SALAMI: Structural Analysis of Large Amounts of Music Information

Annotator's Guide

SALAMI in a nutshell:
Our goal is to provide an unprecedented number of structural analyses of pieces of music for future study. These studies may include training computer programs to automatically do structural analysis themselves, or tracing the evolution of form over centuries, or investigating which forms seem to dominate which genres, or things of that nature. In light of this, we are striving to cover a wide variety of musics, from western popular to Indian classical, including recordings in live and studio settings.

Structural Annotators:
Your job will be to generate the structural descriptions of pieces of music that will be used in this research, and you will strive for both accuracy and speed. Of course, analyzing the structure of a piece of music is hard: it requires skill and judgment, and it can't usually be said that there's a "right answer" for a particular piece. It's also often a very fuzzy process: the exact definition of "form" can be hard to pin down, and even musical processes that are relatively well defined (e.g., a modulation) can be very tricky to locate in the music. [Along these lines, please put quotation marks around any word in this pamphlet that you think is being abused.] Despite this, the analyses you produce will need to be very strictly laid out—in fact, they will need to be expressed in a machine-readable way.

Definition:
What do we mean by "formal analysis"? To start with, since many of you are music theorists, here are some examples of what we don't mean:
  • A classification of the piece into a formal type such as sonata, song, or canon.
  • A Schenkerian reduction of the piece into its Ursatz.
What we do mean can be roughly expressed as: "the organization and division of [the piece] into definite sections, and the relation of those sections to each other." (This is Salzer's definition of 'form' as distinct from 'structure' and 'design,' excerpted from the Oxford Dictionary of Music.) Put another way, we will have you partition each piece into several segments, and then give these segments appropriate labels to describe which are similar to each other or which fulfill a related musical role. This definitely overlaps with the first definition above, the "classification of the piece," except that for a given rondo we wouldn't want you to produce the answer "rondo" but the answer "ABACABA."

How to analyze music:
Before the procedure you will use for annotating music is outlined, it may help to consider a few different approaches one could take to do this kind of segmentation + labelling analysis.

1. One approach could be purely perceptual: you would listen for prominent harmonic or rhythmic boundaries in the piece in order to segment it, and apply labels to the resulting sections by comparing them and determining which were similar to each other. One problem with this approach is that it does not reflect the function of different segments (e.g., an 'introduction' section and a 'transition' section might have exactly the same music, but embody different structural roles).

2. Conversely, you could only pay attention to the functions, and divide the song into its constituent verses and choruses, along with intro, outro, and other sections. Of course, this would result in the opposite problem, where two differently-labelled sections might have the same music; or, where two sections which have the same function have very different musical ideas (for example, an extended two-part introduction).

3. Another option is to attempt to imagine the annotation as a transcription process, whereby musical parameters such as chord patterns and instrumentation are recorded directly. This type of annotation has the benefit of being less subject to subjectivity, but may not provide as much information about the structure of the piece as we would like.

The annotation method we have in mind (described in the next section) draws on aspects of all these three types of analysis listed above, but has the advantage of not conflating them. By separating the organization of instrumentation, musical material, and formal function, the method allows you to analyze a huge variety of pieces of music using a single, highly-constrained vocabulary, in a consistent manner.

Finally, it should be noted that none of the three methods mentioned above seem to do anything to address the fact that musical structure is frequently hierarchical. When analyzing a piece, it can be hard to know what timescale is appropriate: a single description of a piece such as "ABABCAB" might not reflect that each A is composed of two contrasting parts (= "ADBADBCADB"), or that the sequence "AB" has significance at a larger scale (= "DDCD"). Again, although this is an ambiguity that no system can completely resolve, the method described below includes a few markers that will partly address this.

SALAMI Annotation Labels

As stated before, the structural analysis we want of each piece will consist of a partitioning of that piece into sections, and the labelling of these sections. However, there will be three independent layers of labels:

      1. The level of musical similarity;

      2. The level of musical function;

      3. The level of instrumentation.

The labels to be used for each of these layers is described below.

1. Musical similarity:

A, B, C, D, E, ... : these indicate large-scale musical phrases, ideas, or subjects that may be differentiated on the basis of rhythmic, melodic, or harmonic material. The idea is that each particular musical idea gets its own label. We advise limiting your annotation to 5 labels, but if you truly require more letters you are permitted to use *F, G, H*, and so on. Note that every instant in the piece must be labelled with a letter.

Z is a special letter that is used to denote an amusical section that either stands out strongly from the rest of the piece or, more likely, should not even be considered part of the piece. For instance, applause at the beginning or end of a piece, or a brief spoken dialogue in the middle of a piece might be appropriately labelled with *Z*. Note that it in cases where a piece has two such inscrutable sections, they should both be labelled as *Z* even if they are not acoustically similar. *Z* is an exception among letter labels in this respect.

