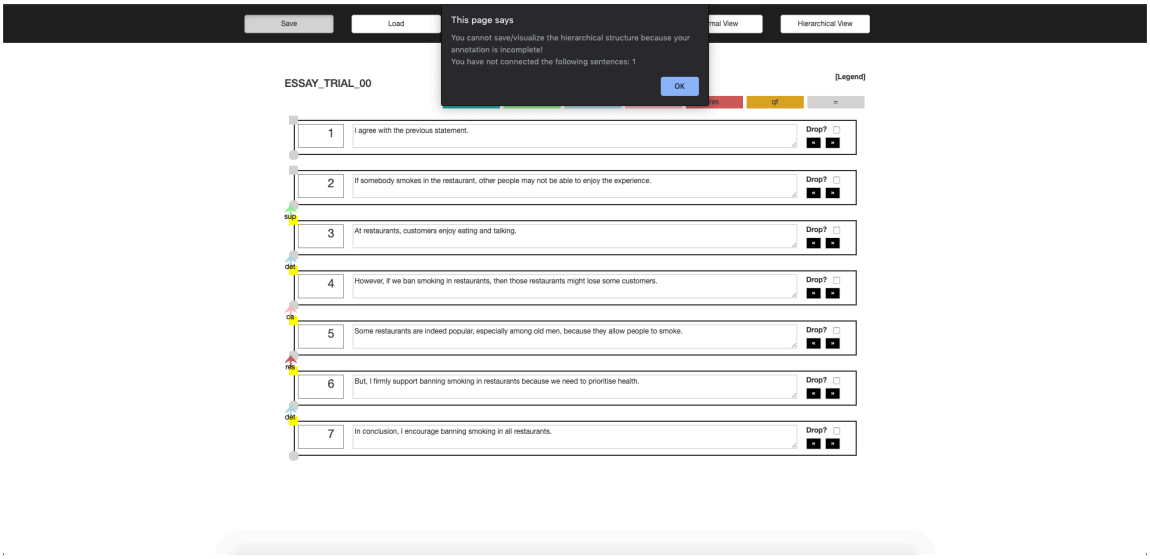


Figure 18. Example of an incomplete annotation

10 Saving Annotated Text to Ms. Excel-readable Format

You can save an annotated text to Ms. Excel-readable format by clicking "`Format relation to excel`" and "`Format file to excel`" on top of the interface (Figure. 19). There are differences between both menus. In "`Format relation to excel`" menu, the tool extracts information on relations of all possible pairs of units, as illustrated in Figure. 20. This is useful for computer inter-annotator agreement. When a pair is not connected by a relation, "`n`" is written in the "`relation`" column. In "`Format file to excel`" menu, the tool converts the annotated file into a human-readable file, as illustrated in Figure. 21.

