NOVA Documentation

Johannes Wagner

06/27/2016

Contents

1 Introduction

The **N(On) Verbal Annotator** (NOVA) offers a graphical interface for machine-aided annotation of large multi-modal databases. It is developed as a side project of the Social Signal Interpretation (SSI) framework to describe databases recorded with SSI. It's key features are:

- **Visualization and playback** of multi-modal data including support for skeleton and face tracking
- Annotation based on **various schemes** (e.g. discrete and continuous)
- **Database back-end** to share annotations between multiple annotators
- Integrated tools to apply machine-aided **cooperative learning**

2 Installation

The latest version of NOVA can be accessed at:

https://github.com/hcmlab/nova

Just download or check-out the release branch and start 'nova.exe' (yes, NOVA is available for Windows systems only).

If you do not see videos on playback, please make also sure you have the right codecs installed. A good choice is the K-lite Codec Pack (You can leave all settings on default, but make sure to not not install any 3rd party software that comes with the installer).

3 Manual Annotation

In the following we will explain the basic steps to manually create and edit an annotation.

3.1 Main Interface

The following image shows NOVA's main interface. On the top, media files are displayed in horizontally aligned boxes. Videos and tracking data (e.g. facial points) are handled as media and displayed frame by frame. Below that other stream data (e.g. audio waveforms) are displayed as time series in vertically aligned tracks. In the same way, annotation data is visualized in vertical tiers. Below the annotation tiers a slider bar allows it to move the displayed clip along the time scale. The clip can be enlarged or shrink by changing the size of the scroll button. Double-clicking the scroll button will zoom out to show the whole session at once (note: if navigation becomes slow try zooming in). Playback buttons are found on the bottom of the window. A red cursor marks the current playback position (it will be moved during playback and when the signal track is clicked). A green cursor shows the current position on the annotation tier and helps to align labels with the signal tracks. A list of annotations items of the currently selected annotation tier is shown on the left. When an item is selected by the red cursor is moved to the according segment in the tier. Only one media, stream and annotation can be selected at a time. When an entity is clicked which is not already selected, it receives the focus (ie.g. its border turns dark gray) and additional information is displayed in the status bar.
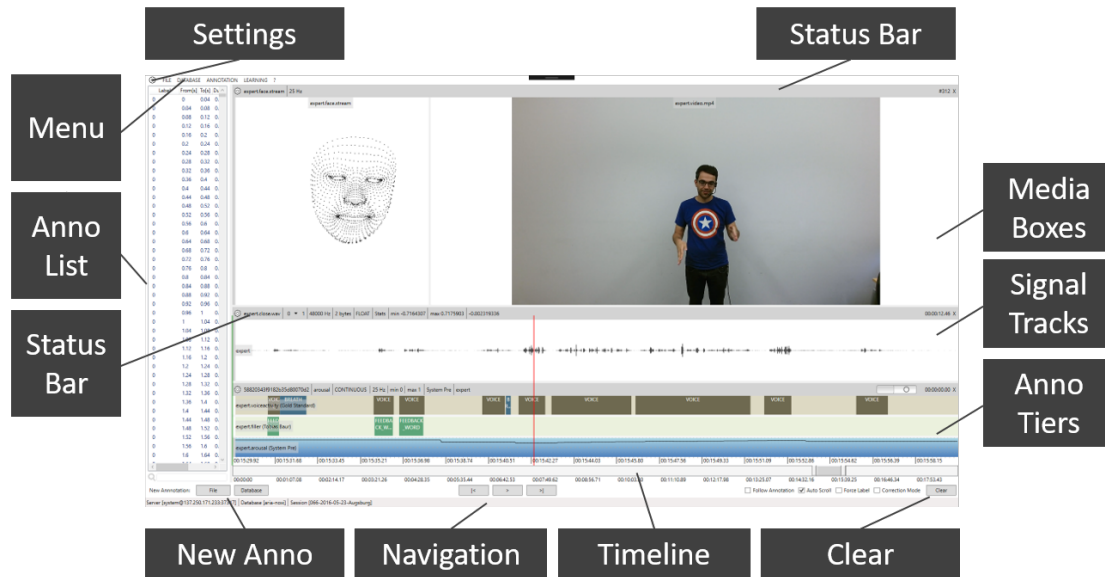
Figure 1: *NOVA's interface: from top down media, streams and annotations are displayed. On the left a list of annotation items of the selected annotation tier is shown. Time-line and the navigation panel are found at the bottom.*

