TENOR BOSTON 2023

8th INTERNATIONAL CONFERENCE ON TECHNOLOGIES FOR MUSIC NOTATION & REPRESENTATION

Department of Music
College of Arts, Media and Design
Northeastern University
360 Huntington Avenue
Boston, MA 02115 USA

Longy School of Music of Bard College
27 Garden Street
Cambridge, MA 02138 USA

New England Conservatory of Music
290 Huntington Ave.
Boston, MA 02115 USA

May 15-17, 2023

Eighth International Conference on Technologies for Music Notation and Representation

Edited by
Anthony Paul De Ritis, Victor Zappi,
Jeremy Van Buskirk and John Mallia

TENOR

INTERNATIONAL CONFERENCE
ON TECHNOLOGIES FOR MUSIC
NOTATION & REPRESENTATION

The Eighth International Conference on Technologies for Music Notation and Representation
Boston, Massachusetts USA May 15-17 2023

Edited by Anthony Paul De Ritis, Victor Zappi, Jeremy Van Buskirk, and John Mallia

Proceedings published by Northeastern University
978-0-6481592-7-8
https://doi.org/10.17760/D20511476

Logo design: Nicolas Taffin, updated for 2023 by Andrea Meger

http://tenor-conference.org
https://tenorboston23.sites.northeastern.edu/

Preface

The 8th International Conference on Technologies for Music Notation and Representation (TENOR) is hosted by the Department of Music at Northeastern University, Boston, MA USA; the Longy School of Music of Bard College, Cambridge, MA. USA; and the New England Conservatory of Music, Boston, MA USA.

For this 8th edition, in addition to the usual topics of the TENOR conference, we propose a focus on Representation and Perception of Music Structure.

In addition to the 22 scholarly papers presented in these proceedings, we held two keynote presentations by Dr. Morwaread Farbood, Associate Director of Music Technology, New York University; and Dr. Holly Watkins, Professor of Musicology, Eastman School of Music; one keynote workshop presented by Tim Perkis, Software Engineer, Researcher and Educator; and Gino Robair, Composer and Visual Artist; and four Concerts, two at the Longy School of Music at Bard College (on May 15, 2023), and two at the New England Conservatory of Music (May 16, 2023), featuring Loadbang, Lisa Mezzacappa and Jason Levis (duoB); and members of the Callithumpian Consort, Stephen Drury, Director.

Many thanks to Craig Vear and the DigiScore Project, a European Research Council (ERC) Project, and to Sandeep Bhagwati for sponsoring the Callithumpion Consort; commissioned composers: Ingrid Laubrock and Kitty Xiao; and awardees of the NeoScore Competition: Xavier Davenport, winner; and Lauren McCall, finalist.

Our greatest thanks goes to Daniel Strong Godfrey, composer, and Chair of the Department of Music at Northeastern University, for his unflagging support of TENOR BOSTON 2023, both spiritual and financial.

Organization

Organizing Committee

Anthony Paul De Ritis, Northeastern University, Department of Music (Conference Chair)
Jeremy Van Buskirk, Longy School of Music of Bard College (Paper Chair)
John Mallia, New England Conservatory (Sonic Works Chair)
Victor Zappi, Northeastern University, Department of Music (Scientific Program Chair)
Aaron Clarke, Longy School of Music of Bard College
Alexandra du Bois, Longy School of Music of Bard College
James Gutierrez, Department of Music, Northeastern University
Rébecca Kleinberger, Department of Music, Northeastern University
Jie Ren, Longy School of Music of Bard College
Lisa Mezzacappa, Independent Artist
Katarina Miljkovic, New England Conservatory

Paper Committee

Jeremy Van Buskirk (Paper Chair)
Victor Zappi (Scientific Program Chair)
Rébecca Kleinberger (Department of Music, Northeastern University)
Andrea Agostini (Scuola di Musica Elettronica del Conservatorio di Torino, SMET, Italy)
Jonathan Bell (PRISM Laboratory; Perception, Representations, Image, Sound, Music; Marseille, France) Jean Bresson (Ableton, Germany)
Clément Cannone (IRCAM, France)
Charles de Paiva de Santana (PRISM-AMU; PRISM Laboratory; Perception, Representations, Image, Sound, Music - Aix-Marseille University, Marseille, France)
Roger B. Dannenberg (Carnegie Mellon University, USA)
Frédéric Dufeu (University of Huddersfield, UK)
Christine Esclapez (PRISM Laboratory; Perception, Representations, Image, Sound, Music; Marseille, France)
Julien Ferrando (PRISM Laboratory; Perception, Representations, Image, Sound, Music; Marseille, France)
Daniele Ghisi (University of California, Berkeley, USA)
Jean-Louis Giavitto (Ircam, France)
Rama Gottfried (Zurich University of the Arts, Switzerland)
Gérard Guillot (Pôle d'Enseignement Supérieur Musique & Danse, PESMD, France)
Georg Hajdu (Hochschule für Musik und Theater Hamburg, HfMT, Germany)
Cat Hope (Monash University, Melbourne, Australia)
Raul Masu (King Mongkut's Institute of Technology Ladkrabang, Thailand)
Peter Nelson (University of Edinburgh, Scotland)
Matthias Nowakowski (Center for Music and Film Informatics, CeMFI, Germany)
Douglas Nunn (Anglia Ruskin University, UK)
Philippe Rigaux (CNAM, Conservatoire National des Arts et Métiers, France)
Ryan Ross Smith (CNAM, Conservatoire National des Arts et Métiers, France)
Luca Turchet (University of Trento, Italy)
Lindsay Vickery (Perth University, Australia)

Sonic Works / Music Committee

John Mallia, New England Conservatory (Sonic Works Chair)
Alexandra du Bois, Longy School of Music of Bard College
Lisa Mezzacappa, Independent Artist
Katarina Miljkovic, New England Conservatory
Ronald Bruce Smith, Department of Music, Northeastern University

Demos / Workshops Committee

Victor Zappi, Department of Music, Northeastern University (Scientific Program Chair)
Rébecca Kleinberger, Department of Music, Northeastern University

Webmaster

Bella Mona Designs

Graphic Design

Rebecca McTeigue

Tenor International Steering Committee

Sandeep Bhagwati: Concordia University, Montréal – Canada
Jean Bresson: Ableton AG, Berlin – Germany
Georg Hajdu: Hochschule für Musik und Theater Hamburg – Germany
Richard Hoadley: Anglia Ruskin University, Cambridge – UK
Cat Hope (General Chair): Monash University, Melbourne – Australia
Craig Vear: University of Nottingham – UK

Table of Contents

STRUCTURING CENSORSHIP IN DIGITAL SCORES

Jaslyn Robertson
Monash University
Melbourne, Australia
jaslyn.robertson1@monash.edu

ABSTRACT

Digital scores allow for new ways to experiment with agency and perception within composition and performance. My recent works *Censoring Experiment* and *Shadow Aria* examine the possibilities for digital scores to incorporate methods of censorship with the aim of highlighting and unpacking it as a social issue. Despite an assumption that censorship is an issue of the past or limited to non-western countries, recent cases of artistic censorship in Australia and North America have brought attention to the ongoing problem, particularly as it affects marginalised artists and composers. In this paper, I discuss my two pieces that attempt to address, complicate, and subvert the issue of artistic censorship through experimental composition. Digital scores are the medium that allows these pieces to exaggerate real censorship and test how performers react creatively to censored environments. I argue that animated notation and mixed-media environments created through technology give me the ability to replicate and change a real-life social issue within a performance, letting my art not just comment on a political question, but work towards new insights through practice-based research.

1. INTRODUCTION

Working with digital scores as a composition medium, my practice-based research explores forms of censorship experienced by underrepresented artists in the field of Western art music, and experiments with ways of incorporating censorship into experimental intermedia works. Early experiences of censorship in my own music and writing career led me to research recent cases of composers and artists across contemporary western society being silenced, particularly other composers who face discrimination for their identities and for the political themes of their work. As a recent example, Black American composer Daniel Bernard Roumain had his opera commission rescinded by Tulsa Opera after writing the work *They Still Want to Kill Us* (2021) because he refused to change the final lyric of the opera, "God Bless America; God Damn America!" His opera reflects on the 1921 massacre of hundreds of Black Americans in Tulsa, and while Tulsa Opera were open to presenting an opera about the massacre, they were only willing to do so without such explicit wording that would shock or offend their audience. [1] A particularly scandalous Australian case of artistic censorship was that of queer artist Casey Jenkins, who had the government arts funding for their work *Immaculate* (2020-2021) revoked by the Australia Council for the Arts. The work in question included streaming video of their process of self-insemination towards pregnancy, letting the viewer into an emotional part of queer life that is often hidden away. [2] These are only two examples among several recent cases of artistic censorship in the USA and Australia, but speak powerfully to the impression that censorship is still an issue faced by composers and artists that represent their marginalised communities.

In light of the understanding that censorship affects contemporary composers and artists in a Western context, and framing ideas of censorship around expanded definitions that Matthew Bunn names New Censorship Theory [3], my compositions test ways of structuring censorship into intermedia works. The two pieces discussed in this paper both create environments of censorship for improvising performers that are mediated through technological forms of communication, or digital scores. While they each include real-time aspects of randomisation, the moments and forms of censorship in the pieces is structured into a composition, with different sections of the pieces triggering different types and speeds of censoring gestures with varying levels of intensity throughout the works. This randomisation within a structure allows the works to play with concepts of surprise and failure within a musical form that unfolds through time. Jack Halberstam's notion that "under certain circumstances failing, losing, forgetting, unmaking, undoing, unbecoming, not knowing may in fact offer more creative, more cooperative, more surprising ways of being in the world" [4] and Linda Candy's method of "reflection-on-surprise" as a way of understanding unexpected outcomes [5] guide my understanding of performer and audience reactions to my work. In my own work, the medium of the digital score provides new possibilities for unpacking the effects of artistic censorship through experimental intermedia composition.

2. DIGITAL SCORES AND NEW POSSIBILITIES

Digital scores for intermedia composition are a provocative medium to explore the concept of censorship. To define digital scores, Craig Vear writes that "when a musician is interested in communicating a musical idea that is only created through and with technology then we can call that a digital score." [6] Some forms of digital scores include animated scores, graphic scores, interactive and AI-based scores as well as digitally enhanced traditional scores. A scanned copy of a paper score, however, is not normally considered a digital score unless it makes use of technology to build upon the original material in some way. [6] Digital scores are classified as such if their core purpose is as "a technically mediated communication interface between the creativity of a composer, the creativity of a performer and the creative mind of the listener." [6] Therefore, the role of technology in digital scores is key to their realisation. As opposed to traditional forms of music notation, digital notation allows for real-time composition and chance processes and different forms of engagement and interactivity between performers and the score. As a composer and practice-based researcher, digital scores give me the potential to simulate environments of censorship using technology. The possibilities of technology to aid in the facilitation of ideas between the composer, performers and audience gives composers the potential to integrate conceptual ideas into the medium of the score as well as the resulting performance. Researchers in the fields of digital and animated notation have pointed out how these forms of notation can facilitate different kinds of interactions between composers and performers, audiences and time. [6][7][8] My recent compositions discussed in this paper experiment with the possibilities for these distinctive mediations to recreate, complicate, and subvert environments of artistic censorship, untangling the effect of censorship on performing artists.

3. *CENSORING EXPERIMENT* (2022)

Censoring Experiment is an audio-visual work for one performer playing through a Max patch on a laptop with a webcam.[1] The work is attended for online performance through live-streaming or pre-recorded distribution on online platforms. This piece is an attempt to recreate and exaggerate forms of censorship faced by marginalised people on websites and social media applications. In the USA, the FOSTA-SESTA law was recently introduced to stop sex trafficking activities online by policing sex work on the internet. Website owners are now held responsible if illegal sex work is advertised on their platform, however because many large websites that host user content have no way of efficiently moderating for that, they use either algorithms or outsourced workers to ban any sexual content. [9] Content that is deemed pornographic because of the inclusion of nudity, body hair or menstruation is routinely banned from platforms including Instagram, Facebook and Twitch. Moderation processes particularly target black people, Indigenous people and people of colour, and non-cis-male bodies. [10] As a response to recent changes in policy and cases of censorship online, I tested the idea of censoring an improvising performer during their performance and observing how they would react. The Max patch processes the incoming video and audio of the performer's improvisation, creating a version that distorts, erases and covers up parts of the visual and audio input, as shown in Figure 1. The censoring events are randomised within a structure that defines the density and kinds of events that occur in different sections of the piece, building to more heavily censored moments. With no other instructions given to the performer except to improvise into the webcam and respond to the processed video and audio played back to them as a score, the resulting piece is the censored mirror the performer sees as they play. A digitally created environment that the performer improvises in becomes both the space of the performance and the score itself, the digital score creating an "integrated cross-disciplinary performance". [6]

Figure 1. One form of randomised visual censoring in *Censoring Experiment*.

Composing in a digital format in this piece reveals possibilities of randomness only available in real-time composition. The randomised structure of the score is important to the work conceptually and practically, giving the performer unexpected moments to respond to. This real-time aspect incorporates a layer of unpredictability into the patch, making it feel more like a malleable environment than a traditional fixed score. Cat Hope posits that "animated notations can assist us to challenge our perceptual notions of time for music and sound". [8] In *Censoring Experiment*, the randomisation of events in time, facilitated by the digital nature of the score, allow the piece to simulate a changing environment of censorship more realistically. The unpredictability of the work to the performer, even after rehearsing or performing the piece multiple times, allows for surprising real-time responses in each iteration. These moments of surprise can lead to observations on how performers respond to unpredictable censorship of artistic expression on online platforms.

In comparison to traditional, fixed forms of music notation, digital and real-time scores change the structures of agency and freedom for performers. Ryan Ross Smith

[1] https://vimeo.com/789422211

writes, “Simply put, agency lies primarily with the performer to activate or dynamize the conventional score, whereas the dynamic score has agency over the performer; movement is perceptible, not of the eye, but to the eye.” [7] This effect is enhanced when the animated score is visible to the audience, letting the viewers in on the process. Lindsay Vickery’s piece *Mueller* (2020) makes a statement about censorship by setting up an animated score that is impossible to keep up with, and projecting that score for the audience as well as the performers. [11] The musicians play fast stabs of sound when the black rectangles of redacted words appear in the Mueller report, flicked through quickly in the projection. The result is that the performers ultimately make mistakes in trying to keep up with the piece, and the report itself is redacted by the speed of the animation. In *Censoring Experiment*, only the resulting video processed through the Max patch is presented to the audience as an online or live-streamed video. Therefore, the performer in this piece has limited agency over the final product seen by the audience, reflecting the nature of censorship on online platforms. They insert themselves into the score and attempt to take back some control over it.

While animated scores change the agency of the performer in relation to the score, and while showing the score to the audience heightens that effect, performers often feel freed, not restricted, by digital scores. Vear has investigated performers’ responses to digital scores and writes that “Digital/generative scores may feel liberating to a performer.” [6] A dichotomy between restrictive and liberating elements in the performance of a digital piece is seen clearly in Elizabeth Jigalin’s performance of *Censoring Experiment*. While the Max patch amplifies restriction and censorship, the underlying theme taunts the performer to work against it and reclaim agency and freedom. Unprompted, Jigalin does exactly that in ways that were unexpected to me as the composer. At the beginning of the piece, when the video and audio processing censors her gently and sporadically, she works with it and allows the score to guide her. As the processing ramps up, she realises that she has to fight to work against it. She plays louder and makes bigger movements, but as these are continually distorted, muted and covered up, she has to take a different approach. She finally finds more playful ways to work against the censoring environment by turning the piece on its head and proving her agency over the digital score. Taking off her shirt, she covers the webcam in red fabric as pictured in Figure 2. This has the multiple effect of heightening the risk of the performance – leaving her underwear visible for a few seconds through the pixelation that could change at any moment – and taking control over the score. Now she can shape the video by moulding it in red fabric. The way that Jigalin takes back control over the censoring score reflects an approach that artists can and do take in real life, using their creativity to push back against algorithmic mass censorship. The possibilities of a digital score and real-time composition allow an experiment like this to be possible, and to condense and replicate a social issue on a small scale to begin identifying ways to address it.

Figure 2. Jigalin reclaiming agency over the digital score in *Censoring Experiment*.

4. *SHADOW ARIA* (2022)

Digital scores take many forms beyond screen-based notation. Vear defines the term broadly:

> From animated graphic scores and projected images to mixed-media environments; from co-located telematics with distributed code to artificial intelligence, thinking machines, robotics and hacked bodies, there is a broad wealth of innovation offered to musicians through the digital score. A defining feature is that they benefit from the usability and functionality of dynamic technological environments at some level and are responsive, evolving as the performance progresses and operating on a level of interactivity more in common with gaming and immersive new-media art. [6]

The idea of a mixed-media environment forming a score is experimented with in *Shadow Aria* (2022).[2] *Shadow Aria* is a short piece for four improvising performers, fixed multichannel tape track and a Max patch controlling four spotlights. It was premiered in late 2022 with double bass, percussion, bass clarinet, viola and a 24.2 speaker spatialisation. The action of the spotlights is randomised within a structure, similarly to the processing in *Censoring Experiment*. Unlike the screen-based *Censoring Experiment*, *Shadow Aria* takes the notions of a digital score beyond the screen by using digital methods to create an analogue, spatialised environment for both performers and audience. The piece is about silenced voices, broaching the topic of censorship in the arts from a different angle. A few years ago, I spoke to Helen Gifford, one of the most interesting Australian composers of the last century. She told me about orchestral pieces and operas she had written over the years that were never performed because of gendered

[2] https://vimeo.com/789426278

discrimination and the restrictions of her illness. [12] I wondered how many pieces have been lost, or never written, due to the lack of opportunities given to marginalised people, and in Gifford's case, women in a male-dominated discipline. The music of *Shadow Aria* is the accompaniment to a silent soloist. Performers are instructed to improvise with the spatialised tape track, drawing from both the sound and its movement in the space. When a spotlight hits them, they become still and silent, representing the ghost of a silenced voice. The score is the environment they improvise in – light, audio and spatialisation are the prompts for the performers.

Figure 3. Spotlight on one performer in *Shadow Aria*.

The lights censor the performers directly (via the computer), in that they must cease to play when they are lit. Because they only play and move in the darkness, their performance is then censored again from the audience. Other multimedia composers have played with this idea of hiding performance in the dark, including Alexander Schubert in his piece *Sensate Focus* (2014) that also makes use of computer-controlled spotlights. [13] Schubert's piece has the performers following a notated score, but their sounds and actions are lit through synced-up lighting cues, giving the audience the impression of digital sampling rather than live performance. While the piece is meant to evoke a digital feel, Schubert also wanted to draw attention to the impossibility of the robotic performance, "it shows the flaws in the body and how we work with them and how it is impossible to fulfil all the demands we put on it". [14] *Shadow Aria* takes a different approach in its use of lighting as a representation of censorship. The presentation of the piece is distinctly analogue, although a Max patch forms a digital real-time score that controls it. The lights are warmer in tone, and turn on and off slowly. Rather than aiming to catch out the performers, the piece gives them time to halt their performance with theatricality, giving the audience a glimpse of each musician's personality even though they never see them play. The tape track, too, is recorded on analogue synthesizers, blending in particularly with the percussion and clarinet sounds in the premiere performance. Instead of being a score that focuses on mistakes and surprise, like *Mueller Report* and *Censoring Experiment*, *Shadow Aria* gives the performers gentle cues to shape their improvisation and complicates the audience's perception of the work. While *Censoring Experiment* tries to exaggerate the restrictive censorship of arts on social media and force a reaction out of the performer, *Shadow Aria* reminds the audience of ghosts and shadows of underrepresented voices.

Spatialisation forms a part of the audio score by providing a moving foundation for the performance. The improvised music is shaped by the tape track, with the combination of the tape track and live performance making up an accompaniment to an imaginary soloist. Because of the suggestion of an accompaniment, and the often-quiet nature of the tape track, the piece feels held-back. Spatialisation allows sounds and gestures of a soft dynamic to fuel the improvisation. As opposed to a stereo track, in multichannel compositions a quiet sound can be heard distinctly above other louder sounds in the same register, either because of the listener's proximity to a speaker or because of a gestural movement. [15] This allowed me to layer a number of synthesizer sounds in the track and have certain gestures available to the attention of one specific performer, giving each musician a different version of the audio score. In the premiere, this was not entirely successful as the spatialisation needed more workshopping to ensure that each musician had an equally interesting experience in their position. However, it did work to some extent, and the performers listened carefully to interact with the sound and movement surrounding them. For the audience and performers, spatialisation also introduced another level of censoring in the work. Having up to eight different audio tracks moving through the speakers in different ways while the four musicians play creates a level of complexity that cannot be fully perceptible. When spatialised sounds are layered in this way, "…all but the most significant characteristics are filtered out of the event, while the rest of the information is more or less ignored." [15] Therefore, each audience member and performer will focus on the sounds loudest, closest to them, or with the most movement perceptible in their position. They each censor out some information to hear their own personal impression of the whole composition.

5. CONCLUSIONS

While *Censoring Experiment* and *Shadow Aria* both explore similar themes of artistic censorship, they test out different ways of incorporating structured censorship into digitally scored compositions. *Censoring Experiment* relies on concepts that are distinctly screen-based and suited for online distribution. The solo performer fights against censorship within the computer screen, reminiscent of the algorithm-fuelled censorship they experience on social media and web-based platforms. This forces them to react creatively, using their body and performance to regain agency over the work or shape their music within the censored environment. *Shadow Aria* asks the four performers to censor themselves in collaboration with the tape track, making space for the silent soloist. The result is a more cautious performance, with a sense that the performers are working together with the lighting and spatialised sound. Both pieces create a structured environment of censorship as different kinds of digital scores – *Censoring Experiment* as real-time animated notation, and *Shadow Aria* as a mixed-media environment of light and moving sound. These pieces begin an exploration for the potential for different forms of digital scores as methods of incorporating censorship into compositions and representing the social issue of censorship to performers and audience. Future experiments will take this concept further by testing the difference between musicians' and audience's perceptions of censorship coming from digital means and coming from a human in live performance, leading to further insight on how we might react to algorithmic moderation on the internet.

Acknowledgments

I would like to thank the DigiScore project for commissioning *Shadow Aria*, and the performers of both pieces – Aaron Wyatt, Louise Devenish, Lindsay Vickery, Helen Svoboda and Elizabeth Jigalin. I would also like to thank my PhD supervisor Cat Hope for her guidance.

6. REFERENCES

[1] J. Wicker, "God Damn America: Tulsa Massacre Opera Drops Black Composer Over Lyrics," *The Guardian*, 2021. [Online] Available: https://www.theguardian.com/music/2021/mar/22/black-composer-tulsa-massacre-opera-daniel-bernard-roumain.

[2] B. Eltham, "Casey Jenkins v Australia Council: When Controversial Art Loses Funding, What Does it Mean For Culture?" *The Guardian*, 2020. [Online] Available: https://www.theguardian.com/culture/2020/oct/17/casey-jenkins-v-australia-council-when-controversial-art-loses-funding-what-does-it-mean-for-culture.

[3] M. Bunn, "Reimagining Repression: New Censorship Theory and After," *History and Theory*, vol. 54, no. 1, 2015, pp. 25-44.

[4] J. Halberstam, *The Queer Art of Failure*. Duke University Press, 2011.

[5] L. Candy, "Reflective Practice Variants and the Creative Practitioner", in C. Vear (Ed.), *The Routledge International Handbook Of Practice-Based Research*, Routledge, 2022.

[6] C. Vear, *The Digital Score: Musicianship, Creativity and Innovation*. Taylor & Francis, 2019.

[7] R. Ross Smith, "An Atomic Approach to Animated Music Notation," in *Proceedings of the International Conference on Technologies for Music Notation and Representation (TENOR'15)*, Paris, 2016, pp. 40-48.

[8] C. Hope, "The Possibilities of Animated Notation," *Computer Music Journal*, vol. 41, no. 3, 2017, pp. 21-35.

[9] D. Blunt and A. Wolf, "Erased: The Impact Of FOSTA-SESTA and the Removal Of Backpage On Sex Workers," *Anti-Trafficking Review*, vol. 14, 2020, pp. 117-121.

[10] G. Faust, "Hair, Blood and the Nipple: Instagram Censorship and the Female Body," in U. Frömming et. al (Eds.), *Digital Environments: Ethnographic Perspectives Across Global and Online Spaces*, Transcript, 2017.

[11] L. Vickery. Mueller. [Online] Available: https://vimeo.com/447707905.

[12] H. Gifford and J. Robertson, "A Brain that Keeps Working: An Interview with Helen Gifford," *Context*, vol. 45, 2019, pp. 75-85.

[13] A. Schubert. Sensate Focus. [Online] Available: https://youtu.be/ojWW9ifdMFs.

[14] Z. Kanga and A. Schubert, "Flaws in the Body and How We Work with Them: An Interview with Alexander Schubert," *Contemporary Music Review*, vol. 35, no. 4-5, 2016, pp. 535-553.

[15] L. Brümmer, "Composition and Perception in Spatial Audio," *Computer Music Journal*, vol. 41, no. 1, 2017, pp. 46-60.

ONLINE SURVEY ON USABILITY AND USER EXPERIENCE OF MUSIC NOTATION EDITORS

Matthias Nowakowski
Center of Music and Film Informatics
Detmold University of Music
Detmold, Germany
matthias.nowakowski@hfm-detmold.de

Aristotelis Hadjakos
Center of Music and Film Informatics
Detmold University of Music
Detmold, Germany
aristotelis.hadjakos@hfm-detmold.de

ABSTRACT

Although one can see a certain convergence between the interaction designs of different notation editors, there is no general consensus or standard since new interaction paradigms keep appearing with most major software updates or new products. In this paper, we present the results from an online survey ($n = 138$) with standardized usability and user experience questionnaires. The users of digital notation editors were asked to fill out the *System Usability Score*, the *AttrakDiff2* and the *Liveness* questionnaire. This provides insights into domain specific design problems with the goal to inform the design of future interfaces. Almost all music notation editors show clear deficiencies in overall usability. Furthermore, a detailed examination of the obtained metrics show specific dependencies of individual qualities, which are helpful to conduct further qualitative research.

1. INTRODUCTION

Usability is one of the key topics of human computer interaction (HCI) research. It describes the property of a system to help a user achieve goals effectively, efficiently and satisfactorily [1]. In the early days of HCI research much fundamental work was done on the psychological and motorical basics of using text editors and graphical user interfaces (GUI) and thereby also determining their efficient use [2, 3, 4]. In music, the digital positioning of graphical objects such as notes, staves, etc. was already problemized since the 1960s [5, 6, 7]. This was not framed as a topic of usability, but rather of automatizing score editing, as demanded by formatting or engraving and contemporary music notation practices [8]. There was no need for music specific interaction paradigms since the interaction was based on text editing.

Before the first graphical user interfaces (GUI) for music notation were developed, GUIs were used to create electronic music and were seen as a creative tool. Novel and more abstract musical representations could be employed, suitable for interactions via mouse and keyboard. First usability considerations in this regard where made in the mid-1980s by analyzing digital workflows [9], describing interactive graphical environments for computer assisted composition [10] and so essentially proposing ideas for a visual programming languages based on parametric manipulations [10, 11].

Today, GUIs to produce sound include visual programming languages such as Max/MSP, Pure Data (Pd), OpenMusic (OM), PWGL, Bach, etc. which in some cases already include modules with notation interfaces. Furthermore, there are Trackers, Sequencers, Digital Audio Workstations (DAW) and score editors. Nash et al. continue to research usability dependent on the creative involvement of the user and developing workflow models to address different use-cases by analyzing flow and cognitive dimension metrics [12, 13, 14]. Hunt et al. [15] even uses the cognitive dimensions approach to design an interactive generative score editor from scratch. Nevertheless, detailed academic examination of music notation software remains scarce. Peterson et al. [16] measure duration of various interactions and analyzed their influence on composer creativity. Compared to handwriting they spent less time on the creative task of writing notes due to menu navigation, which resulted in less musical detail.

Taking an analytical view on the components of score editors, they are mostly based on interaction paradigms better known from other contexts: text processing (e.g., inserting notes with the keyboard), image processing (e.g., changing layout with the mouse) and digital audio processing (e.g., placing sound elements in a temporal order and playing them back, similar to DAWs). Although there are convergences in visual and interaction design, a lack of overarching and consensual metaphors leads to more variation between the applications and may create a barrier to change from one to another.

To also account for the use of musically trained people, we decided to approach the problem from two sides—usability and creativity—using established and standardized questionnaires to develop hypotheses about the importance of design elements, which we will be important to explore in further studies.

2. METHOD

Digital musical notation is often aimed at creative use, be it formatting a score beautifully, or for composing or arranging with simultaneous acoustic verification of the result. The interaction is usually conveyed by screen, keyboard

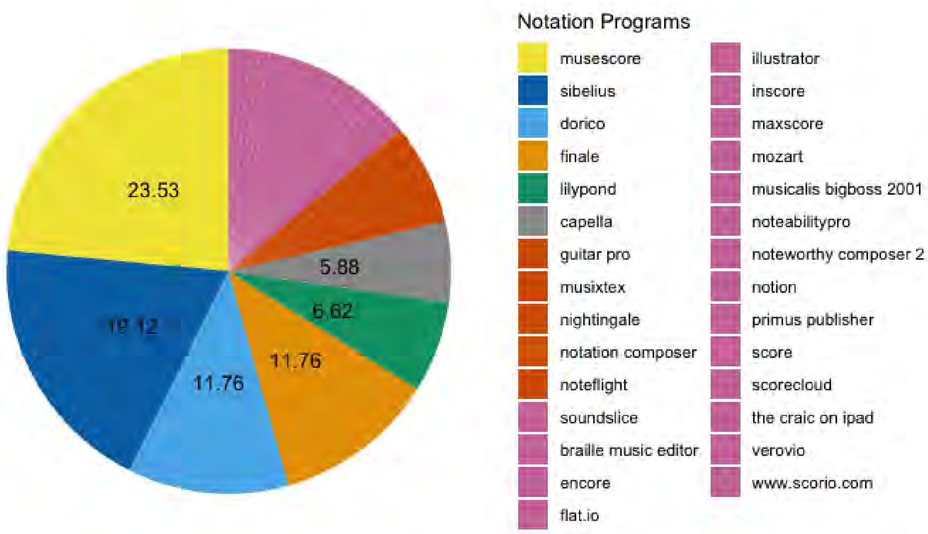


Figure 1. Percentages of all named notation programs in the survey. The dark orange fields equal to 1.47% (named twice) each, while the pink fields equal to 0.74% (named once) each. The first six programs to be analyzed in this paper represent 107 of 136 valid responses.

and mouse, but touch-based solutions are also common. In any case, the design has to ensure good usability and trouble-free use. To cover both aspects, i.e., creativity and usability, we decided to use multiple questionnaires. The System Usability Score (SUS) [17] assesses usability. The AttrakDiff2 [18] assesses pragmatic and hedonistic qualities. The questionnaire by Nash et al. [13] assesses liveness of the interaction.

Pragmatic quality (PQ) measured by the AttrakDiff2 questionnaire is bound to the satisfaction using the software and the feeling of productive impact while using it, whereas hedonistic qualities, which we consider to be desirable goals while working creatively, deal both with:

identity (HQ_I) People express their self through objects. They want to be perceived by relevant others in a specific way. A product can support this by communicating a desired identity.

stimulation (HQ_S) People strive for personal development, i.e., the improvement of knowledge and skills. Products can support this development by having a stimulating effect. Novel, interesting and stimulating functionalities, content, interaction and presentation styles can increase attention, dampen motivation problems or facilitate finding new solutions to existing problems. Thus, stimulation can also indirectly help with task completion.

Considering music production, Nash et al. [13] investigated liveness metrics on trackers and sequencers, which employ different kinds of notation and workflow, whereby score editors can be seen as being similarly manipulation-driven as trackers. Liveness describes generally a sense of subjective sense of intimacy, which can be assessed through system feedback. This is strongly connected to flow, which describes the mental state of immersion while performing a task or using a system [19].

We distributed the survey among mailing lists of interest groups concerning themselves with musical notation in general, digital notation, digital instruments, digital musicology, composition and musical markup languages.[1] By doing this, an international pool of potential participants was addressed. The survey started in July 2022 and lasted until December 2022.

Each participant had to name one editor either by choosing from a preset list or writing into a free text box before answering the questions. Only single mentions of a program were counted as valid entries to ensure differentiated answers. E.g. “musixtex” is valid, while “musixtex and score” is not. The questionnaires could be answered multiple times by the same person, each time answering the questions in relation to another program they have not named before. The participants did not have to be experts with the program but should be confident using it. Further attributes about the participants where not requested. All questions in the survey were mandatory.

We got 138 responses, from which two were not valid. In total, 29 score editors were mentioned. We excluded all editors that were mentioned only once or twice from the further analysis leaving us with six score editors and 107 of the total responses (see Figure 1). The six score editors are: MuseScore, Sibelius, Finale, Dorico, LilyPond and Capella.

MuseScore is available for free, while Sibelius, Finale, Dorico and Capella have to be purchased, but reduced free versions exist. It is worth mentioning that we had nine responses for LilyPond, which is available for free and which follows completely different interaction paradigms than the What-you-see-is-what-you-get (WYSIWYG) in-

[1] The mailing lists were: students at the Detmold University of Music, Music Notation at Ircam (music-notation@listes.ircam.fr), W3 Notation Community Group (public-music-notation@w3.org), NIME Community (nime-community@googlegroups.com), MEI Community (mei-l@lists.uni-paderborn.de)

terfaces we mainly had in mind. It is a notation program, which compiles notations from text files and so is functionally similar to LaTeX.

To make the the questionnaires more comprehensible, we had to add some explanations since concepts condensed in expressions like "viscosity", "diffuseness" or "action-awareness merging", were not always clear in pre-tests in the context of using notation programs. Even after starting the survey we had some remarks about the comprehensibility of some items. Especially for AttrakDiff2 some adjective pairs seemed pointless, as it was communicated to us by English- and German-speaking participants.

Each questionnaire was evaluated as described in their respective papers. To compare the music notation editors, we applied ANOVA on the respective scores for the whole group to determine, if there are any significant differences in the group itself. Additionally η^2 was computed as a metric of confidence about the detected variance. Post-hoc Tukey tests where applied to detect significant differences between pairs of notation editors when ANOVA detected any significant differences. Finally, all metrics were correlated with each other to determine if any metric could be reduced to the outcome of another and what topics might be worth exploring further.

3. RESULTS

3.1 System Usability Score (SUS)

The SUS assesses the general satisfaction of a person dealing with a software. Bangor et al. developed a grading scale that maps SUS scores as follows to school grades [20]:

- below 60: F
- between 60 and 69: D
- between 70 and 79: C
- between 80 and 89: B
- 90 and above: A

Based on 241 studies, the average SUS score is 68 [21, p. 203–204]. Therefore, we added a mark for 68 in Figure 2. Only one of the six programs (Capella) clearly exceeded this threshold. Finale and MuseScore barely exceeded the threshold, while Dorico, Lilypond and Sibelius fall below. Looking at the distribution within the groups we can see large differences. Notation editors, which were mentioned less by our participants (Capella and LilyPond) have smaller ranges between the largest and the smallest scores. Sibelius has an especially wide range.

The one-way ANOVA test shows that there is a significant difference between the classes ($p < .001$). The post-hoc Tukey test shows that there is a significant difference between Capella and all other programs ($p < .01$). Other significant differences are not detected (Table 2).

3.2 AttrakDiff2

The items in this questionnaire consist of polar adjectives, which are ranked on a 7 point Likert scale. A higher number represents a stronger expression of that property.

Overall the one-way ANOVA tests detects significant differences in all three qualities ($p < .001$). Especially HQ_S shows more individual differences and more diverse group parings in the Tukey test (see Table 3). Tukey tests for each quality is characterized by MuseScore and Sibelius having mostly significant differences with most other editors, while Capella, Dorico and LilyPond do not show differences among themselves. HQ_I shows only differences of Capella and Dorico to MuseScore while PQ mirrors the outcome of SUS (see Table 4). In general Capella and MuseScore are considered to be more pragmatic, while Dorico and LilyPond are more stimulating with LilyPond having stronger tendencies towards identity. The scores of Sibelius and Finale are similar to each other across all qualities.

3.3 Liveness

Nash et al. [13, 22] derive their concept of liveness from cognitive dimensions [23] and flow [19] to assess creative work while using programs based on notations. The results in Figure 4 are based on a 5 point Likert scale. Due to the formulations higher values do not always mean a more desirable expression of item. For example "hard mental operations" should be desired to be low, while "no hidden dependencies" is desired to be high, because hidden dependencies may influence prediction of outcomes and may be not controllable by the user. Considering the role of "loss of self consciousness" and "transformation of time" it seems not very clear if lower or higher values are more desirable, but it might be a hint towards more or less rational and controlled handling of the software. However, this did not allow for a meaningful aggregation, so that we evaluated significant differences for every item as demonstrated in Table 5. Due to the relatively high number of examined metrics we decided to choose a significance level of $p < .01$ for one-way ANOVA to reduce random detections of significance. Afterwards Tukey tests were conducted if that significance level was reached. For the Tukey tests, Table 5 reports all group pairings with significance level $p < .05$. For example this tells us, that there are significant differences in "consistency" between Capella & Sibelius and Capella & MuseScore.

It is noticeable that most of the groups in the Tukey test include MuseScore eight times, Capella and LilyPond ar mentioned seven times and Sibelius six times. The most mentioned group is LilyPond & MuseScore with four, followed by Capella & MuseScore, Capella & Sibelus and LilyPond & MuseScore (each three times). LilyPond & Finale and Dorico & MuseScore are mentioned once. MuseScore and Sibelius are often grouped for the same metric (like for "abstraction management" and "intrinsically rewarding") and there are not instances in which significant differences between them can be detected, which is also true for Capella and LilyPond.

Comparing the trajectories of the programs (Figure 5) we can see that Capella is outperforming all others in almost every item. Except for "no premature commitment", "loss of self-consciousness" and "transformation of time" where it is below the level to the best-performing programs. For

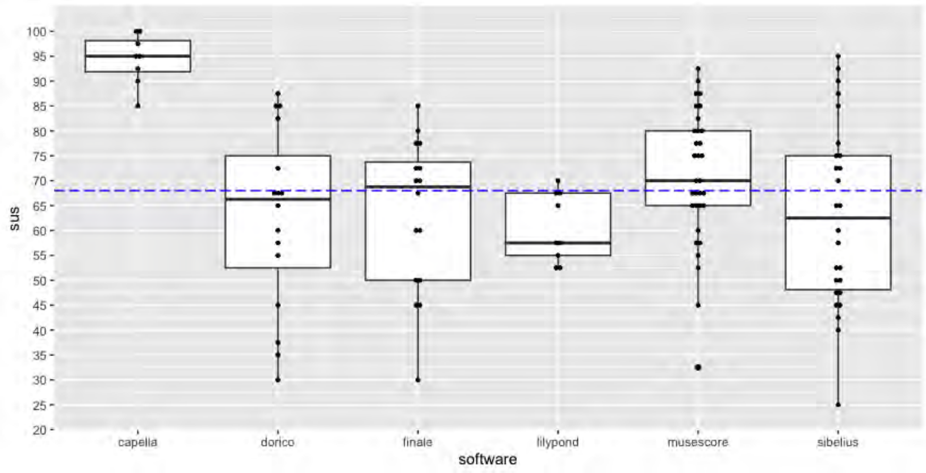


Figure 2. Boxplot of SUS. The points represent the individual results in each group. The dashed line highlights the SUS usability threshold of 68.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	η^2
	5	7738.97	1547.79	6.66	2.12e-05	0.57
Residuals	101	23465.82	232.33			

Table 1. ANOVA for SUS

group1	group2	estimate	conf.low	conf.high	p.adj	p.adj.signif
capella	dorico	-31.88	-51.05	-12.70	7.05e-05	****
capella	finale	-31.09	-50.27	-11.92	1.13e-04	***
capella	lilypond	-33.82	-55.34	-12.30	2.00e-04	***
capella	musescore	-23.75	-41.25	-6.25	2.02e-03	**
capella	sibelius	-31.39	-49.30	-13.49	2.38e-05	****

Table 2. Tukey test for SUS. Only groupings with any significance are shown.

"abstraction management" and "virtuosity" LilyPond performs best, which is also reflected by the significance tests.

4. DISCUSSION

We first discuss limitations and problems of our approach. Then we discuss the results in more detail and how they could be interpreted in the context of this study. All metrics were correlated with each other to develop hypotheses about important features of notation software and their interaction design. Since all the results were correlated separately and were not grouped by software, the effect of outlying and biased results is reduced.

4.1 Limitations and Problems

There are several potential problems with our approach in particular and when using an online questionnaire-based approach in general. These limitations and problems are discussed in the following:

Sampling bias: Sampling bias is often seen as a main objection against online surveys as it is difficult or even impossible to achieve a random sample of Internet users [24]. Our approach also suffers from sampling bias: The questionnaire was sent to a large group using mailing lists. The participants were not selected with the help of a systematic or a random process, but decided themselves to participate in the study. Because of this self-selection, it is possible that only those who are specifically interested in the topic have responded, making it difficult to generalize the results to the general population.

Manipulation: While this is not ethical, some survey participants may have deliberately falsified their responses to push their favorite or to harm a competing product as commercial and personal interests may be involved. Furthermore, participants may have submitted the survey multiple times to skew the results or they may hay have encouraged others to rate a score editor in a specific way. While we do not see obvious patterns of manipulation in our dataset, such manipulations can also not be ruled out completely.

Different target groups: Musical notation editors can be used for different tasks ranging from ideation to music engraving. Since each software has a different set of features, some tasks can be completed more efficiently respectively. This has an impact on the

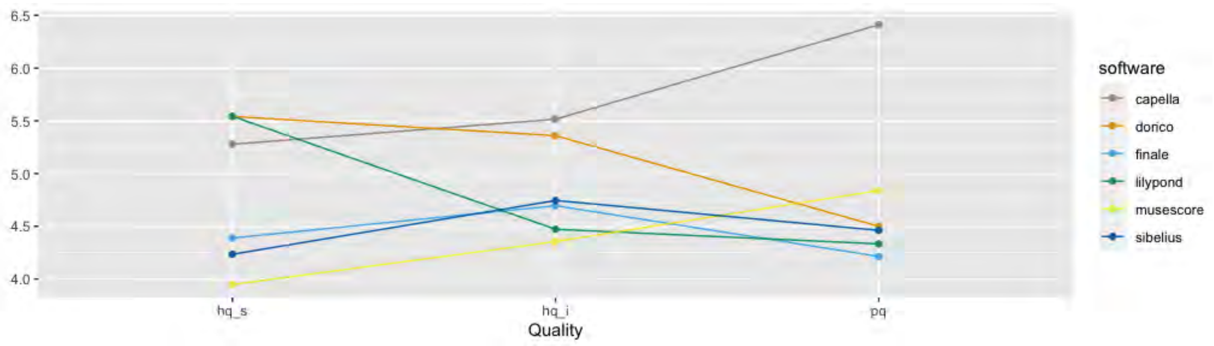


Figure 3. Median values for each quality per software. hq_s = Hedonistisc Quality: Stimulation, hq_i = Hedonistic Quality: Identity, pq = Pragmatic Quality.

Attribute		Df	Sum Sq	Mean Sq	F value	Pr(>F)	η^2
HQ_S		5	62.92	12.58	10.59	3.39e-08	0.72
	Residuals	101	120.00	1.19			
HQ_I		5	23.73	4.75	4.98	0.000405	0.50
	Residuals	101	96.25	0.95			
PQ		5	37.74	7.55	5.32	0.000223	0.51
	Residuals	101	143.37	1.42			

Table 3. ANOVA for AttrakDiff2

Attribute	group1	group2	estimate	conf.low	conf.high	p.adj	p.adj.signif
HQ_S	capella	musescore	-1.66	-2.92	-0.41	2.65e-03	**
	capella	sibelius	-1.33	-2.61	-0.05	3.67e-02	*
	dorico	finale	-1.51	-2.63	-0.39	2.12e-03	**
	dorico	musescore	-1.92	-2.89	-0.95	1.40e-06	****
	dorico	sibelius	-1.58	-2.59	-0.58	1.94e-04	***
	finale	lilypond	1.41	0.09	2.73	2.91e-02	*
	lilypond	musescore	-1.81	-3.01	-0.62	3.64e-04	***
	lilypond	sibelius	-1.48	-2.70	-0.25	8.57e-03	**
HQ_I	capella	musescore	-1.51	-2.63	-0.39	0.00228	**
	dorico	musescore	-1.16	-2.02	-0.29	0.00261	**
PQ	capella	dorico	-2.14	-3.64	-0.64	9.79e-04	***
	capella	finale	-2.47	-3.97	-0.97	8.13e-05	****
	capella	lilypond	-2.19	-3.87	-0.51	3.51e-03	**
	capella	musescore	-1.78	-3.14	-0.41	3.62e-03	**
	capella	sibelius	-2.14	-3.54	-0.74	3.17e-04	***

Table 4. Tukey tests for AttrakDiff2. Only groupings with any significance are shown.

target groups which might have more interest in fast note input, or fine grained layout formatting, etc. depending on proficiency and skill level of the user and purpose of the notation.

Long user history: In their responds participants might refer to a longer or shorter history of using a specific software thereby reflecting advantages and shortcomings which occurred over the years, maybe even changing software in this time period.

4.2 Metrics in detail

The ANOVA for SUS shows that there is no significant difference in usability among programs, except compared with Capella. It is surprising that half of the median scores did not even reach the threshold for usability of 68. Two other music notation editors barely exceeded that threshold, which also corresponds to a school mark of “D” as discussed in Section 3.1. However, having a score below the threshold does not mean, that the application is unusable as shown by a study correlating adjectives with SUS. The adjectives ranged from “worst immaginable” to “best immaginable” of which “OK” occupies the space from 50

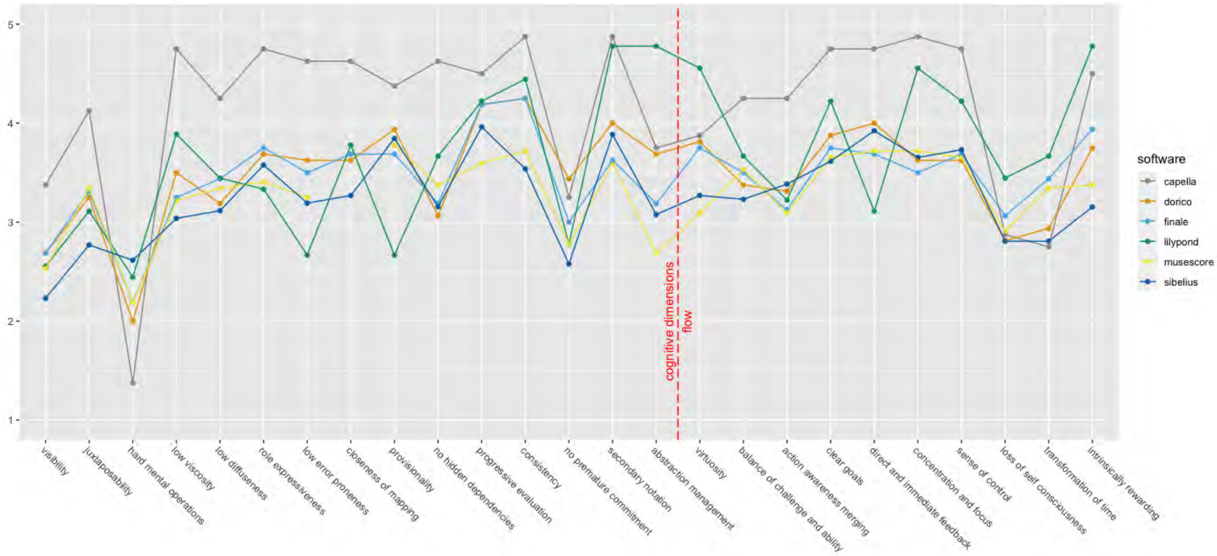


Figure 4. Mean values for each liveness item per software. The dashed line represents the border between items which are derived from cognitive dimensions and flow metrics.

Metric	p-value ANOVA	η^2	Groups Tukey test	p-value Tukey test
Consistency	0.00309	0.44	capella & sibelius capella & musescore	0.0111 0.0355
Secondary notation	0.00513	0.42	capella & musescore lilypond & musescore	0.0250 0.0333
Abstraction management	8.21e-06	0.6	dorico & musescore finale & lilypond lilypond & musescore lilypond & sibelius	2.05e-02 3.68e-03 4.90e-06 4.73e-04
Virtuosity	0.00498	0.42	lilypond & musescore lilypond & sibelius	0.00561 0.02730
Concentration and focus	0.00695	0.41	capella & finale capella & musescore capella & sibelius	0.0239 0.0478 0.0370
Intrinsically rewarding	0.000892	0.47	capella & sibelius lilypond & musescore lilypond & sibelius	0.03620 0.01290 0.00314

Table 5. On the left side, ANOVA tests are shown, if $p < .01$. On the right side, the respective Tukey tests if $p < .05$.

up to 68 [25] and thereby includes lower performing software in our study. It is important mentioning that the SUS was developed for assessing usability of the GUIs specifically. This could explain lower scores for LilyPond in general, since the resulting notation is based on coding text commands. The framing of our survey did not clearly distinguish between a GUI and the process of notating so that when answering the questions the participants relate to the LilyPond language and not the text editor.

There is a noticeable difference in the spread of data points in Figure 2 especially for Sibelius and Dorico, which can be interpreted as heterogenity in the user group. This might be also an effect of the sample size we had in this survey. The values of less mentioned music notation editors such as Capella and LilyPond is usually less spread out. One can assume that different user groups have different technical background. E.g. LilyPond is known to be used by users already familiar with $\LaTeX$or programming. It is not possible to verify such assumptions satisfactorily without qualitative research and knowledge about the circumstances in which the applications are used.