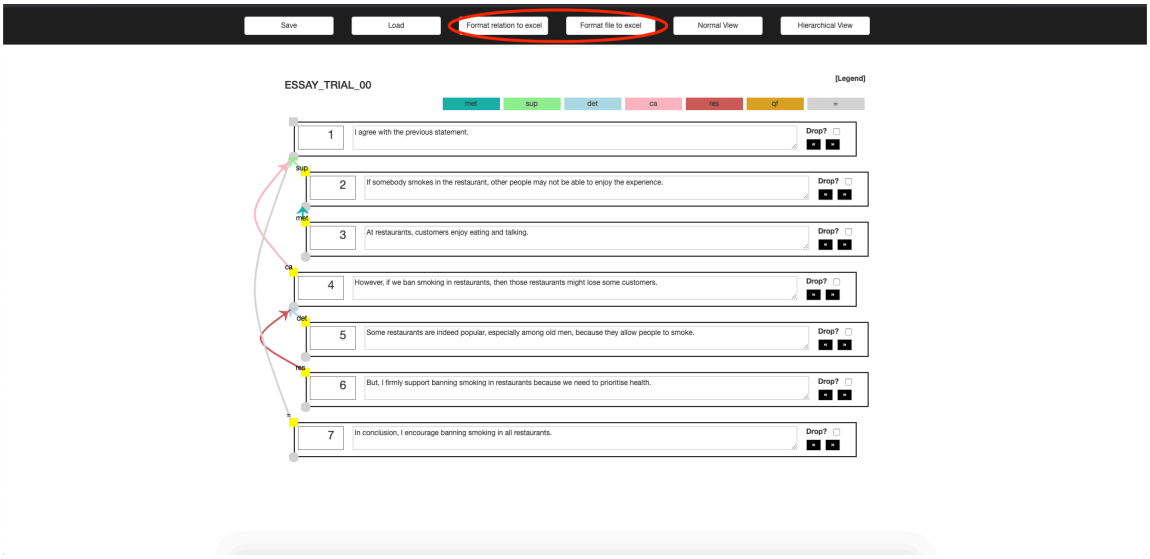
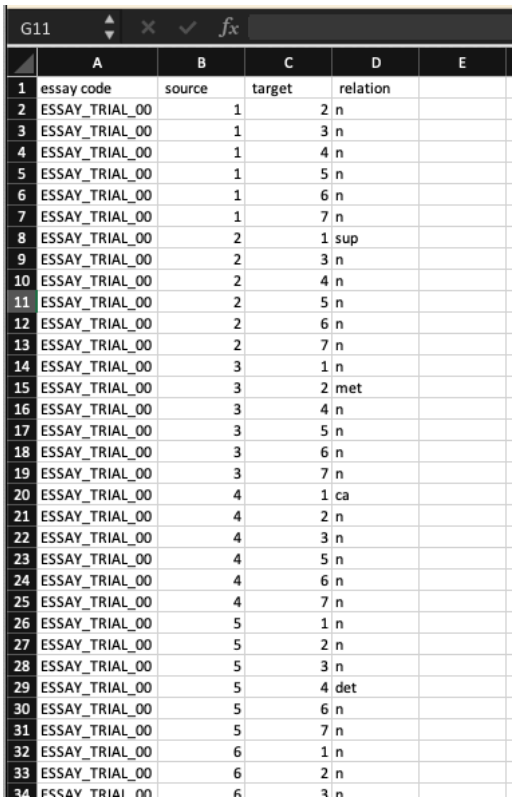


Figure 19. Menus for converting an annotated text to Ms. Excel-readable format



	A	B	C	D	E
1	essay code	source	target	relation	
2	ESSAY_TRIAL_00	1	2	n	
3	ESSAY_TRIAL_00	1	3	n	
4	ESSAY_TRIAL_00	1	4	n	
5	ESSAY_TRIAL_00	1	5	n	
6	ESSAY_TRIAL_00	1	6	n	
7	ESSAY_TRIAL_00	1	7	n	
8	ESSAY_TRIAL_00	2	1	sup	
9	ESSAY_TRIAL_00	2	3	n	
10	ESSAY_TRIAL_00	2	4	n	
11	ESSAY_TRIAL_00	2	5	n	
12	ESSAY_TRIAL_00	2	6	n	
13	ESSAY_TRIAL_00	2	7	n	
14	ESSAY_TRIAL_00	3	1	n	
15	ESSAY_TRIAL_00	3	2	met	
16	ESSAY_TRIAL_00	3	4	n	
17	ESSAY_TRIAL_00	3	5	n	
18	ESSAY_TRIAL_00	3	6	n	
19	ESSAY_TRIAL_00	3	7	n	
20	ESSAY_TRIAL_00	4	1	ca	
21	ESSAY_TRIAL_00	4	2	n	
22	ESSAY_TRIAL_00	4	3	n	
23	ESSAY_TRIAL_00	4	5	n	
24	ESSAY_TRIAL_00	4	6	n	
25	ESSAY_TRIAL_00	4	7	n	
26	ESSAY_TRIAL_00	5	1	n	
27	ESSAY_TRIAL_00	5	2	n	
28	ESSAY_TRIAL_00	5	3	n	
29	ESSAY_TRIAL_00	5	4	det	
30	ESSAY_TRIAL_00	5	6	n	
31	ESSAY_TRIAL_00	5	7	n	
32	ESSAY_TRIAL_00	6	1	n	
33	ESSAY_TRIAL_00	6	2	n	
34	ESSAY_TRIAL_00	6	3	n	

