Calliope: A Co-creative Interface for Multi-Track Music Generation

Renaud Bougueng Tchemeube*
J. Ens
P. Pasquier
rbouguen@sfu.ca
jeff_ens@sfu.ca
pasquier@sfu.ca
Simon Fraser University
Surrey, British Columbia, Canada

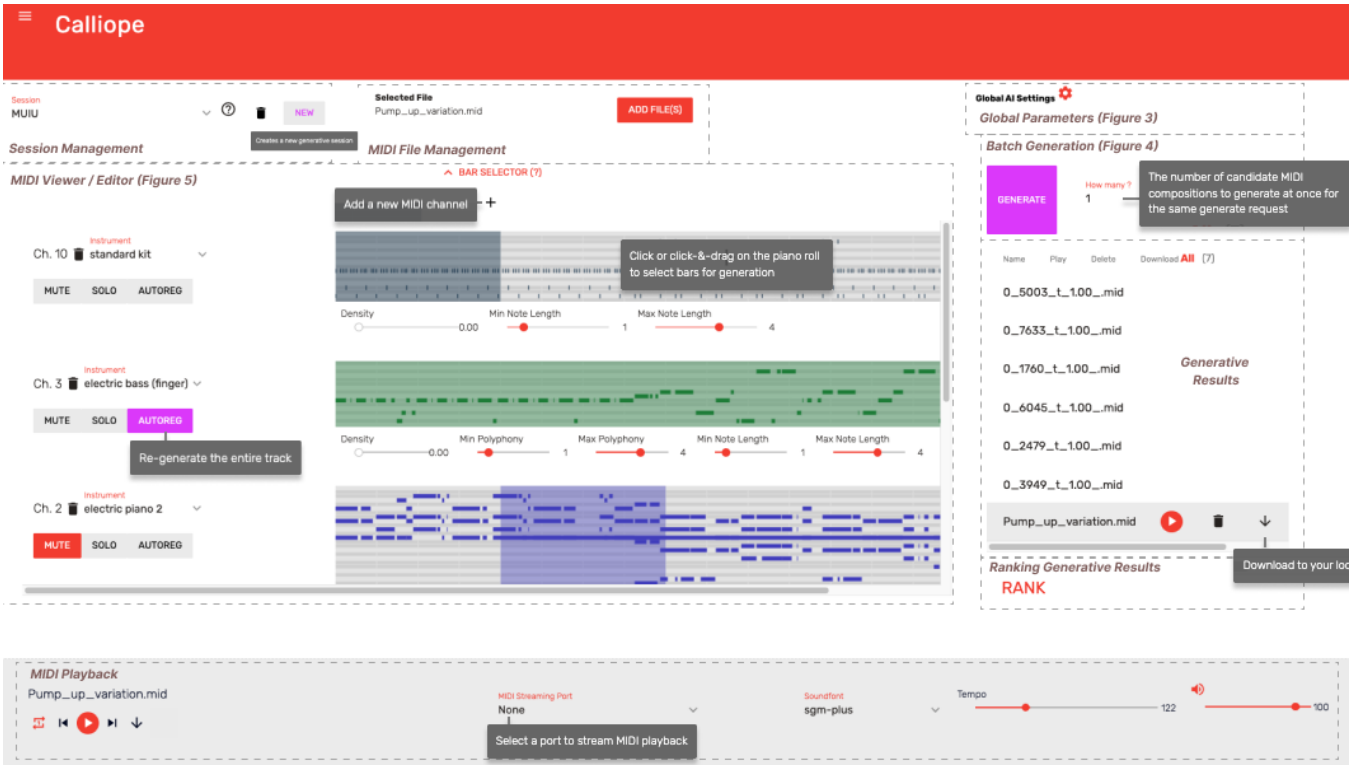

Figure 1: Calliope's Annotated Interface

ABSTRACT

Calliope is a web application for co-creative multi-track music composition (MMM) in the symbolic domain. It is built to facilitate the use of multi-track music machine (MMM). The user can upload Musical Instrument Digital Interface (MIDI) files, visualize and edit MIDI tracks, and generate partial (via bar in-filling) or complete multi-track content using the Multi-Track Music Machine (MMM). Generation of new MIDI excerpts can be done in batch and can be combined with active playback listening for an enhanced Computer-assisted Composition (CAC) workflow. The user can export generated materials as MIDI files or directly stream MIDI playback from the system to their favorite Digital Audio Workstation (DAW). Calliope can be used for creative ideation and for exploring alternatives of musical phrases in composition.

