

CAMAT: Computer Assisted Music Analysis Toolkit

Egor Poliakov¹ and Christon R. Nadar²

¹HMT Leipzig, Germany, egor.poliakov@hmt-leipzig.de

²Semantic Music Technologies, Fraunhofer IDMT, Ilmenau, Germany

Abstract— We introduce CAMAT, a python-based sheet-music parsing and analysis tool based on the auditory model. The toolkit aims to provide computer-assisted analytical methods for musicological research, especially for statistical investigations on larger databases with an educational focus.

Index Terms— sheet music, analysis, python, pandas, auditory model

I. INTRODUCTION

Despite a wide integration of music software in music performance and production, there is still a noticeable gap in the adoption of software tools in various fields of music education. During the one-year scientific and educational project "Computer-assisted Music Analysis" at Hochschule für Musik FRANZ LISZT Weimar (HfM) Weimar, CAMAT was developed, tested, and evaluated several flexibly applicable teaching modules based entirely on open-source software with a goal of providing powerful tools for integration in ongoing musicology and music theory courses. The teaching modules are dedicated, among other things, to the computer-based annotation and visualization of sheet music texts and audio files, the statistical analysis of music corpora, the search for musical patterns (rhythms, melodies, harmony connections, etc.), and the comparison of interpretation. Most of the learning modules are designed as Jupyter Notebooks. In addition, they are available online on a wiki-based resource which also includes a large data bank (over 4800 entries) of MusicXML scores¹.

II. CAMAT

CAMAT (Computer Assisted Music Analysis Tool) was developed as a dedicated tool for parsing and analyzing MusicXML scores. Although there exist various tools like humdrum² [1] and music21³ [2], CAMAT tries to solve some particular design problems that appeared during the conception and realization of certain learning modules. Here is a quick overview of the problem fields we encountered during this process described in the following sections.

II.1. HANDLING OF TIED CHORDS

Analysis of choral and piano music was one of the main topics of developed teaching modules. Therefore, we had to deal with scores that consisted of many tied notes (which is very usual in any polyphonic music). Because the MusicXML is by design a dedicated notation format which is built mainly for engraving purposes, a lot of crucial information about the actual duration of a note is handled as optional information (because it's not an engraved note, but simply an optional character that was added to previous note),

¹<https://analyse.hfm-weimar.de>

²<https://www.humdrum.org/Humdrum/>

³<https://web.mit.edu/music21/>

that in some cases cannot be easily extracted and allocated. That led to severe problems while trying to parse the exact duration of every single tied note and especially note groups in dense chord and polyphonic structures. In CAMAT, we reconsider the weight of ties and handle them on a top priority level as crucial information besides the pitch and duration. This decision also led to the idea of storing all the parsed information in pandas data-frames to preserve the exact duration of every tied note in a persistent rhythmical grid.

II.2. PARSING OF POLYRHYTHM, POLYMETRIC AND UPBEAT STRUCTURES

The tool can parse polyrhythm, polymetric, and upbeat structures while maintaining the unique metric profiling values. Because of the persistent rhythmical grid structure inside the pandas data-frame, we could now correctly parse and synchronize even very complex polyrhythm and polymetric structures while maintaining the unique metric profiling for every given part even if different time signatures are defined. Parsing of various metric positions of upbeats or repeat signs could also be fully integrated. Overall, the way we designed the storage of notation-based data within the pandas data frame, which includes combining and synchronizing all events along a fixed timeline, is very similar to an auditory model, cause the exact pitch-duration information of the actual perceivable musical events can be extracted.

II.3. APPROPRIATE INFRASTRUCTURE FOR NOTE SHEET BASED CORPUS ANALYSIS

Due to the scalability of pandas dataframes we found an easy solution to parse and store the information of multiple scores, that led to efficient integration of corpus based analysis in the learning modules.

III. CONCLUSION

The goal of the CAMAT is to provide a computer-based analysis tool for music analysis. CAMAT also provides parsing, visualizing musical texts, statistical analysis of music corpora, and searching for musical patterns such as melodies and rhythms.

IV. ACKNOWLEDGMENTS

The project Computergestützte Musikanalyse in der digitalen Hochschullehre (computer-aided music analysis within digital higher education) is funded by the Thuringian Ministry for Economy, Science and Digital Change and by Deutscher Stifterverband.

V. REFERENCES

- [1] D. B. Huron, *The humdrum toolkit: Reference manual*. Center for Computer Assisted Research in the Humanities, 1994.
- [2] M. Cuthbert and C. Ariza, "Music21: A toolkit for computer-aided musicology and symbolic music data," in *ISMIR*, 2010.