Figure 20. "`Format relation to excel`" result

	A	B	C	D	E	F
1	essay code	unit ID	text	target	relation	drop flag
2	ESSAY_TRIAL_00	1	I agree with the previous statement.			FALSE
3	ESSAY_TRIAL_00	2	If somebody smokes in the restaurant, other people may not be able to enjoy the experience.	1	sup	FALSE
4	ESSAY_TRIAL_00	3	At restaurants, customers enjoy eating and talking.	2	met	FALSE
5	ESSAY_TRIAL_00	4	However, if we ban smoking in restaurants, then those restaurants might lose some customers.	1	ca	FALSE
6	ESSAY_TRIAL_00	5	Some restaurants are indeed popular, especially among old men, because they allow people to smoke.	4	det	FALSE
7	ESSAY_TRIAL_00	6	But, I firmly support banning smoking in restaurants because we need to prioritise health.	4	res	FALSE
8	ESSAY_TRIAL_00	7	In conclusion, I encourage banning smoking in all restaurants.	1	=	FALSE
9						

Figure 21. "`Format file to excel`" result

11 Hierarchical Visualization

You can visualize the annotated text in a hierarchical view. Simply click the "`hierarchical view`" button located on top of the interface. For example, you want to visualize Figure. 22 hierarchical structure. After you clicked the "`hierarchical view`" button, the visualization looks like Figure. 23.

You might feel overwheled in the hierarchical view since the interface might be packed. You can collapse the boxes by clicking the "`collapse button`" on top of each box in the hierarchical view (Figure. 23). Boxes at the lower levels will be collapsed, as shown in Figure. 24. If you want to unfold the collapsed boxes, simply click the "`collapse button`" again. To go back to the normal view, click the "`normal view`" button located on top of the interface.