CCS CONCEPTS

• **Human-centered computing** → **Sound-based input / output**; *Web-based interaction*; • **Computing methodologies** → *Neural networks.*

*All authors contributed equally to this research.

C&C'22, June 20–23, 2023, Venice, Italy

ACM ISBN 978-1-4503-9327-0/22/06.
https://doi.org/10.1145/3527927.3535200

KEYWORDS

Computer-Assisted Composition, Generative Systems, Co-Creativity, Human-Computer Interaction, Deep Learning, Music Computing

ACM Reference Format:
Renaud Bougueng Tchemeube, J. Ens, and P. Pasquier. 2022. Calliope: A Co-creative Interface for Multi-Track Music Generation. In *Venice '22: The 14th ACM conference on Creativity & Cognition, 20-23, June 2022, Venice, Italy.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3527927.3535200

1 INTRODUCTION

Musical Metacreation (MuMe) [14] is a field of research which addresses the partial or complete automation of creative musical tasks including composition, interpretation, improvisation, accompaniment or mixing. It investigates purely generative systems for music as well as interactive ones. Computer-assisted Composition (CAC) is a subfield that focuses on developing systems for automating music composition processes. Some of the composition tasks include: multi-instrument pattern generation, multi-instrument complete generation, rhythm generation, harmonization, chord progression generation or melody generation. Many machine learning-based (ML) systems have been developed for CAC: OpenMusic [5], Style Machine [4], Magenta Studio [15], Manuscore [12], Morpheus [9]; demo systems such as Sornting [17], DrumVAE [17], DeepDrum [11] and commercial systems such as FlowMachines [13], AIVA [1], Spliqs [3], and Melody Sauce [2]. Typically, the user interacts with these systems by acting on a limited number of accessible parameters, making generative requests and listening to results before repeating the process. This provides less fine-grained user control, less overall possibilities for creative music composition in the concrete musical practice, and adversely impact user's sense of authorship, trust, or self-efficacy. Ryan et al. [8] shows that "AI can overwhelm users with the amount of musical content it generates, and frustrate them with its non-deterministic output". Calliope differs by interfacing a unique style-agnostic generative model (MMM) and offering an interactive composition workflow that is built to integrate with the user's musical practice. It takes an iterative, steering approach with bar infilling, similar to the Cococo system [9], on both track and bar levels, but offers richer attribute-controls such as instrument selection, note duration range and still handles single- and multi-track partial or complete generation. Its contributions are an accessible web interface for effectively controlling the MMM model, performing scalable batch generation and ranking, manipulating full MIDI pieces interoperably with the user's native environment (e.g., DAW) in the context of generative composition. Calliope is built for any user, novices and professionals alike. The project is accessible at https://metacreation.net/calliope.

2 SYSTEM DESCRIPTION

Inspired from the work on Apollo [16], an interactive web environment that makes corpus-based music algorithms available for training and generation via a convenient graphical interface, Calliope focuses on advanced MIDI manipulation in the browser and generative controllability of the Multi-Track Music Machine (MMM) [7]. A Transformer-based music model, for batch generation and ranking of partial or complete multi-track compositions. It is interoperable with the MMM pre-trained model via the Python runtime. The aim is to enable the ability for users to effectively co-create with a generative system. Calliope is built in Node.js, the Web stack (HTML, CSS, Javascript) and MongoDB.

