

MUSIC ANNOTATION INTERFACE

ANNOTEMUS USERS GUIDE

VERSION 0.7.2

**BY OLIVIER LARTILLOT
& ANDERS ELOWSSON**

MIRAGE PROJECT, RITMO, UNIVERSITY OF OSLO

MUSIC ANNOTATION INTERFACE

ANNOTEMUS USER'S GUIDE

VERSION 0.7.2

BY OLIVIER LARTILLOT

& ANDERS ELOWSSON

MIRAGE PROJECT, RITMO, UNIVERSITY OF OSLO

© 2025 by Olivier Lartillot and Anders Elowsson. This work is licensed under a Creative Commons Attribution 4.0 International License

Annotemus is a software developed in the context of the MIRAGE project, funded by The Research Council of Norway under the program IKTPLUSS, grant number 287152.

RITMO Centre for Interdisciplinary Studies in Rhythm, Time and Motion is funded by the Research Council of Norway through its Centers of Excellence scheme, project number 262762.

January 22, 2025
Olivier Lartillot, Anders Elowsson
University of Oslo

ANNOTEMUS

Annotemus is a graphical interface that facilitates the annotation of note positions in audio recordings. The interface has been designed to allow anyone, without any needed expertise in computer science or signal processing, to perform the annotation tasks with minimum time and effort. A time-frequency representations of the audio is displayed to visually guide the annotation. Annotemus provides a wide range of tools to facilitate music annotation with high time resolution.

1.

GENERAL PRESENTATION OF ANNOTEMUS

Annotemus offers a comfortable user experience to facilitate the addition and edition of notes, as well as navigating within the recordings, with relative ease. Individual notes or series of notes can be played back in many different ways. These multiple playback methods offer alternative approaches to facilitate a precise localisation of pitch as well as onset and offset positions. This include: simultaneous sonification of the annotations, timing playback to the annotations so that the start and end point of the playback can be used as perceptual cues, slowing down playback, soloing or muting individual pitches within a superposition of pitches.

Annotemus is available as a standalone compiled application, programmed in the MATLAB environment and available for Windows, MacOS and Linux.

The software has been designed in the context of the MIRAGE project, aimed at developing tools to automate the annotation of Norwegian folk music played on Hardanger fiddle.

2.

INSTALLATION

Download the zip file **Annotemus_Installer.zip** and extract it. You obtain a folder containing an app called **Annotemus_Installer**.

Launch the app:

- * On MacOS, double click the file **Annotemus_Installer.app**
- * On Windows, double click the file **Annotemus_Installer.exe**
- * On Linux, in the Terminal, type:
sudo -H ./Annotemus_Installer

Choose the installation folder for the application. Click **Next**.

If MATLAB® Runtime is not already installed on your machine, choose the installation folder for the MATLAB Runtime libraries and click **Next**.

Select **Yes** to accept the terms of the MATLAB Runtime license agreement and click **Next**.

Click **Install** > to begin the installation.

Click **Finish** to exit the installer.

If you accepted the default settings, you can find the installed application in one of the following locations:

- * On MacOS: **/Applications/Annotemus**
- * On Windows: **C:\Program Files\Annotemus**
- * On Linux: **/usr/Annotemus**

3.

WALK-THROUGH

Once you have launched the application Annotemus, after waiting for a little while, you are asked to **select an audio file**. (For instance *filename.wav*)

Then a new window appears where is displayed a 2D graphical representation of the audio recording you have selected, with time on the X-axis and frequency on the Y-axis. This representation will be called hereafter **Time-Frequency Representation (TFR)**. Cf. Section 5 for more details about the TFR.

NAVIGATION

To **zoom** in a particular time-frequency region, **draw** with your mouse (or trackpad) while pressing the **Ctrl** key at the same time. When drawing, you will create a rectangle which will define the zoomed region. Once you finish drawing, the zoom is carried out.

If you Ctrl-click **without displacing the pointer**, a full **zoom-out** is triggered, revealing the full TFR along the complete time and frequency range.

You can **move forward or backward in time** using the **left and right** arrow keys and to a **higher or lower frequency** region using the **up and down** arrow keys.

It is also possible to **zoom in and out along one axis** using arrow keys. **Shift** and **left or right** key displaces the left border accordingly and the right border in the opposite way; **shift** and **up or down** key displaces the top border accordingly and the lower border in the opposite way.