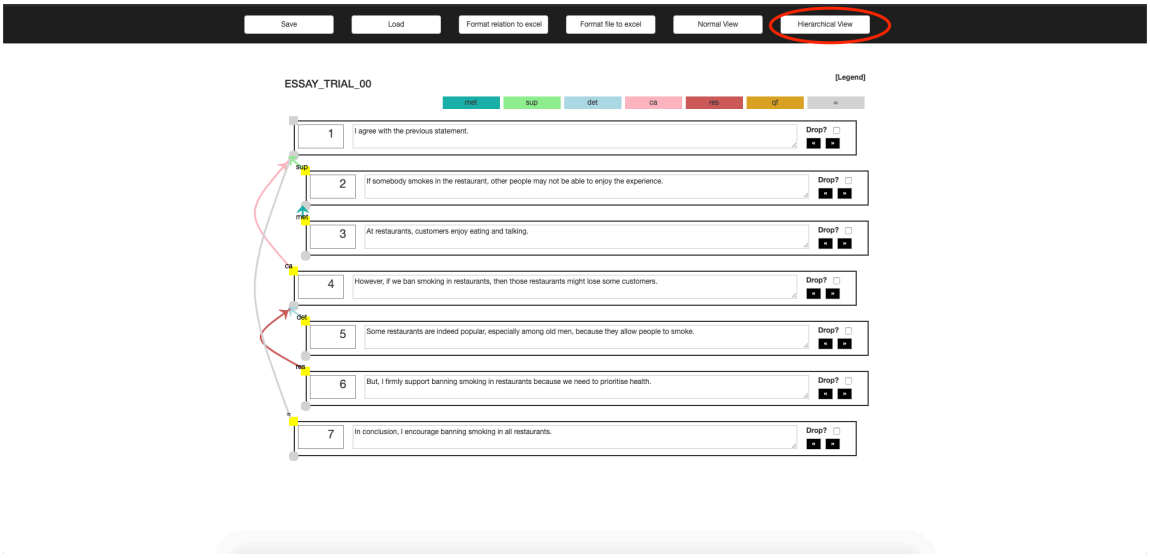


Figure 22. Example of an annotation

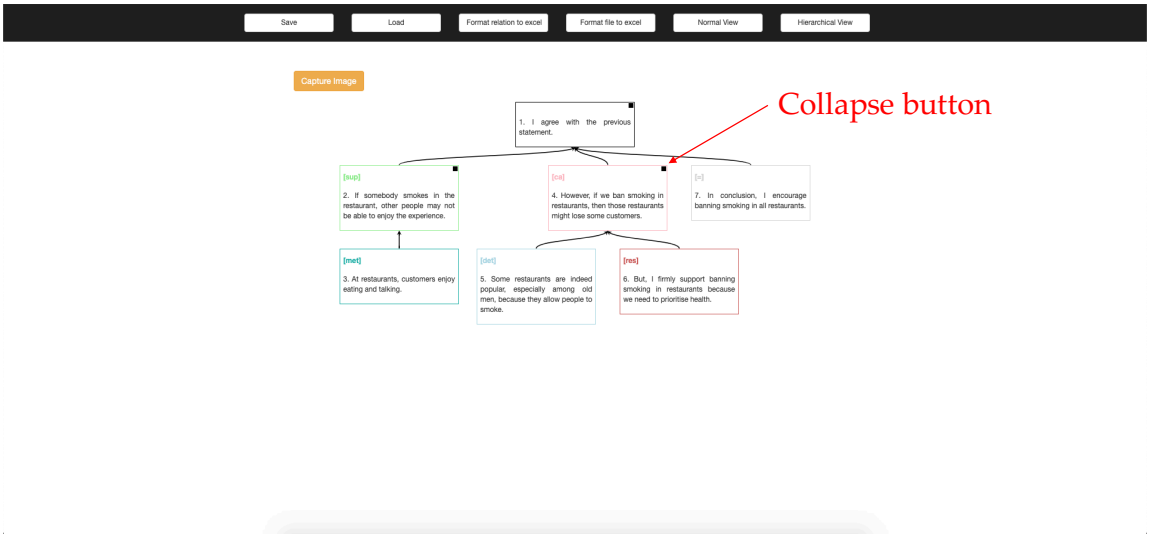


Figure 23. Example of Figure. 22 in "`hierarchical view`" mode

You can capture and save the visualized annotation by using the "capture Image" button located on the top-left side of the interface (Figure 25). The resulting image quality depends on your device, e.g., screen resolution[4]. Please note that large images might be cut-off. However, you can scroll the visualization horizontally and vertically in our tool.