With AttrakDiff2 we wanted to expand the field of interest from pure usability to user experience and emotional responses connected to the application using PQ and SUS as a pivot attribute. In general, we can see that the ranking of the music notation editors matches in both metrics. The exceptional position of Capella in PQ is reflected by significance tests with results similar to SUS (see Table 2 and

Table 4).

Dorico, Capella and LilyPond are considered to be more stimulating than the other programs (see Figure 3). Especially Dorico and LilyPond employ interaction paradigms, which are not found in the other ones. Dorico can be used by opening popovers to create elements. Also notes can be inserted according to beat time divisions within the bar rather than on pre-existing rhythms notated as rests. Furthermore, Note input and layout are strictly separated by different views. The visual output of LilyPond can be completely controlled by manipulating the underlying text files. Future experimental developments might combine fluent transitions between text and GUIs.

Following the individual AttrakDiff2 trajectories for each application (Figure 3), Capella has a similar value for stimulation as Dorico, but overall it is considered more pragmatic than stimulating, which is also true for the trajectory of MuseScore. Finale and Sibelius are balanced over all qualities with slight tendencies towards identity. Noticeable differences from HQ_S to HQ_I is only seen in LilyPond (a drop of 1.1 from 5.5 to 4.4) whereas the remaining programs have differences of between 0.3 to 0.5. Identity and its corresponding items, as described in the original paper can be read in different ways [18]. On the one hand, being perceived by relevant others could be achieved by the software itself through direct communication or collaboration. On the other hand, using it might identify oneself as part of a community. Despite the relative low score, LilyPond is usually used to make scores of higher visual aesthetic quality, by having fine grained control over every visual element. This could be understood as a perceived need for creating shareable scores, contributing to identity.

For Liveness we found six metrics with significant differences, which will may helpful to inform future qualitative research and to examine central design differences. Especially Sibelius and MuseScore are considered to be significantly less consistent than Capella. Secondary notation is the interaction with all graphical objects, which are not the musical notation itself but are rather complementing it, like annotations and coloring. LilyPond and Capella both are rated very high for secondary notation compared to MuseScore. It is however not clear why LilyPond performed that good for secondary notation, since the possibilities of annotating the notation are restricted. However adding secondary notations like comments and formatting the text file is much more straight forward, which might explain the good score. Most significant differences in “abstraction management” are detected by comparison to LilyPond, which shows that knowledge of automated and aggregated actions are valued high for this product although they might be hard to learn. “Virtuosity” is only positively attributed for LilyPond, which is a hint of the value of skillfulness using the program. Capella is significantly different in “concentration and focus” compared to Finale, MuseScore and Sibelius.

Considering the trajectories in Figure 4, we can see again the strong overall performance of Capella. Here we would like to examine some peculiarities in the trajectory of the individual programs and where there are stronger deviations than in others. MuseScore and Sibelius have a very similar profile, while the trajectories of the other programs are much more diverse. For the most of the cognitive dimensions items Sibelius and MuseScore have similar values and are generally underperforming in “consistency”, “abstraction management” and “virtuosity” compared to the other programs. LilyPond tends to have lower values with items, which can be explained from it being text-based and the need to be compiled first, like having less feedback and being more prone to errors.

4.3 Correlating the results

By correlating 107 individual results we hope to find important features to investigate closer in future research. We gain a more general view on informative features by *not* grouping the individual results by music notation editor. As correlations that we consider important we decided so set a threshold at $|c| \geq 0.5$ (see Figure 5).

On first sight the negative correlations of “hard mental operations” stand out, which is an artifact because low values are more preferable in this case. SUS and PQ have the highest correlation, which supports our notion that these metrics are connected to the same attributes.

The values of “transformation of time” and “loss of self-consciousness” have mostly weak or no correlations. In turn, we can see more pronounced correlations in action-related metrics, rather than such describing mental states. In general, the matrix represents the connections of our measured and isolated features above by showing high transitive correlations. PQ, “clear goals”, SUS and “direct an immediate feedback” are the metrics with the most high ranking values. These are measures of control and they are highly correlated with other measures indicating control over the system. As argued by Csikszentmihalyi [19], an application should support automated behavior, which is based on habit and patterns, and always present its current UI state and what inputs are possible. In a state of flow one is absorbed in the task, and therefore has no resources to reflect on the current action. The outcomes must be based on ordered rules and non-contradictory actions to establish an unbroken experience. Transferred to music notation editors this means that notes and chords should be presented and played directly when selected. Also changing input modalities after inserting a note might lead to break of flow.

“Concentration and focus” represents an important pivot feature, which subsumes correlations with many metrics of flow (such as “clear goals” and “direct an immediate feedback”) and cognitive dimensions alike. It is also correlated with usability metrics, but it is mostly associated with pragmatic, rather than hedonistic qualities. This is also true, e.g., for “role-expressiveness”, which expresses whether the user can see how each component of a program relates to the whole [26, 23]: The purpose and condition of the musical structure should therefore be readily visible and the relationships should be easy to see. This can be achieved by overview or analysis functions to make see harmonic relationships (e.g. by proposing chord names) or by having a good view over all instruments while mak-

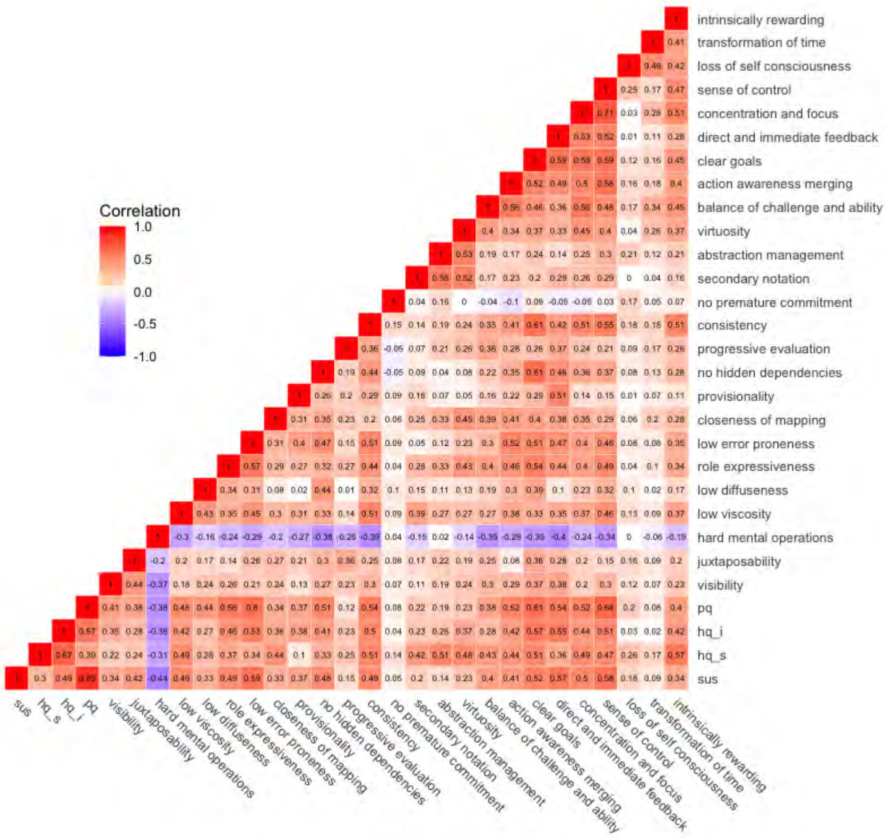


Figure 5. Correlation matrix of all individual results.

ing changes thus pointing to qualities of changing between views of the same score and reduction of distracting and irrelevant elements.

In Table 5 "Virtuosity" has mostly shown differences for LilyPond, but in general it is mostly associated with "secondary notation" and "abstraction management". "Virtuosity" is a widely used term in music and comparing it to HQ_S they both reflect skillfulness. When designing software that helps to develop skills in certain fields, be it composition, musicology, editing or music practice, one might need to identify more detailed or overlapping goals. This is not necessarily bound to musical skills, but also to the learnability of the system itself and knowing the tools and possibilities to create a score with creative needs in mind.

In summary, we consider the following metrics as the most informative for music score editors:

- SUS
- HQ_S
- Consistency
- Secondary notation
- Abstraction management
- Virtuosity
- Concentration and Focus
- Role expressiveness

5. CONCLUSION & FUTURE WORK

In this paper we examined usability and user experience of music notation editors with the help of an online survey with standardized questionnaires. Of course there are limitations to such an approach (sampling bias, manipulation, different target groups and long user history), but this study provides a first starting point for a scientific examination of existing music notation editors. Almost all examined music notation editors show weak results in usability as measured by the SUS. The exception is Capella that achieved a school grade of "A" and an adjective rating of "Excellent" according to the rating scheme of Bangor et al. [25].

User experience questionnaires, like AttrakDiff2 [18] and the liveness questionnaire by Nash et al. [13], examine further important aspects of working with music notation editors. While the pragmatic quality of the AttrakDiff2 was highly correlated with SUS (as expected), the hedonistic qualities stimulation (HQ_S) and identity (HQ_I) were strongly differentiating factors between the programs. It seems that non-standard interaction paradigms such as text input can lead to higher values in both hedonistic qualities. Comparing results for the AttrakDiff2 and liveness metrics, MuseScore and Sibelius are very similar.

We correlated all individual results. Metrics with implications for action and control correlated relatively strongly with each other as opposed to metrics that fit better in creative contexts.

Our study of usability and user experience of music notation editors, can help to create experimental music notation programs. Mixtures of different interaction paradigms, like those used in Dorico and LilyPond, could be developed. Future music notation editors could even be adapted by gradually moving from one paradigm to the other. We also expect interesting insights from cognitive and eye-tracking studies. This could be beneficial to assess specific interaction patterns, also in comparison to non-digital music writing [16].

6. REFERENCES

[1] N. Bevana, J. Kirakowski, and J. Maissela, "What is Usability?" in *Proceedings of the 4th International Conference on HCI*, 1991, pp. 1–6.

[2] T. L. Roberts and T. P. Moran, "The evaluation of text editors: methodology and empirical results," *Communications of the ACM*, vol. 26, no. 4, pp. 265–283, 1983. [Online]. Available: https://doi.org/10.1145/2163.2164

[3] N. S. Borenstein, "The evaluation of text editors: a critical review of the Roberts and Morgan methodology based on new experiments," *ACM SIGCHI Bulletin*, vol. 16, no. 4, pp. 99–105, 1985. [Online]. Available: https://doi.org/10.1145/1165385.317475

[4] S. K. Card, T. P. Moran, and A. Newell, "The keystroke-level model for user performance time with interactive systems," *Communications of the ACM*, vol. 23, no. 7, pp. 396–410, 1980.

[5] L. A. Hiller and R. A. Baker, "Automated music printing," *Journal of Music Theory*, vol. 9, no. 1, pp. 129–152, 1965. [Online]. Available: http://www.jstor.org/stable/843151

[6] D. A. Byrd, "Music notation by computer," Ph.D. dissertation, Indiana University, 1984.

[7] G. Loy and C. Abbott, "Programming languages for computer music synthesis, performance, and composition," *ACM Computing Surveys*, vol. 17, no. 2, p. 235–265, 1985. [Online]. Available: https://doi.org/10.1145/4468.4485

[8] R. B. Dannenberg, "Music representation issues, techniques, and systems," *Computer Music Journal*, vol. 17, no. 3, pp. 20–30, 1993. [Online]. Available: http://www.jstor.org/stable/3680940

[9] D. V. Oppenheim, "The Need for Essential Improvements in the Machine-Composer Interface used for the Composition of Electroacoustic Computer Music," in *Proceedings of the International Computer Music Conference (ICMC)*, 1986.

[10] ——, "The P-G-G Environment for Music Composition: A Proposal," in *Proceedings of the 1987 International Computer Music Conference (ICMC)*, 1987, pp. 40–48.

[11] M. S. Puckette, "The Patcher," in *Proceedings of the International Computer Music Conference (ICMC)*, 1988.

[12] C. Nash and A. Blackwell, "Tracking Virtuosity and Flow in Computer Music," in *Proceedings of the International Computer Music Conference (ICMC)*, 2011.

[13] ——, "Liveness and Flow in Notation Use," in *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*, 2012. [Online]. Available: http://www.nime.org/proceedings/2012/nime2012_217.pdf

[14] ——, "Flow of Creative Interaction with Digital Music Notations," in *The Oxford Handbook of Interactive Audio*, H. T. Kasey Collins, Bill Kapralos, Ed. Oxford University Press, 2014. [Online]. Available: https://doi.org/10.1093/oxfordhb/9780199797226.013.023

[15] S. Hunt, T. Mitchell, and C. Nash, "A Cognitive Dimensions Approach for the Design of an Interactive Generative Score Editor," in *Proceedings of the International Conference on Technologies for Music Notation and Representation (TENOR'18)*, 2018, pp. 119–127.

[16] J. Peterson and E. Schubert, "Music Notation Software: Some Observations on its Effects on Composer Creativity," in *Proceedings of Intercational Conference on Music Communication Science (ICoMCS)*, vol. 127–130, 2007.

[17] J. Brooke, "SUS: A Quick and Dirty Usability Scale," *Usability Evaluation in Industry*, vol. 189, no. 194, pp. 4–7, 1996.

[18] M. Hassenzahl, M. Burmester, and F. Koller, "AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität," in *Mensch Computer 2003: Interaktion in Bewegung*, 2003, pp. 187–196.

[19] M. Csikszentmihalyi, "Flow and the psychology of discovery and invention," *Harper Perennial*, vol. 39, 1997.

[20] A. Bangor, P. Kortum, and J. Miller, "The system usability scale (SUS): An empirical evaluation," *International Journal of Human-Computer Interaction*, vol. 24, no. 6, pp. 574–594, 2008.

[21] J. Sauro and J. R. Lewis, *Quantifying the user experience: Practical statistics for user research*. Morgan Kaufmann, 2016.

[22] C. Nash, "The Cognitive Dimensions of Music Notations," in *Proceedings of the First International Conference on Technologies for Music Notation and Representation (TENOR'15)*, 2015, pp. 190–202.

[23] T. R. G. Green and M. Petre, “Usability analysis of visual programming environments: A ‘cognitive dimensions’ framework,” *Journal of Visual Languages & Computing*, vol. 7, no. 2, pp. 131–174, 1996. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1045926X96900099

[24] M. Van Selm and N. W. Jankowski, “Conducting online surveys,” *Quality and quantity*, vol. 40, no. 3, pp. 435–456, 2006.

[25] A. Bangor, P. Kortum, and J. Miller, “Determining what individual SUS scores mean: Adding an adjective rating scale,” *J. Usability Studies*, vol. 4, no. 3, pp. 114–123, 2009.

[26] A. Blackwell, C. Britton, A. Cox, T. Green, C. Gurr, G. Kadoda, M. Kutar, M. Loomes, C. Nehaniv, M. Petre, C. Roast, C. Roe, A. Wong, and R. Young, “Cognitive Dimensions of Notations: Design Tools for Cognitive Technology,” in *Cognitive Technology: Instruments of Mind*, M. Beynon, C. L. Nehaniv, and K. Dautenhahn, Eds. Springer-Verlag, 2001, pp. 325–341.

BP Sequencer: A Low-Barrier-to-Entry Assessment Tool for Musical Creativity

Psyche Loui
Northeastern University
Boston, USA
p.loui@northeastern.edu

ABSTRACT

This paper introduces a newly designed sequencer of music in the Bohlen-Pierce (BP) scale to assess creative perception and cognition. We begin with a brief overview of scientific studies on musical creativity, leading up to a gap in the field that is unaddressed by traditional studies that compare different forms of musical training and instrument-dependent output for understanding creativity. We then introduce the novel BP sequencer, an experimental interface that affords generating and rating the creativity of novel musical sequences that are uniquely composed in the Bohlen-Pierce scale. We then report three preliminary experiments in which we quantify the number of sequences generated by each individual and isolate the musical-informatic features that were rated as more creative. Participants showed a wide range of creativity in generating novel sequences. Sequences that were rated as more creative were generally longer, had more unique pitches, and had more different interval sizes. Furthermore, a preliminary electrophysiological (EEG) study quantifies three distinct candidate biomarkers for the perception of musical creativity. This novel sequencer of Bohlen-Pierce scale music can provide a useful tool for assessing creative perception and cognition. The code for this tool is freely available, along with video tutorials and documentation.

1. INTRODUCTION

Techniques for the notation, representation, and visualization of music and sound are inextricably linked to the human understanding of musical structure within their broad contexts. These understandings include the cognitive representations that the human mind develops in response to continuous exposure to perceptual input with its statistical regularities in one's environment, that give rise to understanding of musical structure and creative interpretations of said structure in the form of musical outputs. Musical outputs that feature improvisation may, because of their inherently quick-changing nature, require different forms of notation from musical output that are through-composed; this is related to the different forms of mental representation that musicians who practice improvisation may have, relative to musicians who primarily perform through-composed music.

2. STUDIES IN MUSICAL CREATIVITY

Studies in music cognition that compare musicians with different levels of improvisation training have focused on western classical and jazz improvisation genres, mainly due to its relative ease of access to researchers who are conducting neural and cognitive studies of improvisation in lab-based, relatively experimentally controlled environments. In one representative previous study, 38 young adult participants with varying degrees of musical improvisation training completed a novel improvisation continuation task and underwent Magnetic Resonance Imaging (MRI) scans to relate the creativity of musical improvisations to brain structure [1]. In the improvisation continuation task, participants were tested in a lab-based environment with a computer and a MIDI keyboard. The computer presented pre-recorded short musical motifs, and then participants were asked to continue the short motifs on the keyboard and then improvise and extend to the motifs (instructions were to "riff off of" the motifs). Recorded performances were then rated by expert jazz instructors for creativity. Voxel-based morphometric analyses on T1 data showed that creativity ratings were negatively associated with gray matter volume in the right inferior temporal gyrus and bilateral hippocampus. Furthermore the duration of improvisation training, which was significantly correlated with creativity ratings, was negatively associated with gray matter volume in the rolandic operculum. Although experience with a keyboard was positively correlated with creativity ratings, this study notably introduced an improvisation task that could be performed even by participants with no specific musical training. The observation of brain structures that correlate with improvisation experience informs us that the capacity for creativity may be measurable at the level of the brain, and relatable to improvisatory behavior. On the other hand, these anatomical measures are relatively stable over time, changeable only over long periods of practice throughout the lifespan [2]–[4]. More time-sensitive measures of brain activity are needed to observe moment-by-moment fluctuations in the perception and production of creative output. In that regard, another previous study compared the electrical brain potentials (using the electroencephalogram, or EEG) of classical and jazz-trained musicians as they listened to Western chord progressions that were either highly expected (I-IV-V-I), slightly unexpected (I-N6-V-I), or highly unexpected (I-IV-V-N6) [5]. This study found that while jazz-trained musicians, with their well-learned improvisation strategies, were more sensitive to slight harmonic expectancy violations early on in the perceptual pathway (around 100-200 milliseconds after the onset of the unexpected chord), they quickly integrated the unexpected chords into the ongoing musical contexts such that there was less error-related brain activity, as evidenced by a return to baseline in brain potentials by around 800 milliseconds after the unexpected chord onset. In contrast, classical musicians were less sensitive to unexpected chords at 100-200 ms, but showed a persistent significant brain potential difference at 800 ms after the onset of the highly unexpected chord. These results suggest that expectations and contextual information are crucial in the mental representation of musical structure, and that different types of training and enculturation within a musical genre can completely alter the temporal cascade of neural events that give rise to the perception of musical structure.

Using similar techniques of electrical brain potentials coupled with behavioral testing, other studies have shown that differences in mental representation between classical and jazz musicians are localized to the brain mechanisms that generate motor patterns that subserve commonly expected and unexpected chord progressions [6], and are influenced by the cognitive representation of functional categorization on the basis of musical structure as well as on the basis of motoric representations of musical chords and chord progressions [7]. Taken together, this line of research in the neuroscience of music suggests that creativity, as operationally defined by training in improvisation through jazz musical training, involves time-sensitive categorization and reconstruction of functional categories of musical sounds as filtered by the musician's training and experiences. In that sense, studying the effects of training in improvisation is useful for creativity researchers, as improvisation is a form of real-time creativity that can be quantitatively studied [8].

3. CHALLENGES AND MOTIVATIONS BEHIND PRESENT RESEARCH

Although there is much to be learned from these studies, many open questions remain about the nature of creativity that are yet unaccounted for by this line of work. One major limitation lies in the operational definition of creativity as that which is encouraged by jazz training. While common-practice jazz pedagogy often does rely upon improvisation as a core part of the curriculum, this form of training is only available to a select few, and only represents a small part of the broad and diverse musical experiences found around the world. If our goal is to understand creativity through real-time generative musical processes that can reveal some information about our mental representations of musical structures, then studying jazz improvisation only represents a very narrow part of the many possible mental representations.

A more scalable and broadly applicable approach may be to design user-friendly interfaces for musical creativity that rely on minimal training, and can quickly become conducive to studies in the perception and cognition of musical structures as they are being created in real time. To maximize accessibility for use in music cognition studies, and to lower the barrier of entry for individuals with no specific musical training, we make use of novel tuning systems that are different from the world's commonly used musical systems. Specifically, we use the Bohlen-Pierce (BP) scale, a thirteen-tone macrotonal scale that differs from the world's scale structures: while the majority of musical scales around the world rely on octave equivalence [9], with octave being a 2:1 ratio in frequency, the BP scale makes use of the tritave, which is a 3:1 ratio in frequency [10]. Previous innovations in musical notation (presented at TENOR) have considered expansions for non-standard musical systems such as BP, such as by introducing dynamic notation [11]. In the same spirit, here we aim to create a training-independent common ground that affords studies in music perception and cognition, specifically neurocognitive studies of musical creativity, that are relatively independent of musical culture and genre-specific training.

Here we introduce the BP sequencer, a platform-independent tool to assess musical creativity. We describe core features of the user interface and the types of data collection that it affords as a research tool. We also describe variations on the core interface that allow for separate studies in the perception and cognition of creativity in this nontraditional musical system. Finally, we describe a preliminary EEG study that captures several candidates for the neural correlates of creativity as operationalized by listeners' ratings of creativity from musical data obtained using the BP sequencer.

4. THE BP SEQUENCER

The Interface: Inspired by Max/MSP's live.step object and interface, we have created a user interface that iteratively and interactively generates loops, or musical sequences in the Bohlen-Pierce scale, that is intuitive to use with no background in music theory and no previous training, while also being a logical extension of the Improvisation Continuation Task reviewed above for quantifying creativity in musical improvisation [1]. The interface automatically generates tone sequences with fundamental frequencies ranging from 440 Hz to 1320 Hz in 13 possible steps as determined by the Bohlen-Pierce scale [10]. The interface generated tone sequences with a fixed inter-onset interval of 148 ms. Each sequence could be of variable length as determined by the participant. The structure of the sequence (i.e. the specific sequence of pitches in each loop) is repeated continuously, and can be set by the participant in real time using a visual interface as shown in Figure 1.

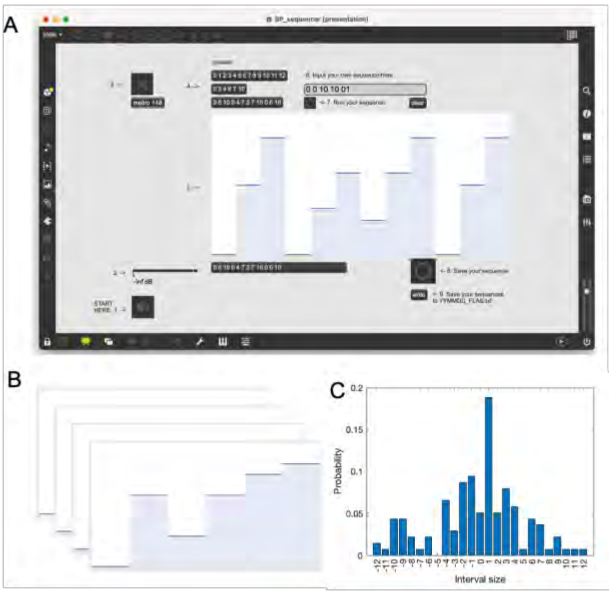


Figure 1. BP sequencer. **A:** Experiment interface for Sequence Generation Task. **B:** Sequences can be concatenated across participants to compute histograms (**C**) from which interval entropy will be computed.

5. EXPERIMENT 1: SEQUENCE PRODUCTION TASK: GENERATING CREATIVE OUTPUT

In a first experiment using this interface, young adult participants from Northeastern University (n = 11) participated in return for course credit. The participant's task is to generate as many creative sequences as possible within five minutes, where a creative sequence is defined as both original and appropriate [12]. Unlike the tasks typically used to measure creativity in music (e.g. improvisation at the keyboard), this sequence generation task is easy and intuitive to individuals across cultures without specialized training. Each participant's generated sequences are saved in a text file, with each sequence represented as a string of integers from 0 to 12 corresponding to the 13 possible steps in the Bohlen-Pierce scale. Sequence files can then be concatenated across all participants to form a larger database of creative outputs (Figure 1B).

These data can be used to extract the following measures from each participant's sequence file for between-subjects analyses: fluency as the number of unique sequences generated by each participant, entropy of interval sizes generated by each participant as a measure of originality, and information content of pitch classes as a measure of surprise (computed from histogram of interval sizes, Figure 1C). Fluency and originality have been previously shown by our lab and others to be related to the neural markers of creativity [13], whereas entropy and information content are known markers of uncertainty and surprise that have been linked to prediction and reward in fMRI and electrophysiological studies [14]–[16].

We expect that participants who generate higher fluency, entropy, and information content will also score higher on fluency and entropy on other non-musical creativity tasks in the lab, such as the Alternative Uses Task [17], supporting the hypothesis that participants who are more creative will generate sequences with higher interval entropy and information content.

6. EXPERIMENT 2: SEQUENCE RATINGS TASK: PERCEPTION OF CREATIVITY

Sequence Rating Task: In a second experiment, another group of young adult participants (n = 10) rated the sequences generated by the first group for creativity. Using a visual slider, participants are asked to rate each sequence for creativity, originality, and appropriateness. Participants are allowed to rate as many sequences as they can within five minutes.

Creativity ratings for each sequence were significantly correlated with the number of unique pitches used in the sequences (r = 0.31, Figure 2A). Among sequences that used all possible pitches, sequences with large intervals and more entropy and information content resulted in a higher creativity rating than sequences with smaller intervals and low information content. For example, Figure 2B and 2C shows two sequences that were generated by participants from Experiment 1. Both sequences used the maximal number of unique pitches (13 in the BP scale). However, the sequence in Figure 2B has larger intervals and a higher entropy of interval sizes, and is rated higher by listeners in Experiment 2. This is consistent with prior research on jazz improvisation [1], [13], [18], in which participants improvised on given sequences on a piano keyboard, and the improvised sequences were subsequently rated by jazz instructors.

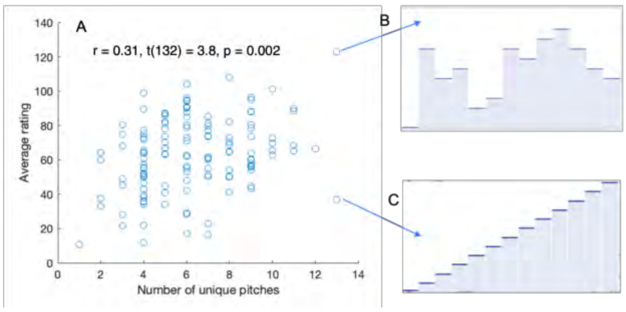


Figure 2. A-C. Preliminary results from Sequence Generation Task.

7. EXPERIMENT 3: EEG SIGNATURES OF CREATIVITY FROM BP SEQUENCER DATA

Experiment 3 aims to relate the perception of sequences generated from the BP sequencer to neural markers of creativity in EEG. Participants were young adults recruited from Northeastern University (n = 6). Each participant listened to each sequence generated from Experiment 1, and rated it for creativity (originality and appropriateness) using the same interface as the Sequence Ratings Task in Experiment 2, with EEG triggers (event time tags) generated by Max/MSP and recorded with the BrainVision system. EEG was recorded with a 64-channel BrainVision actiCHamp system with PyCorder software in a sound-attenuated and electrically-shielded chamber. EEG data was sampled at 1000 Hz and filtered using .5 Hz high pass filter and 60 Hz notch filter for electrical noise. EEG was re-referenced to electrode channels TP9 and TP10, which are relatively stable and commonly used for auditory EEG studies, and corrected for ocular artifacts using Independent Components Analysis (ICA) consistent with most auditory EEG data acquisition protocols [19]. Preprocessing and analyses were done in Matlab with EEGLAB toolbox [19].

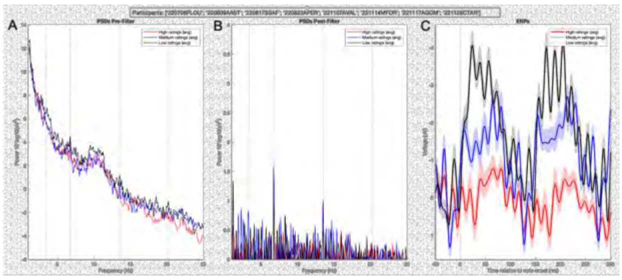


Figure 3. Preliminary results from EEG study in **A.** Frequency-domain, **B.** Steady-state evoked potential, and **C.** Event-Related Potential (time-domain) analyses.

Frequency-based: To assess neural entrainment, power spectral densities (PSDs) were computed over the full length of each EEG channel using Welch's procedure (pwelch in MATLAB) with a 1600 sample window and a 8000 sample overlap. Lower alpha power and higher SSEP at the first harmonic were observed during sequences that received the highest creativity ratings (tertile split) (Figure 3A), suggesting less boredom and higher entrainment to the stimulus.

Steady-state evoked potential (SS-EP): As the frequency-tagging technique has previously been used to assess neuronal entrainment to beat and meter [20], we isolated EEG activity in the form of steady state evoked potentials (SS-EPs) specifically at the note rate, which is set at 148 ms per note (6.76 Hz) for our perception task. Starting from the frequency-based results from above, the broad-band frequency activity was removed by subtracting, at each bin of the frequency spectra, the average amplitude measured at neighboring frequency bins (two frequency bins ranging from −0.15 to −0.09 Hz and from +0.09 to +0.15 Hz relative to each frequency bin). A prominent SSEP was observed at the note rate of 6.76 Hz (corresponding to the inter-onset interval of 148 ms) and its harmonic of 13.5 Hz (Figure 3B), confirming that our task elicits robust neural entrainment to the beat frequency.

Event-Related Potential (ERP): In addition to assessing frequency-based EEG markers of creativity, we computed Event-Related Potentials (ERPs, or electrical fluctuations that are uniquely related to the event of interest [21]) to tones in sequences that were rated as highly creative against sequences that were rated as medium in creativity. Figure 3C shows ERPs of the highest-rated tertile of trials (red) against the medium-rated tertile (black). A persistent difference in ERP was observed between high, medium, and low ratings, from less than 10 ms after tone onset throughout the duration of the time window being analyzed (Figure 3A). Sequences that were rated low in creativity showed the most negative waveforms, whereas sequences that were rated high in creativity showed the least negative waveforms. Negative waveforms during auditory processing, also known as the processing negativity [22], have been identified as the mismatch negativity (MMN), a negative waveform observed in response to unexpected or unpredicted tones given a context [23]–[25]. Results are consistent with the link between unexpectedness and creativity: while highly predictable sound sequences may be perceived as uninspiring, and highly unpredictable sound sequences may be perceived as negative prediction error (in the sense that the participant's mind is unable to form a predictive model of how the sequence would go), it is the slightly unexpected sequences that elicit positive prediction errors, and are linked to higher creativity ratings. These neural markers can thus be coupled with behavioral testing and music technology to enable a better understanding of musical creativity.

8. CONCLUSIONS

We designed a unique user interface for collecting data on musical creativity, using the non-standard tuning system of the Bohlen-Pierce scale. In three experiments, we begin to address questions related to musical creativity as it relates to originality and appropriateness, and we relate the perception of musical creativity to brain-based measures as characterized by EEG. The BP scale is not a commonly used musical system in any culture; thus a sequencer that is built on the BP scale is equally accessible to those with varying levels of experience in different musical systems around the world, rendering it similarly useful for someone from the Western musical tradition and from other traditions, such as Chinese pentatonic or Indonesian heptatonic musical traditions. Furthermore, since Max/MSP is a platform-independent programming language with free runtime versions, a Max/MSP-based user interface that is quick to learn can be scalable in the future for massive online studies. The current studies, though using only small sample sizes, are a step towards this goal. Taken together, we believe that the BP sequencer is a platform-independent, relatively training-independent tool that lowers the barrier-to-entry in neural and cognitive assessments for musical creativity. The code for these tasks are freely available at https://github.com/mind-lab-bos/BPsequencer, along with written instructions and video tutorials.

9. REFERENCES

[1] C. Arkin, E. Przysinda, C. W. Pfeifer, T. Zeng, and P. Loui, "Gray Matter Correlates of Creativity in

Musical Improvisation," *Front Hum Neurosci*, vol. 13, p. 169, 2019, doi: 10.3389/fnhum.2019.00169.

[2] G. F. Halwani, P. Loui, T. Rüber, and G. Schlaug, "Effects of practice and experience on the arcuate fasciculus: comparing singers, instrumentalists, and non-musicians," *Front. Psychol.*, vol. 2, p. 156, Jul. 2011, doi: 10.3389/fpsyg.2011.00156.

[3] E. Moore, R. S. Schaefer, M. E. Bastin, N. Roberts, and K. Overy, "Diffusion tensor MRI tractography reveals increased fractional anisotropy (FA) in arcuate fasciculus following music-cued motor training," *Brain Cogn.*, vol. 116, pp. 40–46, Aug. 2017, doi: 10.1016/j.bandc.2017.05.001.

[4] H. Johansen-Berg, "Behavioural relevance of variation in white matter microstructure," *Current Opinion in Neurology*, vol. 23, no. 4, 2010, [Online]. Available: https://journals.lww.com/co-neurology/Fulltext/2010/08000/Behavioural_relevance_of_variation_in_white_matter.3.aspx

[5] E. Przysinda, T. Zeng, K. Maves, C. Arkin, and P. Loui, "Jazz musicians reveal role of expectancy in human creativity," *Brain and cognition*, vol. 119, pp. 45–53, Dec. 2017, doi: 10.1016/j.bandc.2017.09.008.

[6] R. Bianco, G. Novembre, P. E. Keller, A. Villringer, and D. Sammler, "Musical genre-dependent behavioural and EEG signatures of action planning. A comparison between classical and jazz pianists," *NeuroImage*, vol. 169, pp. 383–394, 2018.

[7] A. Goldman, T. Jackson, and P. Sajda, "Improvisation experience predicts how musicians categorize musical structures," *Psychology of Music*, p. 0305735618779444, 2018, doi: 10.1177/0305735618779444.

[8] P. Loui, "Rapid and Flexible Creativity in Musical Improvisation: Review and a Model," *Annuals of the New York Academy of Sciences*, vol. 1423, no. 1, pp. 138–145, 2018.

[9] J. Mcbride and T. Tlusty, "Cross-cultural data shows musical scales evolved to maximise imperfect fifths," *arXiv: Sound*, 2019.

[10] M. V. Mathews, J. R. Pierce, A. Reeves, and L. A. Roberts, "Theoretical and experimental explorations of the Bohlen-Pierce scale," *J Acoustical Soc Am*, vol. 84, pp. 1214–1222, 1988.

[11] G. Hajdu, "Dynamic Notation–A Solution to the Conundrum of Non-Standard Music Practice," presented at the Proceedings of the International Conference on Technologies for Music Notation and Representation (TENOR'15), Paris, France, 2015, pp. 241–248.

[12] M. A. Runco and G. J. Jaeger, "The standard definition of creativity," *Creativity Research Journal*, vol. 24, no. 1, pp. 92–96, 2012, doi: 10.1080/10400419.2012.650092.

[13] A. Belden, T. Zeng, E. Przysinda, S. A. Anteraper, S. Whitfield-Gabrieli, and P. Loui, "Improvising at rest: Differentiating jazz and classical music training with resting state functional connectivity," *NeuroImage*, vol. 207, p. 116384, Nov. 2019, doi: 10.1016/j.neuroimage.2019.116384.

[14] V. K. M. Cheung, P. M. C. Harrison, L. Meyer, M. T. Pearce, J. D. Haynes, and S. Koelsch, "Uncertainty and Surprise Jointly Predict Musical Pleasure and Amygdala, Hippocampus, and Auditory Cortex Activity," *Curr Biol*, vol. 29, no. 23, pp. 4084–4092, Nov. 2019, doi: 10.1016/j.cub.2019.09.067.

[15] G. M. Di Liberto *et al.*, "Cortical encoding of melodic expectations in human temporal cortex," *Elife*, vol. 9, Mar. 2020, doi: 10.7554/eLife.51784.

[16] D. Omigie *et al.*, "Intracranial Recordings and Computational Modeling of Music Reveal the Time Course of Prediction Error Signaling in Frontal and Temporal Cortices," *J Cogn Neurosci*, vol. 31, no. 6, pp. 855–873, Jun. 2019, doi: 10.1162/jocn_a_01388.

[17] E. P. Torrance, "Examples and Rationales of Test Tasks for Assessing Creative Abilities," *The Journal of Creative Behavior*, vol. 2, no. 3, pp. 165–178, 1968, doi: 10.1002/j.2162-6057.1968.tb00099.x.

[18] T. Zeng, E. Przysinda, C. Pfeifer, C. Arkin, and P. Loui, "White Matter Connectivity Reflects Success in Musical Improvisation," *bioRxiv*, 2018, doi: 10.1101/218024.

[19] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of Neuroscience Methods*, vol. 134, pp. 9–21, 2004.

[20] S. Nozaradan, I. Peretz, and A. Mouraux, "Selective Neuronal Entrainment to the Beat and Meter Embedded in a Musical Rhythm," *Journal of Neuroscience*, vol. 32, no. 49, pp. 17572–17581, 2012, doi: 10.1523/jneurosci.3203-12.2012.

[21] P. Loui, "Neuroscience of Musical Improvisation," in *Handbook of Artificial Intelligence for Music: Foundations, Advanced Approaches, and Developments for Creativity*, E. R. Miranda, Ed. Cham: Springer International Publishing, 2021, pp. 97–115. doi: 10.1007/978-3-030-72116-9_5.

[22] R. Naatanen, M. Simpson, and N. E. Loveless, "Stimulus deviance and evoked potentials," *Biol Psychol*, vol. 14, no. 1–2, pp. 53–98, Feb. 1982.

[23] R. Naatanen and K. Alho, "Mismatch negativity--a unique measure of sensory processing in audition," *Int J Neurosci*, vol. 80, no. 1–4, pp. 317–37, 1995.

[24] S. Koelsch, P. Vuust, and K. Friston, "Predictive Processes and the Peculiar Case of Music," *Trends in cognitive sciences*, vol. 23, no. 1, pp. 63–77, Nov. 2018, doi: 10.1016/j.tics.2018.10.006.

[25] P. Vuust, O. A. Heggli, K. J. Friston, and M. L. Kringelbach, "Music in the brain," *Nature Reviews Neuroscience*, Mar. 2022, doi: 10.1038/s41583-022-00578-5.

Acknowledgments

We acknowledge funding support from NIH R01AG078376, NIH R21AG075232, NSF-CAREER 1945436, and NSF 2240330 to PL. We thank lab members Anjali Asthagiri, Jethro Lee, Catherine Zhou, Kristina Abyad, Carly Monson, Ayla Hadley, Corinna Parish, Eva Wu, and Parker Tichko for data collection and technical support.

NOTATING EXPERIENCES: A NEW SYSTEM FOR VISUAL DOCUMENTATION IN INSTRUMENT DESIGN

Ciaran Frame
SensiLab, Monash University
Melbourne, Australia
ciaran.frame@monash.edu

ABSTRACT

In the community of Digital Musical Instruments (DMIs), documentation surrounding iterative interactions and the creation of mappings is largely absent from DMI projects beyond the recording of a performance or subsequent evaluation with a performer. This is because the performance or interactive experience with the instrument is often viewed as the end point for a DMI project, and the description of a mapping or open-sourcing of software considered the 'score'. This paper outlines the creation of a visual notation based on unique interactions with the AirSticks, a gestural musical instrument. These notations expand on the concept of descriptive notation, creating a form of retrospective score and record-keeping for instrument designers. By capturing the intimate experiences and musical collaborations that contribute to the iterative design of the instrument, it is concluded the notation system provides an avenue for critical analysis that will aid the further development of DMIs.

1. BACKGROUND

1.1 Documentation for Digital Musical Instruments

The need for longevity in Digital Musical Instruments (DMIs) has been highlighted in the instrument design community by the likes of Calegario et al. as an important and often overlooked factor in the design process [1].

There is a shared feeling in DMI communities that the large number of new interfaces presented each year will only result in a few instances of generalised, replicable instruments [2], with the rest falling into the category of what Calegario et al. refer to as "...once-interesting but now-unplayable interfaces..." [1]. Replication thus serves as an important design consideration to validate conclusions drawn across the community and to ensure instruments are played by different people over time.

Increased use of documentation [1] and published construction and design processes [3] are often cited as examples of processes that might aid in the longevity of instruments. However, these processes of documentation continue to be rarely used by instrument designers, and when present, usually take the form of code repositories or raw recordings of a performance [1], both of which are often seen as the end-point of the instrument design process.

Notation presents itself as one overlooked opportunity that might be used to document interactions with Digital Musical Instruments in new ways.

1.2 The AirSticks Community

The AirSticks (seen in Figure 1) are one example of a DMI, with a community of programmers, hardware and software designers, composers, performers and improvisers working around it. They are a custom-designed DMI for gestural music making, combining Bluetooth Low-Energy technologies to give low-latency wireless control over MIDI or OSC, reconnecting movement, sound and visuals with the transparency and expressiveness of acoustic instruments.

The AirSticks community has created a wide variety of musical pieces, interactions and experiences for a diverse array of players. In particular, we are interested in interrogating the iterative and collaborative nature of our design, which draws deeply from practice-based research methods.

Figure 1. The AirStick, a Digital Musical Instrument.

In the context of an instrument used by many musicians in different ways, a key challenge for us has been capturing and notating these diverse experiences beyond the standardised documentation outlined in Section 1.1.

We first draw on past work to investigate what notation might look like in this context.

2. RELATED WORK

2.1 Prescriptive notation

In his book *Sonic Writing*, Thor Magnusson describes a dichotomy of musical scores – "descriptive" and "prescriptive" [4]. Prescriptive scores are those which provide instructions on what to 'do', prescribing actions that may not necessarily reflect the sonic result.

Electronic music has a rich history of prescriptive notation ranging from sets of text instructions (such as Steve Reich's *Pendulum Music* [5]) to code scores [6] to prescribing parameters like dynamics and reverb [7].

In the field of DMIs, prescriptive notation has typically been used to specify how a performer should move to create sound, such as describing interactions with a touch interface [8] or motion controller [9].

Whilst we can draw from vocabularies of prescriptive DMI notation, we seek a notation that describes, not prescribes – that is, instead of creating notation *for* a participant, we aim to create notation *describing* the movements and sound produced while using the AirSticks.

2.2 Descriptive notation

Descriptive notation represents the sonic outcome of a work, typically used for analysis or discussion purposes.

This category of notation has been explored widely in electronic music, whether it be for archiving the creative process and knowledge captured in composing electronic and electro-acoustic works [10] or adding extra detail around a electro-acoustic score [11].

Section 1.1 discussed the lack of documentation in the field of DMIs, and this conclusion can be extended to the lack of descriptive notation. Of course, documentation can be said to be a form of descriptive notation, with examples like documenting creative developments [12] or publishing a clear design processes [3].

Notations that capture the design process and development of a DMI are descriptive in the sense that they provide insight for analysis and discussion around the musical works for and interactions with the instruments, but are not linked to the outcomes of interactions with DMIs – what Small would describe as the process of 'musicking' [13].

This has meant that the question of 'what happened' in interactions with Digital Musical Instruments has been left largely unanswered by notation, instead taking the form of video or audio recordings.

2.3 Describing experience?

Can we expand the notion of the descriptive score, notating what is *heard*, to a broader definition of notating what *happens*?

Magnusson has aptly described Digital Musical Instruments as "epistemic tools" [14] – experience-driven technology that generates its own ways of understanding and communicating musical knowledge. This is at odds with notation and documentation centering around creative developments and processes, which capture the design experience, but often not the playing experience.

In the same way a descriptive electro-acoustic score (such as Luening and Ussachevsky's *Incantation* [15] shown in Figure 2) might be used for analysis and discussion, so too might a descriptive score that captures movement and sound in tandem be used to unpack interactions with a gestural musical instrument like the AirSticks.

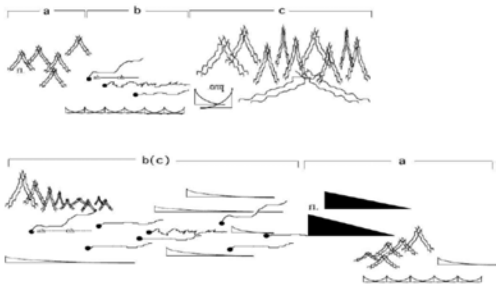


Figure 2. Score from *Incantation* (1953) [15] in [16].

Notation offers us the opportunity to capture and share specific musical work created for individuals in a way that embodies the original interaction, encompassing not just sound generation and basic documentation of a performance, but unique features of a player's movements.

The aim, therefore, is not to produce archival notation that can be used to replicate the experience, though this may be an unintended consequence of the system. Instead, we aim to archive a diverse range of experiences that players have with the AirSticks, conveying 'what happened' in a descriptive notation that offers rich archival content that might be used by others.

There is a lack of approachable and generalised descriptive notation for Digital Musical Instruments, which may be preventing instrument designers from taking full advantage of the potentials of documentation. This project proposes one possible system of documentation by creating a notation system that visualises interactions with a gestural instrument, the AirSticks.

3. THE NOTATION SYSTEM

3.1 Overview

We propose a visualisation system that captures experiences with a gestural DMI, representing movement and sound through a reconstruction of the player's experience.

The term 'experience' is used here as the visualisation does not have to be documenting the act of performance. For instance, the system could be use to capture a player's first time playing the instrument, a new mapping or rehearsal.

The process of creating the visualisations leads to the production of a video, with the end result being a video visualisation of a chosen moment of interaction.

Section 3.2 discusses the variables needed to create the final visualisation, used in the process outlined in 3.3. A case study in Section 3.4 illustrates how this works in practice.

3.2 Capturing AirStick experiences

In order to capture an interaction with the AirSticks, we draw on multiple datapoints to form a snapshot of any given experience. A visualisation of an interaction with the AirSticks can be generated when all of these datapoints are present:

- **Gesture recording** – A recording of the IMU (inertial measurement unit) that includes acceleration and orientation information. The AirSticks receiver software has the ability to record sensor data received from the instrument (seen in Figure 3). This gestural data is saved to a `.json` file, which can be replayed or visualised within the software as if the player was playing in real time.

- **Audio recording** – Audio recording of the sound generated from the AirStick interaction.

- **Video recording** – Video recording of the AirStick interaction.

Figure 3. Gesture recording inside the AirSticks receiver software.

These datapoints all inform the creation of a visualisation system, outlined below.

3.3 Technical process

Central to the visualisation is how movement is mapped to video. In order to form a spatiotemporal representation, the motion of the AirStick is displayed much like a rhythmic gymnastics ribbon – that is, there is a clear sense of movement and change over time, visualising a window of time instead of a discrete coordinate, shown in Figure 4.

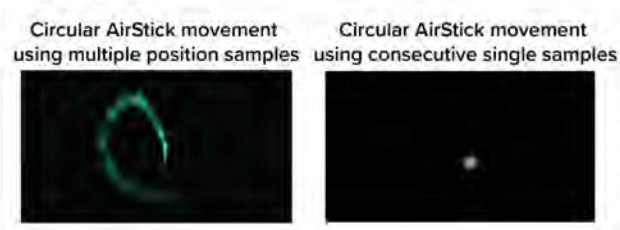


Figure 4. Comparing visualisations of coordinates over time.

This effect is achieved through the Adobe After Effects particle system *CC Particle Systems II*, which generates particles at defined X and Y coordinates over time, providing a 'trail' of movement data in disembodied graphic form.

A gesture recording from the `.json` file is tracked in the AirSticks receiver software, producing a set of of two-dimensional coordinates over time. These coordinates are then assigned to the particle generator in After Effects, and synchronised audio so they may be used in tandem for the visualisation.

To create unique visualisations for different contexts, variables within the particle system are altered based on variables extracted from data collected during interactions with the AirSticks, outlined in Table 1.

Particle variable	Interaction variable
Shape	Brighter sound from audio recording equates to a sharper, line-like particle; duller sound equates to softer, spherical particle (calculated using Spectral Centroid using the Librosa python package [1])
Size	Noisier sound from audio recording equates to larger particle; cleaner sound equates to smaller particle (calculated using Zero Crossings using the Librosa python package)
Colour	Feature colour extracted from video of the interaction using colour picker
Velocity	Average amount of 'energy' in the AirSticks (average acceleration over 50 sensor cycles)

Table 1. Particle system variables controlled by interactions.

Once the particle system variables have been entered into *CC Particle Systems II*, the generator produces particles at the AirStick coordinates over time. The visualisation is then exported alongside the audio that was capture alongside the AirStick movements, creating a holistic spatiotemporal representation of the movement and mapping. This process is outlined in Figure 5.

The technical process and assignment of standard variables means that a visual vocabulary of interactions is formed, with the end result being a series of visually distinct notations that summarise the interactions with the AirSticks, as seen in Figure 7.

3.4 Case study

To illustrate how the system has been used in practice, it is useful to focus on a particular AirSticks mapping.