LOADING A TRANSCRIPTION

Press '**L**' to select a transcription file to **load**. The file can be a **MIDI** file, or a CSV file (exported from Annotemus, cf. below).

ANNOTATION OF NOTES

This first version of Annotemus solely focuses on the annotation of unquantized notes, represented simply as lines. **Drawing a note-line** consists of selecting one extremity (onset or offset with a specific pitch), holding down the click button and moving the pointer, and finally releasing the click button once the pointer reaches the other extremity of the line. A horizontality constraint is used: if the pointer moves up or down while drawing, the first extremity is moved along vertically, so that the resulting **line remains horizontal**.

Note-lines can be modified. But to avoid accidental modifications, they need to be selected first. A line is selected by clicking directly on the line or at its immediate vicinity.

Selected lines are shown in a different color (yellow, instead of the default red color for unselected lines).

To **select multiple lines**, select the lines one after the other while pressing the **Shift** key. All selected lines are deselected by clicking somewhere away from any line, or by selecting another line without the Shift key.

To avoid accidental modification of lines, a line that has just been drawn (or modified through drawing) becomes unselected. Further modification of the line requires an explicit selection of the line. A line can be **modified** only by **selecting one of its endpoint and moving it**. The other endpoint remains static (unless the horizontality constraint forces it to move vertically, as explained above).

It is possible to **automatically set the offset** of one note as the **onset of the next one**, by selecting the two notes and pressing 'H'.

Lines can be moved upward and downward with an increment of 5 cents using **up and down** arrows with **Control and Shift** keys pressed.

Each endpoint of a selected line can also be **moved** (individually) using **left and right** arrow keys, with an increment of 50 ms. Finer shift of the start time (onset) and end time (offset), with an increment of 5 ms, is also possible. Besides, the audio related to that particular time region is automatically played at each step. This audio feedback enables to precisely pinpoint these onset and offset positions. Concerning the onset position, in particular, if the selected point is a little after the actual onset, we can hear in the audio playback that the attack phase is truncated at its beginning. By progressively moving the onset position with 5 ms increment, the attack phase is less and less truncated, and the onset time can be set at the position where the whole attack phase is perceived. Annotemus also features an extensive number of other playback options as described below.

Some notes may in some annotation contexts be **of less importance**. Such notes could correspond to little ornaments, to very short notes of low dynamics, of events considered as somewhat accidental, etc. For that purpose, the interface offers the possibility to flag some of the notes

using the 'u' key ('uncertain'). They appear differently in the graphical representation, with a dotted line instead of a solid line.

PLAYBACK

You can **play back** the temporal region currently displayed, by pressing '**p**'. While playing, a vertical play-head traverses the display from left to right.

If you have selected one or several notes, you can play the audio **from the onset** time of the first selected note **to the offset** time of the last selected note (press '**j**'). This enables you to listen to if the annotation spans the full extent of the note.

Only when selecting one single note: You can play **from 600 ms before the onset** position **up until** the exact time point of the annotated **onset** (press '**i**').

You can play **from 500 ms before** the (first) **onset** position **until** the (last) **offset** position, with a burst of **noise** indicating the **onset** position (press '**o**'). In this way, you can listen to the start of the note, with a bit of what happens before, and check whether the location of the burst of noise corresponds to what we would consider as the onset position.

You can play **from 500 ms before** the (first) onset position **until 1 s after** the (last) **offset** position, with bursts of **noise** indicating **both** the **onset and offset** positions (press '**f**'). Same as for the previous point, but this time checking both onset and offset times.

You can time-shift the audio so that it can be played back at a **slower speed**. Playback speed can be multiplied (by pressing '**r**') or divided ('**e**') by factors of 2, or in a **finer grain** (by pressing also **shift**) a factor of $2^{(1/4)}$, multiple times. The objective is to enable you to decipher complex and fast successions of notes and to track the exact timing with better accuracy.

In order to distinguish between simultaneous notes, you can **isolate** the **partials** of particular notes in the audio playback. One or several selected note annotations can be isolated, **leaving** out the harmonics related to other notes (by pressing **Shift** while pressing any of the 'j', 'l', 'o' or 'f' options). It is also possible to instead **remove** the partials of the selected note annotations, in order to better listen to the other simultaneous notes (**Control and Shift**).

