

Guideline for Annotating “The Black Willow” Novel using MMAX2

1. Purpose

This document provides guideline for annotating “The Black Willow” novel using MMAX2 tool. The main purpose of the annotation is to annotate verbs needing a location, and realized locations in the text and also to make a relationship between them. Below you will find details about how to do the annotation.

2. Markables

There are two layers of markables in this project:

2.1 “Verb” layer:

All relevant verbs for our purpose must be annotated as “verb” markables. More details about the criteria for selecting interesting verbs is provided in section 3.

2.2 “Location” layer:

All the locations with a relationship with a relevant verb must be annotated as “location” markables. More details about the criteria for selecting interesting locations is available in section 4.

2.3 How to Add Markables

To add a markable to a markable layer, you just need to select the word. A list of options will be shown. These options include all the markable layers defined in the project. You should choose the markable layer that you want to add the word to, and the markable is added. Then, you should set its related attributes with appropriate values.

2.4 How to Delete Markables

To delete a markable from a markable layer, you should select the word and click the right mouse button. A list is shown which includes all the markables that contain the selected word. You should choose the appropriate markable layer, and then click on “delete this markable”.

2.5 How to Relate Markables

To make a relationship between two markables, you should select the word from “verb” markable (It will get yellow), and then select the location you want to link to the verb. Click the right mouse button; choose the appropriate markable layer and then “point to this markable”. If

everything is done correctly, you will see a number in front of the related attribute in the attribute window of the verb.

3. Criteria for Selecting Verbs

3.1 FrameNet frames

As the purpose of this annotation is finding a relationship between verbs needing a location and realized locations, the annotated verbs must need a location. Therefore, to specify which verbs must be annotated, we extracted all the frames from FrameNet which had at least one of the five defined location roles (cf. section 5) as their core element.

3.2 FrameNet verbs and their synonyms

From the selected frames, all the verbs and also their synonyms were extracted, and the list is provided in a separate file named “verbs_and_roles.txt” (cf. section 4.4).

4. What must be annotated?

4.1 Relevant verbs

Relevant verbs are verbs which need at least one of the five location roles defined in section 5. To simplify the work, a list of verbs which need location roles is prepared and provided to you in a file called “verbs_and_roles.txt”; however, it does not mean that all the instances of these verbs must be annotated in the text. Each verb can have more than one meaning and not all the meanings need location roles. Therefore, it is the annotator who decides whether the instance used in the text must be annotated or not.

Also, it does not mean that just the verbs of the list must be annotated. The annotator can annotate a verb which is not included in the provided list.

4.2 Relevant locations

We do not annotate all locations in the text – rather, just locations that describe the location of one of the relevant verbs. To see what “location” means in this context, see section 5.

4.3 Relations between verbs and locations

For each verb, it must be annotated that what location roles the verb needs and how they are realized. To see definition of location roles, see section 5.

4.4 How to use “verbs_and_roles.txt”?

This file includes a list of verbs and the possible sets of roles they need. For example, for verb sink, there are two sets of roles:”Goal,path,source” and ”place”. It means that in its different meanings, this verb can need one of these sets of roles. Therefore, for annotating the text, based on the context, the annotator must decide which set of role the given instance of the verb needs. For a complete example, see section 7, and for more examples see section 8.

4.5 Additional considerations

4.5.1 Verbs not in the list

The provided list of verbs is a set of verbs which need one or more location roles. Nevertheless, it is possible to face a verb which needs location but is not included in the list. In this case the verb must still be annotated, its required set of location must be specified and its related locations must be annotated.

4.5.1.1 How to specify roles for verbs not in the list?

To specify location roles of these verbs, you can try to find semantically similar cases in the list. If you can find such a case, you should choose the same roles for the in question verb, and if you cannot find its similar cases, you should choose the required location roles from the possible choices based on your intuition (cf. section 7 for a complete example).

4.5.2 Phrasal verbs

In case you face a phrasal verb, check the phrase in WordNet. If it exists as a separate lexicon, then all the phrase must be annotated together. For example, *sit down*, *lean back* and *leave behind* must be annotated as one verb markable; and if required, you can annotate the particle part as a location role of the verb.

4.5.3 Passives

Passives (in particular adjectival passives) of motion verbs have generally a ‘position’ meaning. We nevertheless annotate passives as “full” motion verbs. This means that many location roles will remain undetermined, but that is in accordance with FrameNet practice (“constructional non-instantiation”). Our reason for doing so is that we can distinguish source, goal, and path arguments in this manner, which would be impossible if we annotated only “Place” for a position verb.