2.1 User Workflow

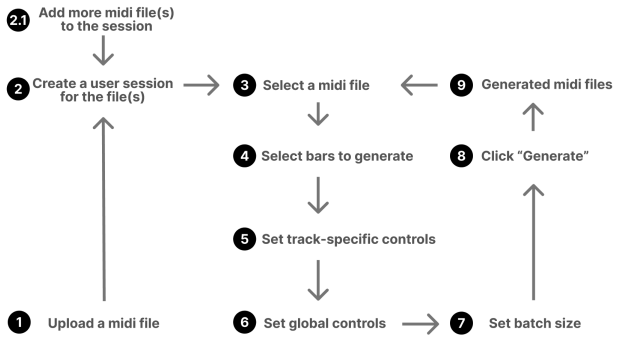

Figure 2: User Workflow

Figure 2 depicts the user workflow enabled by Calliope. The user creates a new user session and uploads a MIDI file to start with. The user can select and visualize the MIDI content of the file using the MIDI Viewer. They can also edit the tracks they are working with or add new empty MIDI track to the file. Then, for each MIDI track, the user can select a partial or complete set of bar sections they wish to generate content for. They adjust track-level controls for each MIDI track as well as MMM's global-level controls. They set the batch size n for generation and click on the "Generate" button. Calliope runs the request and returns a set of n newly generated MIDI files. The user selects each MIDI file and uses the MIDI player to audit the output. They skip through the content until they identify a file whose content they would like to further generate from. The process repeats from there.

2.2 MIDI File Management and Playback

The user can upload collection of MIDI files on the server (Figure 1). Data uploaded is stored into the user account set by the user prior to use. The user can use sessions to represent a conceptual space for exploring generation for a particular set of music files. The user can playback a selected MIDI file with features including adjusting tempo, soundfont, volume, and repeat. The user can also move to the previous or next file in the collection. Playback is an important step in the auditing of generated outputs which influences the user's next steps in the generative process. The user can also playback music while generating new sequences.

2.3 MIDI Viewer and Editor

MIDI notes from any uploaded MIDI file can be visualized in a piano roll format (Figure 5). Metadata info such as the MIDI channel number and assigned MIDI instrument can also be viewed and edited. It is also possible to delete or add MIDI tracks to an existing MIDI file. The MIDI player supports the General MIDI (GM) standard for MIDI playback and the capacity to select from a list of soundfonts.

2.4 MMM's Generative Request

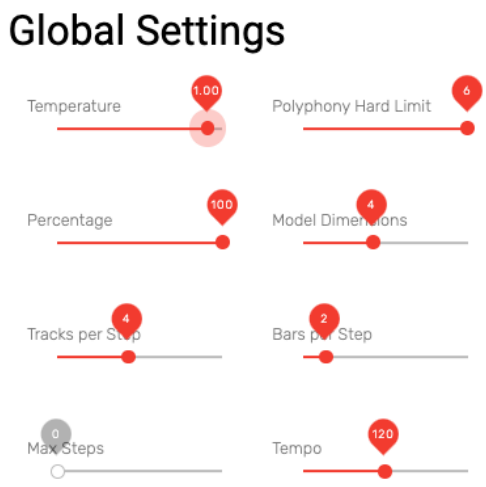

Figure 3: Global Parameters

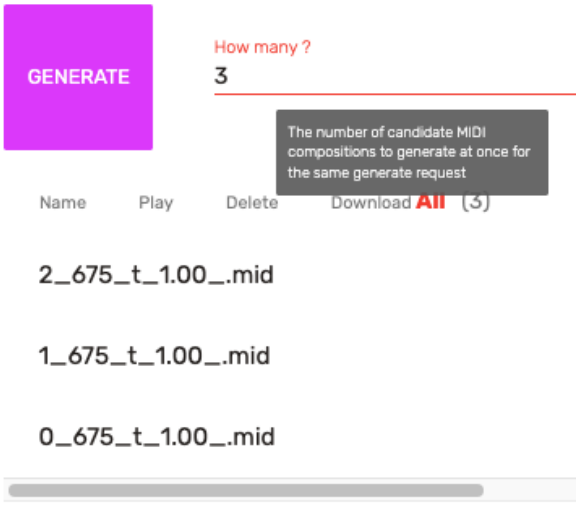

Figure 4: Batch Generation

MMM is trained on the MetaMIDI Dataset (MMD) [8], a half a million MIDI dataset containing various musical styles (pop, folk, electronic, classical, rock, hip-hop, etc) and instrumentation. Its power comes from [7] 1) its data representation which creates a time-ordered sequence of musical events for each track in the MIDI file instead of a single interleaved time-order sequence for all events, and 2) its integration of four different modes of generative usage: unconditioned, continuation, in-painting, and attribute-control. It thus combines a novel music representation approach and the attention-based capability of Transformer models to capture long-term dependency in the musical data. MMM offers both global-level deep learning parameters: temperature, model dimensions, polyphony hard limit, percentage, tracks per step, bar per step, max steps, tempo (Figure 3); and track-level music-based controls: instrument selection, note density, polyphony range, note duration range (Figure 5). Bar selection is used to refine the request for generation to a subset of bars (Figure 5). Generation for a subset of bars of a given track is constrained on musical information that precedes those bars and on musical information found in neighbouring tracks.