You can **combine** the three mechanisms of slowdown, tone separation and temporal focus at the same time, the resulting toolbox enables to observe fine details of the audio recordings.

An alternative approach for combined visualisation/sonification (by pressing 'v') consists in scanning the whole audio recording from start to end, and following the graphical representation on a micro level so as to reveal details in the representation. This corresponds, to use a cinematographic term, to a **following shot**", tracking the playhead in "close-up", abscissa-wise, throughout the TFR. This "scene" can then be stored into a video file.

SONIFICATION

You can **synthesise the annotated lines** in the form of sinusoids, by pressing the **spacebar**. Only the annotated notes visible in the region of the TFR under consideration are played. If some annotated notes are selected, these solely will be sonified.

The **initial audio recording and the sonified annotation** can be **played synchronously in stereo** (by pressing 'b'), with the audio in the right channel and the synthesized waveforms in the left channel.

DATA MANAGEMENT

The annotations are **saved as CSV** files, storing for each note its onset and offset time, its frequency (expressed in logarithmic scale as a MIDI real number) as well as whether it is an important note or not (cf. ‘u’ command above).

These **CSV files can be loaded** back in a subsequent session (by pressing ‘l’ and selecting the file) and the corresponding lines are displayed back again.

As soon as a first edit is made in the annotation during a session (such as annotating a first note or modifying a note from a loaded annotation), the user is prompted to indicate a file and path name to the save file. Besides, after each series of 10 edits, the resulting annotation is saved again. Each time, the saved file has a distinct name, appending the date and time at the end of the file name. In that way, previously saved files are conserved for backup purposes.

4.

METRICAL ANNOTATION

You first need to load a note annotation (of format **.csv**), so that you can annotate each beat onset by selecting one of the corresponding annotated notes. To load a note annotation, press **L** and select the note annotation file.

When playing the sequence using **B**, the beat annotations are played with short noise. To avoid playing the beat annotations, press shift **B** instead.

Notation convention:

the first beat of bar 1 is written “1:1/3” (bar = 1, upper = 1, lower = 3)

the second beat is “1:2/3” (bar = 1, upper = 2, lower = 3)

the third beat “1:3/3” (bar = 1, upper = 3, lower = 3)

then the first beat of bar 2 is “2:1/3” (bar = 2, upper = 1, lower = 3)

To simplify, for the first beat, we can simply write “1:/” (bar = 1, upper left void, lower left void)

How to annotate beats:

For the very first beat, select a note that is on that beat, and press **M**. A new window appears in which you need to indicate the bar number, the upper number, and the lower number, as indicated above. Again if it is the first beat of a bar, just indicate the bar number and leave the other fields (upper and lower) empty.

You will then see on top of the green line a red text indicating the metrical position. If it is the first beat of bar 1, it just shows: “1:/”.

You then just select the note related to the next beat, and press **shift-M**. It will automatically shows: “1:2/3”, meaning bar 1, beat 2 out of 3. Next beat: “1:3/3”. Next: “2:/”. And so on, “2:2/3”, “2:3/3”, “3:/”, etc.

You can remove the beat annotation of any of those notes by pressing **command+shift+M**

If there is a silent beat, without any note related to it, you can go directly to the next beat, and press **shift-M** twice (once for the silent beat, and once again to increment the beat once more).

If for some reasons you need to reset the beat annotation to a particular value, press **M** without shift, like when you started annotating the first note. The subsequent **shift+M** will continue from that point.

By default, the time signature is 3/3, so that each bar is divided into 3 beats. But this can be changed:

- By pressing **2**, **shift+M** will increment by half a bar: “1:/”, “1:2/2”, “2:/”, etc.
- By pressing **1**, **shift+M** will increment by a whole bar: “1:/”, “2:/”, “3:/”, etc.
- By pressing **3**, **shift+M** returns to the default 1/3 increment.

How to compare your beat annotation with just another one:

Load the other beat annotation by pressing **command+shift+L** and select the corresponding file.

- The beat positions that are the same in both annotations are still shown with a red text.
- The beat positions that are different are shown in green (for your annotations) and in blue (for the annotations of the compared file).

