Overview

The Mixin project aims to develop an artificially intelligent DJ system capable of creating seamless transitions between tracks in a user's playlist. Achieving this vision requires precise and consistent labeling of musical structures within tracks. This guide provides detailed methodology and criteria for the data annotation process.

Definition of Chorus

A chorus is a distinctive section of a song that showcases the primary, recurring theme or element, and significantly contributes to the song's identity, memorability, and emotional impact. Acknowledging the diversity of musical genres, some of which may not adhere to conventional verse-chorus structures, we broadly define the chorus as a core thematic segment. The identification of choruses and core thematic segments is based on their:

Distinctiveness: The segment's unique auditory qualities that distinguish it from the rest of the composition.

Thematic Representation: The segment's ability to encapsulate and convey the principal emotional or thematic message of the song.

Structural Alignment: The segment's strategic location within the song's overall structural and progressive framework.

Criteria for Chorus Labeling

The identification of a chorus is highly dependent on the context of each song. Accurate annotation is contingent upon the presence of contextual hallmarks that may precede, coincide with, or follow the chorus, including:

Intensity Shift: Significant variations in volume, intensity, or energy that underscore the segment's prominence.

Complexity Shift: Transitions in the song's complexity, either from complex to simple or vice versa.

Feature Inflection Point: Noticeable changes in vocal, tonal, or percussive features that demarcate a segment's boundaries.

Dynamic Effects: Utilization of strategic silences, crescendos, decrescendos, risers, sweeps, and similar effects to build anticipation leading into the chorus and to enhance its impact.

Additive Layering: Progressive incorporation of sound layers, instruments, or timbral elements as a lead-in to the chorus.

Pattern Repetition: The recurrence or variation of vocal, tonal, or percussive patterns.

Annotation Process

Data Preprocessing

Format and Sampling Rate: Songs should be uniformly formatted (e.g., '.mp3') and processed at a consistent sampling rate (e.g., 44100 Hz).

Genre Diversity: While the current dataset predominantly features electronically produced genres, future efforts should aim to broaden the project's scope by encompassing a wider variety of musical genres.

Audio File Preparation: Trimming of leading and trailing silences, with an option to include metadata relevant to the artist, track, and genre, alongside audio features such as tempo, key signature, and time signature, extracted via Spotify's API.

Processed audio files are to be assigned a unique ID, with their metadata cataloged in a .csv file. Detailed instructions and code for these preprocessing steps are available in this [notebook](#).

Tools and Software

For the purpose of labeling choruses, we utilized Serato DJ for its advanced audio processing features, including playback with spectrogram visualizations and beat/bar quantization. These features are instrumental in making informed decisions based on the structural context of a song, though other DJ software offering similar capabilities may also be suitable.

Guidelines for Labeling

Label Format
Start and end labels for a chorus are to be documented in the format (hh:mm:ss.millisecond), with milliseconds rounded to one decimal place.

Down-beat Alignment
Each chorus annotation must begin on a down-beat—the first beat of a measure in the song's time signature. This ensures that the chorus is contextually positioned within the song's rhythmic structure. The end label should also follow the last beat of the last measure to include the entire chorus section fully.

Heuristic Considerations
The determination of whether a subsegment constitutes part of the chorus relies on personal judgment, guided by the criteria listed above and specific heuristics regarding down-beats, structural patterns, and instrumental effects. Below are some heuristics used during the labeling process:

- **Invisible Down-beats:** Sometimes, a chorus may start or end on an "invisible" down-beat, where the musical phrase begins just before or extends slightly beyond the precise down-beat. Annotators should use their discretion to decide if an "invisible" down-beat aligns with the structural intent of the chorus.

- **Pattern Symmetry and Looping:** When uncertain about the inclusion of a segment as part of the chorus, consider the symmetry of musical patterns. Choruses typically exhibit a symmetrical structure, often in multiples of four bars. If including a segment disrupts this symmetry (e.g., causing an 8-bar chorus to extend to an asymmetrical 9 bars), it likely does not belong to the chorus. Generally speaking, an entire chorus segment should be able to loop by itself.

- **Recurring Motifs:** Choruses that feature specific patterns or motifs should be annotated from the beginning of the first down-beat of the pattern/motif, extending to the end of the last bar, even if the pattern commences slightly before the down-beat or extends beyond the last bar.

- **Tension-building Effects:** Similar to patterns, effects are evaluated based on their contribution to the chorus's overall form and function. Effects such as risers or sweeps that lead into a chorus are generally not considered part of a chorus. A chorus can start or end on an "invisible" down-beat.

Handling Ambiguities
Tracks where distinct choruses or other segments cannot be clearly identified due to ambiguity in musical structure or thematic elements should be noted and skipped to maintain annotation quality.

While most songs will feature at least two choruses, annotators may encounter tracks with only one or possibly three or more distinct choruses. The number of annotated choruses should reflect the track's structural composition, however, songs with more than three choruses should be skipped and their chorus segments re-evaluated.