4.5.4 Non-physical motion

We annotate metaphorical and metonymical motion (“meet an obstacle”, “arrive at a conclusion”) like standard physical motion.

5. Location roles

For each location, we describe the type of relation between the verb markable and the location markable with one of five location roles: “place of happening”, “source”, “goal”, “path” and “direction”. However, not all the roles are available for all verbs.

5.1 Location Roles Definition

Below you can find the definition of the roles from FrameNet:

5.1.1 Place

Place identifies the setting in which the Theme's movement takes place without a specified Path.

5.1.2 Source

If A is the source of a motion event with mover B, then B is at place A at the beginning of the event.

5.1.3 Goal

If A is the goal of a motion event with mover B, then B is at place A at the end of the event.

5.1.4 Path

If A is the path of a motion event with mover B, then B is at place A at sometime during the movement event, but neither at the beginning of the event nor at the end.

5.1.5 Direction

If A is the direction of a motion event from source B to goal C, then A is [...] a straight line from B to C.

5.2 More Practical Definition of the Roles

The above definitions for the location roles are the formal definition. To simplify annotating the text here is more practical definitions for the roles:

5.2.1 Place:

The Place is usually recognized by: a place preposition (in, inside, ...) + place.

5.2.2 Source:

The source is usually recognized by: a source preposition (from, out of, ...) + the starting point of the motion. However, there is a small set of special adverbials which can have the role of source without any preposition (e.g. *He left home*).

5.2.3 Goal:

The Goal is usually recognized by: a goal preposition (to, into, ...) + the end point of the motion. However, there is a small set of special adverbials which can have the role of goal without any preposition (e.g. *He reached home*).

5.2.4 Path:

The Path is usually recognized by: (optional adverbial) + preposition (by, through, past, ...) + a place which is neither a source nor a goal. However, there is a small set of special adverbials which can have the role of path without any preposition (e.g. *He passed home*).

5.2.5 Direction:

The Direction is just recognized by an adverbial indicating direction (up, down, towards, away from, ...).

5.3 Co-referent locations

To relate a location markable to a verb markable, we use the nearest location, even if it is a pronoun referring to a location mentioned earlier. To make it clearer, here is an example:

Allan unfolded another page, this one crowded with letters.

In this example, there is a relation between *crowded* and *this one*, not with *another page*.

6. Attributes

Each markable can have some attributes. Currently, just the “verb” markable has some attributes. Below you can find some information about its attributes.

6.1 Verb_is_in_FN_list

This attribute shows whether the annotated verb is included in the list provided in “verbs_and_roles.txt” or not.

6.2 attr_is_in_FN_list

This attribute shows whether the set of roles which is chosen for the annotated verb is mentioned as one possibility in “verbs_and_roles.txt” for that verb or not. It means that the annotator can choose a set of role for the verb which is not mentioned as possible set of role for that verb. In this case, the value of this attribute must be set to “No”, and the appropriate set of required roles for the verb must be chosen from the drop_down list of required roles based on the intuition of the annotator.

6.3 Required roles

This attribute specifies the required set of roles for each verb. For each verb, the annotator chooses one set of roles from the drop-down list.

6.2 Location role attributes

Each relation between a verb and a location needs to be labeled with one of the relations (roles) described in Section 5. In addition to choosing a label, the annotator needs to choose a *status* for each of the possible roles. We currently distinguish five status groups:

1. **Local realization:** This option is selected when the related location is overtly expressed in the same predicate-argument structure, that is, when there is bona fide syntactic relation between verb and location (including long-distance dependencies). Here is an example:

Ex.2 He was sitting **in the room**.

In this example, “in the room” is the “place of happening” of the “sit” event. Therefore, the “locally realized” option must be selected for this role.

2. **Local inference:** This option is selected when the place of the occurrence of the event is not expressed in the same phrase, but can still be found in the same, the previous or next sentence. Here is an example:

Ex.3 He went **to the room** and sat down.

In this example, “the room” in the previous sentence has the role of “place of happening” for “sat down” event. Therefore, “inferred” option must be selected for this attribute.

3. **Global inference:** This option is selected when the place of happening of the event is present somewhere in the document but not near the predicate. Below you can find an example:

Ex.4 He was **in the room**. [...]. He went towards the window.

In the above example, “the room” is the place of happening for “go” event, even though it is mentioned rather far away, at the beginning of the paragraph.

4. **World knowledge:** This option is selected when the role is not mentioned in the text, but it can be inferred from the meaning of the verb or from the complete context. For example:

Ex.5 He raised his arm.