When playing back using **P**, only beats in red and green (so basically your initial annotations before comparing) are sonified with short bursts of noise. By pressing **shift+P**, the other interpretation (red and blue) is sonified.

You can delete any of your (green) beat annotations by pressing **command+shift+M**, as before.

You don't need to delete the (blue) beat annotations of the compared file as they will be deleted by default (cf. below). But if you think one of your (green) annotation is incorrect, delete it, select the corresponding blue beat annotation and press **K** to confirm its correctness and turn it red.

Once you have checked everything, confirm all the modifications by pressing **shift+K**, press OK to confirm. All the blue annotations are deleted (once you have solved any conflicts concerning changed note positions between the two files, cf. below) and the remaining green annotations are turned red. Save your annotation by pressing **S**.

What happens if you changed note positions, or added/removed notes while annotating beats?

While notes usually appear with **green lines**, any new note that you added (by mistake or by choice) will appear either:

- **In red** if your note cannot be aligned with any note from the compared file: If it was a mistake, you can delete it; else you can turn it into a green note by pressing **K**.
- **In blue** if your note look similar to another note from the compared file, also shown in blue. You can delete anyone of those, and the other one will turn green (else press **K** if it did not turn green.) Make sure that you have not deleted the beat annotation, else add it again manually to the remaining note.

When you finish the beat comparison and press **shift+K**, you will be asked to checked any unsolved note conflict. The first note conflict found in the music will be automatically shown and selected, so that you can quickly delete it for instance. Then press **shift+K** again to continue.

NEW: How to compare multiple beat annotations (could be more than two):

Load the first beat annotation by pressing L as usual. The **beat** annotation is shown in **red**.

Load the second beat annotation by pressing shift+L. This **beat** annotation is shown in **green**. Besides, a new legend explains the color code, showing a red line with the file name of the first beat annotation, and a green line with the file of the second beat annotation.

In the case there is any difference between the note annotations of the first beat annotation and of the second beat annotation:

- If there are couple of notes, one from each beat annotation, that are very closed and similar, they are shown both in **blue** (same as in subsection above). In the same way as above, you can delete one of the line **(Please avoid using K at the moment, delete the other note instead.)**
- Else any additional note from any of the beat annotations is shown in **red**. They can be deleted. **(Please avoid using K at the moment, just leave the note red.)**

If there is any of those discrepancies between files, a message window pops up and shows three pieces of information:

- ‘Extra0’: followed by, if any, list of aligned pair (blue) of dissimilarities, showing one onset time for each of those pairs.
- ‘Extra1’: followed by, if any, list of additional note onset times from the first beat annotation.
- ‘Extra2’: followed by, if any, list of additional note onset times from the second beat annotation.

Press enter to close that message window.

Besides, below the figure, it is also indicated “Misaligned: x, Extra1: y, Extra 2: z”, with x the number of blue pairs, y the number of additional notes from first beat annotation, z for the second annotation.

So before resolving beat annotation discrepancies, please resolve all these Extra notes, by choosing one note for each blue pair, and by deleting unwanted red notes. You should see the count decrementing in the text below the figure (the x, y and z). So at the end, it should be “Misaligned: 0, Extra1: 0, Extra 2: 0” (although y and z could be higher than 0 if there are any red notes you want to keep). **Note however: if you have a conflict between two notes, but at the same time if beat annotation is associated to the two conflicting notes, better not deleting any of them, else you lose the corresponding beat annotation.**

You can then add a third beat annotation file, which will appear in **blue**. Press shift+L. Please note: the current note sequence is considered as resolved, and is considered as the new first sequence; the new loaded sequence is the new second note sequence. Now everything mentioned above is performed once again for the comparison of those two note sequences.

You can then add a fourth beat annotation, which will appear in **black**, and so on.

Now you can look at the disagreement related to the multiple beat annotations, with their respective colors (red, green, blue, black).

5.