[4] Just like normal screenshot function

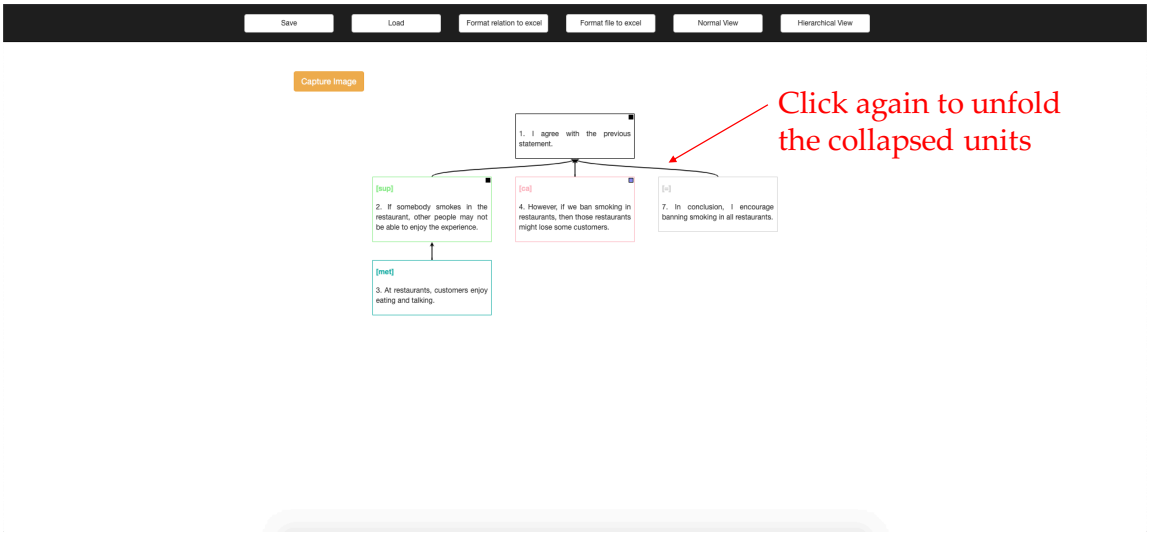


Figure 24. Figure. 22 in "`hierarchical view`" mode (a few units collapsed)

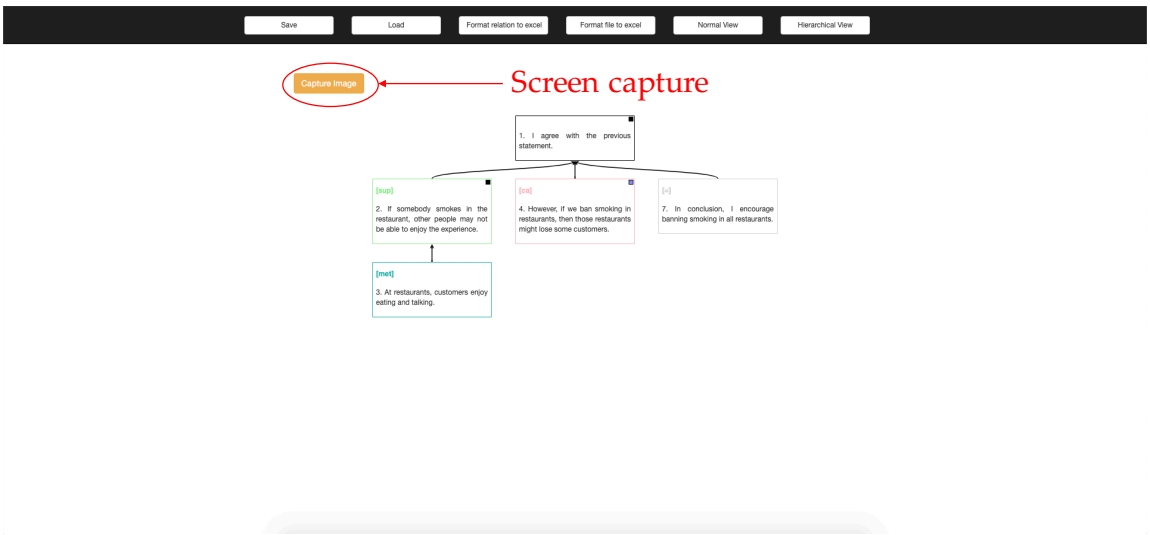


Figure 25. Screen capture

12 Configuring the Tool

The tool's configuration file can be found in "`js/annotation-globalsetting.js`". You can edit the variables in this file[5]. Figure. 26 shows an example. In the configuration file, you can enable or disable dropping (Section 5), reordering (Section 6) and midway-saving (Section 9).