Andrew is a member of a physical theatre group that the AirSticks team produced individual mappings for as part of a theatre production in June 2021.

The interaction designed for Andrew evoked the metaphor of a 'drumming goldfish' – a high-energy drumming mapping that triggered rhythmic sequences of heavy rock drumming mapped to the change in acceleration of the AirStick.

[1] https://librosa.org/doc/latest/index.html

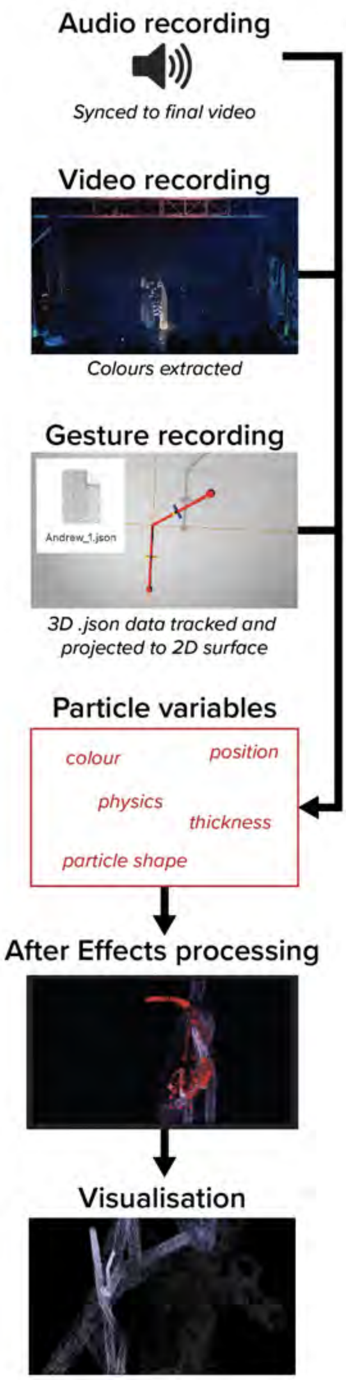


Figure 5. The process of visualisation.

Andrew's final performance was chosen for visualisation as we had audio, video and gesture recordings of the interaction. Combining the recordings, the particle was set to a purple colour extracted from the lighting state in the video, with the bright sound of the cymbals (and thus a higher Spectral Centroid) leading to a sharp line for the particle shape. The extreme range of Andrew's movements meant the velocity of the particle system was high (creating a spraying effect), and the noisier timbre of the drums (and thus more Zero Crossings) meant a large particle size, all shown in Figure 6.

The final visualisation was a high-energy video that showed

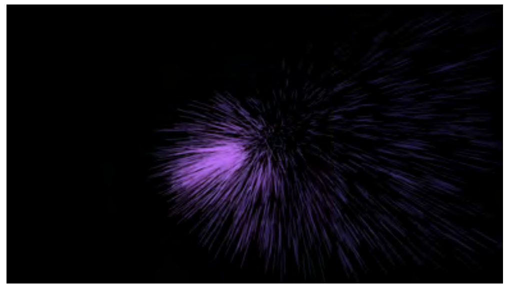


Figure 6. Screenshot from visualisation of Andrew's mapping.

a strong causal link between larger movements with the AirStick and the drumming. It also illustrated Andrew's dynamic range of movement, with large splashes of particles occurring on dramatic waves of the arm.

This visualisation was then displayed alongside snippets of additional information such as photos from rehearsal, video from the theatre performance, and quotes from Andrew himself, drawing on a rich catalogue of documentation (screenshot seen in Figure 8). This complete set of documentation was used to display ten different interactions by ten performers in the physical theatre group, creating a catalogue of interactions that could be compared and analysed further.

The notation of Andrew's experience is an example of how the visualisation system presents a fuller picture of 'what happened' in a series of interactions that involved many players in a theatre group with many different mappings, movements and sound worlds.

4. DISCUSSION

4.1 Utility of new systems

Why is a retrospective notation that visualises interactions with DMIs useful?

By creating a system of documentation and notation that captures interactions with the AirSticks, we are building a body of knowledge that reveals different aspects of 'what happened' when a diverse range of people used the instrument. This is particularly relevant in the context of discussions in Section 1.1, which note the lack of clear documentation that might hinder longevity of DMIs.

High-resolution capture and creative visualisation of of gestural data adds integrity to the design and contribution of the AirSticks, separate from code snippets or recordings of performances. In line with advances in data visualisation and the growing relevance of creative representations of data [17], the system offers a unique archival perspective on a new technology.

These visualisations cannot and should not be used to replicate the interactions themselves – they do not prescribe to a prospective player how they should play the AirSticks, and nor do they offer instruction as to how future interactions should occur.

Instead, using visualisation to diversify the prevalence of richer documentation means that more instrument designers and players might learn from or expand upon an

Figure 7. Screenshots from visualisations of a range of interactions with the AirSticks.

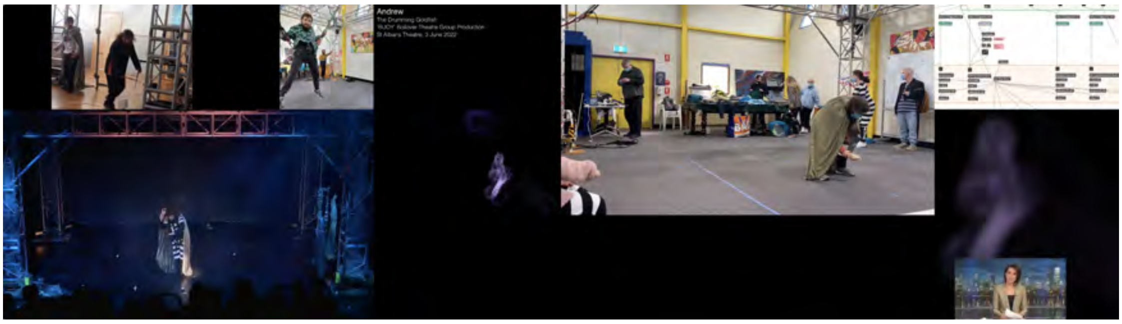


Figure 8. Screenshot of additional documentation shown alongside Andrew's visualisation.

archival system that provides insight into how players interact with new instruments. Expanding the concept of descriptive notation to encompass experiences with a gestural DMI is a logical next step that fills a gap in the iterative process of instrument design, providing an avenue for critical analysis that will aid the further development of DMIs.

4.2 Future work

This paper presents a possible system of documentation by creating a notation system that visualises interactions with a specific gestural instrument.

The natural next step for a system such as this one is to expand its use to other DMIs. The ability to test the notation on other DMIs, and collect evidence as to how it contributes to the longevity of DMIs would be an invaluable addition to the instrument design community.

One barrier to this extension would be data integrity and compatibility problems that DMIs are often faced with – for instance there is no gestural data 'standard' that can be exported from each DMI, and nor do they necessarily interpret or map gestural data the same way [1]. Additionally, not all DMIs are gesture-based, and the need for alternative data that can be captured with other DMIs may arise.

Another natural step for the notation system would be to target development in the area of interactivity so that the documentation could communicate the knowledge embedded within the playing of these 'epistemic tools' [14]. This might involve a more interactive system of documentation, perhaps allowing the viewer to play the mapping whilst also watching what happened when someone else played.

5. REFERENCES

[1] F. Calegario, J. Tragtenberg, C. Frisson, E. Meneses, J. Malloch, V. Cusson, and M. M. Wanderley, "Documentation and Replicability in the NIME Community," in *NIME 2021*, apr 29 2021, https://nime.pubpub.org/pub/czq0nt9i.

[2] F. Morreale and A. P. McPherson, "Design for longevity: ongoing use of instruments from nime 2010-14," in *NIME*, 2017.

[3] A. J. Tom, H. J. Venkatesan, I. Franco, and M. Wanderley, "Rebuilding and Reinterpreting a Digital Musical Instrument - The Sponge," in *Proceedings of the International Conference on*

New Interfaces for Musical Expression. Zenodo, Jun. 2019, pp. 37–42. [Online]. Available: https://doi.org/10.5281/zenodo.3672858

[4] T. Magnusson, *Sonic writing : technologies of material, symbolic, and signal inscriptions*. Bloomsbury Academic, Bloomsbury Publishing Inc, Jan. 2019.

[5] S. Reich, "Pendulum music (for microphones, amplifiers speakers and performers)," 1968.

[6] T. Magnusson, "Scoring with code: Composing with algorithmic notation," *Organised Sound*, vol. 19, no. 3, p. 268–275, 2014.

[7] C. Laurenzi and M. Stroppa, "The Notation of Dynamic Levels in the Performance of Electronic Music," in *Proceedings of the First International Conference on Technologies for Music Notation and Representation (TENOR)*, May 2015.

[8] W. Enström, J. Dennis, B. Lynch, and K. Schlei, "Musical notation for multi-touch interfaces," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, ser. NIME 2015. Baton Rouge, Louisiana, USA: The School of Music and the Center for Computation and Technology (CCT), Louisiana State University, 2015, p. 83–86.

[9] M. Huberth and C. Nanou, "Notation for 3D motion tracking controllers: A gametrak case study," in *Proceedings of the International Conference on New Interfaces for Musical Expression*. Zenodo, Jun 2016, p. 96–105.

[10] G. Boutard and C. Guastavino, "Archiving electroacoustic and mixed music," *Journal of Documentation*, vol. 68, no. 6, pp. 749–771, 2012. [Online]. Available: https://doi.org/10.1108/00220411211277028

[11] S. Lemouton, A. Bonardi, L. Pottier, and J. Warnier, "On the Documentation of Electronic Music," *Computer Music Journal*, vol. 42, no. 4, pp. 41–58, 12 2018. [Online]. Available: https://doi.org/10.1162/comj_a_00486

[12] G. Boutard and C. Guastavino, "Following gesture following: Grounding the documentation of a multi-agent music creation process," *Comput. Music J.*, vol. 36, no. 4, p. 59–80, dec 2012. [Online]. Available: https://doi.org/10.1162/COMJ_a_00147

[13] C. Small, *Musicking: The Meanings of Performing and Listening*, ser. Music / Culture. Wesleyan University Press, 1998. [Online]. Available: https://books.google.com.au/books?id=1lOx9nr0aHkC

[14] T. Magnusson, "Of epistemic tools: musical instruments as cognitive extensions," *Organised Sound*, vol. 14, no. 2, p. 168–176, 2009.

[15] O. Luening and V. Ussachevsky, "Incantation," 1953.

[16] W.-Y. Chan, H. Qu, and W.-H. Mak, "Visualizing the semantic structure in classical music works," *IEEE transactions on visualization and computer graphics*, vol. 16, pp. 161–73, 01 2010.

[17] A. Unwin, "Why is Data Visualization Important? What is Important in Data Visualization?" *Harvard Data Science Review*, vol. 2, no. 1, jan 31 2020, https://hdsr.duqduq.org/pub/zok97i7p.

A NOTATION SYSTEM FOR DISTRIBUTED MEDIA ART

Jean-Michaël Celerier
Concordia University / ossia.io
Montréal, Canada
jeanmichael.celerier@gmail.com

Akané Levy
ossia.io
Talence, France
levy.akane@gmail.com

ABSTRACT

Interactive, intermedia scores are scoring systems which enable composers to specify temporal evolutions and variations of such multimedia systems. We introduce a visual notation for distribution of the interpretation and execution of such scores over a computer network: both the temporal organization of the score and the multimedia data can be distributed through a set of scoring primitives.

For instance, we will show how one can write and execute a score which specifies: part A plays on a first group of computers, followed by part B on a second group of computers, while a parallel part C containing a synchronized video effect is being played on a separate machine with dedicated video hardware. Based on the pre-existing scoring environment *ossia score*, we cover how the extensions interact with the existing score execution algorithm, and present a set of small distributed scores which enables varied behaviours to be defined by the composer from our proposed set of nine temporal and three dataflow primitives for score distribution.

1. INTRODUCTION

Interactive scores are scoring systems which encode interactivity and the set of variations that can take place during the performance of said scores. Intermedia scores are scoring systems which encode not only the actions of musical instruments, but also of video systems, OSC controls, etc. *ossia score* [3, 4] is both an interactive, intermedia scoring language, and the software implementation of said language as an open-source score editor and interpreter.

The present research covers multiple ongoing parallel tracks for the distribution of interactive, intermedia scores, which converged in an implementation in *ossia score*. Simply put, we are interested in all the possible ways in which one can "distribute" an interactive, intermedia score over a computer network, and in how visual scoring languages can be defined to enable composers to write such a score, with a unique visual representation conveying the distribution semantics desired by the composer.

A fundamental example of distribution is the classical orchestra: orchestral scores can specify staves for violins, flutes…which are then distributed to an arbitrary number of performers which will all have their own personal interpretation of the score.

In the case of intermedia scores, it is common for computers and their strict internal clocks to be at the core of the performance: interactive, intermedia artworks can range from pieces that do not incorporate any human element during the performance to intricate collaborations between humans and machines. The goal of this research is to devise a notation to encode distributed semantics inside *ossia score*, which will allow performance to occur from a single document shared over the network, and provide an implementation of the distributed performance of said notation.

Distributed works in media art have existed for a long time: laptop orchestras [14] are a famous instance of musical ensembles revolving around a group of human performers being conducted and themselves using their laptops as musical instruments. Said laptops can optionally be synchronized, for instance to keep a shared musical beat across the piece. The question of conducting in the face of computer networks is discussed by Smallwood in [13]:

> Thus, possibilities for coordination, message-passing, group control, quantization, tempo, dynamics and so on are on the table for all composers working with PLOrk. Should these tasks be given to a conductor? Should the conductor be human, or should it be a program operating over the network? Or should there be both kinds of conductor?

Scoring systems for laptop orchestra pieces are as far as the author could find, bespoke and generally customized for each individual piece. Other systems that embody distributed composition and performance often focus on distributing scores that are closer to traditional western sheet music, such as *Quintet.Net* introduced by Hajdu in [8] and based on Max/MSP. Likewise, *Indra* [1] is another Max-based system for networked live performance with real-time compositional aspects. Other systems focus more on graphical or alternative notations: *Decibel ScorePlayer* [9] "allows for network-synchronised scrolling of proportional colour music scores on multiple iPads". More recently, Drawsocket [7] by Gottfried and Hajdu focused on live generation of graphical scores in the web browser from Max/MSP patches and a NodeJS communication layer, with the main aim being conduction of human performers through the distribution of the visual notation over multiple networked displays.

In the present paper, we will argue for a basic set of visual primitives for writing such scores in the hope of fostering collaboration and a common understanding of the ways performers, be they human or machine, can be synchronized (or not) in such artworks. First, semantics for distribution of both temporal interactions and specifications will be presented, then we will introduce a core primitive for symbolizing the various ways exchange and sharing of multimedia data can be done in a distributed score. Extensions to *ossia score*'s visual language to encode these distribution semantics into the scoring language will be presented, and various examples of usage will be discussed. The implementation is entirely self-contained in *ossia score*. Due to its current use of WebSockets for communication, it works transparently across both desktop apps and the beta version of the WebAssembly port of the software.

2. AN OSSIA SCORE PRIMER

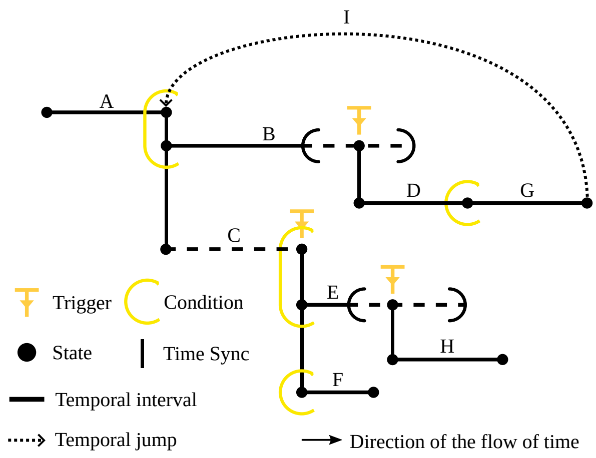


Figure 1: Syntactic elements of *ossia score*'s visual language.

Dubbed "interactive sequencer for the intermedia arts", *ossia score* is a system which combines both non-linear time-lines and the data-flow paradigm to allow artists to create interactive multimedia artworks, musical pieces, museum installations, etc. It supports audio and video playback and effects, as well as communication and control over protocols such as OSC, OSCQuery, WebSockets, Art-Net or even raw serial port communication.

The visual language is showcased in fig. 1. Here is a short description of the syntax elements and their semantics: *time intervals* are represented with horizontal lines proportional to their durations. A full horizontal line means that the time must not be interrupted, while a dashed horizontal line means that the time of the interval can be interrupted to continue to the next part of the score according to an external event. The time during which such interruption can occur can be bounded by the user: these bounds are visible as the small black parentheses around the dashes. Start and end points of a group of intervals can be synchronized. On these synchronizations points, an optional trigger can be used to specify that the synchronization is dependent on an external event, such as an OSC message. Conditions allow to selectively disable parts of the score depending on the value of an external control at the moment execution reaches that point: an OSC parameter, a MIDI CC…

In fig. 1, execution occurs as follows: the interval A runs for a fixed duration. When it ends, a condition is evaluated: if it is false, the branch which contains B will not run. Else, after some time, the flow of time in B reaches a flexible area centred on an interaction point, also called a trigger. If an interaction happens, B stops and D starts. If there is none, D starts when the max bound of B is reached by the flow of time in B. Just like after A, an instantaneous condition will make G execute or not execute. In all cases, C started executing after A. C expects an interaction, without time-outs. If the interaction happens, the two instantaneous conditions which follow C are evaluated: the truth value of each will decide of the execution of E and F. Finally, after G, execution circles back to the end of A thanks to the temporal jump I; the condition will be reevaluated, at least C will start again, and B depending on the value of the condition. Besides this, the system allows varying each interval's execution speed independently during the performance: even if all the external conditions evaluate to the same value at the same time, it is possible to have two very different executions for the same score.

In itself, this only represents an abstract "flow of time": the score language needs to have a way to specify concrete behaviours, such as sounds or videos being played, OSC messages being sent, etc. This is done at two levels: *states* can contain a bundle of messages that are sent at a specific point in time, and *intervals* contain *processes* that generate, filter and output data: automations, MIDI, sound or video file players, VST effects and instruments, PureData patches, JavaScript scripts, video effect processors, etc. Processes can be linked together in a dataflow graph akin to the usual patching or coding environments: PureData [12]…, or communicate directly with the external environment: OSC-like addresses can be specified for their inputs and outputs ports to address the external environment of the software, visualized in 2.

These processes are rarely represented on paper notation. The software implements this visual language along with many such processes; the temporal scoring language is itself defined as such a process, called *Scenario*, and as such can be nested recursively to implement grouping and hierarchy: scenarios describe the organization of intervals, which themselves contain processes.

3. DISTRIBUTING SCORES

A first prototype for distribution in *ossia score* had been introduced in [2]. We recapitulate here the core ideas: the distribution of interactive scores is a semantic that allows defining on which hardware specific parts of scores are going to run. For instance: a computer handles the audio processing of a performance, while another processes video feeds. Through simple modifications to the software model of *ossia score*, this enables varied networked behaviours to occur. Our implementation uses a client-server architecture: one of the machines acts as a server and ensures the coherency of the network session. More interesting peer-to-peer implementations would be possible through distributed consensus algorithms such as Paxos[10], but in

Figure 2: A screenshot of *ossia score*, showing the main system for authoring scores with the visual language of fig. 1. On the left, a tree view shows the OSC addresses to which the score can speak. On the right, an inspector allows to edit the properties of selected elements. The score is in the center.

practice the current implementation has performed sufficiently well for the needs of the scores written with it so far. The success of the client-server architecture in *Quintet.Net* shows that this is a viable long-term approach.

3.1 Abstracting over hardware with groups

The core idea is the addition of a notion of groups to the software model. Groups are the interface and abstraction between physical hardware and the semantic level of the score: we do not want to annotate IP addresses representing a specific computer in scores for obvious maintainability and readability reasons. The way groups are used by the composer is simply by assigning objects of the score to a group defined as part of the score's metadata: a part of the score can be assigned to a group named `audio`, another part to a group named `osc_controls` for instance. These parts can play in parallel, one after each other, or more generally in any temporal arrangement made possible by the scoring semantics given in Section 2. Then, when the performance happens, all the computers supposed to partake in the execution of the score connect to the session, and choose which groups they want to join. During playback, a given client computer will execute only the elements of the score which are assigned to any of its groups. This means that for prototyping, it is trivial to run the entire score on a single computer, simply by assigning the unique score instance to every group defined in the score.

Due to the hierarchic model of *ossia score* processes, where a scenario is itself a process, entire parts of the score can easily be distributed to various groups, simply by changing a property on a parent node of this hierarchical process tree. The group property is inherited across the child objects: if an interval is assigned to a group `A`, all its children processes are recursively assigned to that group too unless one explicitly selects another group. Likewise, all the other properties mentioned in the sections below propagate hierarchically until an element of the score with properties explicitly set by the composer is encountered.

3.2 Distribution of interaction

Remember that *ossia score* allows defining interaction points in the score: in the non-distributed case, this means for instance that someone can write a score where a sound plays until a physical sensor is activated, then a light flash occurs. The interaction point will contain an expression such as: `device:/sensor/1 > 100`. The question we ask is: what are the possible semantics for the distribution of the evaluation of such an expression over the network – and most importantly, what does it enable intermedia composers to do. The main idea is that we can leverage the multiplicity of computer clients to represent consensus in the score: for instance, by assigning the interaction point to a group, we require that all the participants to the group validate the expression for enacting forward progress past the interaction point in the score; client machines outside of the group simply don't participate in the expression evaluation and just follow the result defined by the group evaluating the expression.

We consider two axes for defining the temporal relationships between how the synchronization of expression resolution is achieved across a group of clients executing a given scenario. The temporal properties to balance are:

- Latency: how fast an individual machine reacts to the resolution of an interactive expression on the network.
- Respect of the temporal order: how precisely the overall execution of the score on the network matches the specification given by the composer.

These two properties are at odds: for instance, if in the case given above, different groups and hardware execute sound playback and light control, is it tolerable for the aesthetics desired by the composer that the computers of the first group still play back sound for a few milliseconds if that means that the light flash starts closer to the sensor parameter change? Only the composer can answer this question.

The proposed distribution system enables fine-tuning of the synchronization mechanism by leveraging the pre-existing asynchronous infrastructure in *ossia score*: interactive trigger points. Two cases are possible for interactive trigger points assigned to a group: they can be time-compensated – that is, when all the computers validate the trigger, the network engine will try to define a date in the future at which every computer must actually execute the trigger point, so that this happens simultaneously from the point of view of an external observer looking at all the computers. Or, execution can also happen on an as-fast-as-possible basis, which can be helpful when the artistic trade-off between latency and synchronization tips in favour of the shortest reaction times.

- Asynchronous versus synchronous: in the asynchronous case, the synchronization semantics of trigger points are not respected across the network. The triggering algorithm is:

1. Obtention of a consensus on the value of the expression across the network on the server machine.
2. The server notifies all the clients, which react as soon as they get a message.

This means that if the latency between a client executing a process that precedes the trigger point and the server is greater than the latency between the server and a client that executes a process following the trigger point, to an external observer, an external observer will see both timelines overlap for a short duration, which would be impossible in a non-distributed execution of the score.

In the synchronous case, the semantics of trigger points are respected across the network: the following intervals cannot start before the previous intervals have ended on all clients. The triggering algorithm is:

1. Obtention of a consensus on the value of the expression across the network on the server machine.
2. The server notifies all the clients that the trigger has started executing.
3. When the clients finish executing the intervals that precede the trigger point, they notify the server.
4. The server notifies all the clients that the trigger has finished executing.
5. The clients start executing the following intervals.

This of course increases latency, but ensures that the temporal semantics of the score are respected globally: the score executes on the network in the same order as it would on a single machine, except with greater delays.

- Uncompensated versus compensated: in the uncompensated case, messages are processed as soon as they arrive. In the compensated case, the server tries to derive a timestamp in the future where all clients are supposed to have received a message according to their respective latency with regards to the server, in order to make sure that from an external observer point-of-view, everything happens simultaneously.

  For instance, consider a situation with three machines: one server at time t_0 with latency 0 with regards to itself, one with a latency of 50 milliseconds and one with a latency of 100 milliseconds, the server will send to the entire network session messages that request execution at a minimum of $t_0 + 100$ms. Note that we implicitly assume that the machine's wall clocks are synchronized through an external mechanism such as NTP or PTP.

3.3 Polyphony

Processes of the score can operate either in `free` or `shared` mode. Free means that the properties and execution of a given process is independent across machines; shared means that the process's state will be synchronized across all the machines that execute it. For instance: a scenario playing a sound could play it at two different speeds and have entirely different conditions resolutions on multiple clients in the free case, while in the shared case, all the clients executing it will be kept in sync for this specific scenario's execution: condition resolution and interval speed adjustments will be synchronized across the network. Likewise, for processes with controls, this means that a change of control during execution will not be synchronized across machines in the `free` mode, to enable for instance multiple performers to interact with a virtual musical instrument each in a different way on their respective computer. In contrast, in the `shared` mode, the controls will be synchronized across all the machines.

For instance, for processes such as sound generators or filters, this simply means that a process in `free` mode can execute with different state for its user-interface controls across the session. In `shared` mode, a change of control from a client is applied to all the other machines in the network.

Likewise, in `free` mode, a scenario process will be able to execute with different interval speeds, conditions and interactions in all the clients executing it. In `shared` mode, the interval speeds, expression resolutions, will be synchronized across the network through the mechanisms described above.

4. DISTRIBUTING DATA

An open question which had been evoked in [2] was the distribution of the live, run-time data of the score: the original implementation did not consider the transfer of things such as OSC parameters, sound or video streams: it only focused on the distribution of the temporal semantics, and users had to define the transmission of controls through separate systems such as NetJACK[11].

We introduce here a new set of processes in *ossia score* which leverage the underlying networking session to enable scores to embed media data transfer semantics through user-friendly objects. These processes are called Netpit: Message Pit, Audio Pit and Video Pit. Right now, their implementation focuses on ease of implementation for the sake of prototyping and experimenting with writing scores and as such uses WebSockets for data transfer instead of WebRTC [1]. The final implementation aims to use WebRTC as it is the only real-time audio-video protocol supported by web browsers: there is no other choice if we aim for our system to be fully functional in the web.

These processes have a combining semantic. They have one input port, and one output port. Their operating rule is: the data that a Netpit process instance receives as input is combined with all the clients executing this process with a user-chosen function: summing or concatenating channels for audio, taking the mean for real-valued control sig-

[1] An implementation using WebRTC, more optimized for audio and video transfer, has been started but is stuck on a bug in the underlying GStreamer library implenting the WebRTC protocol: `https://gitlab.freedesktop.org/gstreamer/gstreamer/-/issues/1261`

nals, applying a blend for videos, etc. For instance, if two clients execute a given Netpit audio process at some point during the execution of a score, with each their custom microphone input set in the input port, both clients will by default get a sum of both microphone signals as output of the process. There is as of now no latency compensation in the system.

This, combined with the notion of group, enables to transparently define a set of distributed behaviours: users do not have to fiddle with defining explicit distribution semantics such as one would do in environments such as Max/MSP, by using custom network send / receive objects for each signal they want to transmit; here, the objects scale automatically to the network topology used and allow he desired distribution semantics inside the dataflow graph engine of *ossia score*.

5. VISUAL LANGUAGE EXTENSIONS

To represent the multiple states of synchronization of the elements of *ossia score*, we propose to introduce a set of alternative symbols to denote the various network-related semantics. Table 1 presents the matrix of possibilities. The idea behind this prototypical design language is to represent the concepts described in sections 3 and 4 as follows:

Inactive elements (e.g. those that are not in a group that a given *ossia score* instance executes) are in grey, while active elements are in colour.

The *free–shared* dichotomy is represented by a connecting line: free elements aren't tied together, shared elements are.

We chose to not represent the *uncompensated–compensated* axis yet as it would double the amount of symbols necessary while only providing marginal benefits in local networks with very short latencies ; the toggle is however still available from the user interface.

The *asynchronous–synchronous* dichotomy straightens the round shapes to evoke the stricter constraints of synchronization related to non-synchronization. Frutiger says about the square in [6],

> (…) a symbolic object, boundaried property, also a dwelling place with the feeling of floor, ceiling, walls, protection, etc. (p.43)

The notion of readiness which matches well with our *free* and *uncompensated* and *asynchronous* semantics is also associated with the circle by the same author:

> The most used figures are 1 and 0, two signs that have highly differentiated forms: the straight line and the circle. Once again we come across the binary principle: 1 = notch, cut, hardness (two visible stroke endings); 0 = emptiness, readiness (no beginning or ending). (p.211)

The most restrictive and synchronized case should be the same as the original, non-networked version, as a network with only one computer can be considered as fully synchronized with itself. It would be of course possible to argue the opposite: that a network with a single score instance running is equivalent to the *free* case. However, using the original symbols in the *free* case would likely have implied that the *shared* cases would have to bear additional visual elements to denote the stronger semantics synchronization semantics that they carry: experiments in this direction made the visuals too heavy and harder to read. In short, we opted for deconstructing the existing elements instead of constructing new ones, with the sole exception of the condition becoming "straighter": debate is ongoing about changing its appearance to the squared version in the non-networked mode of *ossia score*.

These visual language extensions are still ongoing a prototyping and experimentation process in *ossia score* and may still evolve in terms of representation, with the goal being making the transmission of our distribution semantics as obvious as possible to the users.

6. IMPLEMENTATION

The system discussed here is implemented in C++, using simple messages over WebSockets through the Qt WebSockets implementation. This enables the system to function over both desktop, embedded, and tentatively through the web with the WebAssembly port of Qt, with a single codebase. The only limitation is that the web version cannot act as a server for the session as web browsers do not allow web page to open network ports.

The latency of simple operations such as a control change is overwhelmingly defined by the network's implicit latency, plus the local latency between the network thread of the software, and the audio or video thread which will end up turning the network message into an observable behaviour, which is at most the duration of an audio buffer or screen frame. Since our current implementation is based on a client-server architecture, the network communication latency is always the latency between the originating client and the server, plus the latency between the server and the other clients that will receive the message.

For the synchronized operations discussed in Section 3.2, the latency will increase due to multiple round-trips between the involved clients and server: the overall latency will always be bounded by the latency of the slowest client as the server waits for all the clients in a group to trigger the advancement of the score.

The system has been tested so far with up to 8 instances on a local network, over Wi-fi and Gigabit Ethernet, and with two instances over the internet. Our naive video and audio transfer implementations however are only useable with correct results on a local network so far, and would require using specific codecs with optimizations (such as GPU encoding / decoding for video) making them suitable for use over the internet.

7. DISTRIBUTION EXAMPLES

We cover in this section a few examples of usage of the new data distribution primitives and distributed scores.

Element	Free		Shared			
			Uncompensated		Compensated	
	Inactive	Active	Inactive	Active	Inactive	Active
Triggers						
Conditions						
Processes						

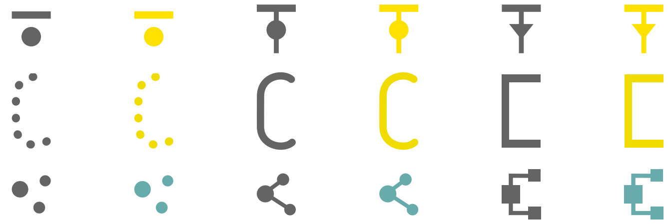


Table 1: Proposed visual syntax for the distribution specification of *ossia score*'s elements. The icons for processes are displayed in their header, and are also used for intervals.

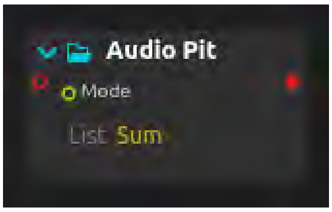


Figure 3: Distribution of a mixed audio stream across machines: this is the entire score.

7.1 Sending data between machines

In this example, we want to showcase the simplest possible example of distributing data: there are two machines, both want to hear a mix of all the musical instruments playing on the network: for instance the scenario is the classic networked band rehearsal. Fig. 3 shows how simple it is: this behaviour is the one that the "Audio Pit" object will provide by default. All the machines can be in the same default group, `all`. They will all execute the process: its input is set to the address `audio:/in/main` which represents the default input of the sound card of the computer. The process will pull the audio input, and sends it to all the clients which then each perform the downmix with the data they received [2] .

7.2 Combining control data across a group of players

In this example, we wish to combine MIDI CCs, gamepad inputs or even GUI widgets which would be controlled by musicians on cheap hardware, in order to generate audio on a computer with a powerful sound card. Fig. 4 shows how such a setup can be achieved.

The groups are defined as follows: there is a `players` group, and an `audio` group. The "Knob" process is associated with the first group: only the clients in this group will execute this process. The second group is associated to the entire bottom interval: it applies recursively to the processes within. During execution, all the clients which registered themselves in the `players` group will have their "Knob" send an input message to the "Message Pit" process, which by default is assigned to the `all` group, which everyone is part of by default. Then, the score instance which registered itself in the `audio` group will receive a list containing all the input frequencies of the clients. It applies it to a simple polyphonic synthesizer and effect chain and renders it on a sound system.

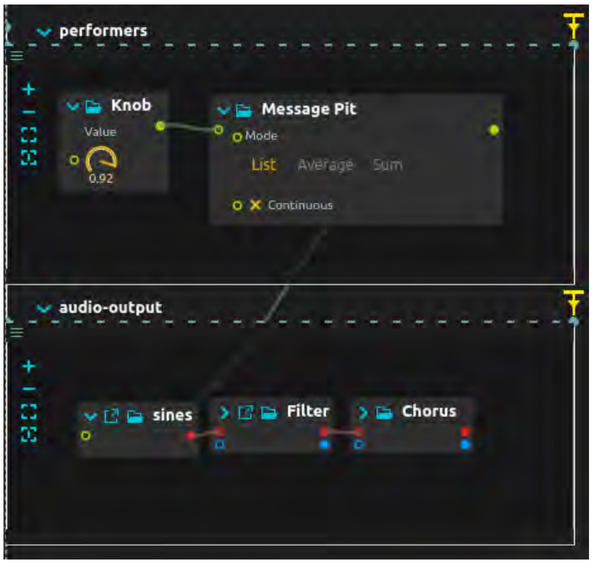


Figure 4: Clients-to-server communication.

7.3 Duplicating an input

This example is a sort of contraposition of example 7.2. A single machine produces an input, which is going to be broadcast. A set of machines will synthesize either sounds or visuals from this input, depending on their groups. The group organization is actually almost the same as example 7.2: the only difference is that a `video` group is added, in order to distribute audio and video rendering on different hardware. Fig. 5 shows the score. Note that the system has no way to specify the number of participants, which is purely a property of the actual performance that will take place: a performance could use one client, two audio machines and two video machines, while the next performance could actually average the input of four clients, and only use a single audio and video output clients, without chang-

[2] Our current research prototype implementation uses a central server; a production implementation using WebRTC would be P2P

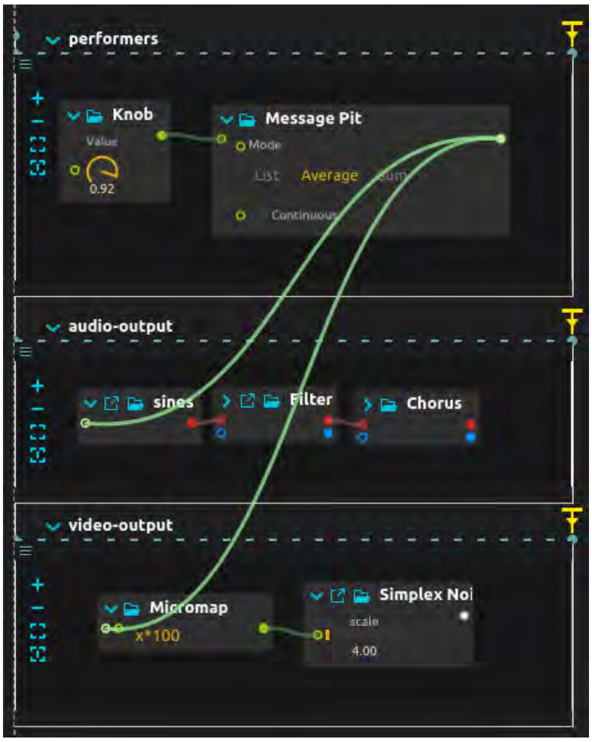


Figure 5: A single client's input data will be broadcast to all the machines in the audio and video groups.

Figure 6: Excerpt of the score demoed over the internet at SMC2022 – this was before the introduction of the visual language.

ing anything to the score. Conceptually, the score specifies only sections and not individual players, just like scores for orchestras or choirs usually does not specify an exact number of second violins or singers.

7.4 Score for SMC2022

A successful prototype demonstration happened for the 2022 Sound & Music Computing Conference (fig. 6), located in Saint-Étienne (France). The score was distributed between this location and the home of the author in Peyriac-Minervois, 357 km away and featured both interactive triggering and distribution of OSC data. Audio and video transmission had not been implemented yet and have so far only been tested on local networks.

This score was before the introduction of the visual language, it featured multiple groups and controls, as well as a shared evolution of the interactive timeline over the network: the automations at the top would for the first part all be synchronized together as the main scenario was in "shared" mode. Then, a sub-scenario *D* was in free mode: inside this hierarchical level, the trigger's triggering time would be independent across machines.

7.5 Polyphony, sharing and visual language

This example (fig. 7) presents some of the visual language extensions introduced in Section 5 The root scenario is shared: execution speed and triggers will be synchronized across the entire network (the scenario is assigned to the special "all" group which encompasses all clients). Then, its score is as follows: intervals *A* and *B* are assigned to respective groups of the same name. Both contain an audio generator.

The generator of group A is shared: if any client changes one of the control, this is passed on all the other machines which execute it. The generator of group B is free: every client can change its controls independently. Clients that are neither part of *A* nor *B* will not execute any of these sound processes. Whenever the interactive trigger point between A and the Video interval is triggered, all the machines start executing the video effects (GLSL shaders) locally. The Scenario which contains *GFX1* and *GFX2* is in `free` mode: all the machines can execute it independently. That means for instance that the speed of execution of *GFX1* and *GFX2* can differ across all clients, and that the interactive trigger point between both isn't synchronized across the network. The output of the video generator processes is connected to an *Echo Trace* process in `shared` mode: its controls will be shared across the network. Finally, in the video effect chain, the *VVMotionBlur* video effect is in free mode: its controls are also independent across the network.

8. CONCLUSION

This work introduces a notation system for intermedia composers to author distributed behaviours in a simple graphical environment. The interactive system has been tested both in local networks and over the internet.

The remaining implementation goals are to make sure that the system works correctly from web browsers through *ossia score*'s WebAssembly port, and improving the data protocol implementations to use more optimized protocols than the current WebSockets-backed implementation, which is not well suited to streamed multimedia data such as audio and video. In particular, a pathway we would like to explore, is to automatically adapt the protocols used for au-

Figure 7: A score with various behaviours across various groups. This uses the visual language discussed in Section 5.

dio/video data transmission to the clients connected and the network situation: if the performance does not use web-based clients and is entirely situated on a local network, it would make sense to automatically use NewTek NDI for video transmission due to its efficiency [5]. In contrast, if the performance is done over internet, it would make more sense to use WebRTC for its P2P communication abilities, and support for various NAT-bypassing technologies which would enable its use across for instance academic institution firewalls. A tentative to use the Opus codec for transmitting audio content has also been done, but so far did not provide an acceptable quality / latency trade-off for musical applications.

References

[1] Drake Andersen. "INDRA: A Virtual Score Platform For Networked Musical Performance." In: *Proceedings of the International Conference on Technologies for Music Notation and Representation (TENOR)*. Hamburg, Germany, 2021.

[2] Jean-Michaël Celerier, Myriam Desainte-Catherine, and Jean-Michel Couturier. "Exécution Répartie De Scénarios Interactifs." In: *Proceedings of the Journées d'Informatique Musicale (JIM)*. Paris, France, 2017.

[3] Jean-Michaël Celerier et al. "OSSIA: Towards a Unified Interface for Scoring Time and Interaction." In: *Proceedings of the International Conference on Technologies for Music Notation and Representation (TENOR)*. Paris, France, 2015.

[4] Jean-Michaël et al. Celerier. "ossia score 3." In: *Proceedings of the International Conference on Technologies for Music Notation and Representation (TENOR)*. Marseille, France, 2022.

[5] Andrew Cross et al. "IP Workflow and Scalable Video Performance." In: *SMPTE17: Embracing Connective Media*. SMPTE. 2017, pp. 1–12.

[6] Adrian Frutiger. *Signs and symbols, their design and meaning*. 1989.

[7] Rama Gottfried and Georg Hajdu. "Drawsocket: A browser based system for networked score display." In: *Proceedings of the International Conference on Technologies for Music Notation and Representation (TENOR)*. Melbourne, Australia: Monash University. 2019, pp. 15–25.

[8] Georg Hajdu. "Quintet. net: An environment for composing and performing music on the Internet." In: *Leonardo* 38.1 (2005), pp. 23–30.

[9] Cat Hope and Lindsay R Vickery. "The Decibel Scoreplayer – A Digital Tool for Reading Graphic Notation." In: *Proceedings of the International Conference on Technologies for Music Notation and Representation (TENOR)*. Paris, France, 2015.

[10] Leslie Lamport. "The Part-time Parliament." In: *ACM Transactions on Computer Systems* 16.2 (1998), pp. 133–169.

[11] Stéphane Letz, Nedko Arnaudov, and Romain Moret. "What's new in JACK2?" In: *Proceedings of the Linux Audio Conference (LAC)*. Utrecht, Netherlands, 2009.

[12] Miller Puckette et al. "Pure Data: Another Integrated Computer Music Environment." In: *Proceedings of the Second Intercollege Computer Music Concerts*. Tokyo, Japan, 1996.

[13] Scott Smallwood et al. "Composing for laptop orchestra." In: *Computer Music Journal* 32.1 (2008), pp. 9–25.

[14] Dan Trueman. "Why a laptop orchestra?" In: *Organised Sound* 12.2 (2007), pp. 171–179.

THE MAGNETIC SCORE: SOMATOSENSORY INSCRIPTIONS AND RELATIONAL DESIGN IN THE INSTRUMENT-SCORE

Nicola Privato
Intelligent Instruments Lab
Iceland University of the Arts
Reykjavík, Iceland
nicola@lhi.is

Thor Magnusson
Intelligent Instruments Lab
Iceland University of the Arts
Reykjavík, Iceland
thor.magnusson@lhi.is

Einar Torfi Einarsson
Department of Music
Iceland University of the Arts
Reykjavík, Iceland
et@lhi.is

ABSTRACT

With the changes occurring in the dialectics between the composer and the interpreter during the second half of the twentieth century, the traditional concept of the musical score has undergone an ontological change. As composers began exploring unconventional notational practices and offering to the interpreter a higher autonomy, the locus of the musical information became less defined, at times merging with that of the instrument. In this paper we explore the dual nature of notation both as score and as instrument from the point of view of non-visual methods of representation. We do this by presenting the Magnetic Score, a system for the inscription and generation of sound that relies on permanent magnetic fields. In magnetic scores, the performative gestures emerge out of the interpreter's embodied interaction with the magnetic fields, and the relational design of the inscriptions together with the interdependence of the symbolic and somatosensory layers offer original insights on the role and situatedness of the musical score in contemporary practices.

1. INTRODUCTION

In his seminal work *Opera Aperta*, Eco describes the changes that characterise the emerging artistic poetics of the 1950s and 60s [1]. If classical musical works consisted of organised sets of information reproducing an enclosed structure as imagined by the author, the practices of composers such as Stockhausen, Berio and Pousseur offered to the interpreter a higher degree of autonomy in relating with the musical material.

With this change in the dialectics between the composer and the interpreter, some composers began approaching music notation as a description of gestural information for the performer rather than as pitch organised in time [2], and the mapping of such relations became a crucial element in designing musical interactions [3]. As a consequence, the acquired freedom in defining and representing the inscribed musical parameters has led to the emergence of a plethora of compositional approaches [4]. Among such, an increasing number redefine the composer's and performer's traditional roles and attributed agencies [5], explore the relational aspects of the inscription [6], or explicitly suggest a dynamic and open idea of the score's situatedness in relation to the instrument [7].

The overlapping of the musical expression with the score is rooted in the very nature of musical instruments, as they inherently embed theoretical models that define the interaction and the musical practices that have developed around them [8]. In the domain of electronic music this becomes particularly evident, with the compositional processes and their technological substrate overlapping with unprecedented fluidity. As a consequence, the *notational space*, within the indeterminacy of the current artistic poetics, escapes the physical constraints of the score and coincides with that of the *dispositif*: an extension of the instrument-score, incorporating all the structural, tangible and virtual components that support the inscription [9].

In this paper we explore the dual nature of music notation both as score and as instrument. We also explore approaches that do not solely rely on visual representation, investigating how the embodiment that characterises contemporary compositional practices favours a holistic and sensuous experience of the inscription. We do this by introducing the *magnetic score*, a system for embodied notation in which the inscription is encoded via permanent magnets, and can be subjectively experienced through somatosensory feedback. As we will observe, magnetic scores combine the tangible features of instrument-scores with the relational and situational qualities of event scores, since, rather than being unilaterally inscribed, the information emerges through the interaction of the components that define the dispositif.

2. BACKGROUND

The use of graphic signs and symbols has been the prominent approach in inscribing music both in the tradition and within the avant-garde movements in the second half of the twentieth century. Works such as Cornelius Cardew's *Treatise* [10] and Earle Brown's *December* [11] adopt graphic notation to convey musical ideas, oftentimes endowing the performer with a high degree of interpretative freedom. In this sense, graphic notation has been a way for composers to develop personal systems of representation and elude the expressive constraints of traditional compositional praxis. Nonetheless, with the change in the artistic poetics that characterises the second half of the twentieth century, approaches that do not necessarily rely on visual notation have emerged, such as the implementation of tangible interfaces and the adoption of haptics [12].

In this section, we explore the relationship between instruments, scores and performers via the notion of non-visual inscriptions. In order to establish a theoretical framework, we look into inherent scores, tangible scores and event scores. We also describe the recent adoption of magnets in music performances and instrument-scores, as it provides a technical context for this work.

2.1 Instruments-Scores and Non-visual Inscriptions

Tomás and Kaltenbrunner propose the concept of *inherent scores* to describe the progressive embedding of inscriptions within the instrument [3], and trace its origin back to Alvin Lucier, who, in describing the practices of the Sonic Art Union, stated that the scores were inherent to the circuits developed by the members of the collective [13]. This approach is not isolated to a single, although particularly influential, group of artists, as it overarches the practices of a large group of composers, such as Pauline Oliveros [14], Gordon Mumma [15] and David Tudor [7], as well as sound artists such as Peter Vogel [16] or Gerhard Trimpin [17].

Among inherent scores, *tangible scores* are a particular subgroup that relies on the tactile interaction with the instrument for the generation of sound as well as for the interpretation of the sign [18]. Tangible scores certainly hold a visual dimension in that they suggest specific gestures through the graphic inscriptions embedded in the instrument, but they complement it with a strong orientation toward tactility. Signs are engraved rather than printed on top of the surface: this adds a tangible layer that informs the performance as well as the generation of sound.

Similarly to that of tangible scores, the concept of *composed instruments* provides a non-visual take on the navigation of the inscription. At the basis of this type of instrument-score is often the decoupling of the sound-producing component with the gestural one [19]. As a consequence, the score is incorporated in multiple, modular mappings, whose features define the interaction between a controller and an arbitrary synthesis engine. In composed instruments, the score is encoded inside the dispositif in the form of a defined set of mappings and constraints, and is navigated through embodiment within the performative act.

Non-visual inscriptions are also particularly effective in contexts in which the interpreters are free to explore the performative space and can not rely on physical supports, or in the cases of "comprovisational" practices in which the instruction is situational [20]. In such cases, cues provided by haptic devices embedded within garments have granted the needed flexibility and at the same time preserved the situational character of the compositions [21]. Haptics is indeed a promising domain, as it allows to dynamically inform the performance without interfering with the interpreter's interaction with the instrument and with the space. Furthermore, it offers compositional and performative control over multiple parameters at the same time, such as frequency, intensity and duration, as well as spectral content and spatial position [22].

Finally, other systems explore non-visual notation with the specific aim of easing the learning of a piece by the visually-impaired [23, 24]. Even though they are relevant as non-visual scoring methodologies, for the most part these approaches are substantially different from the one proposed in this paper, in that they focus on the re-encoding of traditional notation rather than in the exploration of alternative ways to inform the performance.

2.2 Event Scores and Non-visual Inscriptions

Instead of encoding the information in the instrument, other methods of non-visual inscription explore the notational possibilities offered by the performers' embodied knowledge and reciprocal interactions.

Event scores are brief sets of verbal instructions defining rules to follow, actions to take and concepts to be aware of in the act of performing [25]. Among such, Pauline Oliveros' *text scores* [14] focus on the listening experience that emerges in the performance. Even though text scores hold a visual component that is functional to their transmission, in Oliveros' works a different informational layer emerges and unfolds within the relation of the interpreters with each other and with the environment. The practice of *deep listening* becomes the space where information is produced, and sound the domain in which the process operates. Through this, the score acquires a relational dimension, as the musical intake of an agent informs the action of another.

The aural quality that characterises event scores is also a feature of *audio scores*, in which information is presented during the performance through recorded or live-generated sound. Different types of audio scores have been proposed, some providing precise and repeatable sets of instructions [26], others inviting the performers to interact with a set of live generated sounds [27]. In *Pricked and Away* [28], Elisabeth Schimana interestingly explores memory itself as a medium for the inscription and for the re-elaboration of musical ideas: the sound excerpts are presented to the performers long before the performance, and the musicians are required to remember and play them during the concert following a specific timeline.