LIST OF FUNCTIONALITIES (INCOMPLETE)

Zoom in.....Ctrl + click and draw
Full zoom out.....Ctrl + click
Local zoom out Shift + click
Shift selected region Arrow key
Slight zoom in/out.....Shift + arrow

Draw a new line Click and draw
Select an additional line Click near line
Deselect all lines Click outside of line
Modify a selected line Click endpoint and draw
Move onset of a selected line Command + up/down
Move offset of a selected line Command + Shift + arrow
Delete selected lines Backspace
Flag/unflag as uncertain.....U

Play audio in selected region P
Synth render annotated lines Spacebar
Play both audio and annotation B
Stop playing.....X
Louder synth rendering.....+
Softer synth rendering.....-
Slower playback.....T
Slightly slower playback.....Shift + T
Faster playback R

Slightly faster playback Shift + R
 Play audio covered by line(s) with sound indication of onset
 position O
 as well as offset position.....F
 Play audio before onset I
 Play audio during line only J
 Play only part of audio related to pitch corresponding to
 selected line(s) Shift + O, F, I or J
 Play only part of audio **un**related to pitch corresponding to
 selected line(s) Ctrl + Shift + O, F, I or J

 Export annotation to CSV file S
 Change export file name or path.....N
 Load annotation from CSV file L
 Export to video file.....V

6.

TECHNICAL DETAILS

The Time Frequency Representation (TFR) displayed once an audio file is computed using a constant-Q transform (CQT) spectrogram, based on the implementation by Schörkhuber et al. (2014) across the frequencies 37-5000 Hz. The resolution is set to 60 bins per octave and the parameter gamma, which controls the time/frequency trade-off in lower frequencies, is 7.

Sonification is performed with additive synthesis. Sinusoids are generated at each octave of the annotated pitch up to 8 kHz with a randomly selected phase. The magnitude of each sinusoid is set to correspond to the long-term average spectrum of popular music computed by \cite[elowsson2017long]. We use equation X from that paper, but take the square of each magnitude value to reduce higher frequencies, which more often are generated by percussion instruments \citep[elowsson2017long]. For our synthesis, an audible attack is generated by boosting the upper harmonics for the first 100 ms (or half the duration, whichever is shorter) of the tone, falling linearly down to the same magnitude of the steady part. The factor with which the magnitude is boosted in the attack increases linearly from 1 to 4 from the first partial to the highest partial. A fade-out is also applied to the final 10 ms of each synthesized note. The synthesized audio waveform (consisting of one or several notes depending on playback mode) is finally normalized to just below clipping level (+-1).

7.

RELEASE HISTORY

VERSION 0.1

August 5, 2021

VERSION 0.2

March 6, 2022

Merge beat annotations and solve conflicts.

VERSION 0.3

March 30, 2022

Modify the increment for the metrical annotation, adapting to time signatures different from 3/3.

VERSION 0.4

April 3, 2022

Improvement of beat annotation comparison between two annotation files, including resolving conflicts related to changed positions and addition/deletion of notes.

VERSION 0.5

April 4, 2022

Possibility to play different versions of beat annotations separately, when compared together.

VERSION 0.6

April 8, 2022

Beat sonification also for 2/2 time signature

Note and beat alignments are saved and reloadable.

VERSION 0.7.1

December 21, 2024

Improvement of beat alignments

Load MIDI files

Connect successive notes of same pitch: offset synched with onset.

Note synthesis was not working with sampling rate different from 44100Hz.

Pitchogram is not taken into account anymore.

VERSION 0.7.2

January 22, 2025

Fixes a bug where notes previously deleted were reappearing again in yellow when reloading the CSV file.

All previous notes are deselected when a new note is created.

7.

DOCUMENTATION HISTORY

VERSION 0.1

August 5, 2021

VERSION 0.2

March 6, 2022

New section Metrical annotation.

VERSION 0.3

March 30, 2022

Adding metrical increment modification in Metrical annotation section.

VERSION 0.4

April 3, 2022

Adding note conflict resolution in Metrical annotation section.

VERSION 0.5

April 4, 2022

Playing with and without sonification of beats explained.

VERSION 0.6

April 8, 2022

Explaining the loading of the secondary file (TFR).

Update how to play different versions of beat annotation.

VERSION 0.7.1

December 21, 2024

Superposition of multiple beat annotations.

Pitchogram not taken into account anymore.

Waiting a while while launching the Annotemus app.

Loading a CSV or MIDI file.

Graying out some less important text.

Sync offset to next onset with 'h'

VERSION 0.7.2

January 22, 2025

Clarifies a bit how to deselect all notes.