3.2 Loading Files

To add a file the FILE menu can be used. It allows to select multiple media, stream or annotation files at once. Alternatively, files or folders can be dropped from the explorer. NOVA supports all common video formats, audio wave files and SSI stream files, as well as, CSV files, which will be imported as stream files (in the latter case the user has to provide number format and sample rate). The currently selected media, stream or annotation file can be removed by clicking the 'x' in the according status bar. Clicking the 'Clear' button on the right bottom will remove all files at once. The current workspace can be stored to a project file and reloaded at a later point from the FILE menu (or via drag and drop). Note that in the following we will generally refer to any file that is not annotation as stream.

3.3 Creating an Annotation

After loading at least one media or data stream, new annotations tracks can be added to the project. New annotations are created by clicking the 'File' button (bottom left). Alternatively, annotations can be created from the database, too, see here. A window pops up, which allows to select a scheme type:

- Discrete:

Figure 2: *New annotation scheme window.*

Figure 3: *Example of a discrete annotation tier.*

Discrete annotations consist of a list of labelled time segments. Each segment has a start and and end time (in seconds) and a name (label). Segments can be of varying length, may overlap and possibly there are gaps between adjacent segments. An annotator cannot change the names of labels and has to assign exactly one label to each segment. If none of the labels is applicable the label "GARBAGE" is always available.

Scheme name, class names and colours are set in the following dialogue:

- Free:

Like discrete annotations, but this time annotators are free to choose the label names. This is obviously useful if an annotation task can not easily be reduced to a few labels (for example in case of speech transcriptions). Of course there is the risk that the same phenomenon may be labelled differently (either because a synonym is used or due to misspelling).

- Continuous:

In contrast to discrete annotations, continuous annotations assign numerical values (label score) instead of label names. Also, they have a fixed sample rate, i.e. label scores are assigned in regular intervals. For instance, a sample of 2 Hz means that an annotator has to assign two scores each second. All score values have to be within a fixed interval (defined by a minimum and a maximum score). Optionally, it is possible to quantize the interval into limited steps. The value NAN is assigned when no score is applicable.

Scheme name, sample rate, value range and colours are set in the following dialogue:

3.4 Editing an Annotation

- Free:

To place a new segment right click on the start position of the label, keep button pressed and move the cursor to the end position (or vice-versa). Now release the button and a new segment will be added (and selected). If 'Force Label' is active you will be asked to

Figure 4: *New discrete annotation scheme window.*

Figure 5: *Example of a free annotation tier.*

Figure 6: *Example of a continuous annotation tier.*

Figure 7: *New continuous annotation scheme window.*

enter label name and confidence. Otherwise it gets the label that was previously used. To edit an existing segment left click inside the segment (the colour of the selected segment will change to blue). If you now move the cursor near the borders two arrows will be displayed and you can change the position by holding down the left mouse button. If you move the cursor towards the centre of the segment you can move the whole segment along the track. To change the label hit the 'W' key. Note that you can change the position of a label during playback. If 'Follow Annotation' is active a newly added segment will be immediately played back. Alternatively, you can also jump to a segment by selecting it from the list left to the annotation tiers. It is also possible to select several segments in the list and rename them all at once.

- Discrete:

Works in the same way as free annotations, but when entering or changing a label a selection dialogue pops up that allows you to select a name from the loaded scheme. Alternatively, you can change the label name of a selected segment by pressing a key between '0' and '9'. For example, if the selected scheme contains three names 'A','B' and 'C' pressing '1' assigns 'A', '2' assigns 'B', and '3' assigns 'C'. Any other numbers and '0' assign the garbage class.

- Continuous:

To change the values in a continuous track hold down the right mouse button and move the cursor to the desired position within the track. You will notice that the values immediately start to follow the cursor. Hitting 'L' (or switching on the live button in the status bar) turns on the live mode, which no longer requires to hold down the right mouse button. Instead a white button is displayed at the left border of the current track to mark the current value. The height of the white button can either be controlled by the vertical position of the mouse cursor (press key 'M' or activate check box next to live button). Or otherwise by pressing a key between '0' and '9' (see option 'number of levels in live mode' in the SETTINGS). This is especially handy to annotate during playback. Hitting 'L' again turns off the live annotation mode and brings back the default behaviour.