The practices and conventions that characterise musical performances always implicitly involve notions of embodiment and interaction. Nonetheless, in the aforementioned works the score incorporated in the performers' embodied knowledge and interacting subjectivities is amplified, formalised and defined within the system's setup. The relational nature of these compositional approaches, detectable in the situational stance of the performative instructions as well as in the emergent character of the work, underpins a direct involvement of the performers in the compositional process. In describing the intersubjectivity characterising relational aesthetics in the modern work of art, Bourriaud states that "the sense of the work issues from the movement that links up the signs transmitted by the artists, as well as from the collaboration between people in the exhibition space" [29].

As we will see, the system we propose embraces the embodied, relational nature that characterises some event scores and combines it with the tangible materiality of the instrument-score. This is achieved through a set of permanent magnets embedded and creatively displaced on the different components of the dispositif, and whose magnetic fields interact in order to generate the inscription as somatosensory manifestation and sound.

2.3 Permanent Magnets

Permanent magnets constitute a key component of most audio electronics and in the building of all kinds of actuators, and are extensively incorporated in the design of modern musical instruments and amplification technologies. As musicians, we operate with magnets on a daily basis, from pickups to speaker cones. However, the application of magnets as structural elements in a score's interaction design, as gestural control or for the generation of sound, has not been extensively explored as of yet.

Neodymium magnets have been introduced in musical scores quite recently. Michelle Agnes Magalhaes' *Mobile* [1] first explores their use on the piano strings in order to obtain bouncing and glissando effects. Because of the unique sounds they produce, magnets have since then been incorporated in the works of other composers such as Elena Rykova in *Bat Jamming* and *Cositas Diminutas*, [2] and Gustavo Díaz-Jerez in *Metaludios*. [3]

A notable example of the use of magnets as key components in the interaction design of an instrument-score is the *Chowndolo* [4] by Giacomo Lepri: a pendulum whose movement is dynamically controlled through a set of permanent magnets on its base. A different approach is instead explored by David Griffith in the *Pattern Matrix*, [5] a tangible AR live coding environment controlled through the orientation of permanent magnets on a tangible 5x5 matrix.

In the *Marble Machine* [6] the merging of the score with the instrument becomes particularly apparent. In this system, the instrument's sounds are generated through the interactions of ferromagnetic marbles with different surfaces, membranes or strings, and their timing is controlled by a tangible step sequencer made of small magnetic cylinders attracting and repositioning the marbles.

Finally, in NIME's proceedings from 2001 to present we identified three papers describing the application of small magnetic tags for position sensing [30, 31, 24]. The advantages of this approach are the precise representation of the tag's position and the granular control that can be achieved within a circumscribed space. In such cases the sound is defined by the dynamic repositioning of the passive elements in relation to a sensor. As we will describe in the next section, in magnetic scores the paradigm is flipped, and the performer interacts with the passive elements by moving the sensors in space.

[1] https://www.youtube.com/watch?v=xLLctkt14qs
[2] https://www.elenarykova.rocks/
[3] https://www.metaludios.com/
[4] http://www.giacomolepri.com/chowndolo
[5] https://penelope.hypotheses.org/category/pattern-matrix
[6] https://en.wikipedia.org/wiki/Wintergatan

3. THE MAGNETIC SCORE

Magnetic scores (Fig.1) enable the composition of haptic scores within the functionality of the instrument itself, as a specific instance of inherent scores. They comprises a board and a pair of controllers, both embedded with magnets. As the performer navigates the board with the controllers, the magnets attract and repel each other, thus suggesting the performative gesture.

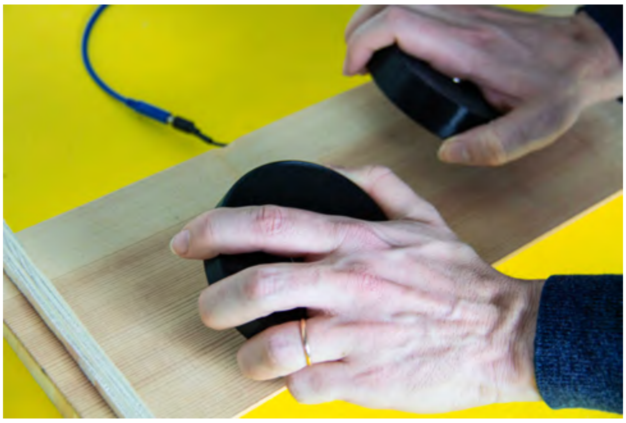


Figure 1. The Magnetic Score.

3.1 Magnetic Board

Magnetic scores encode performative instructions through magnets mounted underneath two-dimensional or tridimensional surfaces of variable dimensions and shape, named *magnetic boards*. The performer holds two controllers containing a series of sensors and a permanent magnet each, as described in 3.2. The interaction between the magnets within the board and the ones mounted on the controllers provides somatosensory feedback in the form of points of attraction and repulsion whose strength depends on the dimension of the magnets and their reciprocal distance. With the controllers, the performer dynamically explores the magnetic board as organised by the composer in the strength, spatial distribution and orientation of the magnetic fields.

The primary function of the magnets is to generate the score's information through their interaction, but in order to further investigate the boundaries between scores and instruments, in the particular instance of the Magnetic Score described in this paper, we turned a number of magnets into sound sources. This feature was easily achieved by selecting magnets and ferromagnetic material of different shapes, and placing them inside wooden tracks and 3D printed boxes mounted underneath the magnetic board. Upon interaction with the controllers, the magnets move against the board, producing sounds that are captured by two piezoelectric sensors. The sound is then routed to a laptop for processing in combination with the data generated by the magnetic discs.

For this first iteration of the Magnetic Score we designed a 50 x 15cm bi-dimensional wooden board (Fig 2). No visual information is inscribed, and the performer relies on somatosensory feedback and sound in exploring the surface. For the sound generation we included two flat, rounded neodymium magnets for punctual, sharp sounds, a series of ferromagnetic spheres of different dimensions for drones of variable pitch, and two sets of small screws for high-pitch, dense clusters of sound.

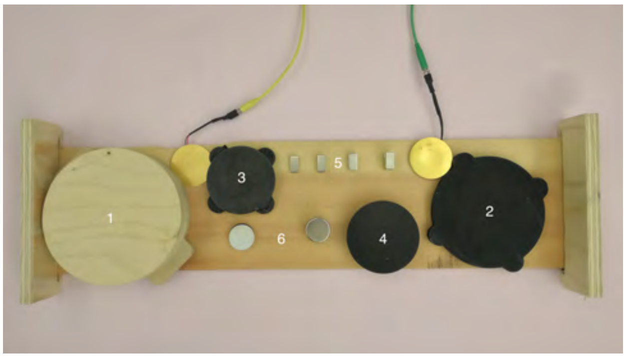


Figure 2. Magnetic Board's underside. 1-2 Ferromagnetic Marbles; 3-4 Ferromagnetic Screws and Marbles; 5-6 Magnets with Alternating Polarities.

3.2 Magnetic Discs

The *magnetic discs* (Fig. 3) are two 3D-printed, PLA cylindrical boxes. With a diameter of 10 cm and thickness of 2 cm, they mount a three-dimensional gyroscope and accelerometer, a three-dimensional magnetic sensor, one ESP32 microcontroller and a 1000 mAh battery. At the centre of the discs, a cavity hosts cylindrical magnets with a diameter of 3 cm. The magnets are loose within the discs, and are held in place by the performer's hand while holding the controller. When the magnet on a disc is approached to an external magnet with identical polarity, it moves and pushes on the performer's palm thus providing a proportional haptic response. At the same time, the resistance of the palm transfers the force of the magnet to the whole arm, thus influencing the performer's proprioceptive perception.

Figure 3. The Magnetic Discs.

Each disc wirelessly forwards to a laptop two data points: one relative to the xyz orientation of the device and one to the xyz strength of the magnetic field it is exposed to. Since the position of the disc and the orientation of the magnetic field are interrelated, the shape and orientation of the magnetic board allow to sensibly change the sound processing parameters. Interacting with a vertical score becomes therefore a very different experience than that of exploring a horizontal one, and curved surfaces allow to smoothly modulate in-between musical parameters.

As a consequence, further implementation of magnetic scores will extend the interaction design to larger and more articulated three dimensional artefacts, or even wider architectonic structures. This may be facilitated by the long-range wireless communication capabilities of the discs: through the *ESP-NOW* protocol, a dedicated wireless network is instantiated between the ESP32 microcontrollers mounted on the discs and a third ESP32 connected to the laptop's serial input and acting as a server. In an open space, the client devices can reach the server within a distance of 320 metres.[7] In addition, the flexibility of this protocol allows to add any number of client devices and even to instantiate parallel communication between them. This feature further expands the possible applications of the magnetic scores to large group performances and to different interaction modes.

Because the sensors transmit position-related data and no switches are embedded in the discs, the activation of specific behaviours at will is not easily achieved by the performer. We consider this as a feature of the system, which partially limits the performer's control and favours the emerging of the composer's intention. Nonetheless, in order to offer to the performer some agency over the individual dimensions, we leveraged the design features of the embedded magnetic sensor[8], whose axes individually saturate when the magnetic field is too close. The magnet's cavity is placed on the disc's lid two millimetres above the back of the sensor. Because of this, when the magnet is entirely inside the disc (i.e. no magnetic field of identical polarity is encountered or resisted) the z axis saturates, returning the maximum value regardless of the presence of an external magnetic field. When the performer encounters a magnet on the board with identical polarity and releases the palm's pressure on the disc, the disc's magnet moves away from the sensor, and the z axis starts reporting correct values. Through this, it becomes possible to momentarily activate the reading of at least two parameters (one for each of the discs) at will.

In this iteration of the Magnetic Score, the sound is provided by the magnets mounted on the board through piezoelectric sensors, and the data resulting from the interaction between the magnets, the performer and the space, is provided by the discs. The visualisation of the data and the sound processing are performed on a laptop. A Max/MSP patch (Fig. 5) returns visual feedback on the orientation of the two controllers, on the presence and position of a nearby magnetic field and on the magnetic sensor's z axis activation and measurements. Even though in performative scenarios we imagine the interaction with magnetic scores mainly as a somatosensory experience, the availability of a

[7] https://www.espressif.com/en/news/ESP-NOW
[8] https://www.adafruit.com/product/4022

possibilities will become accessible.

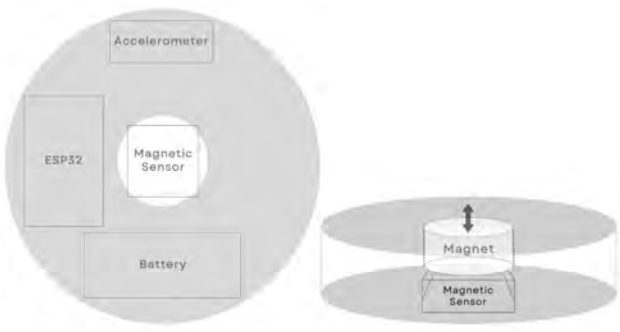


Figure 4. Magnetic Disc's Design.

real-time, three dimensional representation of the data has proven useful in calibrating the sensors and in describing the system to an audience, and might facilitate in the future the training of a performer or the application of machine learning algorithms for gesture recognition.

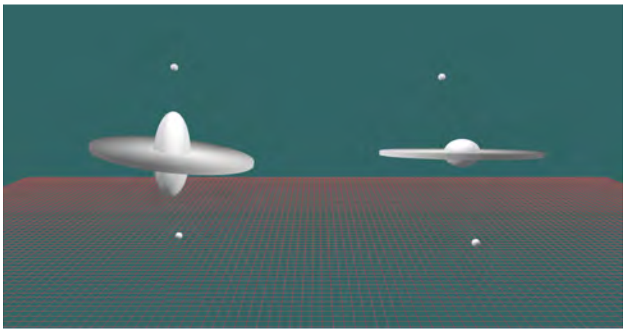


Figure 5. Magnetic Discs' Visualisation.

3.3 Sound Processing

In the occasion of the demo sessions described in section 5, we built a Max/MSP patch with a series of resonant filters, delay effects, and FM (with audio input as carrier) applied to the board's sounds conveyed through the two piezoelectric sensors in stereo configuration. The data forwarded to the laptop from the disc on the left and right hands is used to process respectively the left and right channels. The x and y values from the accelerometers define the centre frequency of the resonant filters and the x and y values from the magnetic sensors the amount of delay that is applied. Finally, the z value of the magnetic sensor, whose readings can be activated by approaching a field of identical polarity and releasing the pressure of the palm, activates and controls the amount of the frequency modulation. Future versions of the Magnetic Score will instead make use of Neural Synthesis [32]. By incorporating all of the sounds of the magnetic board within an AI synthesis model, thus separating the sound generation from the inscription layer, the encoding possibilities offered by the board's design will be highly extended, and new performative and compositional possibilities will become accessible.

4. PRESENTING THE MAGNETIC SCORE

We presented the Magnetic Score on two different occasions: a lecture with master's students in music composition and a discussion with a group of artists and researchers. During both events a demonstrative piece was performed.

In the first presentation, a prototype of the system was introduced to the participants and played by three of them, and a discussion followed. The initial comments centred around the possibility of designing different sound interactions: if the current version makes use of screws, spheres and magnets as sound sources, other designs might include boxes containing loose magnets and strings, or membranes and springs. The participants also suggested using multiple boards at once, each with a particular character defined by shape and interaction design, and to consider the possibility of extending the score to the whole room by embedding magnets inside double walls. The main limitation that emerged was the absence of a visual representation of the discs' position in space and in relation to nearby magnetic fields that could facilitate the initial understanding of the system. In response to this problem, we developed the Max/MSP patch described in 3.1, and presented it together with a more refined version of the Magnetic Score at the successive open event.

In the second session, the visual feedback was overall appreciated and helped clarify the relation between the discs and the surface. One participant noted that the magnetic discs could be separated from the rest of the system and independently explored as a musical instrument in their own right. The design of the discs is indeed articulated enough to generate complex interactions, and even though they do not unilaterally generate sound, the discs may be used to control a synthesis engine. Furthermore, mastering the control of the removable magnets inside each of them requires a good amount of practice, which makes them akin to traditional musical instruments.

During the discussion, a participant noted that when the element that generates the notation is the one that produces the sound, differentiating between the instrument and the score becomes a complex task. Even though the two concepts overlap in contemporary practices, a differentiation could still be observed in the amount of prescriptions that a system provides and in how they change in time. An instrument-score might be more akin to a score if it prescribes specific actions whose effects evolve in time, and more similar to a musical instrument if it provides a set of constraints for the performer to explore.

This final consideration is particularly useful in order to frame a reflection on the specificities of the Magnetic Score. In what ways this system may be seen as a score rather than a musical instrument? What is the experience of relating with magnetic fields as carriers of performative information? In the next chapter we explore magnetic scores as compositional system, focus on the somatosensory experience of the interaction, and frame them as a particular type of inherent score in which the inscription is relationally generated by the interaction of the magnetic fields.

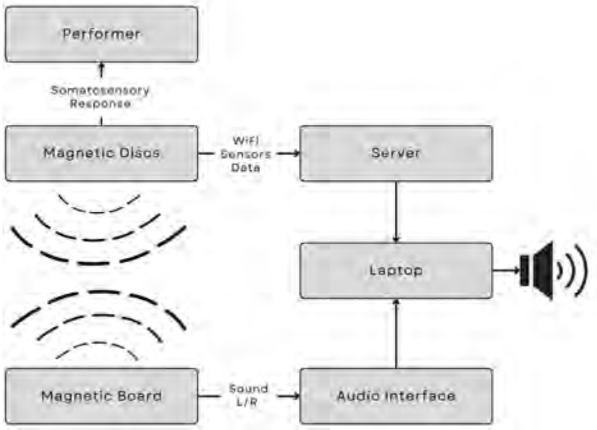


Figure 6. The Magnetic Score System.

5. DISCUSSION

As noted during the second public encounter, the overlapping of the instrument and the score is an apparent feature of this system. Because of the physical decoupling between the board and the two magnetic discs, it may be possible to interpret the Magnetic Score as an instrument (the board) being played with a device that excites it (the disc), similarly to a violin and a bow, a guitar and a pick, or a drum set and a drumstick interact. This is a useful metaphor in describing the generation of the sound in this particular instance of the system, but it does not take into account the variability of the mapping between the discs and the board and above all the articulated information that is possible to encode through the deliberate displacement of the magnets.

As explained above, the Magnetic Score is a compositional environment in which each board and mapping of the disc's parameters functions as an autonomous composition. By organising the invisible, attractive and repulsive forces embedded within, the composer guides the performer along the board. Similarly, by defining the mappings between the discs and the processing of the sound, it becomes possible to design the acoustic environment in which the interaction unfolds.

5.1 Magnetic Inscriptions

Composing a magnetic score appears as a very different experience than that of writing music on paper or other visual supports, as it requires to consider the performer's interactive, sensuous experience with the inscription. It also differs from engraving a tangible score in that, rather than focusing on fine tactility, it entails a more holistic interaction: the performer experiences the score as a force that dynamically pushes and pulls the arms, that unbalances the body, and through this produces specific gestures rather than suggesting them. By rehearsing the score, a performer learns to oppose the strength of the magnets, to follow them on the board, and internalises an abstract representation of the magnets' position and strength.

Through the size and positioning of the magnets, the composer can suggest specific gestures to the interpreter. In performing with the Magnetic Score, we realised that patterns of magnets with alternating polarities (such as in Fig. 2.3) suggest rapid movements over the board, as it becomes complex for the performer to operate with the discs. On the contrary, large magnets (see Fig. 2.7) are better suited for slow, vertical motions, as the magnetic fields interact with each other in a more predictable way. Loose magnets underneath the board are instead easier to control, and the performer's gestures tend to focus on the sound rather than on navigating the magnetic field.

We also realised that the granularity in the perception of the magnetic inscriptions is considerably lower than the one achievable through vision or direct tactility. As a consequence, in order to facilitate the recognition of magnetic patterns and avoid undesired interactions between the magnets, it is advised to use large surfaces as boards. By distancing the magnetic fields, the composer's intention can be interpreted more clearly, and it becomes possible to combine the magnets in order to inscribe simple shapes or symbols.

5.2 The Magnetic Score as Inherent Score

Because inherent scores combine inscriptions suggesting performative gestures with a device that generates the sound, we consider magnetic scores as an instance of this category. Nonetheless, the described system displays notable differences with other typologies of inherent scores, such as tangible scores and composed instruments, as well as similarities with situational practices such as event scores.

Typical tangible scores embed visual information on the surface of the instrument, and despite the overlapping of the score with the instrument in the act of performing, it is still possible to observe the sign from a distance, without interacting with it. Because no visual representation of the magnets is provided, and more broadly no information (except for the board's dimensions) is visually accessible to the performer, in the current version of the Magnetic Score the notation and the instrument further combine into an inextricable unity: in order to be read, the score has to be experienced as a holistic and sensuous encounter within the performative act.

Alternatively, the system may be observed from the perspective of composed instruments (as defined in 3.1). In magnetic scores, because of the interdependence between the controllers and the surface in generating the inscription, the mapping is not completely arbitrary as it would be expected on a composed instrument. Furthermore, even though we do not intend to generalise this feature to all magnetic scores, in this specific instance the gestural component is not decoupled from the sound generation, as the discs directly excite the loose magnets in order to generate the sounds. Nonetheless, through the processing techniques applied to the sound, the composer is capable of architecting time, which is a critical aspect of most music notation.

5.3 Relational Inscriptions

Magnetic scores also mutuate some of their features from event scores. As we have seen in 3.2, in numerous event scores the inscription is dynamically generated through the performers' interactions inside (and sometimes with) the performative space. The idea of the artwork as situationally emerging within the social context in which it is experienced, and out of the complexity of the relations between the people involved in its production and fruition, is at the centre of Bourriaud's relational aesthetics. From this stance, the aesthetic experience becomes a participative process that discursively generates the artwork.

This relational take is a key feature of the magnetic score's design. Rather than functioning one as an encoding and the other as a decoding component, both the board and the controllers have magnets embedded within. As a consequence, the readings of the magnetic fields that carry the inscription are the emergent result of their reciprocal interactions. Because of this, the score could not be considered as inscribed on the board alone, nor is it the unilateral result of the performer's action: it rather dynamically emerges as a series of events, or encounters, between the composer's ideas as inscribed through the displacement of the magnets, and the performer's exploration of the board with the magnetic discs.

Through this, the magnetic score invites the performer to become an active participant in the compositional process, and it does so at the inscription level, by translating the composer's and performer's intentions into a common somatosensory and sonic manifestation. This is a critical aspect to take into account during the compositional process: in designing the magnetic board, it becomes necessary to consider how the magnets' positioning suggests particular gestures to the performer, and how such gestures are affected by the shape, material and orientation of the board.

In this first iteration, we developed a rectangular board whose width is much greater than the height. This suggests a longitudinal exploration of the inscription and allows to operate symmetrically with each disc on one end. The mappings of the discs and the sound processing reflect this symmetry, as they control identical parameters, one on the left, and the other one on the right channel. Other pieces might instead explore asymmetrical mappings, in which one disc influences the parameters of the other. In such cases, the magnet's position and the board's shape might change accordingly, suggesting a whole different set of interactions.

6. FUTURE WORK

As observed by the participants in the evaluation sessions, the Magnetic Score may be further articulated in a variety of ways. In future pieces, it is our intention to increase the physical dimensions of the magnetic surface in order to embed more magnets and extend their distancing. By leveraging the portability and long communication range of the magnetic discs, we envisage to experiment with extended three dimensional surfaces as well as with architectural spaces such as entire rooms and buildings. In such situations, we wish to dig into the diffused character acquired by the musical score, and into the different subjectivities emerging out of the performers' embodied interaction with differently informed spaces.

We also envisage experimenting with additional inscription layers. Through transparent surface revealing the position of the magnets, or through symbols written on the magnetic board, the composer may be able to suggest more articulated interactions and to build upon the incidental relations emerging between the haptic and the visual domains. Additional inscriptions may be also generated by introducing materials that interact with the magnetic fields, such as ferromagnetic powder or ferrofluids. Through this, a dynamic representation layer would be introduced, thus combining the prescriptive nature of the magnetic score with a descriptive one and changing the grounding of the audience and of the performers in relation to the score.

Finally, we anticipate to further develop this system by coupling the permanent magnets embedded on the surface with a series of electromagnets whose polarities and strength are digitally controlled. This would allow to dynamically change the notation, and to introduce new agents in the form of generative algorithms. The information generated by the discs through the interaction with the electromagnets would in turn influence the system, thus instantiating a communication loop between the human performer and the computer. By introducing elements capable of exerting agency such as AI tools [33, 34] or artificial life simulations [35] as in Fig.7, and by exploring their embodied navigation, the roles of the performer and of the composer, the concepts of authorship and creativity and ultimately the cultural function of the musical score may be subject to further change.

Figure 7. *Ferroneural*, Jack Armitage and Nicola Privato.

7. CONCLUSIONS

In this paper we have presented the Magnetic Score, a system consisting of a surface with magnets mounted underneath, and two magnetic discs held by the performer. At the core of the Magnetic Score's compositional approach is the interaction between the magnetic fields of the discs with those of the magnets mounted on the board. Through this interaction, the score's inscriptions are generated as somatosensory feedback and interpreted as data for the processing of the sound.

We explored magnetic scores with the aim of reflecting upon the merger between score and instrument in contemporary musical practices from the perspective of systems that do not solely rely on graphic signs. We argued that magnetic scores are a subcategory of inherent scores that further merges information, sound generation and representation into an inextricable whole: in order to be read, the score has to be experienced.

We also argued that in magnetic scores the nature of the inscription is relational, in that it emerges from the interaction between the controllers held by the performer with the surface as designed by the composer, rather than being unilaterally inscribed. Because of this, the creative intention of the composer inextricably merges and overlaps with that of the performer, supporting the indeterminacy and openness of the modern artistic poetics as postulated by Eco, as well as the transitivity of the aesthetic experience as described by Bourriaud.

We believe that the magnetic score adds to the already pluralistic and heterogeneous nature of contemporary scores and notational practices, in that it explores them from the perspective of the embodied experience, by suggesting performative gestures through the reciprocal attraction and repulsion of the magnetic fields. In this paper we have presented a generalised overview of this system, and defined it as the combination of a bi-dimensional or tri-dimensional surface with embedded magnets, and two magnetic discs that decode the information. Nonetheless, specific instances can be as deterministic and prescribed as the composer desires.

By performing and presenting this system to an audience new research questions have arised: what new performative practices does the embodied, somatosensory manifestation of the score suggest? What types of information is it possible to convey through magnetic inscriptions? What other forms could magnetic scores take? Our intention is to release the hardware and software specifications of the system such that other people can build their own and contribute to the exploration of these questions.

8. ACKNOWLEDGMENTS

This research is supported by the European Research Council (ERC) as part of the Intelligent Instruments project (INTENT), under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 101001848).

INTENT is also supported by an NVIDIA hardware grant of two A5000 GPUs.

9. REFERENCES

[1] U. Eco, *The Poetics of the Open Work. The Open Work*. Cambridge, MA: Harvard University Press, 1989.

[2] T. Magnusson, *Sonic Writing: Technologies of Material, Symbolic and Signal Inscriptions*. New York, NY: Bloomsbury Academic, 2019.

[3] E. Tomás and M. Kaltenbrunner, "Tangible scores: Shaping the inherent instrument score," in *Proceedings of the International Conference on New Interfaces for Musical Expression*. London, United Kingdom: Goldsmiths, University of London, Jun. 2014, pp. 609–614. [Online]. Available: http://www.nime.org/proceedings/2014/nime2014_352.pdf

[4] M. Burtner, "Composing for the (dis)embodied ensemble : Notational systems in (dis)appearances," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, Montreal, Canada, 2003, pp. 63–69. [Online]. Available: http://www.nime.org/proceedings/2003/nime2003_063.pdf

[5] N. Privato and A. Novello, "W.e.i.r.d: Enters the stage," in *xCoAx 2021 9th Conference on Computation, Communication, Aesthetics and X*, D. Research Institute in Art and Society, Eds. Porto, Portugal: School of Fine Arts, University of Porto, 2019, pp. 126–139.

[6] E. T. Einarsson, *Re-Notations: Flattening Hierarchies and Transforming Functions*. Belgium, Leuven University Press, 2017.

[7] D. Tudor and R. Kuivila, in *Sources : Words , Circuits and the Notation-Realization Relation in the Music of*, 2004.

[8] T. Magnusson, "The musical score: the system and the interpreter," in *International Symposium on Electronic Art (ISEA2011)*. ISEA, 2011. [Online]. Available: http://sro.sussex.ac.uk/id/eprint/46866/

[9] V.-R. Carinola and J. Geoffroy, "On notational spaces in interactive music," in *Proceedings of the International Conference on Technologies for Music Notation and Representation – TENOR'2022*, V. Tiffon, J. Bell, and C. de Paiva Santana, Eds. Marseille, France: PRISM Laboratory, 2022, pp. 149–154.

[10] C. Cardew, *Treatise*. Alfred Music, 2022.

[11] E. Brown, "On december 1952," *American Music*, vol. 26, pp. 1–12, 2008.

[12] C. Karpodini and T. Michailidis, "Making graphical scores more accessible: A case study," in *Proceedings of the International Conference on Technologies for Music Notation and Representation – TENOR'2022*, V. Tiffon, J. Bell, and C. de Paiva Santana, Eds. Marseille, France: PRISM Laboratory, 2022, pp. 51–56.

[13] A. Lucier, "Origins of a form: Acoustical exploration, science and incessancy," *Leonardo Music Journal*, vol. 8, pp. 5–11, 1998. [Online]. Available: http://www.jstor.org/stable/1513391

[14] P. Oliveros, *Anthology of Text Scores*. Deep Listening Publications, 2013.

[15] R. S. James and D. Revill, "Mumma, gordon," 2001. [Online]. Available: https://www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000019346

[16] T. Magnusson, "When instruments become architecture: on liquefying frozen music," 2011, a commissioned essay written for Peter Vogel?s Sound of Shadows Exhibition, Brighton, 2011. [Online]. Available: http://sro.sussex.ac.uk/id/eprint/46863/

[17] S. Beal, "Trimpin, gerhard," 2015. [Online]. Available: https://www.oxfordreference.com/view/10.1093/acref/9780199743391.001.0001/acref-9780199743391-e-7727

[18] E. Tomás, *The interface-score: Electronic Musical Interface Design as Embodiment of Performance and Composition*. Linz, Kunstuniversität, 2018.

[19] N. Schnell and M. Battier, "Introducing composed instruments, technical and musicological implications," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, Dublin, Ireland, 2002, pp. 156–160. [Online]. Available: http://www.nime.org/proceedings/2002/nime2002_156.pdf

[20] S. Bhagwati, "Comprovisation–concepts and techniques," *(Re) Thinking Improvisation*, 2013.

[21] S. Bhagwati and M. Giordano, "Musicking the body electric. the "body:suit:score" as a polyvalent score interface for situational scores." in *Proceedings of the International Conference on Technologies for Music Notation and Representation – TENOR'16*, R. Hoadley, C. Nash, and D. Fober, Eds. Cambridge, UK: Anglia Ruskin University, 2016, pp. 121–126.

[22] E. Gunther, G. Davenport, and S. O'Modhrain, "Cutaneous grooves: Composing for the sense of touch," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, Dublin, Ireland, 2002, pp. 73–79. [Online]. Available: http://www.nime.org/proceedings/2002/nime2002_073.pdf

[23] N. Baptiste, "Access to musical information for blind people," in *Proceedings of the First International Conference on Technologies for Music Notation and Representation*. Institut de Recherche en Musicologie, May 2015, pp. 231–235. [Online]. Available: https://doi.org/10.5281/zenodo.923889

[24] L. S. Pardue and J. A. Paradiso, "Musical navigatrics: New musical interactions with passive magnetic tags," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, Dublin, Ireland, 2002, pp. 145–147. [Online]. Available: http://www.nime.org/proceedings/2002/nime2002_145.pdf

[25] G. Ouzounian, "The uncertainty of experience: On george brecht's event scores," *Journal of Visual Culture*, vol. 10, no. 2, pp. 198–211, 2011.

[26] S. Bhagwati, "Elaborate audio scores: Concepts, affordances and tools," in *Proceedings of the International Conference on Technologies for Music Notation and Representation – TENOR'18*, S. Bhagwati and J. Bresson, Eds. Montreal, Canada: Concordia University, 2018, pp. 24–32.

[27] C. Sdraulig and C. Lortie, "Recent audio scores: Affordances and limitations," in *Proceedings of the International Conference on Technologies for Music Notation and Representation – TENOR'19*, C. Hope, L. Vickery, and N. Grant, Eds. Melbourne, Australia: Monash University, 2019, pp. 38–45.

[28] E. Schimana, "Sound as score," in *Proceedings of the International Conference on Technologies for Music Notation and Representation – TENOR'19*, C. Hope, L. Vickery, and N. Grant, Eds. Melbourne, Australia: Monash University, 2019, pp. 33–37.

[29] N. Bourriaud, *Relational Aesthetics*. Les Presses du Réel, 2002.

[30] J. A. Paradiso, K.-y. Hsiao, and A. Benbasat, "Tangible music interfaces using passive magnetic tags," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, Seattle, WA, 2001, pp. 30–33. [Online]. Available: http://www.nime.org/proceedings/2001/nime2001_030.pdf

[31] M. C. Mannone, E. Kitamura, J. Huang, R. Sugawara, and Y. Kitamura, "Cubeharmonic: A new interface from a magnetic 3d motion tracking system to music performance," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, T. M. Luke Dahl, Douglas Bowman, Ed. Blacksburg, Virginia, USA: Virginia Tech, Jun. 2018, pp. 350–351. [Online]. Available: http://www.nime.org/proceedings/2018/nime2018_paper0076.pdf

[32] A. Caillon and P. Esling, "Rave: A variational autoencoder for fast and high-quality neural audio synthesis," 2021. [Online]. Available: https://arxiv.org/abs/2111.05011

[33] N. Privato, O. Rampado, and A. Novello, "A creative tool for the musician combining lstm and markov chains in max/msp," in *Artificial Intelligence in Music, Sound, Art and Design*, T. Martins, N. Rodríguez-Fernández, and S. M. Rebelo, Eds. Cham: Springer International Publishing, 2022, pp. 228–242.

[34] *Scramble Live: Combining LSTM and Markov Chains for Real-time Musical Interaction.* Zenodo, Jun. 2022. [Online]. Available: https://doi.org/10.5281/zenodo.6576266

[35] J. Armitage and T. Magnusson, "Agential Scores: Exploring Emergent, Self-Organising and Entangled Music Notation," in *Proceedings of the 8th International Conference on Technologies for Music Notation and Representation*.

THE REACTIVE SCORE

David Kim-Boyle
Sydney Conservatorium of Music
david.kim-boyle@sydney.edu.au

ABSTRACT

Scores designed to be displayed on screens provide the opportunity for composers to dynamically update information and musical instructions presented to performers during the moment of performance. Such scores may be responsive to the agency of performers or audience members, providing new forms of structural organization or ways in which denoted musical material can be transformed. This paper explores the creative possibilities of such *reactive scores* and situates them within a historical tradition of malleable notation. Two works by the author are discussed in which real-time features of the musical performance drive the notational transformations of the performance score.

1. INTRODUCTION

The concept of a score as an assemblage of malleable parts is not necessarily a new one [1, 2]. Stockhausen's *Refrain* (1959) which features a mobile transparent strip containing various articulations and ornamentations which apply to the printed score, and John Cage's works featuring transparencies and printed notation such as *Fontana Mix* (1956) which are assembled by performers to create a performance score, are two early examples of such approaches to notation.

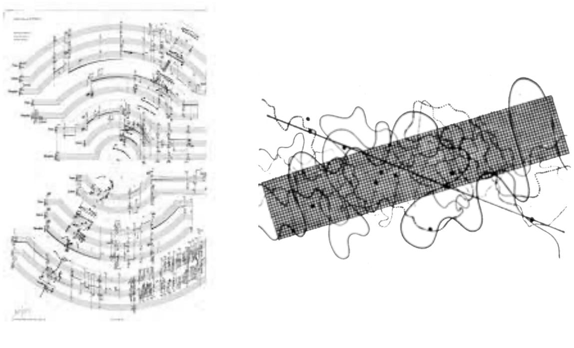


Figure 1. Stockhausen's *Refrain* (1959) with its distinctive circular score (left), and Cage's *Fontana Mix* (1956) which is assembled by the performers from preprinted sheets and transparencies (right).

While the malleability of a score is naturally constrained by the physical properties of the media upon which it is inscribed, the growing adoption of screen-based scores and the rapid development of networking technologies, has afforded new degrees of transformational agency to both audiences and performers. To an extent, such reactive scores are situated within a tradition of participatory art although there is an important distinction to be made in that much participatory art practice is motivated by themes of political and social engagement [3, 4], or a desire to situate the audience in performative, creative roles [5]. The type of reactive scores discussed in this paper, however, are designed to be interpreted by performers with their range of possible transformations affected by the dynamic play of audience or performer agency.

2. AUDIENCE AGENCY

The correlation of audience agency to a performance score's transformation typically follows one of two strategies – synchronous, where audience members are invited to help shape a performance score during the live performance, or asynchronous, where a score's instantiation is contingent on prior decisions or selections made outside of the live performance space. In both modes, audience agency typically operates within a preestablished framework, constrained to affect a select number of musical parameters.

Composer Jason Freeman has adopted both synchronous and asynchronous modes of audience engagement. In his *Saxophone Etudes* (2011), members of the audience are invited to help shape the musical properties of the score displayed to the saxophonist during the live performance [6]. Through a simple interface displayed on a smartphone, see Figure 2, audience members can vote on a range of tempo, dynamic and articulation options. The collective results of this polling are superimposed onto the score read by the performer, helping to guide their live interpretation.

Figure 2. The smartphone interface for Jason Freeman's *Saxophone Etudes* (2011) used by listeners to guide the performer's interpretation.

An asynchronous approach to audience agency is featured in Freeman's earlier work *Graph Theory* (2006), for solo violin or cello. In this work, audience members choose different pathways through short, looping melodic units presented on a dedicated website with these choices affecting the probability weightings used to generate a performance score, see Figure 3. Of particular note in *Graph Theory* is the UI design which adopts simple color mappings and a graph style visualization of a melodic cell's pitch-time structure, to facilitate understanding by users who may not be able to read common practice notation.

Figure 3. User interface for Jason Freeman's *Graph Theory* (2006).

Kevin Baird's *No Clergy* (2005) adopts both an asynchronous and synchronous approach to the real-time creation and transformation of a score [7]. During performance, members of the audience can affect a range of musical properties such as articulation and dynamics which affect subsequent development of the score displayed with GNU LilyPond. Like Freeman's *Graph Theory*, the interface presented to the audience avoids common practice notation, and features simple slider controls and buttons, see Figure 4. Similarly, the user interface for Zhang et al.'s *Open Symphony* web-based interactive system [8], has been carefully designed to present an easily intuited control system for presenting different styles of musical performance, see Figure 4.

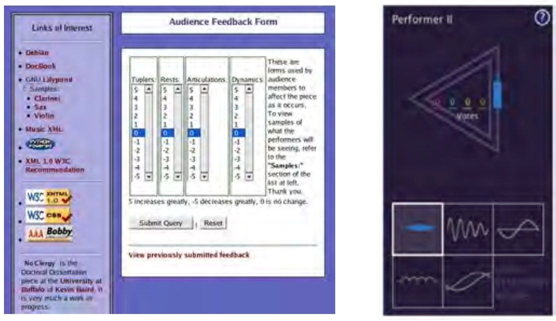


Figure 4. User interface design for Kevin Baird's *No Clergy* (2005) (left) and Zhang et al.'s *Open Symphony* (right).

Rather than interacting through a software interface, Wulfson's LiveScore system features a hardware controller which audience members can manipulate during performance to affect the textural density, pitch, rhythmic, and dynamic properties of a screen-based score [9]. Such an approach to audience engagement, where the actions of audience members are so obviously foregrounded, promotes a very different experience than those works in which the agency exercised by the audience is more concealed.

3. PERFORMANCE AGENCY

Perhaps not surprisingly, works featuring reactive scores responsive to audience agency form a small number. More commonly, score transformations tend to be correlated to decisions made by performers prior to performance, as in works such as *Refrain* or *Fontana Mix*, or synchronously through response to decisions or musical activity made during the moment of performance itself.

A key technical factor in the evolution of performance-driven reactive scores has been the development of score following techniques and softwares such as AnteScofo [10] and software technologies for facilitating the real-time display of musical notation including INScore [11], MAXScore [12], and Bach [13]. While most score following software has been designed to track the nuances of a live performance and provide synchronous electronic accompaniment, see for example works such as Marco Stroppa's *...of Silence* (2007) or Philippe Manoury's recent *Das Wohlpräparierte Klavier* (2021), more simple techniques of providing control over a score's evolution have included the use of foot pedal cues such as those employed in Seth Shafer's *Terraformation* (2016) in which the performer controls progression through the screen-score through various pedal cues [14].

In his work *Semaphore* (2014), composer Richard Hoadley has developed a reactive notation system in which the physical gestures of dancers are tracked with a Kinect interface and used to transform the display of a predetermined text [15]. Spectral features of the spoken text are then used to generate music notation which is in turn presented to performers through the INScore system [11]. The musical performance is, in turn, interpreted by the dancers establishing a complex feedback loop.

A more unusual mapping of performance physicality to notation occurs in Erich Berger and Peter Votava's, performing as the duo Terminalbeach, work *Pixelache* with the Heart Chamber Orchestra [16]. In this work, the heartbeat of each member of the orchestra is monitored with electrocardiogram sensors. This data is then used to generate a musical score which is displayed on screens placed in front of each of the performers, see Figure 5.

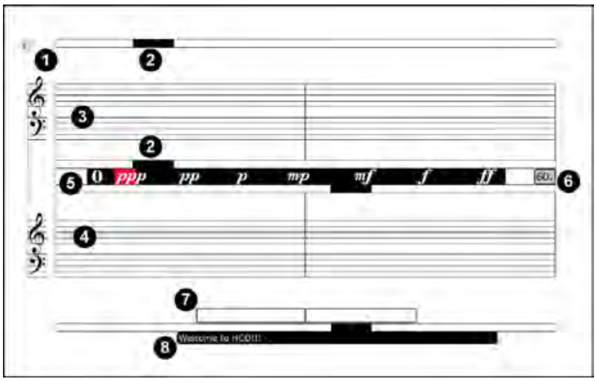


Figure 5. Performance score generated from ECG data in Berger and Votava's *Pixelache* (2012).

Feedback systems, which are a natural paradigm of reactive notations, are more overtly featured in Andrea Valle's *Dispacci dal fronte interno* (2012) for strings and live electronics. In Valle's work, audio features of the live string performance are analysed and used as control material for processes which generate music notation that is printed during performance, see Figure 5, passed on in the form of dispatches to the live musicians and reinterpreted [17].

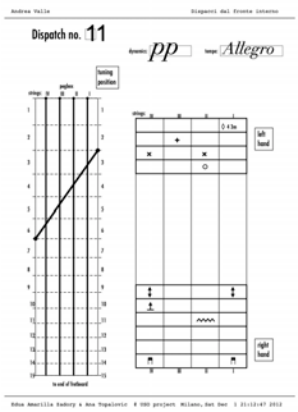


Figure 6. Notation generated in response to performance input for Andrea Valle's *Dispacci dal fronte interno* (2012).

For composers exploring the affordances of notation responsive to performance agency, the extraction of audio features of a live performance is necessary. Composer Sam Hayden and violinist Mieko Kanno have recognized this in current work being undertaken in the development of their live notation system NEXUS [18], which is integrating real-time audio analysis into the generative processes used to create notation. This is being achieved through a range of external Max external objects which provide real-time information on musical parameters including pitch, amplitude, and timbre.

4. FEATURE EXTRACTION AND MAPPING

The author has explored the creative affordances of reactive scores in two recent works, both of which feature three-dimensional graphic performance scores presented to the performers in mixed reality space. Both works rely on the real-time analysis of a live performance to extract audio features which are subsequently mapped to visual properties of the performance score.

4.1 5x3x3 (2019)

5x3x3 (2019) was developed for the ELISION ensemble and uses the HoloLens, an augmented reality head-mounted display developed by Microsoft, to project a three-dimensional holographic score in the performance space for each of the three performers. The score itself is dynamically transformed during performance and consists of a three-dimensional construct of colored nodes connected by lines, or edges in network graph nomenclature, of various colors and curvature. Each of the nodes represents the onset of a musical note with the node's color denoting a type of articulation. Pitches are specified by edge colors and the duration of notes by an edge's spatial length.

As performers physically explore the performance score holographically situated within the performance space, an FFT analysis is performed upon the sonic results within a Max patch. Data derived from this analysis is then scaled and sent via OSC to transform the curvature of the edges connecting nodes within the performance score and the size of the nodes themselves, see Figure 5. Edge curvature is interpreted by performers in the form of timbral transformation with the scale of curvature directly correlating to the scale of transformation while relative node size is mapped by the performers to different dynamic levels.

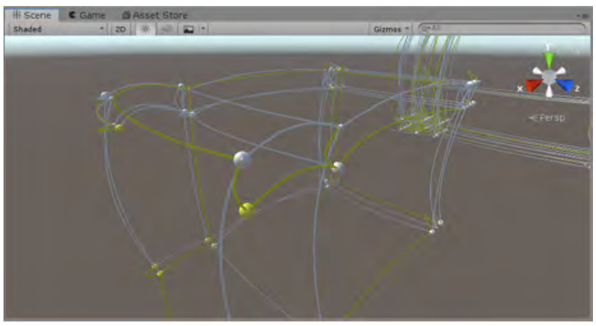
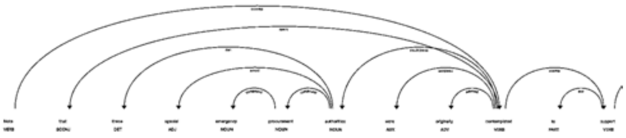


Figure 5. Selection from the score for *5x3x3* (2019) within the Unity 3D development environment showing various edge curvatures between nodes. A real-time FFT analysis is used to map control points of the Bézier curves drawn between nodes.

4.2 reTweets (2022)

5x3x3 was the first work of the author's to explore the musical affordances of reactive notation. Given the technical complexity of the performance system, the mapping of real-time performance data to features of the graphic score was deliberately kept relatively straightforward. *reTweets* (2022) adopts a more sophisticated approach. Like *5x3x3*, it features a three-dimensional score presented in mixed reality space but in *reTweets*, the performance score is generated through a real-time linguistic analysis of tweets posted to Twitter.

reTweets is run from a Jupyter Notebook and uses the Tweepy Python library to access the Twitter public API and return tweet data on predetermined keywords.[1] A dependency parse is performed on the returned tweet with spaCy,[2] a powerful natural language processing library, and the word vectors of each token are called from Gensim using a model trained on Google news.[3] While it is beyond the scope of this paper to delve deeply into the theory of word vectorization and embedding, for this the reader is referred to [19], the basic Natural Language Processing concept is that words may be represented with high dimensional vectors, where words that are semantically similar, e.g. "bathe", "wash", "clean" will return word vectors closer than words semantically dissimilar, e.g. "apple", "laundry", "smoke". In Gensim, word vectors are of default size 100, which in order to be mapped to a Cartesian coordinate, as is required in the *reTweets* visualizations, must be reduced to three. In *reTweets*, this dimensional reduction is performed with scikit-learn,[4] with all returned data subsequently visualized in a three-dimensional node graph within a VR scene. An example of a retrieved tweet, its dependency parse, and subsequent visualization is presented in Figure 6. Each token within the tweet is represented by a node and the dependencies between tweets denoted by white-colored edges.



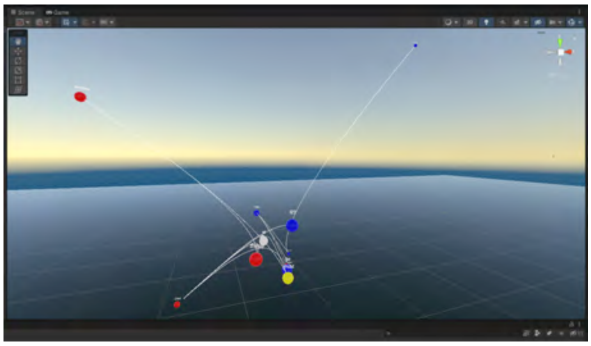


Figure 6. Original tweet (upper), partial dependency parse (middle) and three-dimensional visualization (lower) in the Unity 3D development platform of a tweet retrieved on keyword "Ukraine".

reTweets adopts a similar mapping strategy to previous works by the author where nodes denote the onset of sonic events with their color indicating specific pitches. The duration of these events is denoted by the spatial separation between nodes while connecting edges suggest how performers may determine the sequence of events. Visualized tweets are positioned in a series within a VR scene as shown in Figure 7.

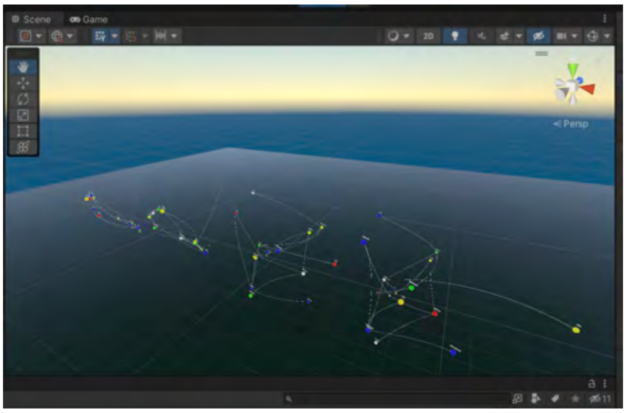


Figure 7. Positioning of tweet visualizations within the VR scene in the Unity 3D development platform.

During the live performance, various audio features of the musical performance (pitch, amplitude, spectral centroid, and spectral flux) are extracted with the zsa.descriptors library [20]. This data is then used to transform the spatial distribution of nodes within each of the tweet visualizations which establishes new temporal relationships between events.

Supplementing the feedback loop established between the score and the performers, *reTweets* features an

[1] https://tweepy.org
[2] https://spacy.io
[3] https://radimrehurek.com/gensim/index.html
[4] https://scikit-learn.org

additional feedback mechanism whereby the spatial repositioning of nodes generates new word vectors and tweets which are subsequently posted back to Twitter. This process is built on the computation of Word2Vec similarities within Gensim [21] and effectively creates a greater pool of tweets which may be sampled for recurrent visualizations. A schematic of the entire performance system for *reTweets* is presented in Figure 8.

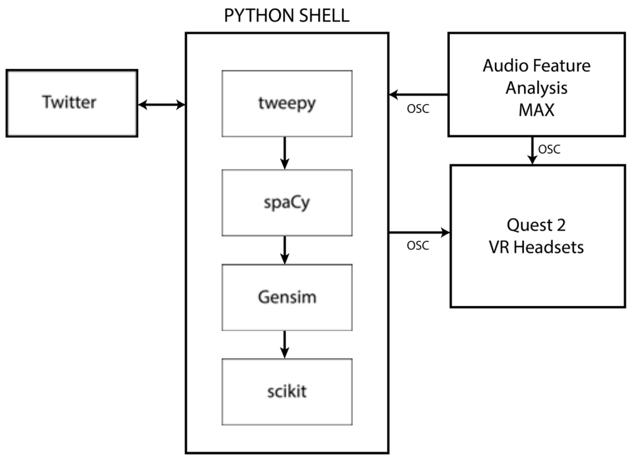


Figure 8. Schematic of the performance system for *reTweets* (2022).