2.5 Batch Generation of Music Outputs

Batch generation of musical outputs can be done by setting a batch size value (Figure 4) and is implemented via the MMM's Python interface which offers batch computation natively. The ability to batch-generate means that the user can quickly explore alternatives, including generating from a previously generated output, for a given set of control parameters. Batch generation is a key feature for enabling rapid computer-assisted exploration of compositional alternatives. We have tested generations of 5, 10, 100, 500, 1000 music samples at a time. These generations can be done within 3 seconds to 10 minutes on an average computer depending on the total number of bars and note density of the music input. This process drives the interactive loop of music generation and playback listening for the user.

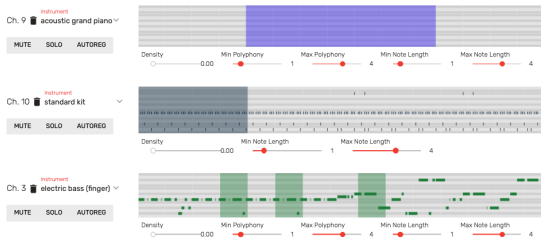

Figure 5: Bar selection for in-filling, Track-level parameters and a Newly created empty track (Ch. 9)

2.6 Ranking of Music Outputs

It is possible to rank a collection of generated MIDI files against a selected target file given their similarity. We employ a ranking algorithm which statistically quantifies the similarity of a generated output MIDI file against other MIDI files [6]. From an interaction point-of-view, it helps the composer explore the variability in similarity among MIDI files and effectively apply filter operations on the set of files. This is especially useful in the context of a large set of generated files (e.g., 50 files and up).

2.7 DAW Integration and MIDI Streaming

Calliope allows for two ways to integrate its workflow with a Digital Audio Workstation (DAW). The first option is to livestream MIDI playback to an input MIDI port on your DAW (e.g., Ableton) or any MIDI-enabled device. This provides a new workflow for the user to interface their native environment with a generative system. It can be achieved by selecting an available port on the "MIDI Streaming Port" option in Calliope's MIDI Player. This offers the ability to audit musical phrases generated by the model given the user's preferred instrumentation (e.g., proprietary virtual instruments or hardware synthesizers). The second option is to download the generated MIDI file and upload it into the DAW. This is for the user to continue the

composition work of generated materials (e.g., advanced editing) or to engage in further musical work taking advantage of a full production environment.

3 CO-CREATIVITY

In terms of co-creative interactions, the design challenge is to make deep learning sequence models such as MMM, accessible and controllable by users towards addressing high-level composition tasks (e.g original ideation, variation, arrangement) in the concrete settings of their musical practice. As a result of MMM's design, the space of parametric user control is versatile enough to accommodate a number of compositional tasks discussed in CAC literature. Single-track, multi-track, bar infilling, conditioned or unconditioned (from scratch) generations all together allow the user to tackle a complex set of creative composition tasks fitting for real world applications. Variations can be generated to explore compositional alternatives in a novel way. Writer's block can also be helped by using the system's multi-track generation from scratch as a "ideation starter". The user can arrange existing compositions with new parts. The user makes use of all the control features to steer the generative behavior of the model. The affordance provided by those controls is powerful enough for the user to effectively guide the creative music composition process.

4 CONCLUSION

We present Calliope, a web application designed to assist users using MMM's generative music capabilities. Calliope's workflow allows for granular control at the generative steps, batch generation, ranking, MIDI support and DAW integration. Additionally, we are working on improving Calliope's support for time signature, MIDI note editing, generation progress control and batch-visualization of MIDI files using disjunctive interfaces. In addition to making powerful generative music models accessible and controllable for users in their musical practice, it can be used as a research context to conduct studies for better understanding AI integration in the music composition practice. It can help explore research questions on creativity support in assisted composition, co-creative design challenges associated with user-AI interaction in music systems, and the automation of creative music processes of interest to the interdisciplinary community of creativity and cognition scholars. Some of our future work involve 1) an ongoing evaluation study of the system along human factors including usability, user experience on feeling of trust, authorship, controllability and measured of technology acceptance for the system among users, 2) a research study on user-AI interaction design using usability, usefulness and creativity support index as main metrics, 3) a co-creative "applied" use of Calliope to submit an entry for the AI Song Contest, revisiting the challenges of "end-to-end" generative composition [10].