Pressing 'Ctrl+Z' and 'Ctrl+Y' allows it to undo and redo the previous changes. To save an annotation go to the ANNOTATION menu or press 'Ctrl+S'.

4 Database

NOVA allows users to manage and share annotations through a database. To connect to a database open the SETTINGS and select the 'Database' panel. Here you can enter the host IP, port number and user credentials. It also allows you to change the folder to which NOVA downloads media and stream files that are used during the annotation process, as well as, the directory where the cooperative learning definitions and models are stored.

4.1 File Format

To manage a database with NOVA, you have to follow a certain file structure. Each database is located in a root folder with the name of the database and may consist of one more sessions. All stream files belonging to a session are grouped in sub-folder within the root folder. The name of the folder defines the name of the session. In a session we distinguish between several users, which take a certain role. E.g. thinking of a dyadic conversation, one user could be the expert sharing the knowledge about a certain topic to a novice user. Hence we have two roles: 'expert' and 'novice'. Each file has a unique name defined by the role and the type of recorded channel. E.g. if we have recorded the interaction using close talk microphones and two webcams, we may use the following file structure:

```
aria-noxi/
    067_2016-05-23_Augsburg/
        expert.close.wav
        expert.video.mp4
        novice.close.wav
        novice.video.mp4
    068_2016-05-23_Augsburg/
        expert.close.wav
        expert.video.mp4
        novice.close.wav
        novice.video.mp4
    ...
```

4.2 Administration

TODO: explain how to setup a database

4.3 Loading a Session

The DATABASE menu allows to load annotation and files for a session (note that this will clear the workspace). In the following window you can select multiple annotations and streams belonging to a certain session of a database. Turn on 'Mine only' to filter only your own annotations and 'Unfinished only' to hide annotations that have already been marked finished. Stream files that are not locally available yet, are displayed in red. If selected NOVA will try to download them before the session is displayed.
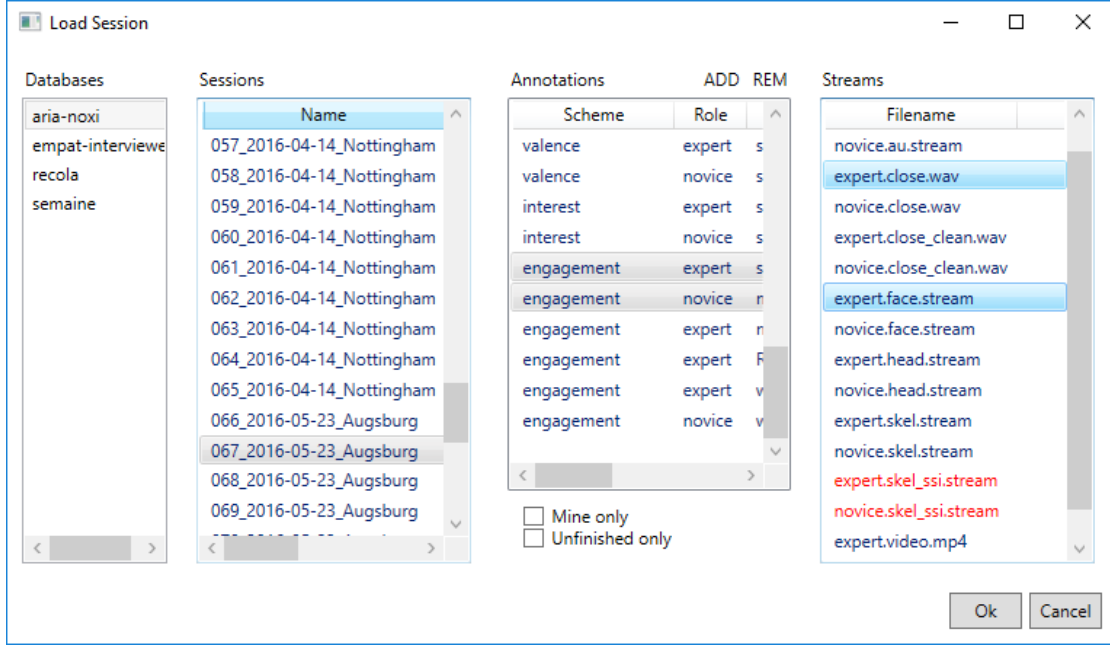
Figure 8: *Load session from database window.*