5. CONCLUDING THOUGHTS AND FUTURE DIRECTIONS

The correlation of audience or performance agency to notational features within a performance score presents numerous compositional challenges. Foremost amongst these, perhaps is the extent to which scores are affected by agency and how this might in turn affect the large-scale organization of musical material. In the author's work, while performance agency affords some flexibility of musical results, this agency is always bound within certain pre-determined constraints allowing consistent musical structure to be realized. In *5x3x3*, for example, the timbral transformations folded into the score are always bounded by certain limits while in *reTweets*, the degree of score transformation becomes its own compositional determinant. Similarly in works in which audience agency is correlated to score transformations such as those described earlier, this agency becomes a carefully mediated action of choice. Through the foregrounding of such choice, members of the audience are encouraged to exercise their agency which, ironically perhaps, discourages any willingness to linger within a moment. Chan frames this dichotomy in terms of the *vita activa* and the *vita contemplativa* arguing that the push away from the contemplative life fundamentally transforms our experience of duration and time towards one of consumption [22]. How audience agency is reconciled within such an aesthetic framework is a particular challenge.

While there are certainly many areas of promising enquiry with the potential to yield fruitful musical results, the rapid evolution of GPT-3 models suggest one particularly interesting area of investigation. Aligned with features obtained through a more sophisticated stylistic analysis of live performance, perhaps through the application of previously developed techniques of learning a musical corpus [23], real-time predictive modelling could be used to create live, reactive scores that facilitate particular types of musical collaboration within members of a small ensemble. Scores could be adaptive to certain musical styles, fostering new types of musical collaboration and indeed enabling collaboration itself to assume structural valence. However these possibilities might play out, it seems inevitable that more sophisticated techniques of feature extraction will be an important contributor to any future development of reactive scores.

How works which feature reactive scores might offer new aesthetic affordances is of course an open question. Deleuze frames the collapsing of the past into the present, an experiential characteristic of any work built on feedback systems, through his concept of the plane of immanence and perhaps that is a helpful place to begin such an analysis [24]. Whatever the outcomes, reactive scores would seem to offer unique pathways for aesthetic investigation and exciting opportunities for creative expression.

6. REFERENCES

[1] A. Clay, "Inventing Malleable Scores: From Paper to Screen Based Scores," in R. Adams, S. Gibson, S. M. Arisona (Eds.), *Transdisciplinary Digital Art, Sound, Vision and the New Screen*. Vol. 7. Springer-Verlag, 2008

[2] R. Ross-Smith, *A Practical and Theoretical Framework for Understanding Contemporary Animated Scoring Practices*. PhD Dissertation. Rensselaer Polytechnic Institute, 2016.

[3] C. Bishop, *Artificial Hells: Participatory Art and the Politics of Spectatorship*. Verso, 2012.

[4] J. Rancière, *The Emancipated Spectator*. Translated by G. Elliott. Verso, 2009.

[5] J. Toelle and J. A. Sloboda, "The Audience as Artist? The Audience's Experience of Participatory Music," *Musicae Scientiae*, Vol. 25(1), 2021, pp. 67-91.

[6] N. Weitzner, J. Freeman, S. Garrett, and Y.-L. Chen, "massMobile – An Audience Participation Framework," in *Proceedings of the New Interfaces for Musical Expression Conference (NIME'12)*, Ann Arbor, 2012.

[7] K. Baird, "Real-Time Generation of Music Notation via Audience Interaction using Python and GNU Lilypond," in *Proceedings of the New Interfaces for Musical Expression Conference (NIME'05)*, Vancouver, 2005.

[8] L. Zhang, Y. Mu, and M. Barthet, "A Web Application for Audience Participation in Live Music Performance: The Open Symphony Use Case," in *Proceedings of the New Interfaces for Musical Expression Conference (NIME'16)*, Brisbane, 2016.

[9] G. D. Barrett and M. Winter, "*LiveScore*: Real-Time Notation in the Music of Harris Wulfson," *Contemporary Music Review*, Vol. 29, 2010, pp. 55-62.

[10] A. Cont, "ANTESCOFO: Anticipatory Synchoniszation and Control of Interactive Parameters in Computer Music," in *Proceedings of the International Computer Music Conference (ICMC2008),* Belfast, 2008.

[11] D. Fober, Y. Orlarey, and S. Letz, "INScore: An Environment for the Design of Live Music Scores," in *Proceedings of the New Interfaces for Musical Expression Conference (NIME'11)*, Oslo, 2011.

[12] N. Didkovsky and G. Hajdu, "MaxScore: Music Notation in Max/MSP," in *Proceedings of the International Computer Music Conference (ICMC2008)*, Belfast, 2008.

[13] A. Agostini and D. Ghisi, "Bach: An Environment for Computer-Aided Composition in Max," in *Proceedings of the International Computer Music Conference (ICMC2012),* Ljubljana, 2012.

[14] S. Shafer, "Terraformation: For Violin or Viola and Computer, in *Proceedings of the International Conference on New Tools for Music Notation and Representation (TENOR'17)*, Cambridge, 2017.

[15] R. Hoadley, "Semaphore: Cross-Domain Expressive Mapping with Live Notation," in *Proceedings of the International Conference on New Tools for Music Notation and Representation (TENOR'15)*, Paris, 2015.

[16] P. Votava and E. Berger, "The Heart Chamber Orchestra," *eContact!*, Vol. 14(2), 2012. Available: https://econtact.ca/14_2/votava-berger_hco.html

[17] A. Valle, "Notazioni Elettromeccaniche, O Forse No," *Musica/Tecnologia*, Vol. 13, 209, pp. 67-84.

[18] S. Hayden and M. Kanno, "*Nexus*: Live Notation as a Hybrid Composition and Performance Tool," in *Proceedings of the International Conference on New Tools for Music Notation and Representation (TENOR'20)*, Hamburg, 2020.

[19] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient Estimation of Word Representations in Vector Space." Available: https://arxiv.org/pdf/1301.3781.pdf

[20] M. Malt and E. Jourdan, "Zsa.Descriptors: A Library for Real-Time Descriptors Analysis," in *Proceedings of the 5th Sound and Music Computing Conference*, Berlin, 2008, pp. 134-137.

[21] R. Rehurek and P. Sojka, "Gensim: Topic modelling for Humans," Available: https://radimrehurek.com/gensim/wiki.html

[22] B. -C. Han, *The Scent of Time: A Philosophical Essay on the Art of Lingering*. Translated by D. Steuer. Polity Press, 2017.

[23] A. Eigenfeldt, "The Human Fingerprint in Machine Generated Music," in *Proceedings of xCoAx2013: Computation, Communication, Aesthetics, and X*, Bergamo, 2013. Available: https://2013.xcoax.org

[24] G. Deleuze and F. Guattari, *What is Philosophy?* Translated by H. Tomlinson and G. Burchell, Columbia University Pres, 1994.

AGENTIAL SCORES: EXPLORING EMERGENT, SELF-ORGANISING AND ENTANGLED MUSIC NOTATION

Jack Armitage
Intelligent Instruments Lab
Reykjavik, Iceland
jack@lhi.is

Thor Magnusson
Intelligent Instruments Lab
Reykjavik, Iceland
thor.magnusson@lhi.is

ABSTRACT

Dynamic scores have gained popularity as an innovative intervention in musical performance, providing novelty for both performers and audiences. In this paper, we discuss agential scores and the implications of their emergent, self-organising, and entanglement affordances for musical performance ecologies. We approach this through practice-based research, introducing Tölvera, an artificial life software library for agential scores. We propose a typology of interaction scenarios for agential scores and investigate a subset of these, presenting the outcomes of early artistic encounters with Tölvera between two improvising guitarists. Reflecting on our experiences, we emphasise the unique challenges that emerge from engaging with scores as real-time agents, suggesting that agential scores promote fluidity of form in notation, which consequently prompts performers to identify with, mirror, and attune to them. Although scores have always possessed agency, we argue that a more explicit and practical focus on agency raises new questions and offers new possibilities for the intra-actions between human and non-human agents within musical ecologies.

1. INTRODUCTION

Musical scores serve as pivotal components within the musical cultures in which they are utilised, influencing the thoughts and actions of composers, performers, concert organisers, audiences, and music as a whole. This impact can be characterised in terms of material agency. Histor-ically, the vigour of this agency has been limited by dis-play and reproduction technologies, such as print media. However, computational multimedia scores are now truly coming to life through innovative mediums encompassing biology, artificial life simulation, and artificial intelligence. In the mediaeval era, music notation evolved into standards that represented principles in instrument making, musical training, and performance contexts. However, notation ex-perienced significant diversification beginning in the twen-tieth century.

Under prescriptive notation, new developments and symbols have been added to the traditional Western notational system. These include graphic scores, enhanced tablatures, verbal notation [1], event scores [2], generative software and live coding [3], instruction pieces [4], realisation scores [5], soundpainting [6], S-notation for DJs [7], and animated notation [8]. Additionally, tangible [9] and haptic [10] scores, artificial life scores [11], and more have emerged.

Generative and dynamic scores have been extensively investigated, and the role of computational technologies is crucial in their implementation: computers enable us to achieve truly "dynamic media" [12] that were impossible with acoustic or purely analogue electronic technologies. These scores exhibit such diversity that their own becomes fluid and interchangeable with that of the composer, instrument or audience. Such systems span from being open to interpretation, where performers interpret as they wish, to more rigid systems where performers precisely execute what the system generates, for instance, by playing generated notes displayed in staff notation.

In our work involving dynamic and agential scores, we endeavour to utilise the emergent and self-organising properties of artificial life as a notational system for composition and performance. This is part of our wider effort to develop a library of technical components for composing musical systems, encompassing real-time MIDI models [13], feedback systems [14], sensors, actuators, and more, which we refer to as the *Organolib* [1] . The Organolib does not target a specific combination of composer, interpreter, instrument, and audience; instead, it aims to promote an ecological mindset regarding the distribution of agencies among technical elements and musical roles. Ultimately, we strive to facilitate the emergence of epiphenomenal structures that could be characterised as musical agency.

This paper presents our preliminary investigations into agential scores through artificial life. In contrast to the recent trends in artificial intelligence that favour machine learning approaches, both in industry and the arts, artificial life remains a relatively marginalised field, albeit one that may be experiencing a resurgence in interest. For us, artificial life serves as an almost literal Petri dish for practical exploration into musical agency [15], providing a diverse array of systems and species that inevitably engage with topics such as biology, evolution, ecology, philosophy, computation, and more.

[1] https://iil.is/research/organolib

2. BACKGROUND

2.1 Perspectives on Agency

The concept of agency frequently arises in the realm of new musical instrument design. These instruments are perceived to possess a certain degree of agency, influencing how we think, play, and behave within a specific performance ecology [16]. However, there is no general consensus on the meaning of agency, and various theoretical fields offer different perspectives.

In this context, we can contrast analytical philosophy, which views agency as a property of an intelligent being, typically the intentional action of an ethical human being [17], with theoretical biologists who regard agency as a characteristic of living organisms, as exemplified in the theory of autopoiesis [18]. Additionally, sociologists and philosophers of technology consider agency as a property of matter and objects [19], or of their intra-action [20].

Frauenberger refers to the convergence of these diverse viewpoints as "entanglement HCI" [21], asserting that "it is through non-human agency that we can create nuanced links between design intent, context of use, and people intra-acting with technology." In response to this notion, Nordmoen and McPherson describe a "decentring from human to more-than-human" approach in their practical explorations of interactive systems, acknowledging the "ecologies of different types of knowledge and raw materials" involved [22].

The autopoietic theory by Maturana and Varela [18] characterises living organisms not as physical entities, but rather as networks of processes [23]. Some have advocated for a "more embodied reformulation of autopoiesis" to address the perceived overly disembodied nature of the autopoietic theory's original formulation [24]. This argument for embodiment resonates with Sarah Kember's cyberfeminist critique of ALife, which challenges earlier conceptions of life as information [25].

2.2 Exploring Agency through Boundary Objects

In the realm of musical instrument theory, agency is characterised as a relational concept [26]. Although sociologist Latour finds it reasonable to attribute agency to a specific door-locking mechanism [27] (thing-agency), such a notion may not be appealing to a biologist (bio-agency). As we engage with intelligent instruments, we are not only confident in our recognition of agency within the technologies we create but also actively investigating the boundaries surrounding the term. This is not a straightforward task, particularly as the concept of agency is rapidly evolving within the context of contemporary AI advancements.

In this regard, we perceive our agential scoring system, delineated in the subsequent section, as a boundary object [28] for examining agency from a multitude of diverse perspectives. Stapleton and Davis [29] discuss agency as distributed and relational, occurring in the interaction between the performer and the instrument. We concur with this ontology; however, in our work with artificial life and machine learning, it is also intriguing to investigate how we and the users of our instruments engage with these terms when projecting our ideas onto the dynamic behaviour of our systems [14]. Our research programme does not aim to rigidly define these terms from the outset; instead, we endeavour to observe how they are applied in the context of use by composers, performers, and audience members.

3. AGENTIAL SCORES

Taking into account the perspectives reviewed in Section 2.1, we propose an approach to musical notation that we call agential scores. Initially, we have been considering an agential score as one that responds to environmental inputs, possesses goals or directions, seeks equilibrium or development based on these objectives, and maintains a self-identical unity that endures over time. In our work involving artificial life as components of dynamic scores and musical instruments, we explore emergent behaviours that arise only when individual parts interact within a larger system. Although attributing goal-seeking characteristics may appear anthropomorphic, recent biological theories have challenged the notion of restricting such terms exclusively to humans [31]. In this spirit, we prefer to maintain openness regarding defining this space.

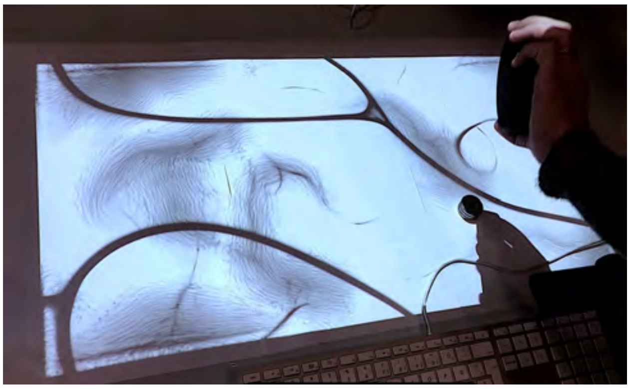
In this section we start from the bottom-up, by describing the agency of raw notational inscriptions, and then move to the top-down, by describing the entanglements [21] between scores, instruments, composers and performers. Then we turn to describing what emerges out of these entanglements in terms of assemblages [32] and intra-action [20].

3.1 Agency of Points and Lines

The historical significance of points and lines can be traced back to the earliest known human writing systems [33], and primitive geometries have been found to possess consistent interpretations across cultures and geographies [34]. Similarly, points and lines are also deemed to be prevalent elements in numerous forms of music notation throughout history [35].

In static, dynamic, and computational musical scores, points and lines exhibit distinct agencies, akin to their differences in paintings, animations, and real-time computer graphics. The painter Kandinsky ascribed concealed human attributes to points, characterising them as "the highest degree of restraint which, nevertheless, speaks." [36] Another artist, Paul Klee, illustrated in his pedagogical sketchbook, "An active line on a walk, moving freely, without goal. A walk for a walk's sake. The mobility agent is a point, shifting its position forward." [37] Although paintings remain static, our eyes move in relation to them, and both Kandinsky and Klee imbue their stationary media with a sense of motion. Conversely, animated points and lines exhibit actual movement. Pioneer Norman McLaren's concept of "pure" cinema emphasised "little or no reliance on factors other than motion", favouring the expressive potential of basic forms over cinematography [38].

Nonetheless, once drawn or rendered, the motion of points and lines in animation remains fixed in time. In this context, we can differentiate between a temporally fixed dynamic score and a procedural dynamic score, where the



(a) *FerroNeural* (2023) featuring gestural control of a Tölvera scene via handheld magnetic disc controllers [30]. Authors: Nicola Privato and Jack Armitage.

(b) *Motherbird* (2023) featuring interactive Tölvera murmuration simulation. Authors: Jessica Shand, Manuel Cherep and Jack Armitage. Photo: James Day, MIT Media Lab.

Figure 1: Photos from the development of the first two collaborative musical works that incorporate Tölvera.

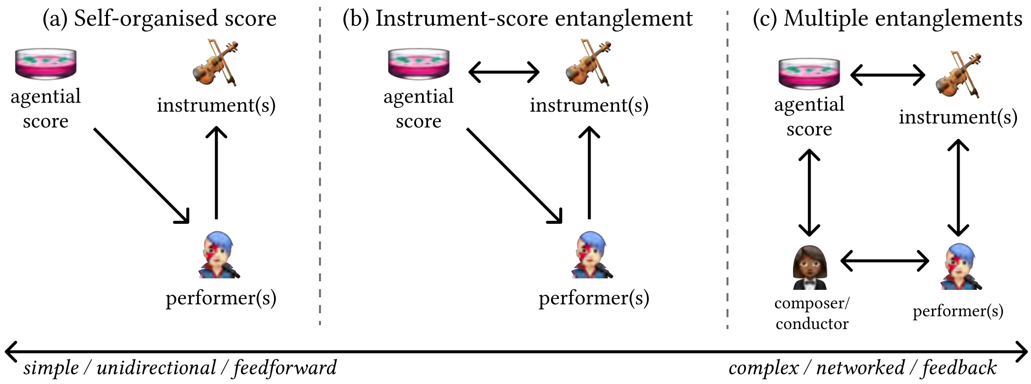


Figure 2: Three examples of entanglements [21] with agential scores, along a continuum of complexity. We investigate the phenomenology of self-organised scores (a, left) in Section 5. We are exploring instrument-score entanglements (b, middle) in the piece *FerroNeural* (Figure 1a), and multiple entanglements (c, right) in the piece *Motherbird* (Figure 1b).

latter is rendered in context, such as during live drawing or when utilising software. Specifically, with software, the preconception of motion is abstracted into the code. Self-organising swarms represent one artificial life approach for delegating notational dynamism in a manner that leads to the emergence of intriguing behaviour. Within this emergence, there exists the potential for points and lines with a fluid materiality, which reside "in the grey area of the continuum and have both the affordance of things and stuff" [39]. Self-organising systems advance the tradition of scores as points and lines, offering novel ways of experiencing fluid points and lines as notation.

3.2 A Typology of Entanglements with Agential Scores

Owing to the potential combinations of mappings between scores, instruments, and musicians, there is plenty to explore. Figure 2 depicts three examples of what we describe as entanglements with agential scores, following Frauenberger's description of entanglement human-computer interaction (HCI) [21]. In this way we suggest a typology of entanglements with agential scores based on identifying the network of relationships between "agents", as one way of decomposing the space of agential scores. There are many more possibilities that could potentially be described in this typological manner. We have found distinguishing types of entanglements can aid in the early stages of the agential score composition process, and when introducing agential scores to collaborators. There is also potential for such a typology to be developed further and used as an analysis tool, since it can also be used to describe performance ensembles and ecologies that do not necessarily make use of dynamic or computational media. We envision this might be useful as one of many "ad-hoc taxonomies" in a heterarchical organological interpretation [40].

3.3 Assemblages and Intra-action

We must also address the limitations of a purely typological approach to agential scores, which can only describe the elements of an entanglement and their relationships in simple terms, and not the emergent properties that arise from the interactions between them. Théberge proposes to consider musical instruments as assemblages, "components within a network" where the instrument maker's role is to "design relationships" rather than objects [32]. Barad

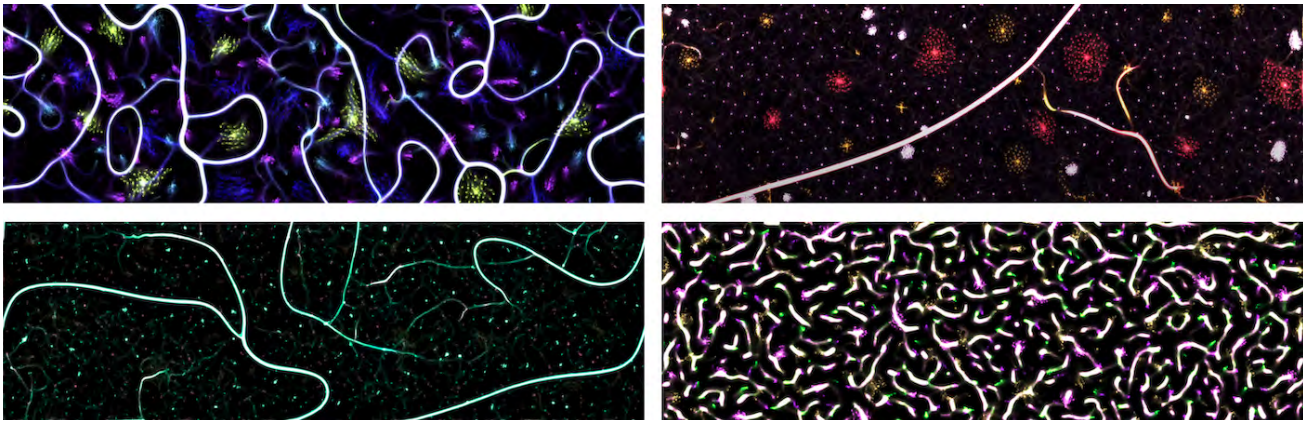


Figure 3: Examples of Tölvera combining Boids particles and Physarum pheromone trails, where the Boids particles also deposit pheromones to attract the Physarum. The variations arise from each species' behavioural parameters and the weighting of the mapping between them.

would describe working with musical scores from this perspective as a form of *intra-action*, "the mutual constitution of entangled agencies", which is contrasted with "the usual 'interaction', which assumes that there are separate individual agencies that precede their interaction". For Barad, "agencies are only distinct in relation to their mutual entanglement; they don't exist as individual elements". [20] Though emergent properties can be described in the reductive context of computer simulations, these are potentially much less determinable in the context of live musical performance. Detailed interpretations and reflections of these do exist in the context of highly entangled musical instruments [26, 29], but less appears to have been said about the role and agency of (particularly dynamic) scores in such entanglements. It is here that we perceive a research need that the framing of agential scores can potentially help to address, if it can be put into practice.

3.4 Agential Scores in Practice via Artificial Life

This work represents our first attempt to practically investigate agential scores, using artificial life, as many have done before although not with our specific aims and motivations in mind. Artificial life has been employed in both the sciences and the arts for various purposes. Christopher Langton, a founding advocate for the field, defines it as "life-as-it-could-be" [41]. Researchers within the field have elucidated and investigated how it can assist biologists and cognitive scientists in exploring the intricate issues of mind, agency, intelligence, and autopoiesis [18]. Although machine learning, particularly deep learning, has recently been the most publicised subdomain of artificial intelligence, there is now a growing interest in incorporating self-organising systems within these fields [42]. Additionally, machine learning is being utilised to train high-level goals in self-organising systems [43, 44]. In the arts, artificial life has been applied in the realm of computational creativity as systems capable of producing creative outcomes, such as in music or computer graphics [45]. In this context, artificial life is not perceived as a tool for solving problems but rather as an instrument that offers innovative, generative patterns and creative pathways [46].

4. TÖLVERA: A LIBRARY OF NUMBER BEINGS

Tölvera[2] is an open-source Python library for designing musical instruments and musical notations using artificial life[3]. Compared with existing tools like *NetLogo* and *Golly*, Tölvera is designed to use a modern and portable implementation, be more accessible and extensible to creative coders, and to be more easily integrated into musical instruments and notation systems. The name is an example of a *kenning*, a metaphorical compound expression found in Old Norse and Old English poetry. Tölvera combines the Icelandic words for computer (*tölvu*, derived from tala - number - and völva - prophetess or oracle) and being (*vera*). Our lab is situated in Iceland, and we have discovered that immersing ourselves in the local culture enhances our work and encourages public participation and collaboration [14]. We are also intentionally engaging with the agency of the Icelandic language itself, juxtaposing artificial life and musical scores with Icelandic folklore and mythology. Combining *number* with *being* also alludes to the discourse surrounding embodiment in artificial life, as discussed in Section 2. Listing 1 shows a simple Tölvera program that renders particles to a window.

```python
import taichi as ti
import tolvera as tol
def main(x=1920, y=1080, n=512, species=3):
    ti.init()
    particles = tol.Particles(x,y,n,species)
    pixels = tol.Pixels(x,y)
    def render():
        pixels.clear()
        particles(pixels)
    pixels.show(render)
```

Listing 1: Example of a Tölvera program that renders particles to a window.

[2] https://github.com/Intelligent-Instruments-Lab/iil-python-tools/tree/master/tolvera

[3] Please see the Tölvera YouTube playlist for demos and work-in-progress experiments https://www.youtube.com/playlist?list=PL8bdQleKUA1vNez5gw-pfQB21Q1-vHn3x.

4.1 Number Beings

Tölvera consists of a small number of simulations (vera, or beings) that are variously inspired by physics, biology, computation, or some mixture of these. Familiar examples include *Boids*, the classic flocking algorithm by Craig Reynolds. [47], and *Physarum*, a slime mould simulation popularised by artists such as Sage Jensen [4]. [48]. We are experimenting with adding many more vera such as *Lenia*, an artificial life continuous cellular automata (CCA) system discovered by Bert Chan, as a smooth generalisation of Conway's Game of Life. [49]. However, so far we have focused on the design of the library, so that it can be easily used by creative coders, and extended by other developers. Listing 2 shows a Tölvera program that renders multiple species of particles to a window.

```
def render():
    pixels.diffuse()
    boids(particles)
    particles(pixels)
    physarum(particles)
    pixels.set(physarum.trail.px)
```

Listing 2: Example of a Tölvera render function that creates compound motion of particles by combining boids and physarum algorithms. In future this will be syntactically simplified via method chaining to e.g. `particles.boids(args).physarum(args)`.

Any real-time Taichi program can potentially be added or contributed to the Tölvera library, and though we do not prescribe a coding style, so far we follow Taichi's objective data-oriented programming paradigm by implementing each species or simulation as a data-oriented class. [5] As in other real-time creative computing contexts, each class aims to provide simple "process" functions that step the simulation forward and update behavioural parameters.

4.2 Mappings and Visualisations

One of the expressed aims of Tölvera is to commingle interfaces and instruments with artificial life. In this way, we envision that instruments provide input and output data streams to Tölvera, becoming its connection to physical reality. At the same time, Tölvera becomes integrated into the instrument for its performer. Listing 3 shows how to send and receive OSC messages from a Tölvera program.

Tölvera can by default run as a window on a regular computer screen. However, we have observed that projecting Tölvera onto horizontal surfaces (either from above or below) encourages intimate engagements, through tactile interventions such as hand-drawing notation, and improvised use of objects as notation (see Figures 1 and 4). Further, Tölvera could in future integrate computer vision to recognise these as inputs to the simulations, resulting in a commingling of physical and computational processes.

[4] https://cargocollective.com/sagejenson/physarum

[5] https://docs.taichi-lang.org/docs/odop

```
p = tol.Particles(x,y,n,species)
# ...
osc_update = OSCUpdaters(osc, client="simple",
    receives={
        "/particles/set/pos": p.osc_set_pos,
        "/particles/set/vel": p.osc_set_vel
    }, receive_count=10,
    sends={
        "/particles/get/pos/all": p.osc_get_pos_all
    }, send_count=60
)
# ...
def render():
    osc_update()
    pixels.clear()
    particles(pixels)
```

Listing 3: Example of OSC messages being sent and received by the OSCUpdaters class in Tölvera. Counters are used to rate-limit how frequently received messages are processed and how frequently messages are sent. This allows OSC messages to be sent and received at a rate that is independent of the frame rate.

4.3 Implementation

Tölvera programs can be run in real-time on the CPU or GPU of a modern laptop, [6] and combined with other Python libraries enabling additional features such as machine learning. Tölvera is implemented in Taichi [7], an imperative, parallel, domain specific language embedded in Python, for high-performance numerical computation. Taichi uses just-in-time (JIT) compiler frameworks such as LLVM, to offload compute-intensive code to native GPU or CPU instructions. We chose Taichi because it enables optimised graphics programming through a familiar high-level language, and because, being embedded in Python, it can interoperate with the rich tapestry of the machine learning ecosystem. Taichi programs can also be run on the GPU without any visual rendering, allowing the programs to be run "headless" in more resource-constrained contexts. They can be compiled to C/C++, and then to JavaScript via Emscripten, and also run in Unity, [8] enabling a wide variety of target contexts to be explored in future (see Section 6.4).

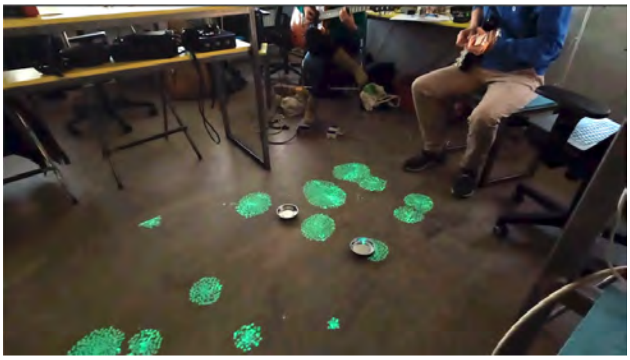
5. MUSICAL ENCOUNTERS WITH TÖLVERA

Tölvera is designed for instrumental and notational design, and we're researching how musicians and audiences interact with the concepts of intelligence, emergence, and agency. We've been conducting informal musical sessions, called *encounters*, with our research prototypes to inform design iterations and inspire future encounters. We present an ethnographic account of these encounters in this section, and offer an interpretation of the emerging themes in Section 6. Two professional guitarists participated in an

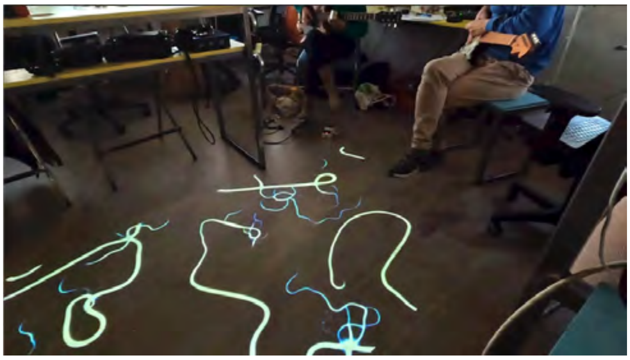
[6] Modern laptop with a GPU supporting backends such as OpenGL, CUDA, Vulkan and Metal: https://docs.taichi-lang.org/docs/hello_world.

[7] https://github.com/taichi-dev/taichi

[8] https://github.com/taichi-dev/Taichi-UnityExample



(a) Boids clustered into spheres, with players 'identifying' as metal bowls on the floor.



(b) Players improvising in response to Physarum pheromone trails.

(c) Boids and Physarum combination, with Player A 'conducting' Player B's attention.

(d) Boids and Physarum combination, with Player B 'conducting' Player A's attention.

Figure 4: Video stills of four setups from the Tölvera improvisation session. Tölvera scenes were projected from above onto the floor. The two guitarists sat in chairs nearby.

informal session with Tölvera in an arts university studio lab space. The primary aim was to determine the extent to which basic configurations of Boids and Physarum particle systems could maintain engagement in an informal improvisation setting. The guitarists played their own instruments, and the Tölvera designer edited and randomised the Tölvera scenes in response to the musicians. The encounters lasted between three and seven minutes each.

5.1 Encounters Summaries

5.1.1 Encounter 1: Boids & Two Guitars

Figure 4a shows a video still from the first encounter between two guitarists and Boids. The Boids were divided into two species, referred to as Species 1 and Species 2. Player A focused on Species 1, which formed lengthy, sinuous streams, while Player B engaged with Species 2, which developed into small, swiftly agitated clusters. Player A alternated between playing phrases akin to those from earlier when Species 1 crossed their bowl and maintaining a steady rhythm of atonal chords for Species 2. Similarly, Player B played as before when Species 2 traversed their bowl and executed loud, heavy chord riffs when Species 1 emerged. The Boids species colours and parameters were randomised several times to generate variations. As more variations emerged, Player B asked, "how do I play this?" The music grew increasingly disjointed and chaotic as they attempted to adapt to new scenarios. The music ultimately ceased following a final randomisation, with both musicians pausing before Player B stated plainly, "I don't know," and they both shared a laugh.

5.1.2 Encounter 2: Physarum & Two Guitars

In the second encounter (Figure 4b), the players responded to the Physarum simulation this time. Player A played similar spacious slide guitar notes as before, while Player B utilised a pen to produce slide guitar sounds as well. Their sound was considerably quieter than in Encounter 1, possibly due to the increased amount of empty space in the scene and the reduced motion of the Physarum. Their flowing, sliding gestures appeared to correspond with the twisting, continuous pheromone trails. After a relatively calm encounter, the players concurred at the end that "this was fun".

5.1.3 Encounter 3: Boids, Physarum, Guitar & Conductor

Encounter 3 (Figure 4c) involved Boids and Physarum combined in a Tölvera scene. The Boids left an additional pheromone trail for the Physarum to follow, resulting in some interaction between them. Player A acted as a conductor by pointing a baton at the projection on the floor to direct Player B, who played guitar. The encounter began with randomisation of Tölvera parameters to find a preferred setting. Player A tracked a Physarum ring of pheromone trail that was slowly closing in on itself while Player B gradually increased their tempo, playing two semitones. As the ring exploded back out into the space, Player

A swept around with their baton before settling on tracing another more stable pheromone trail. Player B adapted their playing accordingly. Player A then followed various Boids clusters from different species, some of which were calmly floating and others that were more animated. Player B responded to each Boid cluster differently, adapting their previous musical ideas. The conducting and playing matched the progression of the Tölvera scene, becoming more dynamic than the previous encounters.

5.1.4 Encounter 4: Reversing Roles from Encounter 3

In the final encounter (Figure 4d), Player A and B switched roles from Encounter 3, with Player B becoming the 'conductor' and Player A returning to playing the guitar. They began by randomising Tölvera parameters to find an appealing scenario, with comments such as "Mmm hmm, this I like!", "It's good.", and "It's beautiful, no?". Player B started by tapping specific points and lifting their baton from the floor, causing Player A to inquire about when to play silence. Both players laughed and shrugged in response. Player B followed the trails of Physarum and collapsing rings, responding with spacious slide guitar phrases that mapped both ends of a trail to low and high pitches. During the performance, Player A accidentally knocked the projector mounting, causing both players to react with surprised laughter as the entire scene shook. Player B then traced a specific Boids cluster around the scene, which suddenly changed direction and darted off the edge of the screen. Both players laughed as Player B exclaimed and stepped back, saying, "I had started to sympathise [with that cluster]... Ah, it comes back this way, right?" upon realizing that the particles wrap around the edges of the scene.

5.2 Post-Encounters Discussion

In a discussion, the researcher and author asked open-ended questions about encounters, leading to a conversation about potential Tölvera development. The players noticed the Boids clusters would occasionally suggest a musical representation, but were sometimes random. Player A talked about their film scoring practices and their awareness of "Mickey Mousing." They found this approach safe and comforting because it allowed for an easy association between the particles and musical gestures. However, they mentioned feeling lost when tracked clusters vanished and would sometimes intentionally try to lose their identity with a cluster. Player A also described experiencing various emotional states, such as feeling safe, lost, bored, and comforted. Player B commented that musicians are not predictable but follow certain expectations, while artificial life is less predictable.

The players compared their relationships to musical scores in classical and jazz contexts, respectively. Player A believed they adhered more strictly to rules in response to classical scores, while Player B felt jazz players had more freedom in interpretation. They both acknowledged that their experiences influenced their reactions to Tölvera scenes and appreciated the contrast in their musical responses. They discussed the limitations of guitar as an instrument, such as difficult fingerings and challenges with microtonality. Player A felt constrained by their instrument and thought a symphony orchestra was necessary to express diverse activity. Having two guitarists alleviated some pressure, but did not influence their reactions to specific encounters. Both players enjoyed playing and conducting, with the latter providing more freedom and the ability to anticipate the conductor's intentions. This suggests a shared attraction to similar types of activity and interpretations.

During a conduct by Player B, Player A wondered if a more formal, rule-based approach was anticipated, akin to their classical training. Player A suggested dynamically morphing or navigating between parameter states for future Tölvera development, instead of instantaneously randomizing all parameters. They often interpreted the distance between objects as pitch intervals and discussed the possibility of a visual demarcation of pitch space through a stave-like overlay. This stave could be microtonal or continuously graded, but it was uncertain whether it should indicate temporality. Player B proposed dividing the space into a grid of cells, each with different scoring instructions or mappings, and assigning players to specific cells. They also discussed explicitly representing players as particles or clusters and providing control over their range of motion and responses through mapping.

6. DISCUSSION

Despite the nascent state of Tölvera, we discovered that simple emergent patterns of Boids and Physarum particles facilitated diverse musical interpretations, evoked visceral emotional responses, and stimulated reflective discussions about the agency of musical scores and their influence on human and non-human interpreters. As discussed in Section 2, we are interested in agency from multiple perspectives, viewing our research systems as boundary objects [28] that can ignite pluralistic dialogue. In that spirit, we explore various themes of agency throughout the encounters to enrich our overall theme of agential scores.

6.1 Fluid Material Agency

The emergent patterns of the Boids and Physarum particles accentuate points and lines in the Tölvera scores. In accordance with the theories of material agency discussed earlier in Section 2, we suggest that real-time software-based dynamic scores alter the agency of points and lines, enabling them to convey the fluidity of these primitive forms through procedurally adaptive motion. This fluidity, experienced as emergent and self-organising music notation, subsequently restructures the musician's perceptive eye and interpretive mind, leading to a more spontaneous and dialogic negotiation of attention.

In the Tölvera scores, Boids are represented as points and Physarum as curved lines by default. As an open notational score, these fundamental forms raise questions about their musical interpretation. The guitarists improvised responses, with point sizes corresponding to volumes and line lengths to pitch ranges, among other mappings. However, certain species of Boids seemed to exhibit fluid-like behaviour, while Physarum trails could adopt more discrete appearances, such as rings, points, and grids. Individual Boids could merge into a cluster, which could then form larger clusters, containing internal oscillations of individual members. Similarly, Physarum rings could converge inwards to a single point before expanding outwards into tree-like structures or dissipating entirely. These morphological transformations between forms increased the demands on musicians' interpretation, requiring them to seamlessly combine rapid chains of diverse playing techniques and musical phrases, while adjusting their mappings in real-time.

The fluid materiality provided by artificial life consequently alters our perception of the agency of a point or line, acknowledging their inherent fluid potential. Barad contends that physical nature itself exhibits greater fluidity and queerness than ontological discreteness and causality, urging us to adopt a perspective of fluidity when observing the world [50]. How does our sense of self evolve when we associate with such fluid material in a symbolic musical relationship?

6.2 Mapping of Self Onto Agential Materials

The typology we described in Section 3.2 led us to describe this encounter as an entanglement between an agential score, two performers, and their instruments, with a simple unidirectional flow between them (Figure 2). However, we observed another kind of entanglement that was not explicitly described in the typology. It became apparent that the life-mimicking behaviour of the Tölvera scores prompted musicians to associate with specific aspects of the scores, even eliciting a subtle sympathetic response in Encounter 4. This projection of oneself onto the dynamic behaviour of the score was incidental and transient, as the life forms within the score continuously evolved, formed new configurations, and subsequently disintegrated.

Extrapolating from this, we envisage that only a slight perception of agency in a score is necessary to trigger human tendencies to identify, empathise, mirror, and attune. This process establishes an entanglement or mapping at a higher level than between notations and gestures and sounds, connecting human selves with agential materials. When unexpected behaviour from an agential material disrupts this mapping, visceral responses can be elicited ranging from surprise and excitement to disgust. Attunement to an agential score and its subversion presents a novel compositional device for musicians to explore and researchers to examine. In future work we seek to elaborate on these psychological and emotional aspects of experience with agential materials through an investigation into *agentology* [51].

6.3 Perceiving the Intra-Actants

Following Barad's agential realism and the concept of intra-action [20], we propose that each of our encounters generated a transient *intra-actant*, emerging from the assemblage of musicians, media, and materials, and possessing its own distinct agency. The qualities of each intra-actant were inextricably connected to each Tölvera score and the moment it took place in, and resist typological decomposition. What language and methods are appropriate for describing and comparing intra-actants, that evade the trap of a decompositional approach? Perhaps a new approach is needed that builds on astute methods like micro-phenomenonology [52, 53] to compare multiple experiences of the same moment, leading to a macro-phenomenonological account.

For inspiration, composers encountering agential scores for the first time might find it advantageous to examine the experiences of feedback instrumentalists [54]. These musicians have dedicated significant time to initially unruly complex musical systems and have developed aesthetics that are compatible with extended and augmented instrumental agencies. Agential instruments also present a natural pairing for agential scores.

6.4 Future Considerations

In this paper, we established a context for discourse on agential scores and recount early experimental artistic applications. We intentionally avoided providing a rigid definition or framework to allow for alternative perspectives and interpretations. Here we outline what will be our focus as we seek to develop the themes of this work further.

Our aim is to develop the Tölvera library through artistic and empirical encounters, generating insights surrounding agential scores and their design. The Tölvera library will expand in multiple directions, including incorporating more number beings, additional inputs, and a broader range of outputs. We also anticipate a vertical development towards goal-oriented and trainable behaviours and instrument mappings through integrating online machine learning processes.

We have yet to explore the roles of the composer and audience within the context of agential scores, and we envisage composers will eventually engage with agential scores through more accessible interfaces such as graphic design, drawing, and tangible or gestural interactions. In this scenario, a composer could create artificial life patterns and parameters using an assortment of familiar media, and subsequently instruct conductors and performers on interpreting these elements as notation (or even conduct and perform themselves).

With regard to audiences, their behaviour in large groups has been extensively modelled in the field of crowd simulation, providing a compelling entry point for integrating them with agential scores. Additionally, there is a growing research interest in enabling audience participation in music performances through networked, mobile, and gestural interfaces, which is highly relevant to this subject [55]. We consider Tölvera as an element in our Organolib system (see earlier footnote), and we anticipate that our community will discover increasingly diverse applications for it in combination with other elements.

As the system matures, we envisage designing encounters where more detailed accounts of intra-action can be captured, not only during performance but also during design and composition [56]. At the micro scale [57], this could involve comparing eye or gesture tracking data with particle data, and analysing their convergence and divergence across various score states. Based on our other studies, we are also interested to know how perception of symmetry and algorithmic pattern [58] affects projection or attribution of agency.

And qualitatively, micro-phenomenology can be used to capture the experience of the musicians and audience members. Such encounters would enable us to further develop linguistic and conceptual tools for the analysis and interpretation of the macro scale [59] agency of intra-actants.

7. CONCLUSION

In this paper, we introduced the concept of agential scores. An agential score is characterised by a self-identical unity that endures over time, responding to environmental inputs, possessing goals or directions, and striving for equilibrium or development in line with these objectives. We explored this idea by examining how points and lines gain fluid material agency through animation via dynamic computational materials and simulations. We compared a typological and decompositional approach to agential scores with a more holistic, macro-phenomenonological approach that situates agential scores as entangled with their material and social contexts. This broader context we described as an intra-actant, inspired by Barad's idea of intra-action, to emphasise the study and comparison of these high-level emergent entities.

We introduced Tölvera, an open-source software library written in Python and inspired by Icelandic kennings and mythology, for exploring agential scores in real-time through artificial life simulations. An initial musical encounter with Tölvera involving two improvising guitarists was conducted to ground the development process in empirical observation. Even at an early stage, Tölvera demonstrated that simple emergent patterns of Boids and Physarum particles facilitated diverse musical interpretations, elicited emotional reactions, and sparked thoughtful conversations about the agency of musical scores and their human and non-human interpreters.

Reflecting on our observations, we proposed that agential scores foreground a particular compositional mapping layer that mediates between musical selves and agential materials. We perceived this as having implications for ecological music perspectives. Lastly, we posed several challenging issues and questions that we anticipate will require addressing as more advanced agential scoring approaches emerge.

Acknowledgments

Thanks to Nicola Privato and Hafdís Bjarnadóttir for their participation in the encounter. Thanks to *Motherbird* collaborators Jessica Shand and Manuel Cherep. Thanks to Victor Shepardson for software development support, and Marco Donnarumma for testing and feedback. Thanks to the anonymous reviewers for their helpful comments.

This research is supported by the European Research Council (ERC) as part of the Intelligent Instruments project (INTENT), under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 101001848).

8. REFERENCES

[1] J. Lely and J. Saunders, *Word Events: Perspectives on Verbal Notation*. Continuum New York, NY, 2012.

[2] L. Kotz, *Words to Be Looked at: Language in 1960s Art*. Mit Press Cambridge, MA, 2007.

[3] G. Cox and A. McLean, *Speaking Code: Coding as Aesthetic and Political Expression*. MIT Press, 2012.

[4] Y. Ono, *Grapefruit: A Book of Instructions and Drawings by Yoko Ono*. Tokyo: Wunternaum Press, 1964.

[5] J. Dack, "The 'open' form – literature and music," in *Scambi Symposium', Goldsmiths College*, vol. 334, 2005, p. 336.

[6] W. Thompson, *Soundpainting: The Art of Live Composition. Workbook I.* Walter Thompson, 2006.

[7] A. Sonnenfeld and K. F. Hansen, "S-notation: A complete musical notation system for scratching and sample music derived from" Theory of Motions"," in *International Conference on Technologies for Music Notation and Representation*. Anglia Ruskin University, 2016, pp. 50–57.

[8] C. Hope, "Electronic scores for music: The possibilities of animated notation," *Computer Music Journal*, vol. 41, no. 3, pp. 21–35, 2017.

[9] E. Tomás and M. Kaltenbrunner, "Tangible Scores: Shaping the Inherent Instrument Score," in *Proceedings of the International Conference on New Interfaces for Musical Expression*. London, United Kingdom: Goldsmiths, University of London, Jun. 2014, pp. 609–614.

[10] S. Bhagwati, M. Giordano, J. Berzowska, A. Bachmayr, J. Stein, J. Browne, F. D. Tredici, D. Egloff, J. Sullivan, M. Wanderley, and I. Cossette, "Musicking the Body Electric. The "body:suit:score" as a polyvalent score interface for situational scores." in *Proceedings of the International Conference on Technologies for Music Notation and Representation - TENOR2016*, R. Hoadley, C. Nash, and D. Fober, Eds. Cambridge, UK: Anglia Ruskin University, 2016, pp. 121–126.

[11] W. V. Rodrigues and W. de Macedo Rodrigues, "Textural Composition in 3D Environment Through Swarm Algorithm," in *Proceedings of the International Conference on Technologies for Music Notation and Representation - TENOR2022*, 2022.

[12] A. Kay and A. Goldberg, "Personal dynamic media," *Computer*, vol. 10, no. 3, pp. 31–41, 1977.

[13] V. Shepardson, J. Armitage, and T. Magnusson, "Notochord: A Flexible Probabilistic Model for Embodied MIDI Performance," in *Proceedings of the 3rd Conference on AI Music Creativity*. AIMC, Sep. 2022.

[14] J. Armitage, T. Magnusson, V. Shepardson, and H. Ulfarsson, “The Proto-Langspil: Launching an Icelandic NIME Research Lab with the Help of a Marginalised Instrument,” in *Proc. New Interfaces for Musical Expression*.

[15] D. Borgo, *Sync or Swarm: Improvising Music in a Complex Age*. A&C Black, 2005.

[16] S. Waters, “Performance Ecosystems: Ecological approaches to musical interaction,” *EMS: Electroacoustic Music Studies Network*, pp. 1–20, 2007.

[17] G. E. M. Anscombe, *Intention*. Harvard University Press, Oct. 2000.

[18] H. R. Maturana and F. J. Varela, *Autopoiesis and Cognition: The Realization of the Living*. Springer Science & Business Media, 1991, vol. 42.

[19] B. Latour, *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oup Oxford, 2007.

[20] K. Barad, *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning*. Duke University Press, 2007.

[21] C. Frauenberger, “Entanglement HCI the next wave?” *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 27, no. 1, pp. 1–27, 2019.

[22] C. Nordmoen and A. P. McPherson, “Making space for material entanglements: A diffractive analysis of woodwork and the practice of making an interactive system,” in *Designing Interactive Systems Conference*, ser. DIS ’22. New York, NY, USA: Association for Computing Machinery, Jun. 2022, pp. 415–423.

[23] N. Virgo, “The necessity of extended autopoiesis,” *Adaptive Behavior*, vol. 28, no. 1, pp. 23–26, Feb. 2020.

[24] M. Villalobos and P. Razeto-Barry, “Are living beings extended autopoietic systems? An embodied reply,” *Adaptive Behavior*, vol. 28, no. 1, pp. 3–13, Feb. 2020.

[25] S. Kember, *Cyberfeminism and Artificial Life*. Routledge, 2003.

[26] A. P. Melbye, “Resistance, Mastery, Agency: Improvising with the feedback-actuated augmented bass,” *Organised Sound*, vol. 26, no. 1, pp. 19–30, 2021.

[27] B. Latour, “The Berlin key or how to do words with things,” *Matter, materiality and modern culture*, vol. 10, p. 21, 2000.

[28] S. L. Star and J. R. Griesemer, “Institutional ecology, ’translations’ and boundary objects: Amateurs and professionals in Berkeley’s Museum of Vertebrate Zoology, 1907-39,” *Social studies of science*, vol. 19, no. 3, pp. 387–420, 1989.

[29] P. Stapleton and T. Davis, “Ambiguous Devices: Improvisation, agency, touch and feedthrough in distributed music performance,” *Organised Sound*, vol. 26, no. 1, pp. 52–64, Apr. 2021.

[30] N. Privato, E. Torfi, and T. Magnusson, “The Magnetic Score: Somatosensory Inscriptions and relational design in the instrument-score,” in *Proceedings of the 8th International Conference on Technologies for Music Notation and Representation*.

[31] M. Levin, “Technological Approach to Mind Everywhere: An Experimentally-Grounded Framework for Understanding Diverse Bodies and Minds,” *Frontiers in Systems Neuroscience*, vol. 16, 2022.