' : the prime symbol is commonly used in describing structure to indicate when a particular section occupies a gray zone between being a repetition of a previous musical idea, and being a new, independent musical idea. It could be called for, for instance, if a particular passage were repeated, retaining its musical identity but being transposed, or converted from the major to the minor mode.

It was previously mentioned that the structural organization of music is hierarchical in nature. Thus we ask you to annotate musical similarity on at least two scales. The large-scale sections indicated by the uppercase letters *A, B, C, ...* should also be divided into subsections using the lowercase letters *a, b, c, ...* (At the shorter time scale, there is no incentive to use only 5 labels; if necessary, the lowercase letters may continue through *y, z, aa, ab,* and so forth, although the need for this many labels is likely very rare.)

Note that all boundaries marked by an uppercase letter should also be marked by a lowercase letter. Note also that lowercase letters have significance across large-scale sections: the label "*a"* indicates the same musical idea even if used in both sections "*A*" and "*B*".

2. Musical function:

Depending on the piece under analysis, many words could potentially be used to describe the function of a particular segment. However, because we want to have a

consistent annotations, we restrict the function labels to a small vocabulary. The main vocabulary terms you will need are:

- *verse*: in a song, a section in which the tune remains the same, but the text changes with each repetition.
- *chorus* (aka refrain): in a song, a part which contrasts with the verse and which is repeated more strictly. Sometimes two distinct chorus-like sections are present in a single song; in this case, they should still both be labelled as *chorus,* since they will be distinguished by different letter labels.
- *bridge*: a secondary section which constrasts with the verse and chorus, often serving as a transition section.
- *intro*: a part that leads into the rest of the piece.
- *outro*: a part that initiates the end of a piece.
- *solo*: a part in which a single instrument or voice comes to the foreground.

Many other labels will be made available to you (they are described in an appendix to this document), but they mostly provide finer distinctions among the labels listed above (including several varieties of "bridge" or "transition").

Although cases where the above labels apply should be very rare, you may occasionally have an instrumental pop song where the terms *verse* and *chorus* feel like a stretch. In such cases, you may resort instead to more generic labels such as *main theme* and *secondary theme*. The terms *intro, outro,* and *transition* will still likely be applicable.

3. Leading instrumentation:

In polyphonic contexts, those segments where there exists a main melodic referent should be labelled with the appropriate instrument. Here the vocabulary is in principle unrestricted: simply label the segment with the instrument or voice that contains the melody.

For instance, in a rock song, you would label those segments featuring the lead vocal with *vocal*. In other parts of the same song, if the guitar takes the melody, label those sections *guitar* (or maybe *electric guitar*). There may be times when no instrument feels like a lead, and that is fine; equally, there may be instances where two instruments appear to lead, in which case simply mark both of them.

Label rules:

Each segment of the music may be tagged with several labels. However, each layer of labels behaves differently. The format may seem idiosyncratic at first, but it is designed to make annotating songs speedy and intuitive. The rules for each layer are:

Letter labels:	The entire piece must be fully labelled with both uppercase and lowercase letter labels. That is, no time-span should be unlabelled at any hierarchical level. The only exceptions are the special labels described below.
a) *Lowercase labels*:	Every boundary must be provided with a lowercase letter label.

b) *Uppercase labels*:   While the entire piece must be fully labelled with uppercase letter labels, not each section needs to be labelled individually: an uppercase label is assumed to persist until the next uppercase label.

Function labels: Not every section needs to have a function label. Each function label is assumed to persist until the next uppercase label, so if the uppercase label changes within a single functional section, the function label must be repeated.

Instrument labels: The extent of a leading instrument's presence is indicated explicitly using opening and ending tags. The ending tag should occur at the last segment with the leading instrument.

Special labels: Two special labels have framing functions for the song. They are *silence*, which should be used for silent portions at the beginning or end of a song, and *end*, which should mark the very end of the song. The *silence* labels may be replaced by *Z* if there is applause, for example, but the *end* label is mandatory.


Format rules:

Because the annotations need to be read by a machine later on, it is imperative that they be properly and consistently formatted. Small mistakes—a misplaced comma, a typo—cause real errors that need to be laboriously fixed by hand later on. Please study the following formatting rules, which are admittedly complex, while consulting the example annotation on the following page.

All labels: All labels must be comma separated.

Letter labels: Letter labels consist of one uppercase or one lowercase letter. It does not matter which letters you use, except that the letter "Z" is reserved for non-music sections.

Function labels: Function labels should be taken from the list given in the appendix. They should be spelled correctly, but capitalization is not important.

