

Guide and Instructions - Annotation

Overview

We aim to compile a dataset of monophonic recurring patterns that is 1) diverse in genre and instrumentation and 2) contains fine grained annotations for categorized music transformation. For this purpose, we plan to annotate about 50 pieces of works spanning from the early renaissance to post-Romantic period covering a variety of genres, especially works and genres that are known to exercise recurring patterns. For each piece, we provide a list of distinct patterns as the reference for transformation labeling. This involves more or less arbitrary choices in terms of lengths, boundaries, and thematic significance. In general, we are more interested in patterns that have diverse appearances and intricate yet recognizable transformations. Collaborators are encouraged but not required to look for additional patterns while they annotate for occurrences of the patterns we provide.

Annotation Procedure

Please annotate one piece (excerpt) at a time.

For each piece, we provide:

1. The PDF score
2. A list of reference patterns, boxed in the provided PDF score.
3. (For some pieces) Supplement materials, including score following YouTube links, analysis articles, wikipedia page, etc.

We ask the annotator to:

1. Annotate (e.g. highlight) sequences that appear similar to each reference. Use different colors to differentiate patterns. Only annotate occurrences that stay within the staff (no voice crossing) and are monophonic. If an occurrence involves chords, please only annotate the specific voice that is in the pattern to disambiguate.
2. Label each annotated sequence according to the transformation table we provide (see page 2). For multiple labels, separate with commas.
3. (Optional) When appropriate, identify new patterns, label the same way as those we provided, and repeat annotation and labeling for the new patterns. See addendum for our criteria for pattern.
4. Submit the annotated PDF score to the shared folder with the format: [original_score_name]_annotated.pdf

Transformation Table

The table serves as the guide for labeling the types of transformations.

Type	Transformation	Label	Definition	Definition notes
Pitch				
	Same Pitch (and transpositions)	(No label)	Pitches form the same diatonic intervals as the reference, including tonal and parallel transpositions	
	Pitch Alteration	PA	Certain pitches form different intervalic relations than the reference (but still recognizably related to the reference)	Including octave displacement, e.g. M2 up becomes down m7 down
	Inversion	I	Inverting all diatonic intervalic relations of the reference	
	Retrograde	PR	Reversing the order of pitch intervals of the reference	
	No Pitch Relation	NP	No obvious pitch relation to the reference	Annotate only if some rhythmic relation is shown
Rhythm				
	Same Rhythm	(No label)	Same rhythm as the reference	
	Rhythmic Transposition	RT	Proportional change of durations	i.e., augmentation or diminution
	Rhythmic Alteration	RA	Certain rhythms form different temporal relations than the reference (but still recognizably related to the reference)	
	No Rhythm Relation	NR	No obvious rhythmic relation to the reference	Annotate only if some pitch relation is shown
Rhythm+Pitch				
	Fragmentation	F	Partial occurrence of the reference	
-	Ambiguous	?	Uncertain of annotations	See notes for ambiguity
-	Comments	{comment text}	Wrap comments in curly brackets whenever needed.	This is more for extracting comment during digitization.

Criteria for Patterns

This section will come handy when deciding whether to annotate new patterns. Here are the rough guideline of how we define patterns and their reference form:

1. A pattern should, in general, contain at least 3 notes, with its exact form and/or strongly-related transformations appear at least twice in disjoint passages in a piece (i.e., not just a local reiteration). The strongly related transformations are: **Same Pitch/Rhythm, RT, I.**
2. In general, the first occurrence of a pattern is identified as the reference, unless there is an altered form (PA, RA, F, PR) showing up later that appears much more frequently with its strongly related transformations.
3. When a fragment of the original pattern keep reappearing in disjoint passages of a piece, we suggest labeling the fragment as a new pattern. (See Beethoven SQ example). If a fragment is labeled as a new pattern, there's no need to label those occurrences as F for the original pattern any more.

Notes for Ambiguity and Edge Cases

We anticipate there would be ambiguous and edge cases. Here are the general strategies to handle: (See Beethoven SQ example for use cases)

Comments:

If there is anything particular about a annotation (no matter if it is ambiguous), you could wrap a comment with { and }. A comment should also be entered in the digitized annotation.

Ambiguous Label (?) :

Whenever an annotation appears ambiguous, add the "?" label in addition to the other labels that you decide on with your best judgement. Include a comment to explain why the annotation is ambiguous if you may.

Still unsure:

Send an email.

Here are some specific handling cases:

Grace notes: ignore those. (Ok to include in annotation/digitization. They will get ignored in post-processing)

Sequence: Do annotate for existing patterns, but do not count sequential repetition as a new pattern (because they are too local and too similar (i.e., not interesting)

Tied notes: we count duration from the note onset, so don't annotate from a tied-over note, or stop at a note that is still tied further, even if it means changing the rhythmic notation from exact rhythm (no label) to RA

Enharmonic spelling: some scores might not have correct enharmonic spelling, and sometimes those can be compromised for sight reading too, especially for chromatic passages. Therefore, if reasonable respelling can make an occurrence have the same diatonic intervals as the reference, assume they have the same pitch (instead of PA).

Recommended Workflow

Despite the separated tasks of annotation and digitization, we recommend a workflow of annotate-> digitize & on-site fixing manual annotation if needed -> create/submit annotated pdf. Thus digitization can actually serve as a second-check to ensure annotation correctness and consistency.