[32] P. Théberge, “Musical Instruments as Assemblage,” in *Musical Instruments in the 21st Century*, T. Bovermann, p. u. family=Campo, given=Alberto, H. Egermann, S.-I. Hardjowirogo, and S. Weinzierl, Eds. Springer Singapore, pp. 59–66. [Online]. Available: http://link.springer.com/chapter/10.1007/978-981-10-2951-6_5

[33] B. Bacon, A. Khatiri, J. Palmer, T. Freeth, P. Pettitt, and R. Kentridge, “An Upper Palaeolithic Proto-writing System and Phenological Calendar,” *Cambridge Archaeological Journal*, pp. 1–19, Jan. 2023.

[34] A. Ćwiek, S. Fuchs, C. Draxler, E. L. Asu, D. Dediu, K. Hiovain, S. Kawahara, S. Koutalidis, M. Krifka, and P. Lippus, “The bouba/kiki effect is robust across cultures and writing systems,” *Philosophical Transactions of the Royal Society B*, vol. 377, no. 1841, p. 20200390, 2022.

[35] T. Ingold, *Lines: A Brief History*. Routledge, 2016.

[36] W. Kandinsky and H. Rebay, *Point and Line to Plane*. Courier Corporation, 1979.

[37] P. Klee, *Pedagogical Sketchbook*. Introduction and translation by Sibyl Moholy-Nagy. Praeger: New York, 1953.

[38] D. McWilliams and S. Huycke, “Creative Process: Norman McLaren,” 1991.

[39] M. Le Goc, L. H. Kim, A. Parsaei, J.-D. Fekete, P. Dragicevic, and S. Follmer, “Zooids: Building Blocks for Swarm User Interfaces,” in *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, ser. UIST ’16. New York, NY, USA: ACM, 2016, pp. 97–109.

[40] T. Magnusson, “Musical Organics: A Heterarchical Approach to Digital Organology,” vol. 46, no. 3, pp. 286–303. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/09298215.2017.1353636

[41] C. G. Langton, “Studying artificial life with cellular automata,” *Physica D: Nonlinear Phenomena*, vol. 22, no. 1-3, pp. 120–149, 1986.

[42] D. Ha and Y. Tang, “Collective intelligence for deep learning: A survey of recent developments,” *Collective Intelligence*, vol. 1, no. 1, p. 26339137221114874, Aug. 2022.

[43] S. Sudhakaran, E. Najarro, and S. Risi, “Goal-Guided Neural Cellular Automata: Learning to Control Self-Organising Systems,” Apr. 2022.

[44] G. Hamon, M. Etcheverry, B. W.-C. Chan, C. Moulin-Frier, and P.-Y. Oudeyer, “Learning Sensorimotor Agency in Cellular Automata,” Jan. 2022.

[45] M. Whitelaw, *Metacreation: Art and Artificial Life*. Mit Press, 2004.

[46] M. Sato and J. McKinney, “The Enactive and Interactive Dimensions of AI: Ingenuity and Imagination Through the Lens of Art and Music,” *Artificial Life*, vol. 28, no. 3, pp. 310–321, Aug. 2022.

[47] C. W. Reynolds, “Flocks, herds and schools: A distributed behavioral model,” in *Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques*, 1987, pp. 25–34.

[48] J. Jones, “Characteristics of Pattern Formation and Evolution in Approximations of Physarum Transport Networks,” *Artificial Life*, vol. 16, no. 2, pp. 127–153, Apr. 2010.

[49] B. W.-C. Chan, “Lenia and Expanded Universe,” in *The 2020 Conference on Artificial Life*, 2020, pp. 221–229.

[50] K. Barad, “Nature’s queer performativity,” *Qui Parle: Critical Humanities and Social Sciences*, vol. 19, no. 2, pp. 121–158, 2011.

[51] J. Armitage, “Agentology for Organology: Cyber-, Bio-, and Ecosemiotic Perspectives on Instrumental Agency,” in *Instruments, Interfaces, Infrastructures: An Interdisciplinary Conference on Musical Media*. [Online]. Available: https://sites.harvard.edu/instruments-interfaces-infrastructures/

[52] C. N. Reed, C. Nordmoen, A. Martelloni, G. Lepri, N. Robson, E. Zayas-Garin, K. Cotton, L. Mice, and A. McPherson, “Exploring Experiences with New Musical Instruments through Micro-phenomenology,” in *NIME 2022*, jun 16 2022, https://nime.pubpub.org/pub/o7sza3sq.

[53] M. Prpa, S. Fdili-Alaoui, T. Schiphorst, and P. Pasquier, “Articulating Experience: Reflections from Experts Applying Micro-Phenomenology to Design Research in HCI,” in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI ’20. Association for Computing Machinery, pp. 1–14. [Online]. Available: https://doi.org/10.1145/3313831.3376664

[54] T. Magnusson, C. Kiefer, and H. Ulfarrson, “Reflexions upon Feedback,” in *NIME 2022*, Jun. 2022.

[55] A. Xambó and G. Roma, “Performing audiences: Composition strategies for network music using mobile phones,” in *Proceedings of the International Conference on New Interfaces for Musical Expression*, R. Michon and F. Schroeder, Eds. Birmingham, UK: Birmingham City University, 2020, pp. 55–60.

[56] J. Armitage, T. Magnusson, and A. McPherson, “Design Process in Visual Programming: Methods for Visual and Temporal Analysis,” in *Proc. Sound and Music Computing Conference*.

[57] ——, “Studying Subtle and Detailed Digital Lutherie: Motivational Contexts and Technical Needs,” in *Proc. New Interfaces for Musical Expression*.

[58] ——, “Sculpting Algorithmic Pattern: Informal and Visuospatial Interaction in Musical Instrument Design,” in *Proc. Sound and Music Computing Conference*.

[59] ——, “A Scale-Based Ontology of Digital Musical Instrument Design,” in *Proc. New Interfaces for Musical Expression*.

ENUMERATING LEFT HAND FORMS FOR GUITAR TABLATURES USING NON-DECREASING FINGER NUMBERS AND SEPARATORS

Gen Hori
Department of Data Science, Faculty of Business Administration
Asia University, Tokyo 180-8629, Japan
hori@asia-u.ac.jp

ABSTRACT

We introduce a guitar fingering decision method based on HMM and a tree diagram which we call "note-tablature-form tree" that can handle fingering decision of polyphonic pieces. To construct note-tablature-form tree for a given polyphonic chord, we need to (i) enumerate tablatures for a given chord, and (ii) enumerate left hand forms for each tablature. For the former, we introduce an enumeration method of tablatures for a given chord using a concept of permutations. For the latter, we introduce a new idea for exhaustive enumeration of left hand forms for a given tablature based on non-decreasing finger numbers and two kinds of separators that assign fingers to string-fret pairs.

1. INTRODUCTION

The process of determining optimal guitar fingerings for a given musical passage is a complex and subjective task that has long challenged guitarists of all levels. In recent years, researchers have explored the use of computational models to aid in this process. (See Sayegh[1], Radicioni et al.[2], Radisavljevic and Driessen[3], Tuohy and Potter[4], for example.) One such approach is the use of hidden Markov model (HMM), a statistical modeling technique that can capture the underlying patterns and structures in time series. As for applications of HMM to fingering decision, Hori et al.[5] applied input-output HMM to guitar fingering decision and arrangement, Nagata et al.[6] applied HMM to violin fingering decision, and Nakamura et al.[7] applied merged-output HMM to piano fingering decision. Hori and Sagayama[8] and Hori[9] proposed extensions of the Viterbi algorithm for fingering decision.

The purpose of the present study is to extend guitar fingering decision method based on HMM[5] and a tree diagram[10] for monophonic cases to one that can handle polyphonic cases. We cast guitar fingering decision as a decoding problem of HMM whose output symbols are musical notes and hidden states are left hand forms. In this case, we need to enumerate left hand forms for each chord in a given piece to perform fingering decision for polyphonic pieces. To do that, (i) we enumerate tablatures for a chord, and then (ii) enumerate left hand forms for each tablature.

For the former, we introduce an enumeration method of tablatures for a given chord using a concept of permutations in Section 3.1. For the latter, we introduce a new idea for enumeration of left hand forms for a given tablature using non-decreasing finger numbers and separators in Section 3.2, which provides a new insight for exhaustive search for all the possible left hand forms for a given tablature.

The rest of the paper is organized as follows. Section 2.1 reviews guitar fingering decision method based on HMM[5] and Section 2.2 note-tablature-form tree for enumeration of left hand forms using monophonic cases[10]. Section 3 extends note-tablature-form tree for polyphonic cases where Section 3.1 introduces an enumeration method of tablatures for a given chord and Section 3.2 our new idea for exhaustive enumeration of left hand forms for a given tablature. Section 4 concludes the paper.

2. NOTE-TABLATURE-FORM TREE FOR MONOPHONIC CASES

This section reviews the guitar fingering decision method based on HMM whose output symbols are musical notes and hidden states are left hand forms[5] and note-tablature-form tree for enumeration of player's left hand forms[10]. Although we limit our attention to a monophonic case to simplify the explanation in this section, the results extend to polyphonic cases. See [5] for HMM for polyphonic cases. Note-tablature-form tree is extended for polyphonic cases in the following section.

2.1 Fingering decision based on HMM

To play a single note on a guitar, a guitarist holds down a string-fret pair,

$$p_i = (s_i, f_i),$$

with a finger h_i of the left hand and picks the string s_i with the right hand, where $s_i = 1, \ldots, 6$ is a string number (from the highest to the lowest), $f_i = 0, 1, \ldots$ is a fret number where $f_i = 0$ means an open string, and $h_i = 1, 2, 3, 4$ is a finger number where 1,2,3 and 4 mean the index, middle, ring and pinky fingers, respectively. Therefore, a left hand form q_i for playing a single note can be expressed in a triplet q_i,

$$q_i = (s_i, f_i, h_i).$$

The MIDI note number of the note played by the form q_i is calculated as follows where o_{s_i} denotes the MIDI note number of the open string s_i,

$$n(q_i) = o_{s_i} + f_i.$$

We cast fingering decision as a decoding problem of HMM whose output symbols are musical notes and hidden states are left hand forms, where a fingering is obtained as a sequence of hidden states given a monophonic phrase as a sequence of output symbols.

The difficulty levels of the moves from forms to forms are implemented in the probabilities of the transitions from hidden states to hidden states; a small value of the transition probability means the corresponding move is difficult and a large value means easy. We assume that the four fingers of the left hand are always put on consecutive frets in this section for simplicity. This lets us calculate the *index finger position* (the fret number the index finger is put on) of form q_i as $g(q_i) = f_i - h_i + 1$. Using the index finger position, we set the transition probability from hidden state q_i to hidden state q_j as

$$a_{ij}(d_t) \propto \frac{1}{2d_t} \exp\left(-\frac{|g(q_i) - g(q_j)|}{d_t}\right) \times P_H(h_j) \quad (1)$$

where $\propto$ means proportional and the left hand side is normalized so that the summation with respect to j equals 1 for all i. The first term of the right hand side is taken from the probability density function of the Laplace distribution that concentrates on the center and its variance d_t is set to the time interval between the onsets of the $(t-1)$-th note and the t-th note. The second term $P_H(h_j)$ corresponds to the difficulty level of the destination form q_j defined by the finger number h_j.

As for the output probability, because all the hidden states have unique output symbols in our HMM for fingering decision, it is one if the given output symbol n_k is the one that the hidden state q_i outputs and zero if the given output symbol is not,

$$b_{ik} = \begin{cases} 1 & (n_k = n(q_i)) \\ 0 & (n_k \neq n(q_i)) \end{cases} . \quad (2)$$

2.2 Note-tablature-form tree

To perform fingering decision as a decoding problem of HMM described in the previous section, we need to enumerate left hand forms for each note in a given sequence of notes, which is done by drawing note-tablature-form tree

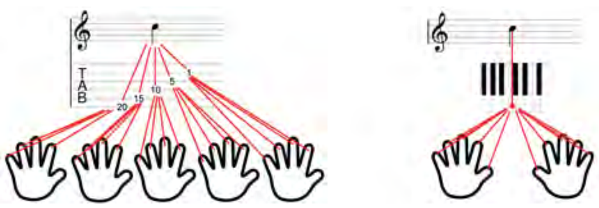


Figure 1. Note-tablature-form tree for guitar (left) and corresponding diagram for piano (right) illustrating difference between fingering decision of string instruments and other instruments

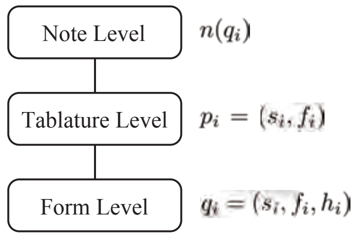


Figure 2. Three-level model for fingering decision of string instruments

that describes the difference between fingering decision of string instruments and other instruments as follows.

For example, on the piano, there is only one key on the keyboard to press for each note, and therefore fingering decision for a given sequence of notes is a matter of deciding which finger to press on the key for each note (Fig. 1, right). On the other hand, with the guitar, each note corresponds to several string-fret pairs that play it, and in addition, we have a matter of which finger to press for each string-fret pair (Fig. 1, left). In other words, fingering decision for the piano is simply a matter of finger assignments while fingering decision for the guitar consists of string assignments followed by finger assignments. This situation with the guitar is illustrated in a tree diagram (Fig. 1, left) which we call "note-tablature-form tree." While a note-tablature-form tree for a monophonic case in Fig. 1 is easy to draw, we tackle the problem of drawing corresponding trees for polyphonic cases in the following section.

The above-explained situation with fingering decision of string instruments is described by a three-level model for string instruments that consists of (1) note level, (2) tablature level, and (3) form level (Fig. 2). In relation to the notation introduced in Section 2.1, the note level contains the information of $n(q_i)$, the tablature level $p_i = (s_i, f_i)$, and the form level $q_i = (s_i, f_i, h_i)$, respectively. In guitar scores, the score and the tablature contains the information of the note level and the tablature level, respectively. From the viewpoint of fingering decision based on HMM, the hidden states correspond to the form level and the observed symbols to the note level.

3. NOTE-TABLATURE-FORM TREE FOR POLYPHONIC CASES

We discussed fingering decision method based on HMM and note-tablature-form tree limiting our attention to monophonic cases in the previous section for the sake of simplicity. However, in order to prove our fingering decision method practical, we need to consider polyphonic cases where multiple notes are played simultaneously. To extend our fingering decision method for polyphonic cases, it is enough to extend note-tablature-form tree to one for polyphonic cases which we construct in this section, and then we can perform fingering decision for polyphonic cases in the same manner for monophonic cases. (See [5] for HMM for polyphonic cases.) The construction of note-tablature-form tree for a given chord consists of searching for tablatures for a given chord followed by searching for left hand forms for each of the obtained tablatures.

3.1 From chord to tablature

Searching for tablatures for a given chord is relatively easy compared to searching for forms for a given tablature. It consists of assigning the chord notes to the strings followed by confirming that the notes are within the pitch ranges of the assigned strings. If we are given a chord consisting of n notes ($n \leq 6$), we have ${}_6P_n$ permutations of n out of 6 that give possible assignments of the n notes to the 6 strings. (If we can assume that the pitches of the notes in the chord is monotonic with respect to the string numbers, we can reduce the search to a number of combinations rather than permutations, but this is not the case because we consider left hand forms including open strings that can break the monotonicity of the pitches.) In Algorithm 1, $chord$ is a variable length array of MIDI note numbers while tab is a fixed length array of fret numbers, 0 or $None$ with length 6, the number of the strings, where 0 means an open string while $None$ means that the string is not played. The array $open$ keeps the MIDI note numbers of the open strings. Subtracting the MIDI note number of the open string from the MIDI note number of the note yields the fret number for playing the note on the string. If all the fret numbers obtained by such subtraction are between 0 and F, the number of the frets of the instrument, then the tablature is valid and is added to the list of tablatures $tabs$. Fig. 3 illustrates an example note-tablature tree for a simple C chord consisting of three notes, C, E and G, generated by Algorithm 1 for a standard guitar with $F = 21$ frets. Out of ${}_6P_3 = 120$ permutations, only 8 permutations give valid tablatures.

Algorithm 1

```
procedure CHORD2TABS(chord)
    tabs ← ∅
    n ← length of chord
    for perm ← perms of n from {0, 1, 2, 3, 4, 5} do
        tab ← [None, None, None, None, None, None]
        for i = 0 to n − 1 do
            tab[perm[i]] ← chord[i] − open[perm[i]]
        if ((f ≥ 0 ∧ f ≤ F) ∨ f = None) for all f ∈ tab then
            tabs ← tabs ∪ {tab}
    return tabs
```

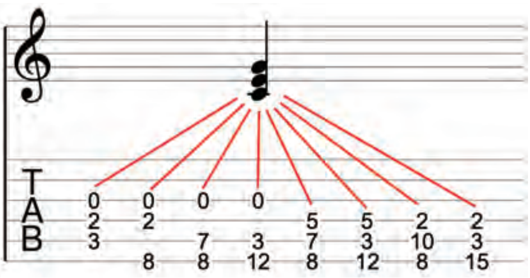


Figure 3. Note-tablature tree for C chord

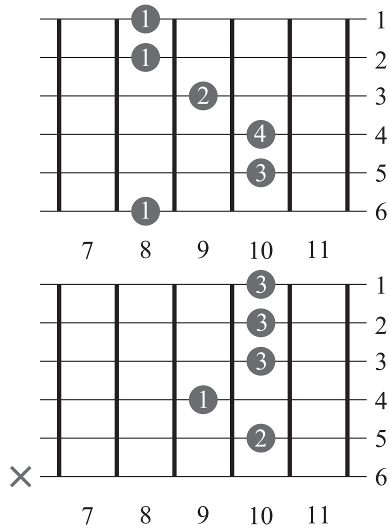


Figure 4. Left hand forms represented by finger numbers

3.2 From tablature to form

Searching for left hand forms for a given tablature is relatively difficult because the search space of left hand forms is huge and it is difficult to enumerate all the possible left hand forms in an orderly manner like possible tablatures as permutations. Two main strategies for searching for left hand forms for a given tablature can be considered here. One is to collect known forms from guitar chord books to build a database of left hand forms and then search for forms most suitable for a given tablature in the database. The other is to exhaustively search for all the possible left hand forms for a given tablature. They have their own pros and cons. The former is easy to implement and is guaranteed not to output strange forms but may miss out on some novel new left hand forms. The latter may not miss out on any forms but is difficult to implement because it is difficult to enumerate all the possible forms. We introduce a new idea for enumerating possible left hand forms to implement the latter approach in the following.

3.2.1 *Representing forms by finger numbers*

Fig. 4 shows two examples of standard guitar chords, C and G7(9), and their standard fingerings where the finger numbers 1, 2, 3 and 4 indicates the index, middle, ring and pinky fingers, respectively. The index finger (1) in the upper chord and the ring finger (3) in the lower chord hold down multiple strings. A method of playing in which one finger holds down multiple strings is called "barre" or "ceja." In the way the chords and their fingerings are displayed in Fig. 4, left hand forms can be represented by assigning finger numbers to all the string-fret pairs to hold down.

3.2.2 *Numbering string-fret pairs*

We consider the numbering of the string-fret pairs on the fretboard as shown in Fig. 5 where a string-fret pair (f, s) is numbered by an integer $6f - s$ in hexadecimal where $f = 1, 2, \ldots$ is a fret number and $s = 1, 2, \ldots, 6$ is a string number (from the highest to the lowest). We show

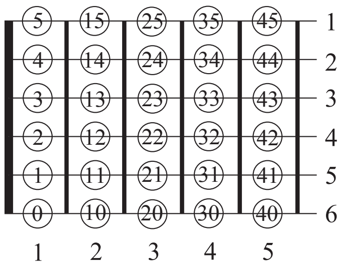


Figure 5. Numbering of string-fret pairs by integer $6f - s$ in hexadecimal

the numbering in hexadecimal notation in Fig. 5 just for ease of reading and it is not essential in the following discussion whether the numbering is expressed in decimal or hexadecimal.

3.2.3 Non-decreasing finger numbers

Then it holds with very few exceptions that the finger numbers representing a single left hand form are monotonically non-decreasing with respect to the numbering. This is because (i) a finger with a larger finger number is put on a higher or the same fret (and therefore a string-fret pair with a larger or equal digit in its 6's place), and (ii) a finger with a larger finger number is shorter (except that the middle finger (2) is longer that the index finger(1)) and is put on a string with a smaller string number (and therefore a string-fret pair with a larger digit in its first place). Fig. 6 illustrates that the finger numbers representing a single left hand form are monotonically non-decreasing with respect to the numbering using two cases of the chords in Fig. 4. The upper chord in Fig. 6 has six string-fret pairs to hold down with numberings 70, 74, 75, 83, 91 and 92 that are assigned with the finger numbers 1, 1, 1, 2, 3 and 4 in Fig. 4, which are monotonically non-decreasing with respect to the numberings. The lower chord in Fig. 6 has five string-fret pairs to hold down with numberings 82, 91, 93, 94 and 95 that are assigned with the finger numbers 1, 2, 3, 3 and 3 in Fig. 4, which are again monotonically non-decreasing with respect to the numberings.

3.2.4 Enumerating left hand forms

Having observed that the finger numbers are non-decreasing with respect to the numbering of the string-fret pairs, we can enumerate left hand forms for a given tablature without any leaks by inserting three separators separating four fingers into the sequence of the string-fret pairs representing the given tablature. Fig. 7 represents two left hand forms of Fig. 4 using string-fret pairs and separators, where long and short lines indicate mandatory and optional separators which are explained in the following sections. The separators are inserted between string-fret pairs where the finger numbers change. The index finger is assigned to the string-fret pairs to the left of the left separator, the middle finger to ones between the left and the middle separators, the ring finger to ones between the middle and the right separators, and the pinky finger to ones to the right of the right separator. If a separator is at the leftmost or the rightmost as in the right form of Fig. 7 or two separators are next to each

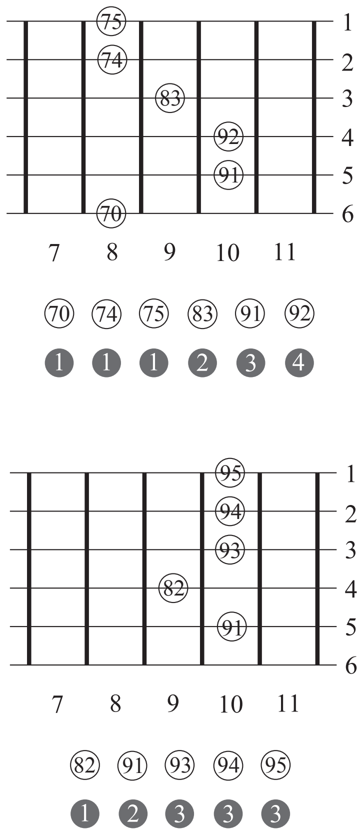


Figure 6. Non-decreasing finger numbers representing left hand forms

other, it means that some fingers are not used in the form. When we have n string-fret pairs to hold down, there are a total of ${}_{n+3}C_3$ ways to insert three separators, but only a small fraction of these are actually available.

3.2.5 Inserting mandatory separators

When we insert three separators into a sequence of string-fret pairs, there are positions where separators must first be inserted. First, because it is impossible to hold down different frets with one finger, a separator must be inserted between two string-fret pairs with different fret numbers, that is, with numberings with different digits in their 6's places. Second, because it is impossible for one finger to hold down string-fret pairs on a single fret separated by a string played at a lower fret, such as 91 and 93 sepa-

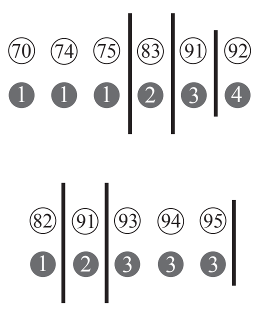


Figure 7. Left hand forms represented by string-fret pairs and separators

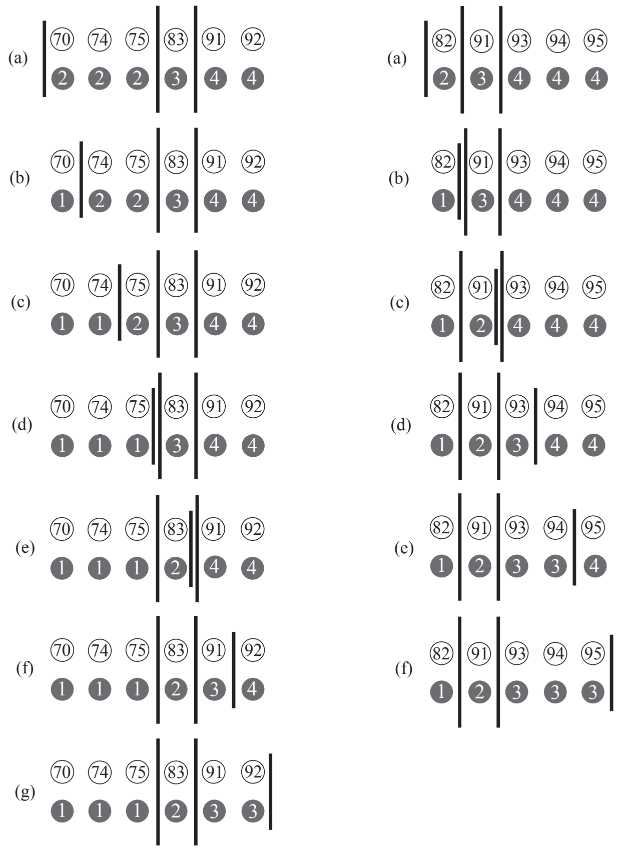


Figure 8. Exhaustive enumeration of left hand forms by inserting optional separators

rated by the fourth string played at 82 in the lower form of Fig. 7, a separator must be inserted between such two separate string-fret pairs. We call separators inserted to such positions "mandatory separators." In Fig. 7, long lines indicate mandatory separators while short lines indicate optional separators which are explained in the following section.

3.2.6 *Inserting optional separators*

When we have less than three mandatory separators, we insert optional separators until we have three separators in total. For example, both forms of Fig. 7 have two mandatory separators, which make us insert one more optional separator for each. The upper form of Fig. 7 has six string-fret pairs thus seven positions to insert an optional separator while the lower form has six positions to insert as shown in Fig. 8, where mandatory and optional separators are indicated by long and short lines. As the positions of the separators change, so do the finger numbers under the string-fret pairs and thus the left hand forms. We note that some forms of Fig. 8 are very difficult or impossible to play and not all of those forms are available. Here we have introduced a new idea for exhaustive enumeration of left hand forms for a given tablature and we still need to discuss how to eliminate the unplayable ones, which we leave to our future study.

4. CONCLUSION

We have reviewed a guitar fingering decision method based on HMM and note-tablature-form tree for monophonic cases and tried to extended the tree diagram to polyphonic cases. For that purpose, we have introduced an enumeration method of tablatures as permutations for a given chord. Furthermore, we have introduced a new idea for exhaustive enumeration of left hand forms for a given tablature based on non-decreasing finger numbers and two kinds of separators that assign fingers to string-fret pairs. We have noted that some left hand forms enumerated by our proposed method are very difficult or impossible to play and need to be eliminated. We leave elimination of such unplayable forms to our future study.

Acknowledgments

This work was supported by JSPS KAKENHI Grant Number 21H03462.

5. REFERENCES

[1] S. I. Sayegh, "Fingering for string instruments with the optimum path paradigm," *Computer Music Journal*, vol. 13, no. 3, pp. 76–84, 1989.

[2] D. Radicioni, L. Anselma, and V. Lombardo, "A segmentation-based prototype to compute string instruments fingering," in *Proceedings of the Conference on Interdisciplinary Musicology (CIM04)*, vol. 17, Graz, Austria, 2004, pp. 97–104.

[3] A. Radisavljevic and P. F. Driessen, "Path difference learning for guitar fingering problem," in *Proceedings of International Computer Music Conference (ICMC2004)*, vol. 28, Miami, USA., 2004.

[4] D. R. Tuohy and W. D. Potter, "A genetic algorithm for the automatic generation of playable guitar tablature," in *Proceedings of International Computer Music Conference (ICMC2005)*, 2005, pp. 499–502.

[5] G. Hori, H. Kameoka, and S. Sagayama, "Input-output HMM applied to automatic arrangement for guitars," *Journal of Information Processing*, vol. 21, no. 2, pp. 264–271, 2013.

[6] W. Nagata, S. Sako, and T. Kitamura, "Violin fingering estimation according to skill level based on hidden Markov model," in *Proceedings of International Computer Music Conference and Sound and Music Computing Conference (ICMC/SMC2014)*, Athens, Greece, 2014, pp. 1233—-1238.

[7] E. Nakamura, N. Ono, and S. Sagayama, "Merged-output HMM for piano fingering of both hands," in *Proceedings of International Society for Music Information Retrieval Conference (ISMIR2014)*, Taipei, Taiwan, 2014, pp. 531–536.

[8] G. Hori and S. Sagayama, "Minimax Viterbi algorithm for HMM-based guitar fingering decision," in *Proceedings of International Society for Music Information Retrieval (ISMIR2016)*, New York City, U.S.A., 2016, pp. 448–453.

[9] G. Hori, "Extension of decoding problem of HMM based on L^p-norm," in *Proceedings of 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP2018)*, 2018, pp. 3989–3993.

[10] ——, "Three-level model for fingering decision of string instruments," in *Proceedings of the 15th International Symposium on CMMR*, 2021, pp. 93–98.

SUPEROM: A SUPERCOLLIDER CLASS TO GENERATE MUSIC SCORES IN OPENMUSIC

Claudio Panariello
Sound and Music Computing group
KTH Royal Institute of Technology,
Stockholm, Sweden
claudiop@kth.se

Emma Frid
IRCAM STMS Lab,
Paris, France
KTH Royal Institute of Technology
Stockholm, Sweden
emmafrid@kth.se

ABSTRACT

This paper introduces SuperOM, a class built for the software SuperCollider in order to create a bridge to OpenMusic and thus facilitate the creation of musical scores from SuperCollider patches. SuperOM is primarily intended to be used as a tool for SuperCollider users who make use of assisted composition techniques and want the output of such processes to be captured through automatic notation transcription. This paper first presents an overview of existing transcription tools for SuperCollider, followed by a detailed description of SuperOM and its implementation, as well as examples of how it can be used in practice. Finally, a case study in which the transcription tool was used as an assistive composition tool to generate the score of a sonification – which later was turned into a piano piece – is discussed.

1. INTRODUCTION

Automatic generation of notation is a complex topic [1]. The design of computational algorithms to convert acoustic music signals into some form of music notation, so-called Automated Music Transcription (AMT), is a challenging task in signal processing and artificial intelligence [2]. Typically, AMT systems take an audio waveform as input and compute a time-frequency representation, after which a representation of pitches over time is outputted [2]. This process involves, among other subtasks, multi-pitch estimation (MPE), onset and offset detection, beat and rhythm tracking, interpretation of expressive timing and dynamics, as well as score typesetting. A comprehensive overview of signal processing methods for music transcription was presented in [3]. Discussions of challenges for AMT have been published in [4, 2]. While the technological aspects of AMT tools are highly relevant to the work presented in this paper, they differ somewhat from Computer-Aided Composition (CAC) tools (software that allows composers to design computer processes that generate musical structures and data [5]) in the sense that they are primarily designed to aid the transcription process, not necessarily the composition process, which is the purpose of the `SuperOM` presented in this paper. More specifically, the purpose of the current work is to translate code into scores, not to generate scores directly from audio files [1].

Sound synthesis tools used in assisted composition usually do not include automatic music notation features, mainly because these tools were not designed with that particular use case as a primary motivation. However, having the possibility to visualize sounds in standard Western music notation can be useful in many composition contexts. Different attempts have been made to fill this gap, but the efforts have often been characterized by a lack of documentation, making it difficult to present a full review of the tools and methods used. A reoccurring strategy in this context seems to be to use SuperCollider [2] [6] classes that can bridge with LilyPond [7], either directly or through third-party software. SuperCollider is a programming language for audio synthesis and algorithmic composition. LilyPond [3] is a free system to write music.

One of the oldest attempts to bridge sound synthesis software with automatic notation tools is LilyCollider [4], developed by Bernardo Barros [8]. LilyCollider is an interactive software that can build sequences of music notation in an interactive way, extending the SuperCollider programming language (sclang). LilyCollider wraps the LilyPond music notation software, meaning that it can be used to generate a LilyPond score from SuperCollider code. However, the system has some limitations when it comes to rendering time; it requires you to wait until the score is engraved [9]. Today, LilyCollider has been abandoned in favor of SuperFomus [5], which was also developed by Barros. SuperFomus relies both on LilyPond and FOMUS [6] to generate a music score. FOMUS is an open-software application developed by David Psenicka which allows the automation of many musical notation tasks for composers and musicians. It was designed with composers who work with algorithms and computer music software languages in mind and facilitates the process of creating professionally notated scores. An interesting aspect of SuperFomus

[1] Although that would be possible using the SuperOM, for example using spectral analysis features.
[2] https://supercollider.github.io/
[3] https://lilypond.org/
[4] https://github.com/smoge/LilyCollider
[5] https://github.com/smoge/superfomus
[6] https://fomus.sourceforge.net/

is that it allows for musicXML export. In other words, it can be used to generate a file that easily can be edited in any scorewrite software. However, SuperFomus has some limitations, specifically when it comes to more advanced rhythm algorithms, and it may not be the optimal solution for working with metric structures [9]. In other words, it can be somewhat unreliable when dealing with complex music notation structures. In addition, it appears as though is no longer maintained, which may add some troubleshooting time when installing it.

Another system is Fosc [7], which stands for FO-rmalised S-core C-ontrol (FO-r S-uperC-ollider). Fosc is an API that ports much of the Python code base of Abjad [8], a system that supports composers in building complex pieces of music notation in iterative and incremental ways, to SuperCollider. Since Abjad wraps the LilyPond music notation package, Fosc can be used for the generation of musical notation in LilyPond. Despite being powerful, Fosc does not allow for musicXML export, thus limiting the quantity and quality of information that can be preserved in the score.

Finally, a custom-made SuperCollider class called SonaGraphLily was included in the SonaGraph framework, a harmonic spectrum analyzer suitable for assisted and algorithmic composition, developed by Andrea Valle [10]. SonaGraphLily manages mapping from sonographic data (i.e. spectrum over time) to music notation, using LilyPond code. It creates LilyPond source files that are rendered as graphic files. However, besides the fact that this class is optimized to work on a SonaGraph instance, it doesn't support musicXML export.

2. MOTIVATION

Although the above-described software solutions may be useful for certain use cases, they all fall short when it comes to generating musically complex scores while allowing for musicXML export, which is the main aim of the `SuperOM`, described in the forthcoming sections. In other words, the goal of `SuperOM` is to enable generation of scores from SuperCollider in the fastest way possible with as few dependencies as possible, while at the same time preserving extremely high precision in the notation, and allowing for musicXML export.

The `SuperOM` is a SuperCollider class that produces music scores in the form of OpenMusic (OM) files, i.e. in the form of `.omi` files. OpenMusic [11] is an open-source environment [9] dedicated to music composition, developed at IRCAM in the end of the 1990s (see e.g. [12, 13]). It is a visual programming language based on Lisp allowing the user to design processes for the generation and manipulation of music material, but it can also be used for other applications [13]. Similarly to other graphical environments such as Pure Data (Pd) [10] and Max/MSP [11], the workflow to code programs in OpenMusic is based on patching together different modules. However, as opposed to such tools, the output of the OM processes is also visualized using conventional music notation. A screenshot of the OpenMusic interface is displayed in Figure 1. OpenMusic has been used by a wide community of composers and computer musicians throughout the years. Notable composers include, among others Kaija Saariaho, Marco Stroppa, Brian Ferneyhough, Philippe Manoury, Fabien Lévy, and Mauro Lanza. [12]

There are several reasons why OpenMusic was adopted in this project. Firstly, OpenMusic's powerful capabilities, especially in terms of handling very complex music structures, as well as its exporting features, were important features. OpenMusic allows for export into many different file formats, e.g. MIDI, bach [14], and musicXML. As such, it enables the generation of files that can be opened and edited in most common scorewriter software (such as MuseScore, Sibelius, Finale, or Dorico, just to name a few). Another reason is the simplicity of the overall installation: `SuperOM`'s only dependency is, in fact, OpenMusic, which installation is quite straightforward. Moreover, no prior knowledge of OpenMusic is actually required, since the only OpenMusic features required by a SuperCollider user are import/export file. Finally, yet another reason is the aim to bridge two open-source software solutions as well as their communities. In fact, `.omi` files generated by the `SuperOM` in SuperCollider are completely legit and working OpenMusic files. This means that they can be manipulated directly by OpenMusic users. This aspect can encourage collaborative frameworks in which SuperCollider and Open Music users can work together, exchanging their own material.

However, as for all software solutions, the adoption of OpenMusic might have some drawbacks. For example, the playback functionality in OpenMusic could be considered an obstacle, since it requires a third-party software synth player to work, which is out of the scope of this paper. Nevertheless, it seems that future versions of OpenMusic will have an embedded synth available. [13] On the other hand, if a SuperCollider user decides to use OpenMusic merely as a way to export a musicXML file for a notation software, the playback part can indeed be skipped.

3. CLASS DESCRIPTION

The main method of `SuperOM` is `.writeOMfile`, which takes the following arguments: `fileName, midicents, magnitudes, rhythmTree, metronome, quantization, threshold, dynamics` (see Listing 1).

Once an `.omi` file has been produced, it can be imported in an OpenMusic patch and opened from there. The file can then be edited directly in OpenMusic and exported as an XML file (for example using the POLY object).

[7] `https://github.com/n-armstrong/fosc`
[8] `https://abjad.github.io/`
[9] It can be downloaded for free from `https://openmusic-project.github.io/openmusic/`
[10] `https://puredata.info/`
[11] `https://cycling74.com/`
[12] `https://en.wikipedia.org/wiki/OpenMusic`
[13] Please see the OpenMusic IRCAM Forum `https://discussion.forum.ircam.fr/t/open-music-midi-player/39930`, accessed 14 December 2022.

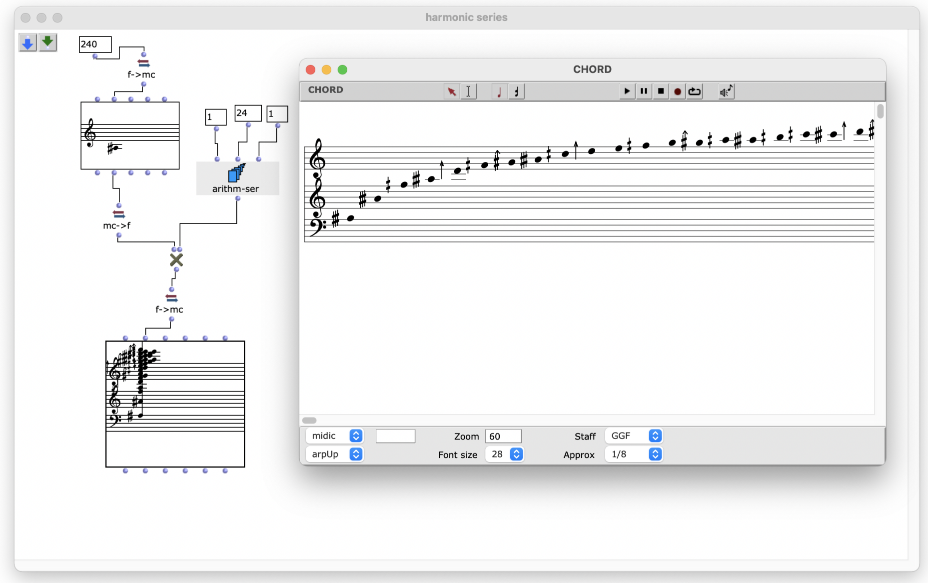


Figure 1. A basic example of an OpenMusic patch. In this example, several modules are connected to generate a harmonic series on the fundamental frequency 240, up to the 24th harmonic. The window CHORD is displaying the result using conventional music notation.

```
o = SuperOM.new;
o.writeOMfile(fileName:, midicents:,
    magnitudes:, rhythmTree:, metronome
    :, quantization:, threshold:,
    dynamics:);
```

Listing 1. SuperOM

A more detailed description of the arguments taken by `.writeOMfile` follows below.

3.1 fileName

The argument `fileName` is the name for the output file, including the `.omi` extension (e.g. `"test.omi"`). The output will be produced in the same folder where the SuperCollider file is located.

3.2 midicents

The argument `midicents` is an array of notes, expressed in midicents. This argument is always passed as an array of staves, thus leaving to `SuperOM` the task to interpret how many staves should be written in the OM output file. For example, in the case of $[[l_0, l_1, l_2, ...], [m_0, m_1, m_2, ...], [n_0, n_1, n_2, ...], ...]$ the rows are interpreted as subsequent separated staves; the first staff will contain notes $[l_0, l_1, l_2, ...]$, the seconds staff will contain notes $[m_0, m_1, m_2, ...]$, the third staff will contain notes $[n_0, n_1, n_2, ...]$, etcetera. Chords can be specified with additional brackets, e.g. $[[n_0, n_1, [n_2, n_3], n_4, ...]]$. In this example, there is one staff and the notes n_2 and n_3 are interpreted as a bichord.

3.3 magnitudes

The argument `magnitudes` is an array (or an array of arrays) of the notes' magnitudes. `magnitudes` can be expressed either in decibels (dBs) or as MIDI velocities (i.e. values from 0 to 127), the `SuperOM` will take care of interpreting the given array in the correct way with the method `.dbvelSmart`. `magnitudes` must match the size of `midicents`. If no magnitudes are given, all the notes will automatically be set to velocity $= 100$.

3.4 rhythmTree

`rhythmTree` is an array of rhythms, provided as floats (or fractions), i.e. 0.5 or $1/4$ for a quarter note, 0.25 or $1/8$ for an eight note, etc. A positive value represents the duration of a note, while a negative value represents a pause. For example, $-1/16$ represents a rest of a sixteenth note. A rest, however, can also be specified using the SuperCollider's operand `Rest()`. If this argument is `nil`, the `magnitudes` argument will be used as a source to create a rhythm tree (see Section 4.2).

3.5 metronome

This argument specifies the metronome of the score. If only one value is given, the metronome will be the same for the entire score. As an alternative, an array of metronomes can be specified, matching the rows size of the `midicents` array, thus assigning a different metronome to each staff. The default metronome value is 60 bpm.

3.6 quantization

Amount of quantization of the notes, expressed in MIDICENTS. The default value is 50 (i.e. a quarter-tone).

3.7 threshold

This argument, expressed in dB, sets a threshold for the magnitudes, meaning that values below this level will be

Figure 2. Instance of an OpenMusic POLY class. The figure shows the content of the file produced by the code shown in Listing 2.

considered as silence, i.e. pauses. The default value is −36 dB.

3.8 dynamics

The dynamics flag can be set to `true` or `false`. If `true`, the output file will display the notes' magnitudes as music dynamics (i.e. from "*ppp*" to "*fff*") in the score. The MIDI velocities are converted into music dynamics using the method `.veldyn`. The default value is `false`.

4. EXAMPLES

In the following section, a number of examples are provided to demonstrate potential use cases for `SuperOM`. All of the example .omi files, as well as the corresponding XML output, are available for download from here: `https://tinyurl.com/9j2fba5p`.

Listing 2 shows the most basic example of `SuperOM`: the array `pitches` contains a list of notes and a chord, everything expressed in midicents; the array `rhythm` contains a list of durations and pauses. Line 5 of Listing 2 produces an `.omi` file with the name `"example0.omi"` that will contain the music material specified in the two arrays. The result is a file that, once imported in OM, will look like the one displayed in Figure 2.

```
1 var pitches = [6000, [6200, 6550,
      6800], 7000];
2 var rhythm = [1/2, -1/6, 2/6, -2/5,
      3/5];
3
4 o = SuperOM.new;
5 o.writeOMfile("example0.omi", [pitches
      ], rhythmTree: rhythm, metronome:
      144);
```

Listing 2. Basic usage of SuperOM.

Listing 3 shows another simple usage of the class, this time producing a chromatic scale with eighth-tones, starting from C4, as 32th notes. Please notice that in order to create a score with eighth-tones, we have to specify the correct quantization in `writeOMfile`, using `quantization: 25`.

```
var pitches = (6000, 6025..7200);
var rhythm = {1/32}.dup(pitches.size);

o = SuperOM.new;
o.writeOMfile("example1.omi", [pitches
    ], rhythmTree: rhythm, metronome:
    144, quantization: 25);
```

Listing 3. A chromatic scale with eighth-tones starting from C4.

Listing 4 shows a simple variation of the previous example, adding random magnitudes and printing them in the score, with the flag for the argument `dynamics` set to `true`. The result imported in OM will look like the one displayed in Figure 3.

```
var pitches = (6000, 6025..7200);
var mags = {rrand(-18, -3)}.dup(
    pitches.size);
var rhythm = {1/32}.dup(pitches.size);

o = SuperOM.new;
o.writeOMfile("example2.omi", [pitches
    ], magnitudes: mags, rhythmTree:
    rhythm, metronome:144, quantization
    : 25, dynamics:true);
```

Listing 4. A chromatic scale with eighth-tones starting from C4, with dynamics.

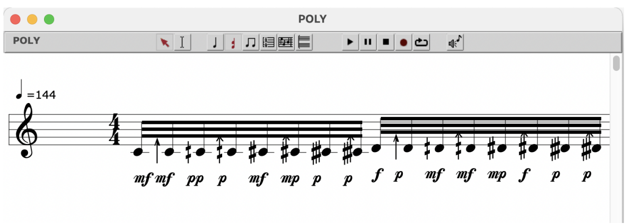


Figure 3. Instance of an OpenMusic POLY class showing the content of the file produced by the code shown in Listing 4.

Using standard features provided in SC, a more complex score can be generated with a few lines of code. Listing 5 shows the code to produce a score with the following properties: five staves with different chromatic scales; rhythms and pauses chosen by a given set, with each staff having a different metronome.It is worth noticing that in this example the argument pitches is passed without extra brackets (see line 7) since it has already been created as an array of staves (see line 2). The result imported in OM will look like the one displayed in Figure 4.

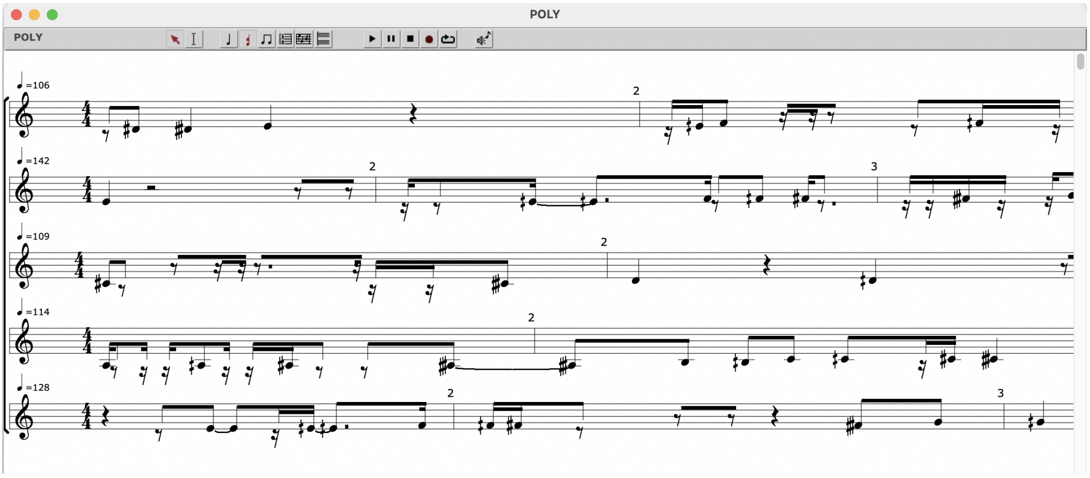


Figure 4. Instance of an OpenMusic POLY class showing the content of the file produced by the code shown in Listing 5.

```
var staves = 5;
var pitches = {(6000, 6050..7200)+(
    rrand(-5, 5)*100)}.dup(staves);
var rhythm = {{[1/4, 1/8, 1/16].choose
    *[-1, 1].choose}.dup(pitches.shape
    [1]*2)}.dup(staves);
var metronomes = {rrand(102, 144)}.dup
    (staves);

o = SuperOM.new;
o.writeOMfile("example3.omi", pitches,
     rhythmTree: rhythm, metronome:
    metronomes);
```

Listing 5. Snippet of code to produce a score with five staves with different chromatic scales, rhythms with pauses chosen by a given set, and five different metronomes.

Listing 6 shows code that produces a score with eight staves, each of them containing random frequencies taken from a harmonic series, random magnitudes, and random rhythms, quantized to 32th notes. The frequencies are conveniently translated into midicents thanks to the `.cpsmidicents` method.

```
var notes = 200, staves = 8;
var pitches = {{Array.fill(24, {|i|(i
    +1*rrand(50, 51))}).choose}.dup(
    notes)}.dup(staves).cpsmidicents;
var mags = {{rrand(-18, -3)}.dup(notes
    )}.dup(staves);
var rhythm = {{rrand(0.1, 1).softRound
    (1/32, 0, 1)}.dup(notes)}.dup(
    staves);

o = SuperOM.new;
o.writeOMfile("example4.omi", pitches,
     mags, rhythm, 84, 25, -36, true);
```

Listing 6. Snippet of code to produce a more complex score with eight staves.

4.1 Writing a score from patterns

Patterns are typical SuperCollider data structures that allow for the management of events in time, specifying the rules for the production of such events [1]. As demonstrated in Listing 7, a musical piece expressed through a pattern can then be translated into a score using the `SuperOM`. In order to do that, events must be conveniently stored into separate arrays, for example using the method `.collect`. These can then be used as arguments for generating the output score.

```
var length = 50, pitches, rhythm;
p = Pbind(
        \midinote, Pxrand([60, 62, 64,
            66, 68, 70], inf),
        \dur, Prand([1/16, 1/8, Rest
            (1/16), Rest(1/8)], inf));

e = p.asStream;
pitches = length.collect({e.next(()).
    midinote})*100;
rhythm = length.collect({e.next(()).
    dur});

o = SuperOM.new;
o.writeOMfile("example6.omi", [pitches
    ], rhythmTree: rhythm, metronome
    :144);
```

Listing 7. Writing a score from a pattern.