4.4 Creating an Annotation

After loading at least one stream file it is possible to add new annotations to the database by pressing the 'Database' button on the bottom left. You will now be prompted to select a scheme and a role. If an according annotation already exists it will be loaded, otherwise an empty one is created. When the annotation is saved it is automatically written to the database. In case you want to store it on disk use 'Export Selected As' from the ANNOTATION menu. You can also load an existing annotation from another user. As soon as you save the annotation for the first time it will be stored under your name.

5 Cooperative Learning

NOVA features tools to considerable speed up the annotation process. The following image illustrates the overall system. (A) Database with recordings of human interaction. (B) NOVA's interface allows to distribute and accomplish annotation tasks among human annotators. (C) At times, Cooperative Learning (CL) can be applied to automatically complete unfinished fractions of the database. Two strategies are available (bottom right box): (I) A session-dependent model is trained on a partly annotated session and applied to complete it. (II) A pool of annotated sessions is used to train a session-independent model and predict labels for the remaining sessions. In both cases, confidence values
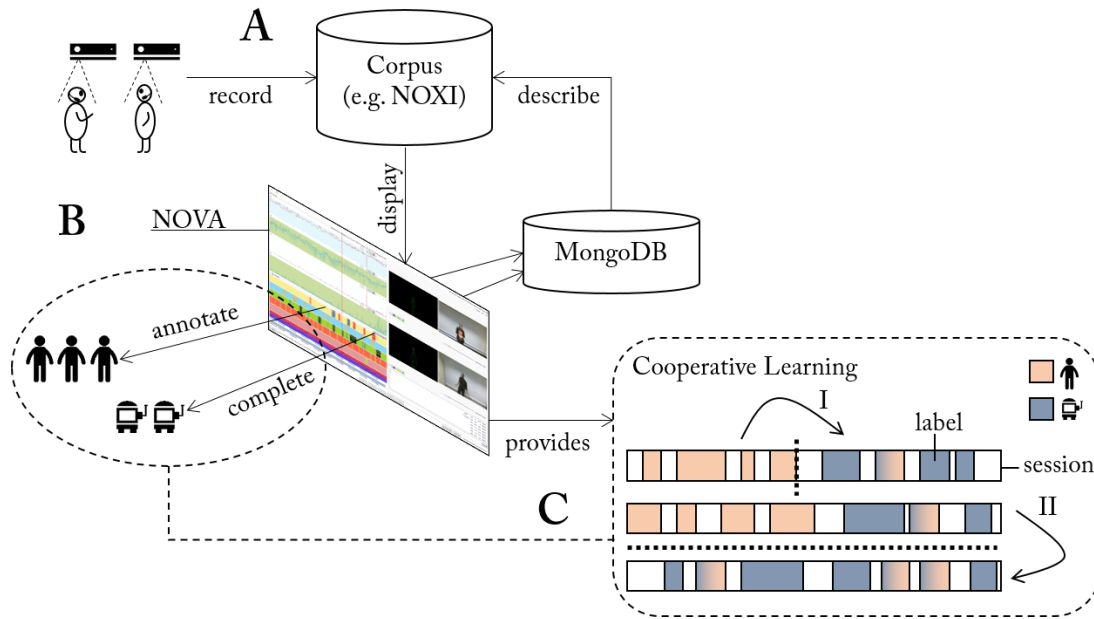
Figure 9: *Overview of the cooperative learning system integrated in NOVA.*

guide the revision of predicted segments (here marked with a colour gradient).