In this configuration file, you can define the relations you want to use, with the corresponding color and relation type (directed or undirected). In terms of normal view visualization, directed relations will have an arrow from the source to the target unit (example in Figure. 6). Undirected relation, however, does not have an arrow. This means, you can semantically think the undirected relation as symmetrical. However, the hierarchical view always display an arrow from child to parent node, denoting which one is the source and target unit.

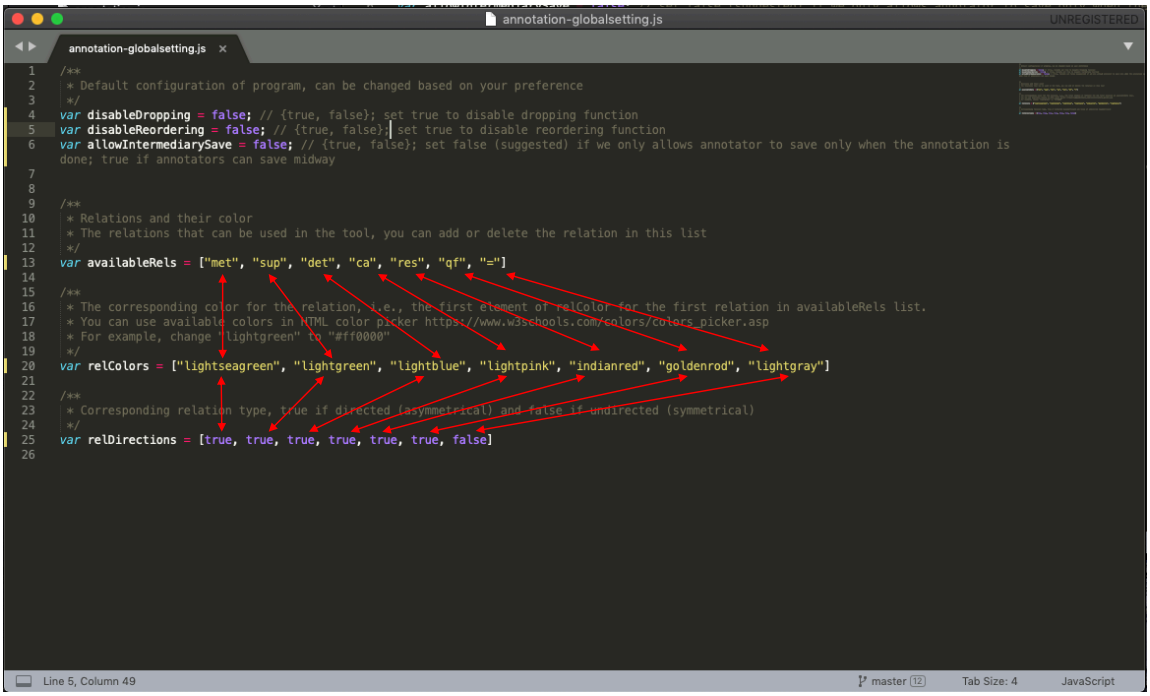


Figure 26. Example of the tool's configuration

Read the instruction carefully when changing the configuration file (especially defining the relations). Improper variable assignment causes error and the tool may crash. Furthermore, you need to make sure the "`availableRels`", "`relColors`" and "`relDirections`" have the same number of items. Figure. 27 shows an example of another (proper) setting. Figure. 28 shows the interface when you changed the configuration to Figure. 27. Notice that the dropping checkboxes have vanished from your view. Also, the legend (showing available relations you can use to annotate) changed compared to Figure. 5.

We suggest you to configure the tool before starting your annotation, and then keep the configuration until the end of the annotation study. The configuration file reflects your annotation scheme, and the tool cannot process properly the saved annotated-texts of different schemes.