4.2 Writing a score from spectral data

Another interesting application of `SuperOM` is the ability to write a score from spectral data, as seen in Listing 8. This feature can be useful in situations when you want to create a score from a series of frequencies and magnitudes from a spectral analysis process. The strategy used here is based on grouping subsequent notes that have the same frequency and magnitude. Grouping in this context means that rhythmic values are summed together (for example, two 1/8 notes are replaced by one 1/4 note). Notes that have different magnitudes will not be grouped. Magnitudes below the given threshold value are interpreted as silence, meaning that such notes are transformed into pauses.

```
var freqs = {rrand(400, 500)}.dup(50).
    cpsmidicents;
var mags = {[-18, -12, -9, -6].choose
    }.dup(freqs.size);
freqs.postln;
mags.postln;
o = SuperOM.new;
o.writeOMfile("example5.omi", [freqs],
     magnitudes: mags, rhythmTree: nil,
     metronome:144, quantization: 100,
    threshold: -12, dynamics:true);
```

Listing 8. Writing a score from spectral data

5. CASE STUDY – GENERATING A PIANO PIECE USING SUPEROM

The `SuperOM` has been intensively used by the first author in his artistic practice. The tool has been very useful especially since it allows for the creation of transcriptions of material derived from spectral analysis. Below is a short account of a case study in which `SuperOM` was used in a composition process involving sonification of video material.

`SuperOM` was successfully used in a joint study carried out together with the pianist and researcher Johan Fröst. This project made use of sonification of a video that was created from multiple video recordings of Fröst playing Debussy's *"Reflects dans l'eau"*. More specifically, the aim of the project was to sonify moving images using a Disklavier piano, a self-playing piano designed by Yamaha [14]. Sonification is defined as the use of nonspeech audio to convey information [15] [15].

The video to be sonified was created starting from multiple video recordings of Fröst's performance. The videos were edited together, highlighting musical events and the musical narrative of Debussy's piece. The final merged video was then sonified, and the sonification was used as starting material in the composition process. The aim of the composition process was to create a new piano piece to be played in synch with the original video in a live performance at a concert series in the Spring of 2023. The incpit of piano piece is shown in Figure 5.

The overall sonification workflow involved three stages and two different softwares: 1) the video was loaded into Max/MSP, in which video processing took place using, among others, the cv.jit package [16]; 2) the data was sent to SuperCollider via OSC [17], where the mappings for the actual sonification was implemented; 3) MIDI messages were generated in SuperCollider and sent to the Disklavier, which played the newly generated piano piece. To establish a strong connection between the video and the generated sound, the sonification used harmonic content derived from the same Debussy piece that drove the creation of the video.

Once the sonification was realized, `SuperOM` was used to generate an actual score of it. In order to do that, the mappings relative to pitches, note durations and velocities were stored into separate arrays, and used to initialize an instance of an `SuperOM`. The output score was an accurate representation of the sonification: as a matter of fact, the complex metric structures produced with the sonification were completely captured in the score. Once exported in musicXML, the file could be opened without errors in a commercial music notation software, where the composition process continued by selecting and merging musically interesting materials, and by reducing their complexity in order to make them playable. However, this workflow had some limitations. Firstly, despite the score containing all the note velocity information, thus making its MIDI playback sound correctly, it did not have any music dynamics printed on the staves. This was solved by improving the code and adding the `dynamics` flag presented in Section 3. A second obstacle was that the score produced by the `SuperOM` contained a large number of staves (namely 25), resulting from how the sonification data was stored. This aspect made the score quite impracticable to read and difficult to work on, and it required a staves merging operation, done by hand, in order to achieve a typical piano-looking score. In the future, such a problem could be solved by carefully designing `SuperOM` methods to collapse many staves into one.

It should be noted that the final product of the process outlined above was not a literal sonification mapping the video input directly to output (i.e. the final piece) in an objective way; the author interacted with the generated material, merging and adapting different generated parts into a final piece, thus disrupting the direct connection between input and output. In other words, sonification was used as a subtask within the composition process, assisting the author's composition process by providing ideas for the final piano piece. This process is somewhat similar to what is referred to as mixed-initiative interaction in the field of Human Computer Interaction (HCI), in which each agent (human or computer) contributes what it is best suited at the most appropriate time [16]. Using this mixed strategy outlined above, the piece closes the circle of re-mediation from a piano piece by Debussy, to: 1) a video made with the intent to visualize the musical material of Debussy's piano music; 2) a sonification based on this video, using material from the original piano music; 3) the use tonal material created in step 2), which in turn served as material for a new composition for the piano; and finally 4) a performance played in real-time by a pianist.

[14] `https://www.disklavier.com/`

[15] This definition has later been expanded by Thomas Hermann to: *"(...) data-dependent generation of sound, if the transformation is systematic, objective and reproducible (...)*, see `https://sonification.de/son/definition/`.

[16] `https://jmpelletier.com/cvjit/`

[17] `https://ccrma.stanford.edu/groups/osc/index.html`

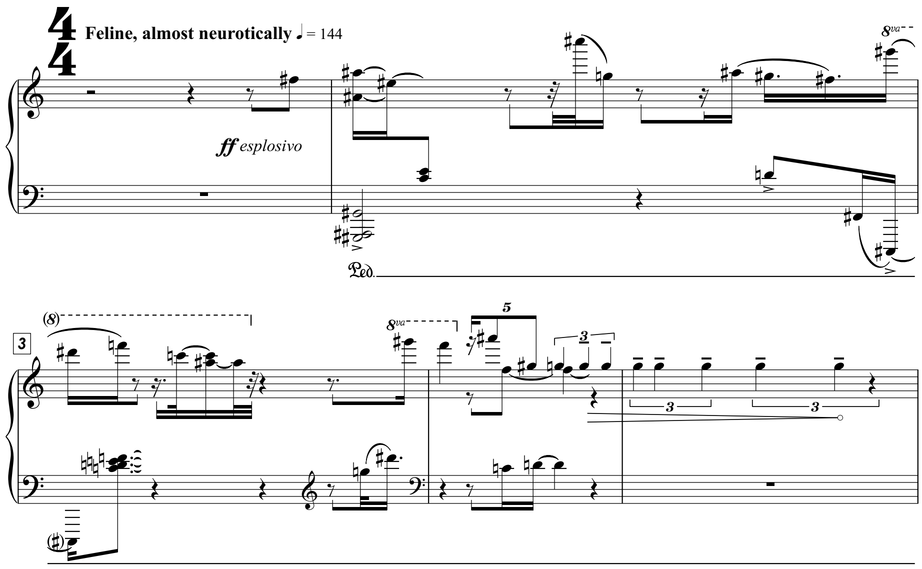


Figure 5. Excerpt from the beginning of the piano piece composed starting from a sonification of video images.

6. LIMITATIONS

A limitation of the current version of the `SuperOM` concerns combining midicents staves of different lengths. As a matter of fact, midicents arrays should all have the same length, in order to create a final rectangular matrix. One workaround to solve this issue is to fill the shorter arrays with zeros (a sort of zero padding), thus matching the size of the longest array, see Listing 9. In this way, the additional 0 pitches will be ignored, as long as the rhythm tree doesn't contain rhythm information.

```
var pitches1 = [7200, 7400, 7500,
    7600];
var rhythm1 = [1/6, -2/6, 1/4, 1/4,
    1/4];

var pitches2 = [6000, 6200, 6550,
    6800, 7000, 6800, 5300, 5625, 6378,
     6740];
var rhythm2 = [1/2, -1/6, 0, 0, 2/6,
    -2/5, 3/5, 1/6, -1/6, 0, 0, 1/6,
    1/4];

var pitches3 = [5500, [5600, 5950],
    5700, 6050];
var rhythm3 = [-1/8, 1/8, 1/8, 1/8];

pitches2.do({pitches1=pitches1++0});
pitches2.do({pitches3=pitches3++0});

o = SuperOM.new;
o.writeOMfile("example9.omi", [
    pitches1, pitches2, pitches3],
    rhythmTree: [rhythm1, rhythm2,
    rhythm3], metronome: 144);
```

Listing 9. Generating a score from midicents arrays with different lengths.

Interestingly, rhythm trees that contain zeroes make subsequent notes collapse, thus creating chords. This effect is demonstrated in the example in Listing 10, in which two ways of writing the same musical score are compared. As before, also here we need to zero pad the shortest arrays.

```
var pitches1 = [6000, [6200, 6550,
    6800], 7000, 6800, [5300, 5625,
    6378], 6740];
var rhythm1 = [1/2, -1/6, 2/6, -2/5,
    3/5, 1/6, -1/6, 1/6, 1/4];

var pitches2 = [6000, 6200, 6550,
    6800, 7000, 6800, 5300, 5625, 6378,
     6740];
var rhythm2 = [1/2, -1/6, 0, 0, 2/6,
    -2/5, 3/5, 1/6, -1/6, 0, 0, 1/6,
    1/4];

pitches2.size.do({pitches1=pitches1
    ++0});

o = SuperOM.new;
o.writeOMfile("example10.omi", [
    pitches1, pitches2], rhythmTree: [
    rhythm1, rhythm2], metronome: 144);
```

Listing 10. Comparison of two different ways of producing the same score in output.

7. CONCLUSIONS AND FUTURE WORK

This paper presents `SuperOM`, a class for bridging SuperCollider with OpenMusic thus enabling generation of complex music scores with a high level of precision. Files generated with `SuperOM` can be imported and edited in OpenMusic, which allows for collaborative frameworks between the two software.

The SuperCollider code of the implementation of `SuperOM` is still in progress and is continuously improved. The code is readily available on from here: `https://github.com/claudiopanariello/SuperOM`. The material provided includes the `SuperOM` files and a tutorial file. At the point of writing, the first author is using `SuperOM` in several projects, for example in the generation of scores from algorithmic compositions realized in SuperCollider, or in the creation of music transcriptions of material derived from spectral analysis of audio recorded material (especially audio feedback recordings). This ongoing work will continue to inform the design of `SuperOM`, allowing it to be iteratively improved over time. The use case presented in Section 5 serves as a formative evaluation of the `SuperOM`, carried out by the author from a first-person perspective. In the future, there is a need for empirical evaluation with actual users, to identify weaknesses and areas of possible improvement.

It is worth mentioning that the music scores discussed in this paper refer to traditional Western score. There are many situations in which other score notations might be more appropriate. Therefore, a possible future direction might be to design other classes that could allow for non-standard notations, similarly to what Ghisi and Agostini did in extending bach by introducing the dada library [17].

Acknowledgments

The first author would like to express his gratitude to Mauro Lanza, who kindly helped out with the OpenMusic code.

8. REFERENCES

[1] A. Valle, *Introduction to SuperCollider*. Logos Verlag Berlin GmbH, 2016.

[2] E. Benetos, S. Dixon, Z. Duan, and S. Ewert, "Automatic music transcription: An overview," *IEEE Signal Processing Magazine*, vol. 36, no. 1, pp. 20–30, 2018.

[3] A. Klapuri and M. Davy, *Signal processing methods for music transcription*. Springer Science & Business Media, 2007.

[4] E. Benetos, S. Dixon, D. Giannoulis, H. Kirchhoff, and A. Klapuri, "Automatic music transcription: Challenges and future directions," *Journal of Intelligent Information Systems*, vol. 41, no. 3, pp. 407–434, 2013.

[5] D. Bouche, J. Nika, A. Chechile, and J. Bresson, "Computer-aided composition of musical processes," *Journal of New Music Research*, vol. 46, no. 1, pp. 3–14, 2017.

[6] S. Wilson, D. Cottle, and N. Collins, *The SuperCollider Book*. The MIT Press, 2011.

[7] H.-W. Nienhuys and J. Nieuwenhuizen, "LilyPond, a system for automated music engraving," in *Proceedings of the XIV Colloquium on Musical Informatics (XIV CIM 2003)*, vol. 1. Citeseer, 2003, pp. 167–171.

[8] B. Barros, "LilyCollider and rhythmic structures," *Revista Vórtex*, vol. 2, no. 2, 2014.

[9] Bernardo Barros, "Music Notation with SuperCollider." [Online]. Available: http://bernardobarros.com/files/lilycollider-sc2013/slides.pdf

[10] A. Valle, "SonoGraph. a cartoonified spectral model for music composition," in *Proceedings of the 16th Sound & Music Computing Conference*. SMC, 2019, pp. 462–469.

[11] G. Assayag, C. Rueda, M. Laurson, C. Agon, and O. Delerue, "Computer-assisted composition at IRCAM: From PatchWork to OpenMusic," *Computer Music Journal*, vol. 23, no. 3, pp. 59–72, 1999.

[12] C. Agon, "Openmusic: Un langage visuel pour la composition musicale assistée par ordinateur," Ph.D. dissertation, Paris 6, 1998.

[13] J. Bresson, C. Agon, and G. Assayag, "OpenMusic: Visual programming environment for music composition, analysis and research," in *Proceedings of the 19th ACM international conference on Multimedia*, 2011, pp. 743–746.

[14] A. Agostini and D. Ghisi, "A Max library for musical notation and computer-aided composition," *Computer Music Journal*, vol. 39, no. 2, pp. 11–27, 2015.

[15] G. Kramer, B. Walker, T. Bonebright, P. Cook, J. H. Flowers, N. Miner, and J. Neuhoff, "Sonification report: Status of the field and research agenda," University of Nebraska - Lincoln, Tech. Rep., 2010.

[16] J. E. Allen, C. I. Guinn, and E. Horvtz, "Mixed-initiative interaction," *IEEE Intelligent Systems and their Applications*, vol. 14, no. 5, pp. 14–23, 1999.

[17] D. Ghisi and A. Agostini, "Extending bach: A family of libraries for real-time computer-assisted composition in Max," *Journal of New Music Research*, vol. 46, no. 1, pp. 34–53, 2017.

MUSASSIST: A DOMAIN SPECIFIC LANGUAGE FOR MUSIC NOTATION

Ilana Shapiro
Pomona College
Claremont, USA
issa2018@mymail.pomona.edu

ABSTRACT

MusAssist is an external, declarative, domain specific language for music notation that bridges the abstraction gap between Western music theory and composition. Users can describe unique high-level templates for chords and arpeggios (all triads and seventh chords), scales (diatonic, chromatic, and whole tone), the five primary cadences, and the four primary harmonic sequences with desired length. Distinctively, MusAssist matches the level of abstraction of a template to the theoretical musical structure it describes (e.g. users can specify a harmonic sequence without needing to manually expand it to the chords and notes it comprises). Thus, users can write out specifications precisely at the conceptual levels of the musical structures they would organically conceive when composing by hand. In MusAssist, users can also change key signatures, start a new measure, and describe fundamental musical objects such as notes, rests, and customized chords. The Haskell-based MusAssist compiler expands high-level templates (thus lowering the level of abstraction to individual notes) and translates the program to MusicXML, a language accepted by most major music notation software, for further manual editing and playback.

1. INTRODUCTION

When writing music in the framework of Western music theory, composers manually transition from theoretical musical concepts to notes on a page. This process can be slow and tedious, requiring the composer to expand complex musical structures by hand, such as cadences and harmonic sequences, to the individual notes they define. Therefore, the level of abstraction of the musical structure is higher than what the composer writes.

Domain specific languages (DSLs) are programming languages highly specialized for a specific application and thus characterized by limited expressiveness. An *external* DSL has custom syntax that is separated from the primary language of its application.

This paper presents MusAssist, an external, declarative DSL for music notation that bridges the abstraction divide between Western music theory and notation. Users describe a composition in MusAssist's straightforward, high-level syntax, modeled around the musical elements composers organically conceive when writing by hand, and the MusAssist compiler automates the expansion of these elements to their constituent notes. MusAssist's declarative programming paradigm was chosen to correspond with the lack of control structures in handwritten music.

MusAssist is unique in that users can encode specifications for complex musical templates at the same level of abstraction as the theoretical musical structures they describe. Specifically, users can specify high-level templates for chords and arpeggios (major, minor, dominant, augmented, half diminished, and diminished triads and seventh chords in any inversion), scales (major, natural/harmonic/melodic minor, chromatic, and whole tone), cadences (perfect authentic, imperfect authentic, plagal, half, deceptive), and harmonic sequences (ascending fifths, descending fifths, ascending 5-6, descending 5-6) of a desired length. MusAssist also supports individual notes, rests, and customized chords consisting of user-defined collections of notes and enables the user to change the key signature or start a new measure. All high-level templates are expanded, lowering the abstraction level to notes, by the Haskell-based MusAssist compiler.

The target language of the MusAssist compiler is MusicXML, itself a DSL that is an extension of XML (Extensible Markup Language). MusicXML is accepted by most major notation software, such as MuseScore. Thus, users can open the resulting MusicXML file of a compiled MusAssist composition in MuseScore or another program for further customization, editing, and playback. MusAssist thus serves as a professional music compositional aid, filling in an integral part in the composition process where Western composers must make the transition from the theoretical, high-level musical framework with which they describe their piece to the notes that these structures constitute. The subsequent ability of users to manually edit the compiled MusAssist composition speaks to MusAssist's present role as an assistant in the creative process, rather than a fully expressive means of music creation. MusAssist may also be particularly helpful to Western music theory students as an educational tool, enabling them to visualize the relationship between a Western theoretical musical structure and its expanded form, such as in conceptualizing the chords resulting from the expansion of a cadence.

This paper first summarizes related work in music notation DSLs. Then, features of the MusAssist language are outlined, followed by the presentation of a use case. Finally, the MusAssist compiler structure and the central logic behind its automated template expansions are described.

2. RELATED WORK

The era of music programming languages began in 1957 with Max Mathews' MUSIC-N languages at Bell Labs, developed on individual IBM mainframes. The first of these languages, Music I, generated a single, equilateral, triangular waveform with identical rise and decay characteristics. It was capable of playing only melody: the user specified pitch, amplitude, and duration for individual notes. Music II supported four independent voices of sound and a choice of 16 waveforms, and Music III introduced the concept of the $unit\text{-}generator$, the building block for sound synthesis programming languages that corresponds to the functions of analog synthesizers. Music IV was a more computationally sophisticated version of Music III, and machine independence was finally reached in 1966 with Music V, which became the first publicly distributed music programming language [1].

Since then, computer scientists have taken advantage of the increased flexibility afforded to DSLs via their limited expressiveness to create music DSLs tailored towards notation, algorithmic composition, sound synthesis, live coding with music performance, and more. Aside from MusAssist, MusicXML, LilyPond, and PyTabs are commonly used music DSLs also specializing in notation.

Michael Good's MusicXML is an Internet-friendly, XML-based, declarative DSL that represents standard Western music notation and scoring practices. Similar to how the popular MIDI format helped create a standardized format for electronic instruments, MusicXML provides a standardized format for online sheet music to create a consistent method for representing complex, structured musical data. MusicXML thus introduces smooth interchange between musical applications specializing in notation, performance, analysis, music information retrieval, and more [2].

To achieve this, Good derives MusicXML from XML due to its Internet-friendly nature, straightforward use in document creation, and human readability. In contrast, MIDI is very difficult to read and write, and is also less powerful and expressive than XML [3].

MusicXML is more expressive than MusAssist, but the abstraction level of all musical elements is extremely low (i.e. chords must be written out as individual notes) and though easily readable, its syntax is cumbersome and tedious to write by hand. However, its enhanced expressiveness makes MusicXML an excellent target compilation language for MusAssist's user-friendly syntax and high-level Western theoretical musical templates.

LilyPond, an external, declarative DSL created by Han-Wen Nienhuys and Jan Nieuwenhuizen, is similar to MusAssist. It features a "modular, extensible and programmable compiler" written in Scheme to generate Western music notation of excellent quality, and supports the mixing of text and music elements. Text-based musical expressions, or fragments of music with set durations, are compiled to an aesthetically formatted score [4].

LilyPond and MusAssist are both music notation DSLs tailored to non-programming audiences. However, they differ in two fundamental areas: (1) MusAssist supports complex music templates at the levels of abstraction of the musical structures they represent, whereas LilyPond only supports granular, low level composition of individual notes and chords, and (2) the output of the MusAssist compiler is intentionally editable via notation software, unlike LilyPond's compiler, which produces a static, printable PostScript or PDF file by taking in a file with a formal representation of the desired music [4].

Simic et al.'s external, declarative DSL PyTabs is also geared toward music notation, but in a different domain than MusAssist. Specifically, the authors attempt to solve the visual problem of tablature notation and the lack of standardization in specifying tablature-based note durations by consolidating these issues into a formal language. However, tablature notation is outside the scope of MusAssist's text-based Western theoretical musical structures [5].

3. LANGUAGE FEATURES

3.1 Low-Level Fundamentals

On the most basic level, MusAssist supports individual rests and notes. Rests are given a duration from sixteenth to whole note, and notes are further defined by note name (A to G), accidental (double flat to double sharp), and octave (1 to 8, after the range of a piano). Just as in traditional Western notation, the absence of an accidental indicates natural quality. Users can also define customized chords, or user-defined lists of individual notes. These are not considered templates as the high-level description of the chord is not given, and the granularity is at the note level.

3.2 High-Level Templates

MusAssist supports templates for chords, arpeggios, scales, cadences, and harmonic sequences, specified uniquely at the abstraction levels of the Western theoretical musical structures they represent.

Following the principles of Western music theory, chords are specified by their root note, quality (major, minor, augmented, dominant, diminished, or half diminished), inversion (root, first, second, or third), and chord type (triad or seventh). Half diminished, dominant, and third inversion options can only apply to seventh chords. The root note cannot have a double accidental, as this can introduce triple accidentals in the chord, which MusAssist does not support.

Arpeggios are defined with a similar specification to MusAssist chords, since according to Western music theory, an arpeggio is simply the notes of a chord played individually in sequence. However, unlike with chords, for arpeggios the user also supplies direction (ascending or descending).

Diatonic scales are given by scale type (major or natural/harmonic/melodic minor) and key, while non-diatonic scales are simply specified by their type (chromatic or whole tone). A scale is either ascending or descending, must be given a length, and does not necessarily begin on the tonic – a start note must be supplied. Following convention, chromatic scales are notated with sharps when ascending and flats when descending.

Cadences are specified by cadence type (perfect/imperfect authentic, half, plagal, or deceptive) and key.

Currently, MusAssist only supports a single treble clef line. Thus, cadences are written out in the upper voices only, in keyboard voice leading style, and incorporating principles of smooth voice leading.

Based on the principles of Western functional harmony, there are several ways to represent a cadence. In MusAssist, the representations in Table 1 were chosen. In each row of Table 1, the major version is presented first, with the minor version following in parentheses.

Perfect Authentic	IV-V-I (iv-V-i)
Imperfect Authentic	IV-vii$^{o}{}^{6}_{4}$-I$^{6}_{4}$ (iv-vii$^{o}{}^{6}_{4}$-i$^{6}_{4}$)
Plagal	IV$^{6}_{4}$-I (iv$^{6}_{4}$-I)
Deceptive	IV-V$^{6}_{4}$-vi$^{6}_{4}$ (iv-V$^{6}_{4}$-VI$^{6}_{4}$)
Half	IV-ii^{6}-V (iv-ii^{o6}-V)

Table 1: MusAssist Cadences Summary

All cadences except perfect authentic are built exclusively with triads. Although MusAssist does not currently support a bass line, in order to simulate the requisite 4-5-1 bass line for perfect authentic cadences, the root in the final chord is doubled. This also allows for the uppermost voice to follow the requisite 2-1 downward step in the final two chords of the cadence without compromising the root position of the final chord. The perfect authentic cadence is demonstrated in Figure 1, produced with the MusAssist syntax `(Perfect Authentic Cadence, Eb5 minor, sixteenth)` compiled and loaded into MuseScore notation software.

Figure 1: Perfect Authentic Cadence in E♭ minor

Finally, harmonic sequences are specified by harmonic sequence type (ascending fifths, descending fifths, ascending 5-6, or descending 5-6), key, duration of each chord, and length of the sequence. Since MusAssist does not yet support multi-line composition, harmonic sequences are written like cadences in keyboard-style voice leading.

In Western music theory, harmonic sequences can be implemented in several ways depending on the desired inversion scheme. Though the upper-voice harmonization of a harmonic sequence need not follow the direction in the sequence's name, MusAssist chooses a chord inversion and voice leading pattern such that each sequence does so. For instance, the Ascending Fifths sequence will ascend, and the Descending 5-6 sequence will descend. Each pattern also maximizes smooth voice leading.

The chosen patterns for each MusAssist sequence are summarized in Table 2. All sequences are shown in major in this demonstration, but their minor counterparts are also supported. Each sequence consists of fourteen distinct chords before repeating in the subsequent octave.

Ascending Fifths	I$^{6}_{4}$	V	ii$^{6}_{4}$	vi	iii$^{6}_{4}$
	vii^{o}	IV$^{6}_{4}$	I	V$^{6}_{4}$	ii
	vi$^{6}_{4}$	iii	vii$^{o}{}^{6}_{4}$	IV	
Descending Fifths	I	IV$^{6}_{4}$	vii^{o}	iii$^{6}_{4}$	vi
	ii$^{6}_{4}$	V	I$^{6}_{4}$	IV	vii$^{o}{}^{6}_{4}$
	iii	vi$^{6}_{4}$	ii	V$^{6}_{4}$	
Ascending 5-6	I	vi^{6}	ii	vii^{o6}	iii
	I^{6}	IV	ii^{6}	V	iii^{6}
	vi	IV^{6}	vii^{o}	V^{6}	
Descending 5-6	I$^{6}_{4}$	V	vi$^{6}_{4}$	iii	IV$^{6}_{4}$
	I	ii$^{6}_{4}$	vi	vii$^{o}{}^{6}_{4}$	IV
	V$^{6}_{4}$	ii	iii$^{6}_{4}$	vii^{o}	

Table 2: MusAssist Harmonic Sequences Summary

3.3 Additional Features

Beyond compositional elements, users can set the key signature at the start of any measure up to seven sharps or flats by specifying note name, accidental, and quality (sharp or flat). Users can also start a new measure or create a blank measure. Finally, users can assign MusAssist expressions to string labels and reuse them later in the program (the labels are syntactic sugar for the expressions). MusAssist comments are designated with the double slash `//`.

The tempo for all MusAssist programs is set to ♩= 80 BPM and cannot currently be changed. This also applies to the time signature, which is set to $\frac{4}{4}$.

All compiled MusAssist programs adhere to standard notation conventions. Notes and rests are broken over barlines as well as over the strong beat (beat three) of the measure. They are divided greedily into valid rhythmic units (from sixteenth to whole note) ordered either least to greatest, or greatest to least in the case of spillage of a tied note over the barline into the following measure.

4. SAMPLE PROGRAM

The full breadth of MusAssist's syntax is demonstrated in Figure 2, and Figure 3 presents the resulting compiled MusicXML code when opened in MuseScore.

Several features of MusAssist are clarified in Figures 2 and 3:

- The key signature can be changed consecutively as many times as desired, but only the last will take effect (as seen m. 1 and m. 5 of Figure 3). Changing the key signature also triggers a new measure.
- Empty measures can be created by consecutively applying the `NEW_MEASURE` command (as seen m. 3 of Figure 3).
- Note and rest durations are automatically broken by the compiler both on the strong beat and on the barline (such as in mm. 1-2 of Figure 3). However, following standard Western notation convention, notes and rests that begin on a beat and fit in the remainder of the measure are not broken on the strong beat (such as in m. 9 of Figure 3).

- Labeled phrases are not notated until the label is referenced, rather than defined.
- The difference between a customized chord and a chord template is exemplified on lines 7 and 11 of Figure 2.

```
1 SET_KEY A major
2 SET_KEY A major
3 SET_KEY G major
4 (D4 sixteenth) (F#4 sixteenth) (A4 eighth)
5 (D#4 eighth) (F##4 eighth) (A#4 quarter) (rest sixteenth)
6 notes = (E3 dotted_eighth) (G4 dotted_eighth)
7 custom_chords = ([Bbb5, Db6, Fb6] half) ([G#5, C#6, E#6] quarter)
8 (Descending Fifths Sequence, G5 minor, eighth, length=7)
9 NEW_MEASURE
10 NEW_MEASURE
11 (Eb4 diminished seventh chord, first inversion, quarter)
12 SET_KEY D minor
13 SET_KEY C# major
14 (C# harmonic minor descending scale, startNote = E6, sixteenth, length=10)
15 (D#4 major triad, first inversion, dotted_quarter) (rest eighth)
16 (E4 major triad, first inversion, eighth) (B4 dotted_eighth)
17 (C5 dotted_eighth) (rest sixteenth) (rest sixteenth) (rest eighth)
18 ([B4, A4] eighth) (C5 dotted_eighth) ([G#4, B4] dotted_eighth)
19 (rest eighth) (Perfect Authentic Cadence, D#5 major, dotted_eighth)
20 (rest sixteenth) notes (F3 dotted_eighth) (Ab4 dotted_eighth)
21 (F#3 dotted_eighth) (A4 dotted_eighth)
22 custom_chords (Ascending Fifths Sequence, C#4 major, eighth, length=8)
23 (G#5 minor arpeggio ascending, root inversion, sixteenth)
24 (C#5 augmented arpeggio descending, second inversion, sixteenth)
25 (Deceptive Cadence, Eb5 minor, eighth) (rest quarter) custom_chords
```

Figure 2: MusAssist Syntax

Figure 3: Compiled MusAssist Program in MuseScore

5. COMPILER STRUCTURE

The MusAssist compiler is written in Haskell. Its high-level structure is as follows:

1. MusAssist's concrete syntax is parsed into abstract syntax, represented as Haskell algebraic data types (ADTs). Parser combinators were chosen for their flexibility and ease of customization. Parsec, an industrial strength parser library, is used, and Parsec's helper module Token handles lexing. The parse preserves the abstraction level of all templates.
2. All templates, now represented as ADTs, are expanded using the logic in Section 6 until the granularity reaches the note level. The result of this intermediate stage is abstract syntax whose abstraction level matches that of the target language, MusicXML.
3. The low-level abstract syntax resulting from the fully expanded templates is translated to MusicXML. This step contains the temporal logic that subdivides notes and rests across barlines and strong beats.

The resulting MusicXML file can then be opened in standard music notation software like MuseScore for viewing, further editing, and playback.

6. TEMPLATE EXPANSION LOGIC

MusAssist's distinguishing feature is its ability to automate the expansions of Western theoretical musical templates given user-supplied specifications at a higher level of abstraction. The logic underlying the expansions is summarized below.

6.1 Backbone Logic

6.1.1 Generating Notes in a Diatonic Scale

Most MusAssist templates are built upon the diatonic scale. To automate the expansions of these templates, we must first be able to generate a note in a desired diatonic scale given a positive interval within one octave of the specified tonic. Recall that a MusAssist note is defined by note name, octave, and accidental. Given the target interval n, to determine the note name we begin at the tonic and travel n steps up MusAssist's custom Haskell ADT for note names, a circular Enum instance ordered as the C major scale is.

The desired octave is either the same as that of the tonic, or one greater if the desired note name (disregarding accidental) comes before the tonic note name in the C major scale. For instance, as seen in Figure 4, the red note names D, E, and F come before G in the C major scale, and the octave number of each is one higher than the tonic in a G major scale.

Figure 4: Octave Analysis of G Major Scale

To determine the desired accidental, first realize that in any key, a perfect interval will generally have the same accidental as the tonic, with two exceptions: the perfect fifth above B is F♯, and the perfect fourth above F is B♭. To work out the logic behind the accidental of a desired imperfect interval, consider Figure 5. Here, we enumerate all single-accidental key signature names (even invalid ones that contain double sharps or flats) in order to establish the pattern. Key signature names are grouped under the accidental of the note that is the desired interval from the tonic. For instance, in the key of A♭, the major second interval from the tonic is B♭. The accidental of B♭ is ♭, so A♭ falls under the ♭ column in Figure 5.

	Major Seconds		
♭	♮	♯	𝄪
F♭	E♭	E	E♯
C♭	B♭	B	B♯
G♭	F	F♯	
D♭	C	C♯	
A♭	G	G♯	
	D	D♯	
	A	A♯	

Figure 5: Accidental of Major Second from Tonic per Key

From Figure 5 we see that given any key, the major second above the tonic has the same accidental as the tonic, except for any key with E or B in its name. In these keys, the accidental is "lifted" (i.e. ♭→ ♮, ♮→ ♯, and ♯→ 𝄪).

A similar pattern emerges for minor thirds, major sixths, and minor sevenths from the tonic. Using the result of this analysis, we can determine the accidentals of the inverse qualities (i.e. major versus minor) of the imperfect intervals by either lowering the computed accidental when going from major to minor, or lifting it otherwise. Augmented and diminished intervals from the tonic are not considered since they do not appear in diatonic scales.

6.1.2 Generating Chord Templates in a Diatonic Scale

Generating chord templates in a diatonic scale becomes relevant for the expansions of the highest-level MusAssist templates – namely, cadences and harmonic sequences – that are defined in Western music theory by lists of chord templates rather than notes.

The goal here is to automate the generation of chord templates for triads in a diatonic scale given a specified tonic tone and quality for the scale (major or minor), inversion, and positive interval within one octave of the tonic for the chordal root. If we need to generate a chord template for a seventh chord, we simply generate the base triad and add the fourth note afterwards.

In order to complete the triad template definition from the supplied information, we simply need to determine the desired chord quality. Western music theory dictates that the major diatonic scale contains the triads I-ii-iii-IV-V-vi-vii°, and the minor diatonic scale contains the triads i-ii°-III-iv-v-VI-VII. Using this, we can compute the triad quality given the tonic quality and the supplied interval between tonic and desired chordal root.

6.2 Template Expansions

6.2.1 Scales

The expansion of all major and natural/harmonic/melodic minor scales is derived from the logic in Section 6.1.1. The scale is generated in relation to its tonic, rather than the specified starting note. The tonic octave is always set so that the tonic falls below the start note. This ensures that the initial interval between tonic and start note is positive, with the interval then increasing for ascending scales and decreasing for descending scales until the desired scale length is reached. If the tonic is reached in the scale generation, we reset the tonic to be one octave higher or lower (corresponding to scale direction) so that the interval between the next note in the scale and the current tonic is always positive and within a single octave. Finally, consider that all minor scales are treated as natural when generating their notes. If needed, the notes are modified afterwards in order to appropriately raise the sixth and/or seventh scale degree(s) for harmonic and melodic minor scales.

For the non-diatonic scales (chromatic and whole tone), the C below the start note is set as the "tonic" in order to determine the octave of each note in the scale generation. As with diatonic scales, the tonic octave is appropriately shifted one octave higher or lower, corresponding to scale direction, if it is reached during the scale generation.

As seen in Figure 6, for 10 of the 12 tones, chromatic scales "double" the note name, with the directional accidental (sharp for ascending, flat for descending) falling on the second occurrence. The two exceptions marked in red in Figure 6 are E and B in the ascending version, and C and F in the descending version.

Figure 6: Chromatic Scales

If we are on a single note, we simply move up the scale. Otherwise, we repeat it and insert an accidental on the second occurrence.

Whole tone scales are constructed with a similar model. As seen in Figure 7, the ascending whole tone scale is missing the note B, while the descending is missing C.

Figure 7: Whole Tone Scales

To determine the next note name, we simply traverse up or down the C major scale, excluding the appropriate tone. To determine the next accidental, we accordingly lift or lower the current accidental if we reach one of the red boundaries in Figure 7, otherwise leaving it unchanged.

6.2.2 Chords and Arpeggios

Each note in a chord or arpeggio is generated with the logic from Section 6.1.1 based on its interval from the tonic (i.e. the chordal root). Chordal thirds, fifths, and sevenths have respective intervals of 2, 4, and 6 from the tonic.

The imperfect chordal intervals are initially set to major for major, dominant, and augmented chords, and to minor for minor, half diminished, and fully diminished chords. Thus, after the note is generated, for augmented chords the chordal fifth accidental must be lifted, and the seventh must be lowered. For dominant seventh chords, the seventh must be lowered. Finally, for diminished chords, the fifth and seventh must be lowered, and for half diminished seventh chords, the fifth alone must be lowered.

To handle inversions, notice that the generated chord (whether triad or seventh) starts out in root position. Let n be the desired inversion value (0 for root, 1 for first inversion, 2 for second, and 3 or third). By incrementing the octaves of the first n tones of the chord in root position, we obtain the correct inversion. This process is demonstrated in Figure 8 with a C dominant seventh chord in third inversion.

If we instead have an arpeggio, we first generate the notes in the underlying chord using the above logic. If the arpeggio is specified as a triad, we then double the first note of

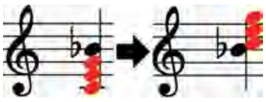


Figure 8: Chord Inversion Analysis

the arpeggio (which is determined by the inversion) an octave higher or lower, according to the specified direction, in order to have a complete arpeggio. For instance, a C5 major descending arpeggio in first inversion would have E6 as its first note and E5 as its last note. Finally, we sort the tones in the arpeggio so that they strictly ascend or descend depending on the specified arpeggio direction, and return them as a list of individual tones, rather than a simultaneous cluster as in a chord.

6.2.3 *Cadences*

A cadence must undergo two levels of expansion: an intermediate step from the cadence to the list of chord templates it defines, and a final step from chord templates to notes.

Recall that the chosen chords for each MusAssist cadence are defined in Table 1, which reveal the interval of each chordal root from the tonic. Using this, for each chord in the cadence, we first employ the logic from Section 6.1.1 to generate the root note of each chord. Then, we use the logic in Section 6.1.2 to generate a template for each chord, which then undergoes a second expansion in Section 6.2.2.

The scale quality supplied for the chord template generation in Section 6.1.2 is generally that of the cadence. However, there are exceptions. All V chord templates are generated within a major scale, no matter the cadence quality, since cadences always have major V chords. Similarly, we want the seventh triad in the imperfect authentic cadence to be diminished – i.e. built on the major seventh scale degree – no matter the local key quality, since we raise the leading tone in minor keys when moving towards the tonic. Thus, this chord template must also be generated within a major scale in Section 6.1.2.

The tonic octave supplied to Section 6.1.2 is also usually that of the cadence. However, in order for the cadences to follow smooth voice leading, the tonic octave must be lowered once in relation to the specified cadence octave when generating second inversion triads, which appear in all cadences except perfect authentic. This is demonstrated in the B major deceptive cadence in Figure 9. After converting to root position for clarity, note that the chordal root in the cadence octave is in blue and the roots an octave below (i.e. those of the second inversion triads) are in red.

Figure 9: Deceptive Cadence Octave Analysis

6.2.4 *Harmonic Sequences*

Like cadences, harmonic sequences are initially expanded to the chord templates they comprise, which then undergo a second expansion to notes. Recall from Figure 2 that each sequence consists of 14 chords, after which it cycles an octave above or below, depending on the direction of the sequence.

In order to generate chord templates for a sequence, we need to determine:

1. The interval between each pair of chordal roots, which dictates how to proceed from one chord to the next in the sequence
2. The inversion of each chord
3. The octave number of each chord (given by the chordal root octave) relative to the tonic

To determine (1) and (2), consider the analysis in Figure 10, which determines the interval pattern for the chordal roots of the descending 5-6 sequence (as defined previously in Figure 2), given the zero-based index of each chord in the sequence.

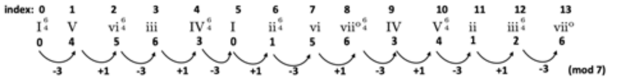


Figure 10: Descending 5-6 Interval Analysis

The top row of Figure 10 is the chord index, the second row is the chord progression from Table 2, the third row is the interval between each chordal root and the tonic, and the bottom row is the interval between each chordal root and the previous (modulo 7). A clear pattern for inversions (in the second row) and interval changes (in the fourth row) emerges based on the parity of the index. Identical analyses are applied to the remaining sequences to formalize their interval and inversion patterns.

Finally, we need to determine (3), the octave number of each chordal root relative to the tonic, or the root of the first chord in the sequence. Consider Figure 11, which presents an octave analysis for the descending 5-6 sequence.

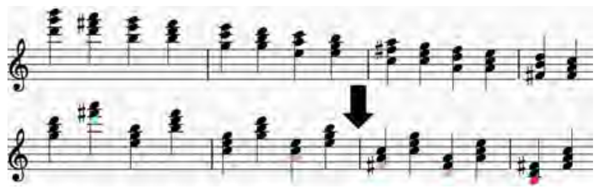


Figure 11: Descending 5-6 Octave Analysis

In Figure 11, all chords are converted to root position in order to visualize the octave numbers of their roots in relation to the tonic. The chordal roots in blue are an octave number above the tonic, the roots in yellow are an octave below, and the root in red is two octaves below. This same analysis is applied to determine the octave numbers of the chords in the remaining sequences.

Importantly, the interval analysis in Figure 10 holds for any representation of a harmonic sequence, as this is what defines the theoretical musical structure. However, the inversion and octave analyses in Figures 10 and 11 hold only for MusAssist's chosen representation of the sequences in Table 2, as a different inversion scheme would alter their outcomes.

7. CONCLUSION

This paper presents MusAssist, an external, declarative DSL for music notation that closes the abstraction gap between Western music theory and written composition. Users can uniquely write specifications in MusAssist's simple and high-level syntax for scales, chords, arpeggios, cadences, and harmonic sequences at the precise levels of abstraction of the Western theoretical musical structures they describe. Fundamental musical elements such as notes, rests, custom note collections, new measures, and key signatures are also supported, along with the ability to reuse labeled expressions and indicate comments. MusAssist programs are translated by its Haskell-based compiler to MusicXML, enabling the composition to be loaded into notation software for further manual editing and playback. Thus, MusAssist serves the role of a tool for Western composers to use after they have mapped out the harmonic framework of their piece and must realize the structures they have described. MusAssist also has the potential to help Western music theory students easily generate expanded forms of musical structures as they learn, such as cadences, sequences, and scales.

Optimally, in the future MusAssist would support templates for additional non-diatonic structures like octatonic and pentatonic scales. Templates for key modulations are also planned, which would provide specifications that generate a sequence of chords modulating from a start key to a target key. Furthermore, future versions of MusAssist will allow for increased customizability of existing templates that are not currently fully expressive. The present MusAssist template expansions for cadences and harmonic sequences are limited to the representation schemes outlined in Tables 1 and 2. In order for MusAssist to fully close the abstraction gap between Western theoretical musical structures and their low-level notational forms, all functional harmonic variations should be supported. Additionally, support for two-clef, multi-staff composition would improve the implementation of cadences and harmonic sequences by including the essential baseline, and support for changing meter and tempo would provide users with increased compositional flexibility.

MusAssist would also benefit from veering beyond Western tonal theory and into other realms such as jazz by supporting all flavors of suspended, ninth, eleventh, and thirteen chords. Furthermore, MusAssist would ideally extend its support for non-diatonic structures beyond those encountered in Western music, such as the microtonal modal systems found in the Arabian-Persian $maq\bar{a}m$, and the intervallic patterns of the Indian $r\bar{a}ga$. Finally, additional user studies of MusAssist would give insight into potential improvements for language design.

Acknowledgments

I am very grateful to Professor Ben Wiedermann of Harvey Mudd College for his invaluable mentorship throughout this project.

8. REFERENCES

[1] C. Roads and M. Mathews, "Interview with max mathews," *Computer Music Journal*, vol. 4, no. 4, pp. 15–22, 1980. [Online]. Available: http://www.jstor.org/stable/3679463

[2] M. Good, "Musicxml: Introduction," Apr 2013. [Online]. Available: https://www.musicxml.com/publications/makemusic-recordare/notation-and-analysis/introduction/

[3] ——, *MusicXML: An Internet-Friendly Format for Sheet Music*, Dec 2001. [Online]. Available: https://michaelgood.info/publications/music/musicxml-an-internet-friendly-format-for-sheet-music/

[4] H.-W. Nienhuys and J. Nieuwenhuizen, *LilyPond, A System for Automated Music Engraving*, May 2003. [Online]. Available: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.100.6160&rep=rep1&type=pdf

[5] M. Simić, Z. Bal, I. Dejanovic, and R. Vaderna, "Pytabs: A DSL for simplified music notation," *ResearchGate*, 2015. [Online]. Available: https://www.researchgate.net/publication/312607043_PyTabs_A_DSL_for_simplified_music_notation

BEYOND THE BASIC-SPACE OF TONAL PITCH SPACE: DISTANCE IN CHORDS AND THEIR INTERPRETATION

Hiroyuki Yamamoto
JAIST
Ishikawa, Japan
yamamoto@kusuli.com

Satoshi Tojo
JAIST
Ishikawa, Japan
tojo@jaist.ac.jp

ABSTRACT

Tonal Pitch Space (TPS) defines a numerical distance between two chord interpretations. Although it is based on musical knowledge and theory, the structure and values are not defined in an objective manner. Preceding works have addressed this problem, and TPS has been revised and optimized the definitions of distance, in the interpretation of chord paths, given chord names. But, because of the property of the task they used, they failed to reassess one of the three subelements of TPS, basic-space. In this study, we modify the task to incorporate pitch class (PC) information so that we can not only train other distance models that concern PC but also compare their performance with that of basic-space. We show that the data-oriented approach improves the accuracy from the original basic-space, especially when we add a distinction of major and minor keys.

1. INTRODUCTION

Tonality identification is an attractive but hard issue; although human listeners, often without any difficulty, fix one key to understand/ recognize tonal music, the exact process is still unknown. To determine a key, we need to consider the relationship between chords, considering cadences or tension/ relaxation structure. Moreover, we also need to model the relationship between each chord and pitch class (PC). But, when we represent this cognitive process in computers, we are required to assess the relationship objectively, excluding our subjectivity, so that the numeric distance in chords should be an intrinsic clue; *Tonal Pitch Space* (TPS) [8] has been one of the most convincing theories to give such a numeric distance between two chords.

Thus far, we have employed TPS to measure the distance in chords, however, some definitions of TPS look arbitrary. For example, the notion of *basic-space* (Figure 1) gives different hierarchical importance among 12 tones in an octave, diatonic (scale) tones, the third, the fifth, and the root in this order; but, is the difference of importance always one?
Yamamoto And Aojo [15] dared to Avoid employing the numerical definition of APS, Aut Anstead, they tried to make machines learn the similar distance model, giving a sequence of Berklee chord names. Since a chord name already includes rich information, it can be limitedly interpreted into pairs of key and degree without needing to consider the relationships to each PC. Therefore, they cannot reduce the definition of distance to each PC; that is, the adequacy of the basic-space in Figure 1 was untouched.

We reconsider the importance of the basic-space so that we once abandon chord names, excluding the bias offered by chord names, and employ chroma vectors, which directly mention each PC. With this, we try to obtain a statistic model which behaves similarly to TPS, to give a plausible interpretation for a sequence of pitch events.

2. PRELIMINARIES

2.1 Tonal Pitch Space

TPS is the quantitative harmony analysis proposed by Fred Lerdahl [8]. It is proposed to complement Lerdahl's original music theory, *A Generative Theory of Tonal Music* (GTTM) [7] which applies generative grammar to extend the Schenkerian theory. In TPS, a chord (e.g., C major triad) is interpreted as a pair of a key and a degree (e.g., interpretations of C major triad are as follows: I/C, III/a, V/F, IV/G, VI/e, and VII/d), then distances are defined between these chord interpretations. The distance between chord interpretations x and y can be calculated as (1).

$$\delta(x, y) = \mathrm{region}(x, y) + \mathrm{chord}(x, y) + \text{basic-space}(x, y) \tag{1}$$

where $\mathrm{region}(x, y)$ is a distance between keys, $\mathrm{chord}(x, y)$ is a distance between degrees.

$\text{basic-space}(x, y)$ is a distance on a structure called basic-space which concerns the importance of each PC relating to the chord interpretations. Basic-space is composed of five levels (i.e., root, fifth, triad, diatonic, and octave) and each level contains the PCs reflecting the chord interpretations. Figure 1 shows the example when $x = \mathrm{I/C}$ and $y = \mathrm{iv/d}$. For each chord interpretation, the root PC of the chord has four circles (i.e., up to the root level), the fifth PC has three circles, the third PC has two circles, and every other diatonic PC has one circle[1]. Then the distance between two chord interpretations is defined as the number

[1] We omit the octave level in Figure 1 and Figure 3 because it does not affect the results.

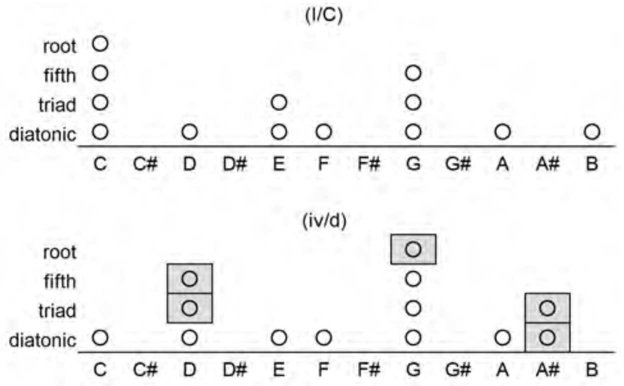


Figure 1. Basic-space.

of circles that exist only in the destination (the boxed circles in Figure 1). In this case, $\mathrm{basic\text{-}space}(\mathrm{I/C}, \mathrm{iv/d}) = 5$. The details are explained in [8].

The calculation above is applicable only when x and y are in relative keys which are defined as follows:

$$\mathrm{C}(R) = \begin{cases} \{\mathrm{I}, \mathrm{i}, \mathrm{ii}, \mathrm{iii}, \mathrm{IV}, \mathrm{V}, \mathrm{vi}\} & \\ \qquad\qquad \text{if } R \text{ is a major key} & \\ \{\mathrm{i}, \mathrm{I}, \mathrm{bIII}, \mathrm{iv}, \mathrm{v}, \mathrm{bVI}, \mathrm{bVII}\} & \\ \qquad\qquad \text{otherwise} & \end{cases} \tag{2}$$

where $\mathrm{C}(R)$ is the set of all relative keys of key R.