Instrument labels: Periods with a leading instrument should be demarcated with opening and closing instrument tags. Opening tags begin with an open parenthesis, e.g.: "(trumpet". Closing tags end with a close parenthesis: "trumpet)". If an instrument only leads for one section, a single open-close tag may be used, as in: "(trumpet)". A closing tag attached to one boundary indicates that the instrument leads until the following boundary.

Special labels: Sections marked as "silence" or "end" should not have any other labels associated with them.

Example

To give a sense of how this all works, here is an example annotation for the song "Think For Yourself," by the Beatles, which you can listen to here: <http://www.youtube.com/watch?v=yXGSBgr8sbg>.
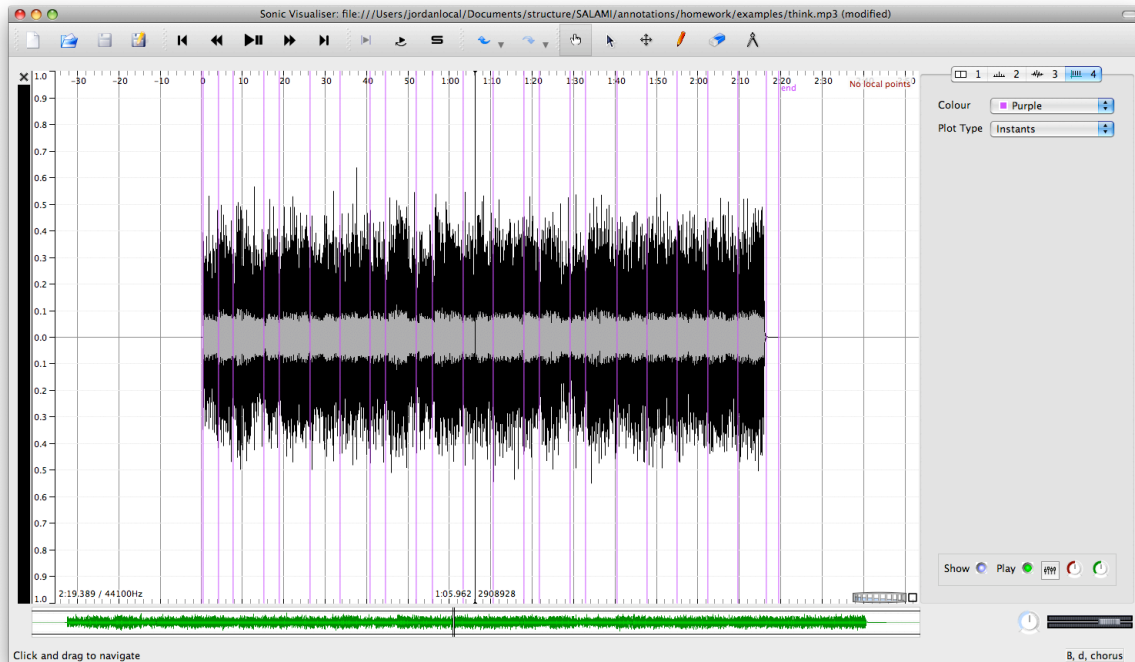
Above is a screenshot of the user interface for Sonic Visualizer, zoomed out so that the entire song is in view. The annotated boundaries are indicated as vertical purple lines, and the "end" label is visible at the end.

On the following page, you'll find the completed structure description for this song. The time of each structural boundary in seconds is given in the first column, the label in the second. The segment to which each label applies thus extends from the time given to the left of the label until to the time given on the following line.

Think For Yourself
by The Beatles

```
0.000000000    silence
0.429569160    C, c, intro
4.109931972    A, a, verse, (vocal
7.783401360    b
15.153628117   a
18.890045351   b'
26.284988662   B, d, chorus
33.523809523   e
40.857528344   A, a, verse
44.563242630   b
52.006893424   a
55.805986394   b'
63.245351473   B, d, chorus
70.530612244   e
77.929659863   A, a, verse
81.653061224   b
89.042721088   a
92.834535147   b'
100.327619047  B, d, chorus
107.647709750  e
114.956190476  B, d, chorus
122.258730158  e
129.544126984  e, vocal), outro
136.434648526  silence
139.334648526  end
```

Annotation Procedure

Annotations shall be produced using the Sonic Visualiser software developed at Queen Mary, University of London. You may download the program online (version 1.7.1 appears to be slightly more stable than 1.7.2) and read the documentation here: <http://sonicvisualiser.org/doc/reference/1.7/en/index.html>. Note in particular the use of the time instants layer (section 6.5) and the instructions on annotation by tapping (section 10). After having oriented yourself with Sonic Visualiser, you may find the following workflow efficient.