[5]For people without programming background: we recommend using `Notepad++` or `Sublime Text` editor

```
/**
 * Default configuration of program, can be changed based on your preference
 */
var disableDropping = true; // set true to disable dropping function
var disableReordering = false; // set true to disable reordering function
var allowIntermediarySave = false; // set false (suggested) if we only allows annotator to save only when the annotation is done; true if
annotators can save midway


/**
 * Relations and their color
 * The relations that can be used in the tool, you can add or delete the relation in this list
 */
var availableRels = ["collection", "comparison", "description", "causation", "response"]


/**
 * The corresponding color for the relation, i.e., the first element of relColor for the first relation in availableRels list.
 * You can use available colors in HTML color picker https://www.w3schools.com/colors/colors_picker.asp
 * For example, change "lightgreen" to "#ff0000"
 */
var relColors = ["lightgray", "lightpink", "lightblue", "lightgreen", "goldenrod"]

/**
 * Corresponding relation type, true if directed (asymmetrical) and false if undirected (symmetrical)
 */
var relDirections = [false, false, true, true, true]

```

Figure 27. Example of changing the tool's configuration

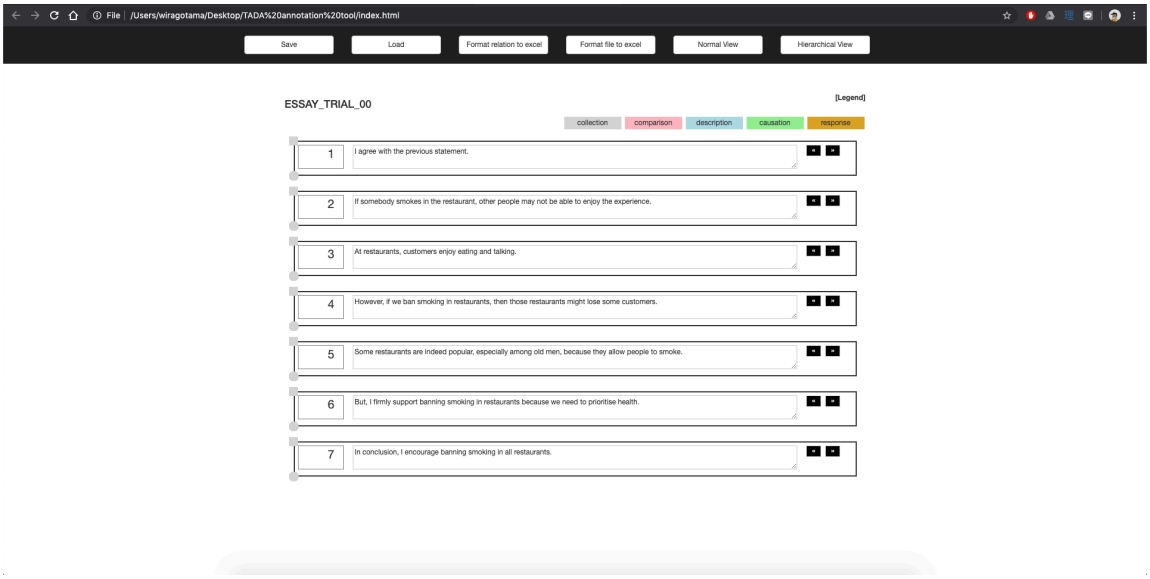


Figure 28. The tool's interface after changing the configuration from Figure. 26 to Figure. 27

13 Citation

We appreciate your patience reading on this manual. If you want to use this tool for research purposes, kindly cite (Putra et al., 2019). We have not published a specific paper for this tool yet, so please kindly check the author's github[6] for the latest citation (star is appreciated).

References

Jan Wira Gotama Putra, Simone Teufel, and Takenobu Tokunaga. An argument annotation scheme for the repair of student essays by sentence reordering. In *Proceedings of Annual Meeting of Association for Natural Language Processing Japan*, pages 546–549, 2019. URL https://www.anlp.jp/proceedings/annual_meeting/2019/pdf_dir/P3-9.pdf.

[6] https://github.com/wiragotama