ACKNOWLEDGMENTS

We would like to acknowledge SSHRC, NSERC and MITACS for their contribution in supporting this research. We would also like to acknowledge Tara Jadidi as a code contributor to the Calliope MIDI Viewer/Editor feature.

REFERENCES

[1] 2022. Aiva. https://www.aiva.ai
[2] 2022. Evabeat. https://www.evabeat.com
[3] 2022. SPLIQS AI. https://www.spliqs.com
[4] Christopher Anderson, Arne Eigenfeldt, and Philippe Pasquier. 2013. The Generative Electronic Dance Music Algorithmic System (GEDMAS). In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE)*, Vol. 9. 5–8.
[5] Gérard Assayag, Camilo Rueda, Mikael Laurson, Carlos Agón, and Olivier Delerue. 1999. Computer-Assisted Composition at IRCAM: From PatchWork to OpenMusic. *Computer Music Journal* 23 (1999), 59–72.
[6] Jeffrey Ens and Philippe Pasquier. 2019. Quantifying Musical Style: Ranking Symbolic Music based on Similarity to a Style. *ArXiv* abs/2003.06226 (2019).
[7] Jeffrey Ens and Philippe Pasquier. 2020. MMM : Exploring Conditional Multi-Track Music Generation with the Transformer. *ArXiv* abs/2008.06048 (2020).
[8] Jeffrey Ens and Philippe Pasquier. 2021. Building the MetaMIDI Dataset: Linking Symbolic and Audio Musical Data. In *Proceedings of 22nd International Conference on Music Information Retrieval (ISMIR 2021)*.
[9] Dorien Herremans and Elaine Chew. 2016. MorpheuS: automatic music generation with recurrent pattern constraints and tension profiles. *Proceedings of IEEE International Technical Conference (TENCON) - 2016 IEEE Region 10 Conference* (2016), 282–285.
[10] Cheng-Zhi Anna Huang, Hendrik Vincent Koops, Ed Newton-Rex, Monica Dinculescu, and Carrie J. Cai. 2020. AI Song Contest: Human-AI Co-Creation in Songwriting. In *Proceedings of 21st International Conference on Music Information Retrieval (ISMIR 2020)*.
[11] Dimos Makris, Maximos A. Kaliakatsos-Papakostas, and Katia Lida Kermanidis. 2018. DeepDrum: An Adaptive Conditional Neural Network. *ArXiv* abs/1809.06127 (2018).
[12] James B. Maxwell. 2014. Generative Music, Cognitive Modelling, and Computer-Assisted Composition in MusiCog and ManuScore. In *Ph.D. Dissertation, School of Interactive Arts+Technology, Simon Fraser University, Home Dept & Faculty of Senior Supervisor: Special Arrangements*.
[13] François Pachet. 2004. On the Design of a Musical Flow Machine. *A Learning Zone of One's Own 111–134*.
[14] Philippe Pasquier, Arne Eigenfeldt, Oliver Roland Bown, and Shlomo Dubnov. 2017. An Introduction to Musical Metacreation. *Computers in Entertainment (CIE)* 14 (2017), 2:1–2:14.
[15] Adam Roberts, Jesse Engel, Yotam Mann, Jon Gillick, Claire Kayacik, Signe Nørly, Monica Dinculescu, Carey Radebaugh, Curtis Hawthorne, and Douglas Eck. 2019. Magenta Studio: Augmenting Creativity with Deep Learning in Ableton Live. *Proceedings of 7th International Workshop on Musical Metacreation (MUME 2019)*.
[16] Renaud Bougueng Tchemeube, Jeffrey Ens, and Philippe Pasquier. 2019. Apollo: An Interactive Environment for Generating Symbolic Musical Phrases using Corpus-based Style Imitation. *Proceedings of 7th International Workshop on Musical Metacreation (MUME 2019)*.
[17] Vibert Thio, Hao-Min Liu, Yin-Cheng Yeh, and Yi-Hsuan Yang. 2019. A Minimal Template for Interactive Web-based Demonstrations of Musical Machine Learning. *ArXiv* abs/1902.03722 (2019).