5.1 Feature Extraction

TODO: explain how to extract features

5.2 Trainer Templates

TODO: explain how to create trainer templates

5.3 Annotation Completion

NOVA allows it to complete a partly finished annotation. Note that NOVA expects that everything before the last annotation item on the tier has been manually annotated. The following image shows a partly finished annotation before completion is applied:

Now, call 'Complete Current Annotation' from the LEARNING menu. In the dialogue that pops up you can now select the feature stream to which machine learning should be applied. Only streams are displayed for which at least one trainer template is available. Depending on the scheme, there may be additional options that allow you to tune the
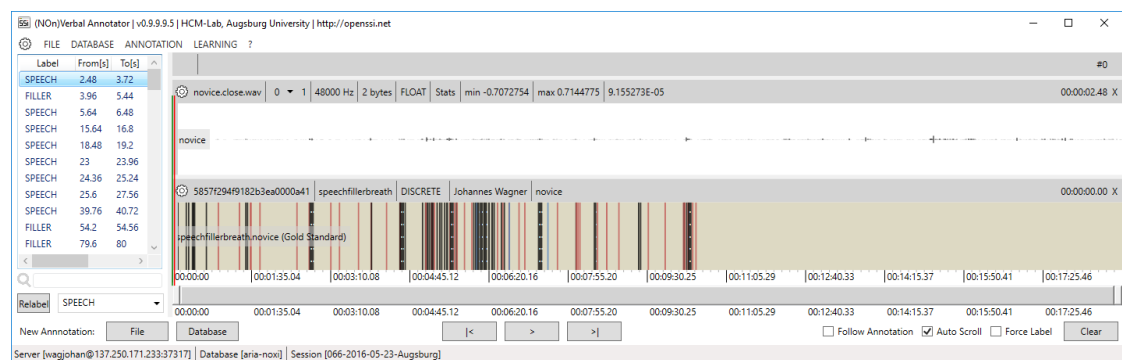
Figure 10: *Session before completion.*

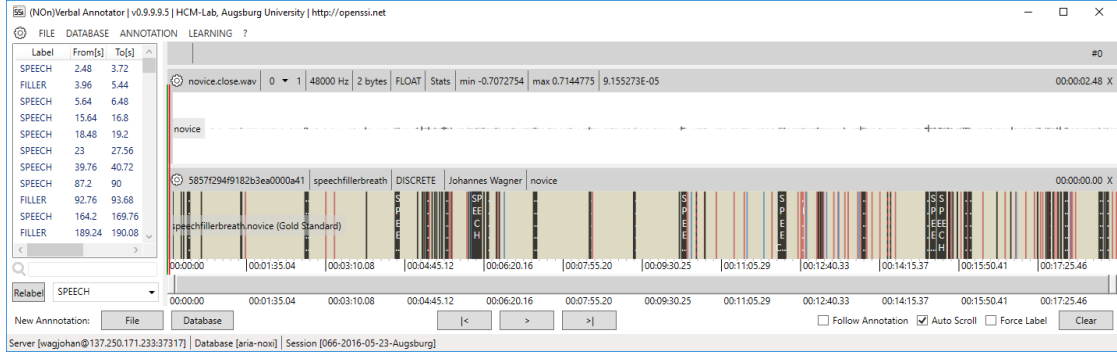Figure 11: *Annotation completion dialogue.*
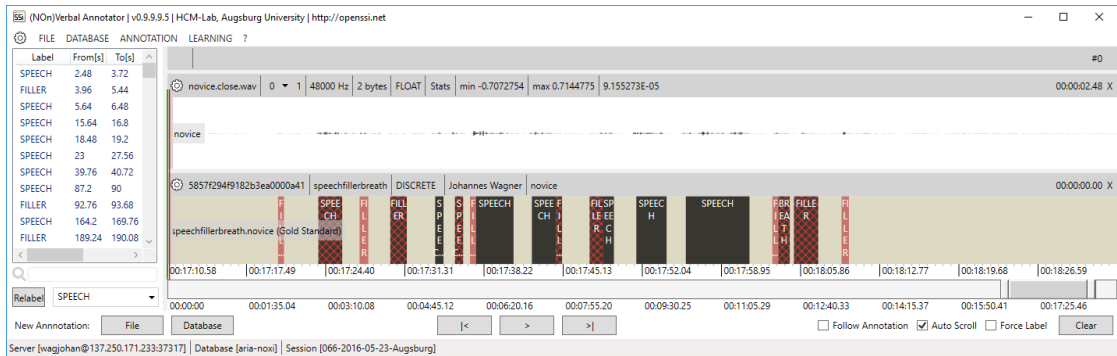
Figure 12: *Session after completion.*

Figure 13: *Predictions with a low confidence are marked for revision.*

generation of the prediction. E.g. in case of a discrete scheme you can automatically fill small gaps or remove segments below a threshold.