Here, the direction in which the arm was raised is not specified, but it is part of the lexical meaning of *raise* that the movement is upwards.

5. **Undetermined:** This option is selected when the place of happening is not mentioned in the text, and when it cannot be inferred from the meaning of the verb or the context either with any certainty. This frequently happens for the place and/or path roles.

6.2.1 World knowledge vs. undetermined?

To annotate the realization status of a location role as “world knowledge” you must be able to name one specific place/path/direction (rather than a set) to annotate the role. Having a location role in the form of “anywhere except X” is not an acceptable reason to annotate the location role as “world knowledge”.

For example, in *He left the room*, you cannot consider “anywhere out of the room” as the “destination” of the event and annotate its realization status as “world knowledge”.

In contrast, if there are multiple possibilities for one location role, its realization status must be annotated as “undetermined”.

Also, you must not use the background location of the text as the world knowledge location

role of verbs which need more specific locations. For example, if the entire story is taking place in a jungle and this sentence exists in the text:

He got tired and sat down.

It is **not** acceptable to say the person sat down “somewhere in the jungle” and to mark the realization status for “place of happening” as “world knowledge”. The reason is that the jungle is the background place of the story and “sat down” needs a more specific place as its “place of happening”. In this case, the “place” role of the verb is “undetermined”.

7. A Complete Example

Assume this sentence:

Allan sat down at his desk and pulled the chair in close.

To annotate this sentence, firstly, we should determine which roles each verb needs. If we take a look at the file “verbs_and_roles.txt”, we can see that there are two “sit_down”s with different sets of roles. According to the context, the “sit_down” we face just needs “place”. Therefore, from the top-down list for “required roles” we should use “place”. Choosing any of the possible options causes the related attributes to appear to specify how the roles are realized in the text. By default, “place” is chosen.

For attribute “place” we choose local realization; because “place of happening” for verb “sat down” is overtly expressed in the same clause as the verb.

When we relate markable “sat down” and “at his desk”, “related place” attribute will automatically change to the identity number of markable “at his desk”.

And here is the annotation for the second verb:

If we look up *Pulled* in the “verbs_and_roles.txt” file, we will see that this verb does not exist in the file. Therefore, attribute “is_in_FN_list” must be set to “No”. Nevertheless, it needs some location roles. To find the required location roles, we try use the “verbs_and_roles.txt” file. For example, according to my intuition the location roles of this verb will be similar to “push”. Thus, I find “push” in the file. According to the file, “push” needs four location roles: goal, path, place and source. Thus, I choose this set of roles for “pull”.

Therefore, the realization of location roles for “pull” will be like this:

Place: local inference (because it is not overtly expressed in the same clause as the verb, but from the previous sentence we can guess the place of happening)

- **Related_place:** at his desk (*at his desk* in the previous sentence is the place where *pulling* has taken place)
 - o **Place_of_occurrence:** previous sentence

Source: undetermined (because the source of the movement is not mentioned in the text at all)

Goal: local realization (because the goal is overtly expressed in the same clause as the verb)

- **Related_place:** close
 - o **Place_of_occurrence:** same_sentence

Path: Undetermined (because the path of the movement is not mentioned in the text at all and it can not be inferred from any part of the text)

8. Some examples of verbs' location roles

Ex.1 *They WALKED slowly, silently, in the grounds.*

In this reading, *walked* needs just “place”.

They WALKED from the balcony across the expanse of the chamber to a marble table which stood against one wall.

In this reading, *walked* needs five location roles (place, source, goal, direction and path) among which three roles are mentioned locally.

Ex. 2 *Tommy was CLIMBING all over it.*

In this reading, *climbing* needs just “place”.

We could CLIMB to the church rock.

In this reading *climb* can have five location roles (place, source, goal, direction and path) among which one role is mentioned overtly.

Ex. 3 *Flora put her hands over her ears and began to JUMP about all over the kitchen.*

In this reading, *jumping* needs just “place”. However two phrases have this role for the verb. The first one is *about* and the second one is *all over the kitchen*.

She JUMPED off the bed heavily.

In this reading *jumped* can have four location roles (source, goal, direction and path) among which one role is mentioned overtly.

Ex. 4 *When I see my postman DRIVING around in his van with the radio on, I crave his job.*

In this reading, *driving* needs just “place”.

If you can DRIVE an articulated truck to the main floor from the street outside, you know you're in business.

In this reading, *drive* can have three location roles (source, goal, and path) among which one role is mentioned overtly.