1. Load the song into Sonic Visualiser (SV), and press the semicolon key to make the first boundary at 0.00 seconds (a time instants layer is automatically created this way). Then skip to the end and mark another boundary.
2. Skip back to the beginning, press play, and mark a boundary wherever you perceive the section or subsection boundaries to be. You will probably want to at least mark a boundary every four measures or so, since that is the usual standard for the chord annotations. Try to anticipate where boundaries are so that you can press the key exactly when they occur; if you know you missed it by a small amount, pause the song and adjust the boundary.
3. Once all the boundaries are marked after one or two listenings, save your work!
4. Press the "E" key to "Edit Layer Data." A new window will open up that shows the segmentation as a spreadsheet. You will now need to fill the right column of this spreadsheet with the section labels. You can start by including the "silence," "end," and "Z" labels where needed.
5. Add both letter layers and the function and leading instrument layers. It is probably easiest to map out the piece with the lowercase letters first, one of which will be assigned to each boundary; to add the uppercase letters and functions in the same step second; and then thirdly to add the leading instrument tags. However, no order is mandatory.
   a. Tip: play the song while looking at the spreadsheet, and use the "tab" key to move the cursor to the rightmost column; if you type fast enough, you can add the annotations while the song is playing.
6. Review your work, and check that your spelling and formatting are correct, and that all the of the rules for each label have been followed. You can give the song a quick re-listen by pressing play and skipping through the sections with the "page-down" key. Alternatively, you can speed up the song significantly (under the "Playback" menu) and play the full song.
7. When you're satisfied with your work, select "Export Annotation Layer" from the menu and save your annotation in .txt format.
8. Upload the text file to the SALAMI website, including your estimate of how long the annotation process took for that song.

Appendix: SALAMI function label list

This appendix provides the entire list of recommended function labels. All of the acceptable function labels are listed in boldface. Following the list, they have been placed into groups of similar functions, such as "transition functions" and "ending labels," with some subtle differences pointed out. While the list relies heavily on popular music terminology, a handful of terms specific to other genres are included.

bridge	**chorus**	**coda**	**end**	**fadeout**
instrumental	**interlude**	**intro**	**main theme**	**outro**
pre-chorus	**pre-verse**	**silence**	**solo**	**(secondary) theme**
transition	**verse**			

Other acceptable function labels:
> jazz: **head**
> classical: **exposition, development, recapitulation**

Basic functions:
   intro (or **introduction**)
   verse
   chorus
   bridge
   outro

Transition functions: the following all indicate intermediary material of some kind. A "pre-verse" may use the same musical material as the verse, and may sound like a vamp (i.e., less transitional than stalling). A "pre-chorus" is that sometimes hard-to-delineate section where you can't decide whether it's the end of the verse or the beginning of the chorus. An "interlude" connotes a pause or break from the regular flow of the music, and it encompasses the terms "break" and "suspension". The last term, "transition", can denote all other intermediate sections that seem designed to lead from one section to another. Note that a "bridge" also has a somewhat transitional nature, but will stand out more as an independent, stand-alone section than any of the terms below.
   pre-verse
   pre-chorus
   interlude
   transition (or **trans**)

Instrumental functions:
The following two functions indicate instrumental breaks in the song. The "solo" label indicates that in that break, an instrument has come to the foreground to deliver a solo, as in a cadenza. On the other hand, "instrumental" suggests that no instrument is foregrounded.

instrumental
solo

Ending functions:
"Outro" will be our generic "conclusion" label, encompassing "ritornello," "closing," and most other concluding section types. By contrast, we reserve "coda" to indicate material that in some sense comes after-the-ending. "Fadeout" is a special term that can be used alone or *in addition* to another function label and refers to the artificial fading out of a recording.
coda
outro
fadeout (or **fade-out**)

Alternative labels:
In instrumental pop, prog rock, or other genres of popular music, using the words "verse" and "bridge" may seem ambiguous or contrived. In such cases, you may want to rely on the following labels which evoke classical terminology.
main theme
theme (or **secondary theme**)
transition

Genre-specific labels:
While the terms "exposition" and "recapitulation" are quite narrow terms mainly used with reference to sonata form, the term "development" has a broader applicability that may be useful in other types of pieces. In jazz, the term "head" may be used as a synonym for "main theme" or "chorus."
exposition
development
recapitulation
head

Special labels:
Each song should begin with a "silence" tag and end with an "end" tag, both marking the extreme ends of the audio file. If the music begins with less than 10 milliseconds of silence, *still label the region* as silence! If there is absolutely no silence gap at the beginning, then you may omit the label. The end tag is mandatory for all songs.
silence
end