If you are not happy with the completion you are still given the opportunity 'Undo' the changes now. Otherwise, you can now continue to work with the automatically completed annotation.

Note that predictions with a low confidence will be marked with a special pattern to guide manual revision. You can change the threshold in the SETTINGS ('Correction certainty level'). Once a label is revised the confidence is automatically set to 1 and it is not longer marked.

5.4 Model Training

Once annotations for a number of sessions have been accomplished a 'strong' classification model can be trained. This requires that features have been extracted for the sessions and a compatible trainer template is available. Now choose 'Train Model' from LEARNING and a dialogue pops up. Here you can select a scheme, one or more roles and an annotator.
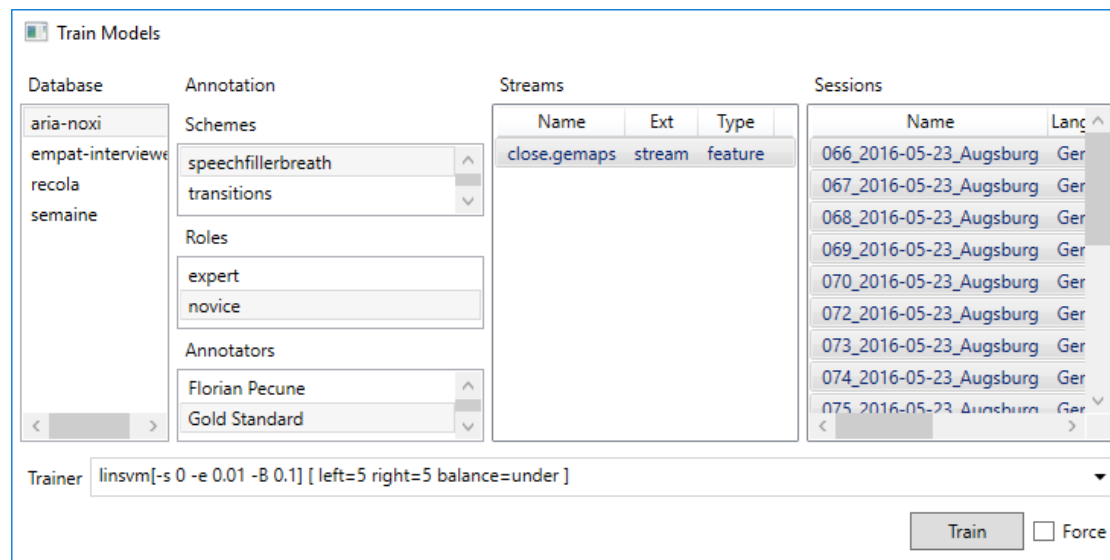
11

Figure 14: *Model training dialogue.*

Sessions for which an annotation exists that satisfies the selection are now shown and can be selected. Finally, choose the stream for which a machine learning model should be trained. Note that only those streams are listed for which at least one trainer template is found. By pressing 'Train' model training is started. Switch on 'Force' if you want to override an existing model.

5.5 Annotation Prediction

With a trained model it is now possible to predict annotations for one or more sessions. Therefore choose 'Prediction Annotations' from the LEARNING menu. Once you have selected a scheme, a role and an Annotator those sessions will be listed, which do not have a matching annotation yet. You can turn of the filter by selecting 'Show existing', however, make sure you are not occidentally overriding existing annotations (no undo available!). Finally, select the stream for which a machine learning model has been trained (only streams are listed for which at least one model is found). Depending on the scheme, there may be additional options that allow you to tune the generation of the prediction. E.g. in case of a discrete scheme you can automatically fill small gaps or remove segments below a threshold. After pressing 'Predict' annotations will be created based on the chosen model for all sessions that have been selected and stored under your name (if you are an administrator you may also generate annotations for another user).

Figure 15: *Annotation prediction dialogue.*