If x and y are not in relative keys, distance between x and y can be calculated as:

$$\delta(x, y) = \min \begin{pmatrix} \delta(x, T_{R_1}) + \Delta(R_1, R_n) + \delta(T_{R_n}, y) \\ |R_1 \in \mathrm{C}(R_x), R_n \in \mathrm{C}(R_y) \end{pmatrix}$$
$$\Delta(R_1, R_n) = \min \begin{pmatrix} \sum_{i=1}^{n-1} \delta(T_{R_i}, T_{R_{i+1}}) \\ |R_{i+1} \in \mathrm{C}(R_i) \end{pmatrix} \tag{3}$$

where T_R is key R's tonic, R_z is chord interpretation z's key. In other words, the transition from x to y must be considered as a combination of transitions within relative keys, and calculate the tonal distance for each combination, and then the shortest of these total distances is taken as the distance between x and y.

2.2 Distance Models concerning Harmonic Features

There have been a lot of approaches to applying some kinds of space to model harmonic features and utilizing the distance to calculate plausibility. Heinichen [5], Kellner [6], and Weber [14] tried to define the space to express the positional relationships of each key area (region). Riemann [10] applied the Tonnetz, which had been invented by Euler [3] as a way of representing just intonation, to analyze harmonic relationships from the viewpoint of PC. Bharucha and Krumhansl [1] proposed a model of tonal hierarchy which has an empirically defined value to express relationships between chords within the same region. Randall et al. [9] explored the similarities with Lerdahl's TPS [8], which is defined rather theoretically, as a metric space, and proposed another distance model. Tymoczko et al. [12] formalized the levels of abstraction when we try to interpret harmony. Yamamoto and Tojo [15] generalized the structure of TPS and applied machine learning to train several distance structures. Here, we try to extend their approach further to complement their study.

3. OUR APPROACH

In this study, we aim to obtain optimal distances between PCs and chords, through the task of finding the most plausible path in chord interpretation, from chroma vectors. First, we review the issue of chord interpretation (§3.1). Second, we introduce some distance models which can calculate the distance between a chroma vector and a chord interpretation (§3.2). Then, we explain the way how to embed the models of §3.2 into the method of §3.1 to enable the method to receive chroma vectors instead of chord names (§3.3).

3.1 From Chord Names to Chord Interpretation Paths

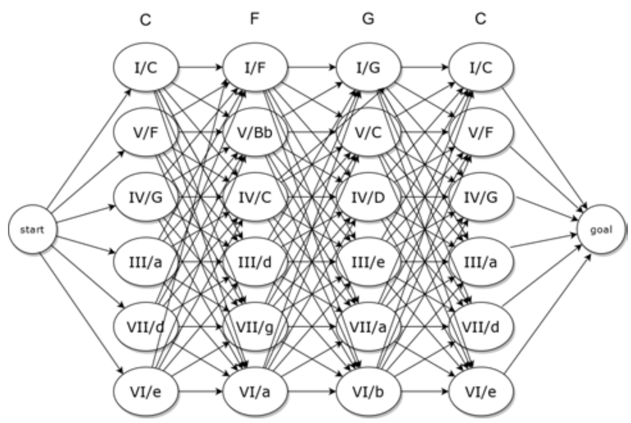


Figure 2. Interpretation graph.

Sakamoto et al. [11] have proposed a method to find the most plausible interpretation of a given chord name sequence. Given a chord name sequence, first, their method extends each chord to its interpretations and constructs a graph whose edges have weights that correspond to the distances on TPS. Then it applies the Viterbi algorithm to find the shortest interpretation paths from the start to the goal. Figure 2 shows an interpretation graph for chord name sequence $\mathrm{C} \to \mathrm{F} \to \mathrm{G} \to \mathrm{C}$. One of the shortest interpretation paths in Figure 2 is $\mathrm{I/C} \to \mathrm{IV/C} \to \mathrm{V/C} \to \mathrm{I/C}$.

Yamamoto and Tojo [15] have tried to generalize TPS and proposed several functions called "distance elements (DEs)" and a way to train them with annotated datasets. Based on the method of [11], their method replaces the TPS with the proposed generalized TPS then convert path distance to path probability such that the shortest path should have the highest probability, and finally apply SGD to update parameters. Their best model (i.e., a DE or combination of DEs) achieved over 86% accuracy while the original TPS was about 40%, and they also found a model with just 58 learnable parameters could achieve more than 80%.

In this study, we pick one of the most effective DEs proposed in [15] [2] as the base model on which we extend the structure in §3.3.

3.2 Between Chroma Vectors and Chord Interpretations

In this section, we introduce chroma distance models which are inspired by the structure of basic-space (and the basic-space itself is one of them).

A chroma vector is a 12-dimensional vector, the element of which represents its membership (1/0) or graded salience of the corresponding PC. We define the distance between a chroma vector and a chord interpretation as the sum of all distances between PCs and the chord interpretation.

Firstly, we can calculate this distance using basic-space. For example, the distance between the chroma vector [1, 0, 0, 1, 0, 0, 0, 1, 0, 0, 1, 0] and the chord interpretation I/C can be calculated as the inner product of the chroma vector and the vector generated from the basic-space (as the number of gray boxes) as in Figure 3. This means basic-space divides PCs into five categories, namely, root (i.e., C in this case), third (i.e., E), fifth (i.e., G), diatonic (i.e., D, F, A, B), and the others then gives the predefined PC-level distance values as in Table 1.

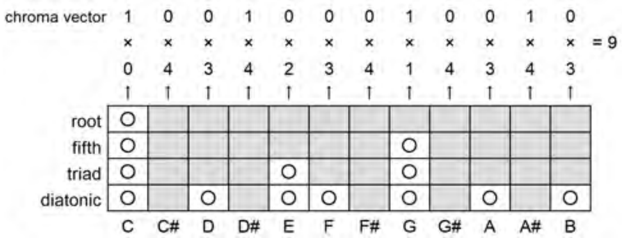


Figure 3. The distance between [1, 0, 0, 1, 0, 0, 0, 1, 0, 0, 1, 0] and I/C based on basic-space.

root	third	fifth	diatonic	other
0	2	1	3	4

Table 1. The PC-level distance values from basic-space.

But, how to classify PCs and what distance values to apply are not obvious. So we try other possible models. Although the distance values are predefined in the original basic-space, the distance values in the following models will be learned by machine learning given an annotated dataset.

The first model, ch_dist_2, simply considers if the PC is in the chord note (i.e., root, third, and fifth) or not. The next model, ch_dist_3, distinguishes whether the PC is diatonic or not in addition to the distinction by chord membership. The next one, ch_dist_5, uses the same categories as those of basic-space. And finally, ch_dist_10 uses the same five categories but also distinguishes major or minor. We also define a dummy model, ch_dist_0, for comparison. This one always returns 0 regardless of what input is given. Table 2 shows the chroma distance models defined above.

[2] DE 8.1. This one achieved 86.25% accuracy with 686 parameters.

	PC classification	params
basic-space	root/third/fifth/diatonic/other	0
ch_dist_0	-	0
ch_dist_2	chord/other	2
ch_dist_3	chord/diatonic/other	3
ch_dist_5	root/third/fifth/diatonic/other	5
ch_dist_10	(root/third/fifth/diatonic/other) ×(major/minor)	10

Table 2. Chroma distance models. **params** is the number of learnable parameters.

3.3 From Chroma Vectors to Chord Interpretation Paths

The method explained in §3.1 receives chord names as the input, but now we modify it to receive chroma vectors. The graph structure becomes like Figure 4. The layer width becomes 24 (keys) × 7 (degrees) = 168 because all interpretations should be considered at every layer. Then every layer is duplicated to accept chroma vector inputs. Nodes in duplicated layers are connected by horizontal edges whose weights express the distances in the models introduced in §3.2.

The learnable parameters are trained to maximize the path probability of the ground truth paths. However, the formula of path probability is revised as follows because of the modifications in the interpretation graph.

$$\mathrm{P}(X_{0:s} = x_{0:s} | c_{0:s}, G_{0:2s}) \triangleq \begin{cases} 1 & \text{if } s = 0\,^{3} \\ (\prod_{t=0}^{s-1} \frac{\exp(-\mathrm{CD}(c_t, x_t) + \mathrm{GTPS}(x_t, x_{t+1}))}{Z_{c,G,t}}) & \\ \times \mathrm{CD}(c_s, x_s) / Z^{(2)}_{c,G,s} & \text{otherwise} \end{cases} \tag{4}$$

where

$$Z_{c,G,t} \triangleq \sum_{l \in G_t} \sum_{m \in G_{t+1}} \mathrm{P}(X_t = l | G_{0:2t-1}) \times \exp(-\mathrm{CD}(c_t, l) + \mathrm{GTPS}(l, m)),$$

$$Z^{(2)}_{c,G,t} \triangleq \sum_{l \in G_t} \mathrm{P}(X_t = l | G_{0:2t-1}) \exp(-\mathrm{CD}(c_t, l)),$$

x_t is a chord interpretation at t, c_t is a chroma vector at t, CD is a chroma distance model, and GTPS is a generalized TPS proposed in [15]. This formula is designed to convert a path distance (i.e., $\sum_{t=0}^{s-1}(\mathrm{CD}(c_t, x_t) + \mathrm{GTPS}(x_t, x_{t+1}))$) to a path probability so that the shorter (shortest) path has higher (highest) probability.

The new graph (Figure 4) is a lot more complex than the original graph (Figure 2). But, all layers (except for the start and end layers) have the same set of nodes so all edges are the same too. Especially, even if a set of fully-connect edges become $168 \times 168 = 28,224$ (originally it was $6 \times 6 = 36$), it can be utilized repetitively. Moreover, the edges of chroma distances (i.e., the horizontal edges

[3] 0th layer contains only one node, that is, the start node

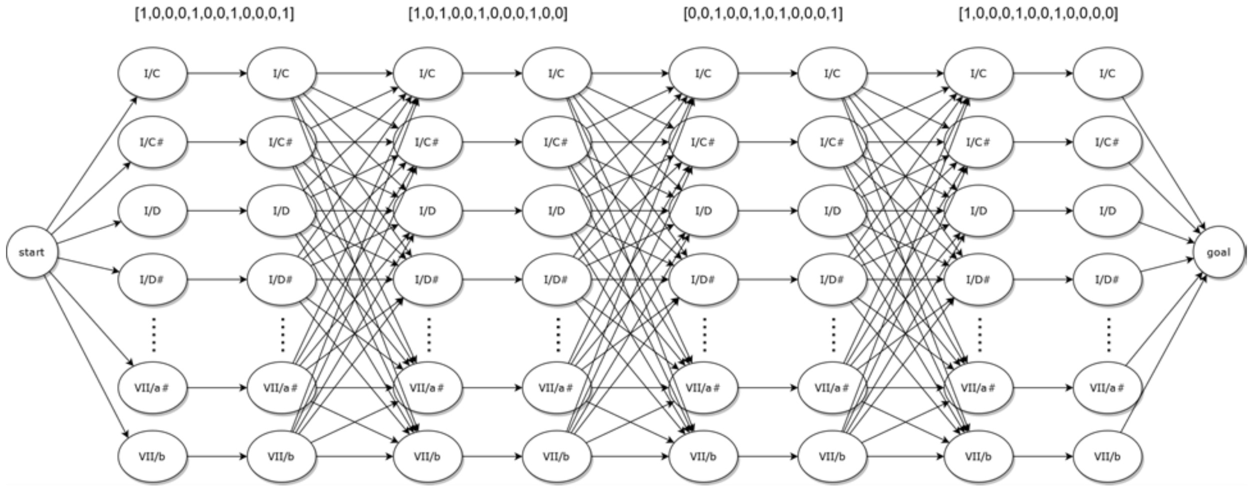


Figure 4. Revised interpretation graph.

below chroma vectors) can be calculated by a matrix product. Therefore, the computational cost does not increase as it looks.

4. EXPERIMENTS

4.1 Dataset

We use the dataset annotated in `rntxt` format [13], published at [4]. There are 384 pieces (1,905 phrases, 68,463 chords) and we regard every phrase as a unit (i.e., to which we predict the interpretation sequences) but when a phrase exceeds 50 chords we divide it into units each of which does not exceed 50 chords. Then we use 80% for training, 10% for validation, and the remaining 10% for the test.

We extracted key, degree, and applied chord information from `rntxt`, then omit all repetitions of the same chord interpretations. About applied chords, a tonic chord is added at the end of every local key section to express pivot chord modulation. Chroma vectors are obtained from `rntxt` using music21 library [2].

We set all the initial parameter values to be zero and train them by mini-batch stochastic gradient descent with batch size=100 and learning rate=0.001. We continue training at least 10 epochs and until no accuracy update in the validation set for an epoch [4] then pick the parameter which gives the highest validation accuracy.

4.2 Results

Table 3 shows the performance of, and Tables 4, 5, 6, and 7 shows the resulting PC-level distance values of the chroma distance models defined in §3.2, and Table 8 illustrates distance values between chroma vector [1, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 1] and some chord interpretations by the distance models. The performance is evaluated by how frequently the found path goes through the ground truth node (i.e., chord interpretation) in the revised interpretation graph. **acc** shows the accuracy the method could estimate ground truth chord interpretation for each chroma vector. **key acc** shows the accuracy the method could estimate at the ground truth key for each chroma vector.

[4] We loosened the stopping condition because the original condition in [15] was too costly to conduct an exhaustive evaluation.

	acc	key acc
basic-space	50.21%	60.14%
ch_dist_0	3.58%	11.22%
ch_dist_2	49.30%	57.07%
ch_dist_3	49.00%	58.67%
ch_dist_5	53.52%	62.86%
ch_dist_10	55.53%	65.67%

Table 3. Performance of chroma distance models.

chord	other
0	2.4576

Table 4. Resulting PC-level distance values of ch_dist_2.

chord	scale	other
0	2.0414	2.6578

Table 5. Resulting PC-level distance values of ch_dist_3.

root	third	fifth	scale	other
0.8525	2.8191	0	3.1986	4.2753

Table 6. Resulting PC-level distance values of ch_dist_5.

	root	third	fifth	scale	other
minor	1.2041	2.5036	0	3.3633	3.6971
major	0.7070	2.6557	0.0956	3.1862	6.2754

Table 7. Resulting PC-level distance values of ch_dist_10.

Even if we used a fairly strong model (i.e., DE 8.1 from [15]), it is almost impossible to narrow down the candidates without a hint from chroma vector (i.e., ch_dist_0) [5].

[5] Having said that, 3.58% is much better than $1/168 \approx 0.60\%$

	I/C	i/e	vi/C	VII/d
basic-space	6	6	9	7
ch_dist_0	0.0000	0.0000	0.0000	0.0000
ch_dist_2 (Table 4)	2.4576	2.4576	4.9152	2.4576
ch_dist_3 (Table 5)	2.0414	2.0414	4.0828	2.6578
ch_dist_5 (Table 6)	6.8702	6.8702	9.2163	7.9469
ch_dist_10 (Table 7)	7.0710	6.6445	18.2302	9.7337

Table 8. Distance values between chroma vector [1, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 1] and some chord interpretations.

Compared to ch_dist_0, ch_dist_2 performed very well with only distinguishing the membership of chords. Also, separating diatonic PCs (ch_dist_3), relative positions in triad (ch_dist_5), and major or minor key (ch_dist_10) all contributed to improve accuracy to some extent. Moreover, the result shows that basic-space worked quite well. It went below the same category model (i.e., ch_dist_5) but outperformed fewer category models (i.e., ch_dist_2 and ch_dist_3). This result we think indicates the importance of distinguishing that five categories. The most complex model (ch_dist_10) can be thought as a combination of two revised basic-spaces for major key and minor key respectively. And it achieved the best performance.

Looking at the learned PC-level distance values, "other" category has the largest and "scale" category has second-largest values. This is consistent with basic-space (Table 1). But within the "chord" category, "root" has the smallest value in basic-space while "fifth" has the smallest in the learned values. It is surprising that giving "fifth" smaller value than "root" enables the method to find better interpretation paths. We think this needs to be investigated further.

5. CONCLUSION

In this research, we have reconstructed the theory of distance between chords, motivated by Tonal Pitch Space (TPS), and proposed a model to guess interpretations (pairs of key and degree) to a sequence of chords, represented by chroma vectors. Since chroma vectors do not refer to human-recognizable interpretation but mention only pitch classes, we can objectively compare the relation between the role of the basic-space in a key and distances in notes.

We have compared six different sets of parameters, including the raw basic-space, and proved that these stochastic models outperformed the original TPS. We have experimented music pieces upon an open database, and the data-driven distance learning improved the accuracy by five percent or so, especially when we added a distinction of major and minor keys.

Our future work includes further refinement of this stochastic TPS, adding other features such as musical *genre* or difference of age, and so on.

Acknowledgments

This research was supported by JSPS Kaken 20H04302 and 21H03572.

6. REFERENCES

[1] J. Bharucha and C. Krumhansl, "The representation of harmonic structures in music: Hierarchies of stability as a functions of context", *Cognition*, 13(1), pp.63-103, 1983.

[2] M. S. Cuthbert, C. Ariza, "music21: A toolkit for computer-aided musicology and symbolic music data", in *Proc. of the 11th International Society for Music Information Retrieval Conference (ISMIR)*, pp.637-642, 2010.

[3] L. Euler, *Tentamen novae theoriae musicae*. St. Petersburg Academy, 1739.

[4] M. Gotham, R. Kleinertz, C. Weiss, M. Müller, S. Klauk. "What if the 'When' implies the 'What'?: human harmonic analysis datasets clarify the relative role of the separate steps in automatic tonal analysis", in *Proc. of the 22nd International Society for Music Information Retrieval Conference (ISMIR)*, pp.229–236, 2021

[5] J. D. Heinichen, *General-bass in der composition*. Dresden: J. D. Heinichen, 1728.

[6] D. Kellner, *Treulicher Unterricht im General-Bass*. Hamburg: C. Herold, 1737

[7] F. Lerdahl, R. Jackendoff, *A Generative Theory of tonal music*. Cambridge, MA, 1983.

[8] F. Lerdahl, *Tonal pitch space*. New York, Oxford University Press, 2001.

[9] R. R. Randall, B. Khan, "Lerdahl's Tonal Pitch Space Model and Associated Metric Spaces", in *Journal of Mathematics and Music*, 4, pp.121-131, 2010.

[10] H. Riemann, *Grosse kompositionslehre, Vol. 1*. Berlin: W. Spemann, 1902.

[11] S. Sakamoto, S. Arn, M. Matsubara, S. Tojo, "Harmonic analysis based on tonal pitch space", in *Proc. of the 8th International Conference on Knowledge and Systems Engineering (KSE)*, pp.230-233, 2016.

[12] C. Callender, I. Quinn, D. Tymoczko, "Generalized Voice-Leading Spaces", in *Science*, 320, pp.346-348, 2008.

[13] D. Tymoczko, M. Gotham, M. S. Cuthbert, C. Ariza, "The romantext format: a flexible and standard method for representing Roman numeral analyses", in *Proc. of the 20th International Society for Music Information Retrieval Conference (ISMIR)*, pp.123-129, 2019.

[14] G. Weber, *Versuch einer geordneten theorie der tonsetzkunst*. Mainz: B. Schotts Söhne, 1821-24.

[15] H. Yamamoto, S. Tojo, "Generalized tonal pitch space with empirical training", in *Proc. of the 18th Sound and Music Computing Conference (SMC)*, 2021, pp.300-307.

HARMONIC MAPS: INTERACTIVE VISUALIZATION OF TRIAD SPACES BASED ON SPECTRAL STRUCTURES

Manuel Gaulhiac
Expressive E, IReMus
Paris, France
manuel.gaulhiac@protonmail.com

Xiao Xiao
Léonard de Vinci Pôle Universitaire
Research Center
Paris La Défense, France
xiao.xiao@devinci.fr

Jean-Marc Chouvel
Sorbonne Université, IReMus
Paris, France
jeanmarc.chouvel@free.fr

ABSTRACT

We present Harmonic Maps, a visualization of three-note chord spaces and an interactive application that allows users to explore in real-time the connection between the visualization and its mapped sounds. While typical harmonic analysis is based only on notes or on an audio signal, our analysis takes a hybrid approach by quantifying different types of interactions between the spectra of notes. These quantifications, which we call Harmonic Descriptors, are derived from acoustic or perceptual models. Three such descriptors are defined and mapped: concordance, third order concordance and roughness.

Harmonic analysis based on spectral structures opens new possibilities beyond traditional note-only or signal-only approaches. They can be applied to a continuum of frequencies, independent of the tuning system, as well as historical and stylistic constraints. Harmonic Maps based on spectral structures can be especially relevant to study the relationship between timbre and harmony. Our interactive exploration of harmonic spaces can have applications for analytical, compositional and educational purposes.

1. INTRODUCTION

A chord is a simultaneous combination of notes, and 3-note chords (triads) are considered to be the basic building block in tonal harmony [1]. A common approach to chord analysis is functional harmony, which defines a role for each chord in a sequence (e.g. tonic, dominant), but its use is limited to chords within a musical context. To analyze chords independent of musical context, three main approaches have been taken, each considering the chord in a different way. One, which we call the combinatorial approach, considers chords as abstract objects and defines ways of classifying them using mathematical formalisms [2]. A contrasting approach, audio descriptors, considers a chord as a single audio signal, and uses signal processing techniques directly on the audio without considering the abstraction of notes [3]. Finally, considering a chord as a sensation in the listener, perceptive models have been created through studies with human subjects [4, 5, 6].

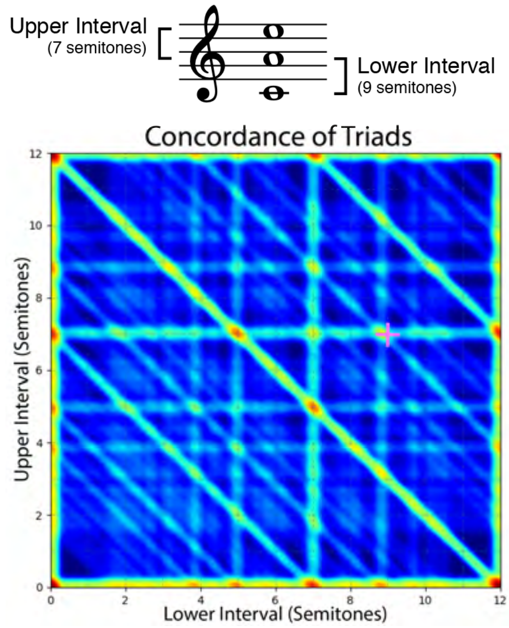


Figure 1. A triad and its equal-temperament position in a harmonic map for concordance for a synthesized sound with 11 partials. See Section 3.2 for the definition of concordance and section Section 4 for the implementation of Harmonic Maps.

Our work examines chords independent of harmonic context but differs from previous approaches in its hybrid nature that unites physical and symbolic approaches. We begin by computing the spectrum of each note, and analyze harmony based on the interactions between the spectra of different notes, with computational methods derived from acoustic and perceptual models. Our analyses are directly interpretable based on symbolic notions derived from the score, such as notes, intervals, and chords.

This article begins with a historic overview of interval, dyad and triad classifications (Section 2). We then present the notion of harmonic descriptors (Section 3), which are different ways of quantifying the spectral interactions between notes. Three harmonic descriptors are discussed, derived from acoustic and perceptual models: concordance, third order concordance, and roughness. Building on the notion of harmonic descriptors, we then focus on a specific type of visualization, the Harmonic Map, taking advantage of the fact that all possible three note chords based on a

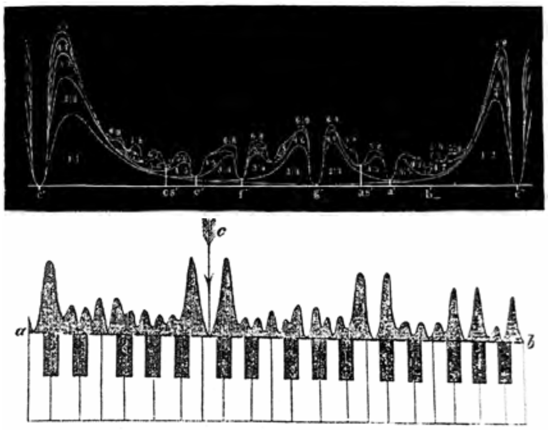


Figure 2. Description of intervals. (Top) Helmholtz's historical roughness curves, from [8]. (Bottom) Mach's vulgarisation of Helmholtz's curves, from [9].

chosen root note can be represented in a 2D plane, with the X and Y dimensions showing the lower and upper intervals. Assumptions and constraints of Harmonic Maps are discussed (Section 4), followed by the design and implementation of an interactive application enabling users to explore the connection between a visualizations and its sounds (Section 5). Finally, to illustrate the usefulness of our representation, Section 6 presents examples that shed light on topics in music theory.

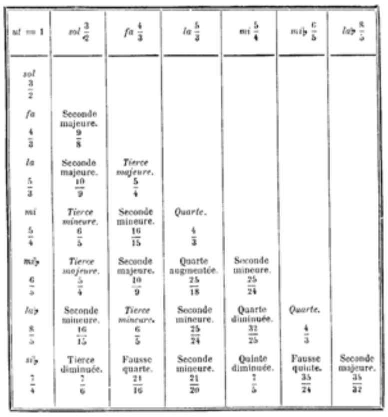
2. BACKGROUND

To situate the analysis of triads, we begin with the classification of intervals. An interval is defined as the set of dyads (two-note chords) whose two notes have the same logarithmic distance on the frequency axis. The classification of intervals is not new, and since the advent of polyphony in Western music, intervals have been classified into two categories: consonant and dissonant, whose meanings and criteria have evolved over history [7].

Historically, intervals were categorized based on discrete frequencies [7]. It is only with Helmholtz that an organization of intervals as a continuum was envisaged, via a psychoacoustic characterization, that links acoustic properties of sounds with their perception [8]. For example, the principle of "roughness" was developed through experiments where Helmholtz listened to different combinations of frequencies and counted the number of "beats" that arose in the frequency interactions. Figure 2 shows Helmholtz roughness curves, where the continuum of frequencies are shown on one axis, above which are the roughness values of the intervals for sounds containing different numbers of partials.

Three years after the publication of Helmholtz's seminal *Theory of Music*, Ernst Mach wrote a text popularizing Helmholtz's theory [9], with a new representation of Helmholt's roughness curve over two octaves, aligned with a piano keyboard with keys centered on their frequency (Figure 2). This representation shows the correspondance between the discrete and the continuous approach.

Despite his innovation in the description of intervals in continuous space, Helmholtz took a reductionist and categorical approach for the study of triads. The reductionist approach considers the interval as the basic building block of harmony, the study of which, both perceptual and theoretical, would suffice to understand higher cardinality chords [10]. Figure 3 shows a table of triads organized according to two constituent intervals: the lower interval and the one between the two extreme notes. The third interval is indicated in the table, and allows Helmholtz to determine a property of the triads, namely its consonance. Mach, in turn, takes the chord chart from Helmholtz by mapping it to the keys of the keyboard [9], with a code to indicate consonant intervals and chords (Figure 3).



ut = 1	sol $\frac{3}{2}$	fa $\frac{4}{3}$	la $\frac{5}{3}$	mi $\frac{5}{4}$	mi♭ $\frac{6}{5}$	la♭ $\frac{8}{5}$
sol $\frac{3}{2}$						
fa $\frac{4}{3}$	Seconde majeure. $\frac{9}{8}$					
la $\frac{5}{3}$	Seconde majeure. $\frac{10}{9}$	*Tierce majeure.* $\frac{5}{4}$				
mi $\frac{5}{4}$	*Tierce mineure.* $\frac{6}{5}$	Seconde mineure. $\frac{16}{15}$	*Quarte.* $\frac{4}{3}$			
mi♭ $\frac{6}{5}$	*Tierce majeure.* $\frac{5}{4}$	Seconde majeure. $\frac{10}{9}$	Quarte augmentée. $\frac{25}{18}$	Seconde mineure. $\frac{25}{24}$		
la♭ $\frac{8}{5}$	Seconde mineure. $\frac{16}{15}$	*Tierce mineure.* $\frac{6}{5}$	Seconde mineure. $\frac{25}{24}$	Quarte diminuée. $\frac{32}{25}$	*Quarte.* $\frac{4}{3}$	
si♭ $\frac{7}{4}$	Tierce diminuée. $\frac{7}{6}$	Fausse quarte. $\frac{21}{16}$	Seconde mineure. $\frac{21}{20}$	Quinte diminuée. $\frac{7}{5}$	Fausse quinte. $\frac{35}{24}$	Seconde majeure. $\frac{35}{32}$

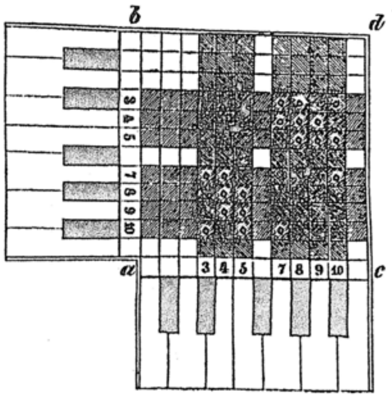


Figure 3. Classification of triads (Top) Helmholtz's table of triads. (Bottom) Mach's vulgarisation of Helmholtz's curves, from [9].

From Mach's representation to a continuous representation of triads, there is only one step, which consists in reducing the size of note subdivisions by taking the limit. This step was only taken at the end of the 20th century, in the works of Chouvel and Sethares (Figure 4) [11, 12]. Chouvel takes the upper and lower intervals as axes, and uses level lines to represent concordance, a physical quantity that measures the energy of interaction. Sethares, meanwhile, uses the same axes as Helmholtz, with a 3D representation of roughness. Both Chouvel and Sethares base their work on theorical spectra and do not use real sounds.

The description of triads based on their constitutive intervals remains the majority approach in the psychoacoustic literature since Helmholtz, in spite of some proposals, to use the expression of Cook, that advocate a "psy-

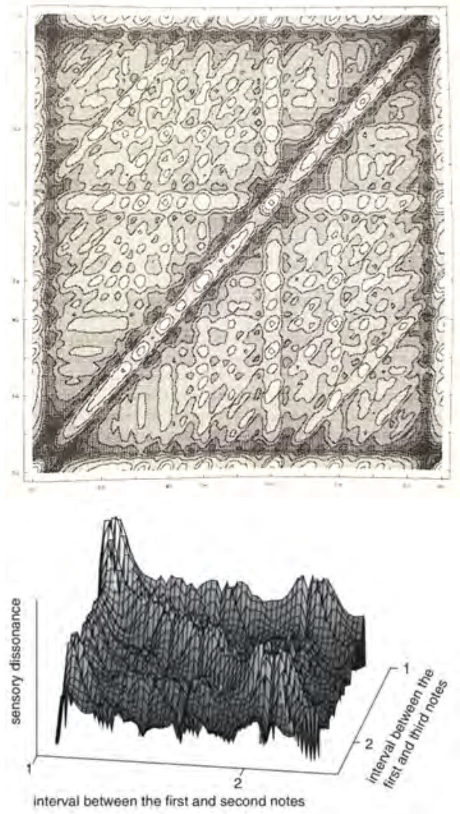


Figure 4. Continuous description of triads (Top) Chouvel, concordance map, from [11]. (Bottom) Sethares, 3D representation of sensory dissonance, from [12].

chophysics with 3 notes" [10]. Several models, however, exhibit acoustic and psychoacoustic properties unique to triads. This is the case of harmonicity [13], which measures the degree of similarity of a spectrum to a harmonic spectrum, of Cook's tension [14], which takes into account the differences between the upper and lower intervals, and of Gaulhiac's third order concordance [15], which quantifies the energy shared in the frequency ranges common to the three notes.

3. HARMONIC DESCRIPTORS

Harmonic descriptors are a set of notions and corresponding computation methods that aim to reconcile signal processing tools and perceptual models with a musicological understanding of music. The incorporation of signal processing tools that make calculations directly on acoustic reality allows us to take into account the timbre of different instruments, the articulation of playing, room acoustics, and to perform analysis on synthesized sounds. However, analysis based purely in signal-processing, such as audio descriptors can be difficult to interpret musically because they are often computed on a time-scale smaller than the duration of notes [3].

The novelty of harmonic descriptors is the association of each note with a spectrum, retaining a note-level description common in musicology while adding a signal-based layer. Each harmonic descriptor describes a type of interaction of a spectral nature between simultaneous notes. They are therefore ideal for the study of chord spaces, such as the triad space. The following sections first describe the general framework of how harmonic descriptors and how the spectra of a note is computed. We then present 3 specific descriptors: concordance, third order concordance and roughness.

3.1 Implementation & Spectra Computation

The implementation of our harmonic descriptors differs from the usual implementation of audio descriptors in that they are not calculated on the analysis windows but on the temporal duration of the chord. The values of the models are then associated with a chord, which can be more easily linked with a musical score. This paradigm shift makes the results readable and usable by musicologists.

Various ways of computing harmonic descriptors have been proposed, depending on the purpose of the analysis [15, 16]. For the purpose of computing our Harmonic Map of triads, each harmonic descriptor takes in three spectra, each associated with a note. A spectrum is computed from an initial reference sound, the sound of a note representative of the timbre that we wish to study. For the moment, the three spectra are generated from the same reference sound. This sound is projected in a time-frequency space by a constant-Q transform (CQT) [17]. This transform has the advantage, compared to the short-term Fourier transform, of preserving the same frequency precision on the logarithmic axis of the frequencies, to the detriment of the temporal precision in the bass register. Here we set a frequency precision of 1/32 of a tone for the analysis, i.e. each tone is divided into 32 equal frequency bands. The next step is to associate this sound with an "average spectrum" in amplitude, obtained by temporally averaging the amplitudes, frequency range by frequency range, over its entire duration. By carrying out such an averaging, the initial sound is approximated by a stationary sound, which puts aside the temporal monitoring of the spectral components' decrease. On the other hand, associating an amplitude spectrum with a note creates a bridge between the signal and symbolic notions, which makes it possible to calculate spectral models on chords. For the calculation on a triad, each note is associated with a spectrum, obtained by translation of the "average spectrum," and the harmonic descriptors are calculated on these spectra.

It is possible to imagine a simpler spectra computation method whereby the Fourier transform is taken on the initial sound without any windowing. However, our method was designed to ensure a constant precision in the frequency-logarithmic domain.

3.2 Concordance

The concordance between two stationary sounds is defined as the scalar product of their amplitude spectra. It quantifies the spectral components common to these two sounds. In the signal vector formalism, concordance is identified with the Hilbertian product in the frequency domain, in the particular case of real spectra. Bonnet interprets it as an energy of interaction between two sounds [18].

In all generality, for two complex spectra X_1 et X_2, there is the relation:

$$\|X^{(1)}+X^{(2)}\|^2 = \|X^{(1)}\|^2+\|X^{(2)}\|^2+2\Re\langle X^{(1)}, X^{(2)}\rangle, \quad (1)$$

where $\Re$ denotes the real part.

The term "harmonic concordance" was introduced by Chouvel, who uses it as a guide to composition, primarily in microtonal spaces [11]. The concordance $\mathcal{C}(X^{(1)}, X^{(2)})$ of two notes with amplitude spectra $X^{(1)}$ and $X^{(2)}$ is expressed as the scalar product of the spectra:

$$\mathcal{C}(X^{(1)}, X^{(2)}) = \langle X^{(1)}, X^{(2)}\rangle = \sum_k X_k^{(1)} X_k^{(2)}, \quad (2)$$

where $X_k^{(i)}$ is the k^{th} bin amplitude of $X^{(i)}$ and k goes through the frequency bins.

From there, the concordance $\mathcal{C}$ of a triad with notes' amplitude spectra $X^{(1)}$, $X^{(2)}$ and $X^{(3)}$ is defined as the sum of the concordances of all the pairs of notes:

$$\mathcal{C} = \mathcal{C}(X^{(1)}, X^{(2)}) + \mathcal{C}(X^{(1)}, X^{(3)}) + \mathcal{C}(X^{(2)}, X^{(3)}). \quad (3)$$

3.3 Third Order Concordance

The concordance of a triad measures the spectral components common to two notes of the triad. Therefore, it is natural to want to measure the spectral components common to the three notes simultaneously. This is the role of Gaulhiac's third-order concordance [15]. Third-order concordance is more restrictive than concordance, since it requires the simultaneous coincidence, and no longer two-by-two, of the partials resulting from the three notes. The term "third order" or "order 3" is added to refers to the simultaneous coincidences of the three notes, and to distinguish it from the notion of previously defined notion of concordance, which is of order 2.

The third order concordance of a triad with notes' amplitude spectra $X^{(1)}$, $X^{(2)}$ and $X^{(3)}$ is defined as follows:

$$\begin{aligned}\mathcal{C}_3 &= \langle X^{(1)}, X^{(2)}, X^{(3)}\rangle \quad (4)\\ &= \sum_k X_k^{(1)} X_k^{(2)} X_k^{(3)}.\end{aligned}$$

3.4 Roughness

The roughness of a sound is the sensation due to the beats between partials that it comprises. When two sinusoidal sounds are played simultaneously, they interact to form frequency beats equal to the difference in the frequencies of the initial sounds. When the frequencies are close enough, the beats are clearly audible, but when they move a little apart, they become too fast to remain audible but still create a feeling of harshness, called roughness. Moving further away, the frequencies are perceived as distinct, and the feeling of roughness disappears.

Research in psychoacoustics has shown that roughness is one of the causes in the judgment of dissonance. If Sauveur already explained dissonance by beats, Helmholtz, to whom we owe the term "roughness," is the first to experimentally study the link between the two, with a psychoacoustic approach that revolutionizes the theory of consonance and dissonance [8]. A century later, based on perceptual experiments, Plomp and Levelt relate roughness to the notion of the critical band, which gives an order of magnitude of the frequency difference below which two sinusoidal sounds are close enough to interact [4]. They propose a dissonance curve for simple sounds where the maximum dissonance corresponds to a frequency difference equal to a quarter of the critical band.

The results of Plomp and Levelt mark the beginning of a long series of models refining that of Helmholtz. We use a model largely inspired by that of Sethares [12], which is based on the curve of Plomp and Levelt.

The roughness $\mathcal{R}(\alpha_1, \alpha_2, f_1, f_2)$ of two sinusoidal sounds of frequency f_1 and f_2 and amplitude α_1 and α_2, with $f_1 \leq f_2$, is expressed as :

$$\begin{aligned}\mathcal{R}(\alpha_1, \alpha_2, f_1, f_2) = \alpha_1\alpha_2 \Big(&e^{-b_1(f_1-f_2)s(f_1)} \quad (5)\\ &-e^{-b_2(f_1-f_2)s(f_1)}\Big),\end{aligned}$$

with

$$s(f_1) = \frac{x^*}{s_1 f_1 + s_2}. \quad (6)$$

The constants b_1 and b_2 control the shape of the roughness curve (location of the roughness maximum and decay after the maximum), while s_1 and s_2 are linked with the variation of the critical band as a function of the lower frequency. We fix $b_1 = 3.5$, $b_2 = 5.75$, $s_1 = 0.021$ and $s_2 = 19$. The location of the maximum roughness in proportion to the critical band is indicated by $x^* = 0.24$.

From this, the roughness $\mathcal{R}$ of a triad with notes' amplitude spectra $X^{(1)}$, $X^{(2)}$ and $X^{(3)}$ is obtained by summing the roughness of all combinations of pairs of bins from different notes :

$$\mathcal{R} = \sum_{i<j}\sum_{k,l} \mathcal{R}(X_k^{(i)}, X_l^{(j)}, f_k, f_l), \quad (7)$$

where $X_k^{(i)}$ is the k^{th} bin amplitude of $X^{(i)}$, f_k the central frequency of bin k, and k and l go independently along the frequency axis.

4. FROM HARMONIC DESCRIPTORS TO HARMONIC MAPS

Harmonic Maps are a way of visualizing the values of a harmonic descriptor for triads, providing an acoustic or psychoacoustic description of the chord space depending on the descriptor used. A Harmonic Map is derived from a single reference sound, from which the values of a chosen descriptor are calculated for all chords with the reference note as the root (lowest) note. Input sounds with little noise and minimal reverberation are preferred. Values of the descriptor can be represented either with the addition of a third dimension or through colors (Figure 1 and 5). While harmonic descriptors can be calculated between intervals of any size, we limit the lower and upper intervals to one octave in the examples we present. The values of our descriptors are normalized between 0 and 1. Our implementation is based on several assumptions that pose constraints on the type of sounds used, to ensure that the Harmonic Maps we generate reflect an acoustic reality.

4.1 Stability of Sounds

The temporal average of the spectrum is representative of the acoustic reality of the basic sound only if the latter has a certain temporal stability. The model instrument is the organ, with its stable and sustained sounds, which offers a wide variety of timbres through its stops.

Temporal instability can arise from the decrease in the partials for sounds that are not maintained, from the attack phases, or from playing techniques such as tremolo or vibrato. When using sounds without temporal stability, the instability of the sound must be taken into account in the interpretation of the Harmonic Maps.

For example, to include the vibrato effect, one must ensure that the starting sound is long enough to contain enough periods of the vibrato. For sounds with a prominent attack, as for the piano, one can eliminate the transients of the attack by processing the type separation of the harmonic part and the percussive part upstream [19]. For a sound that is not sustained, such as the bell (Figure 7), the generated spectra will not reflect the decay in the sound, and the resulting map may be less reflective of acoustic reality. To evaluate the degree of relevance of a Harmonic Map, one can re-synthesize the averaged sound from the generated averaged spectrum. By comparing the averaged sound with the original sound, it is possible to judge by ear its degree of artificiality.

4.2 Timbral Considerations

The implementation of a Harmonic Map takes only one sound as input, and from it generates the spectrum which serves as a model for all notes. However, on acoustic instruments, the shape of the spectrum varies with pitch, and the visualization may be less accurate for notes further away from the reference note. This is another reason we limit our maps to the range of one octave.

Each Harmonic Map is currently limited to a single timbre, where a timbre is associated with a mean amplitude spectrum. A single Harmonic Map cannot take into account chords whose notes are played by different instruments, as can occur in an instrumental ensemble.

5. INTERACTIVE HARMONIC MAPS

We developed an application that enables users explore the connection between a Harmonic Map and the sounds that it describes. Our application supports several interfaces, including the mouse, the stylus of a graphic tablet, and MIDI or MPE (Midi Polyphony Expression) controllers. With a mouse or stylus, a user can hear what sounds correspond to different regions of the map by clicking on points or drawing a continuous trajectory (See figure 6). With a MIDI/MPE controller such as a keyboard, a user can play

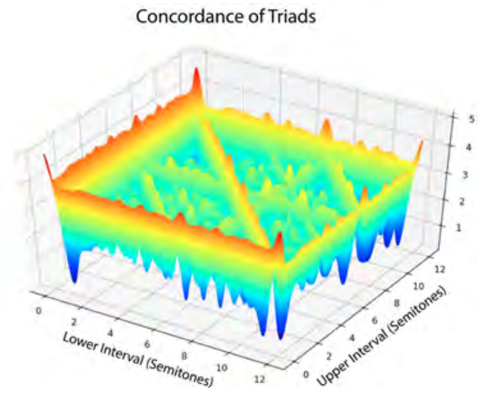


Figure 5. Harmonic Maps can also be visualized in 3D. Here is the 3D version of the concordance map shown in Figure 1, where both the color and height are determined by the concordance. 3D Harmonic Maps can also be used to visualize two descriptors at once.

a chord or sequence of chords to see their corresponding location on the Harmonic Map.

Designed as an interactive tool for exploring a harmonic space, our application can be particularly useful for the composer or the musicologist who works with harmonic material for which there is not yet well-established theories, such as microtonal scales or inharmonic sounds. In addition, the comparison of the different structures associated with different timbres makes it possible to study and explain the timbre-harmony relationship, taking into account, for example, the influence of the instrument, the nuances, the modes of play.

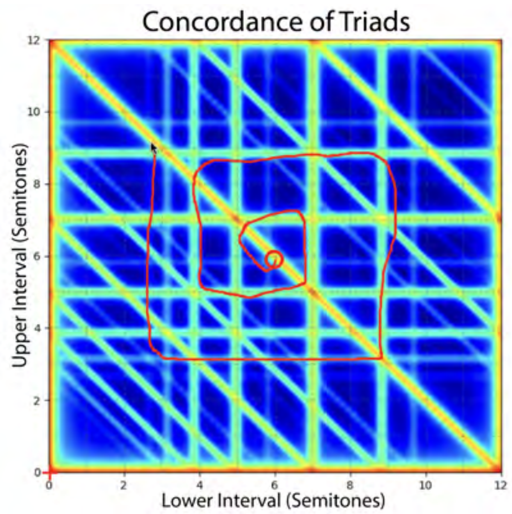


Figure 6. A trajectory on a Harmonic Map

5.1 Implementation

The Harmonic Maps application consists of two parts. First, a Python script manages the calculation of the descriptors and generates images of Harmonic Maps from an input sound. A Max patch then manages the real-time interaction, displaying a selected map and synthesizing sounds based on information from input devices. For the moment, the calculation of the descriptors and the generation of Harmonic Maps is not done in real time. The Python script for a pre-determined set of input sounds is run prior to running the Max patch, which then allows the user to select from a list of pre-computed maps.

For sound synthesis, we use the collection of sampled instruments from the Ircam Solo Instruments 2 library [1], due to the many playing techniques available and its support of the MPE (MIDI Polyphonic Expression) protocol. For each timbre considered, a sound-model of a few seconds of a note is extracted, then injected into the Python script to build the associated Harmonic Maps.

5.2 MPE Control & Harmonic Trajectories

While traditional MIDI keyboards can be used with the Harmonic Maps application, they do not allow access to all regions of a map because the classic pitch-bend wheel transposes the entire keyboard. In order to control the harmonic trajectories by a MIDI controller, it is necessary to have an instrument supporting the MPE protocol. MPE keyboards such as the Roli Seaboard [2] or the Haken Continuum Fingerboard [3] enable each note to be modulated independent of other notes, allowing all points on a Harmonic Map to be visited. We tested the harmonic trajectory feature with a the Osmose piano from Expresse E [4], in which each key can move laterally to modulate its pitch.

6. EXAMPLES

We describe a several examples of Harmonic Maps to show how they can inform the understanding of sound material. The first three examples (Sections 6.1-6.4) focus on concordance, and show the influence from the numbers of partials, timbre, playing mode, and degree of harmonicity. We then show a roughness map example (Section 6.5) and present a theoretical application of third order concordance concerning the minor chord (Section 6.6). All the maps presented below use middle C (261 Hz) as their input and show the space of triads with this note as the root.

6.1 Influence of the Number of Partials

The spectral structures of triad spaces depend on the spectral properties of the considered sounds, and first of all on the number of partials. Figure 10 shows concordance maps with harmonic sounds composed of 2, 3 and 5 partials respectively, as well as their associated spectra. A sound is said to be harmonic when these partials are multiples of the fundamental frequency. The sounds were generated by additive synthesis using the FAUST language [5]. A decrease of $1/\sqrt{k}$ in the amplitude of the partials was applied, where k is thte order of the partial.

The color code is as follows: cold colors correspond to low values of concordance, warm colors to high values, with red representing the maximum concordance values. We are interested in the positions of the local maxima, formed by segments and the intersections between these segments. Each segment corresponds to a coincidence between partials from different notes. The structure becomes more complex with more partials, since the number of interactions increases. With two partials separated by an octave, the triads highlighted are those containing unison or octave intervals, located on the periphery of the map, as well as those whose interval between the bass and the highest note high is the octave. Adding the third partial, an octave and a fifth above the root, emphasizes chords with a fifth or an octave plus a fifth. The addition of partial number 5, located two octaves and a major third above the fundamental, reveals the major chords (fundamental and tight position of coordinates 4-3) and minor (3-4). The highlighted chords are in just intonation, i.e. formed from the intervals naturally present in harmonic sounds. This is why the highlighted major chord has a major third slightly less than 4 semitones. Note the symmetry of the maps according to the large diagonal starting from the origin, a property of concordance, which does not distinguish major and minor chords in root position with closed voicing (within an octave).

6.2 Influence of Timbre

Figures 11 show the concordance maps of three different organ stops: the vox humana, the unda maris, and the tutti. The sounds come from UVI's Orchestral Suite library [6]. Differences in timbre are reflected in the harmonic structures.

The sound of unda maris has few partials compared to the others, so its concordance map structure is simpler. The set of reed pipes tuned slightly higher than the others has the particularity of producing slight beats, which are reflected on the map by the greater width of the segments. The structure is more complex for the vox humana. That of the tutti is essentially distinguished by more contrast, the local maxima of concordance being more marked.

From the point of view of the organist or the composer, each stop calls for a specific writing, and we do not play the same types of chords with the vox humana as with the tutti. Harmonic maps are a tool to study and deepen the understanding of the relationship between timbre and harmony.

6.3 Influence of Dynamics & Playinng Techniques

Different playing techniques on an instrument has an effect on the timbre and therefore on the harmonic structures. Figure 12 shows the concordance maps of the cello with differences in dynamics and vibrato: first piano playing without vibrato, then piano with vibrato, finally forte with vibrato. The sounds come from the Ircam Solo Instruments 2 library.

As with the unda maris, the vibrato has the effect of widening the segments. It also widens the partials on the spectrum. The strong nuance has the effect of enriching the

[1] https://www.uvi.net/ircam-solo-instruments-2
[2] https://roli.com/
[3] https://www.hakenaudio.com/
[4] https://www.expressivee.com/
[5] https://faustdoc.grame.fr/
[6] https://www.uvi.net/en/orchestral/orchestral-suite.html

spectral content by increasing the number of partials, which results in an enrichment of the map.

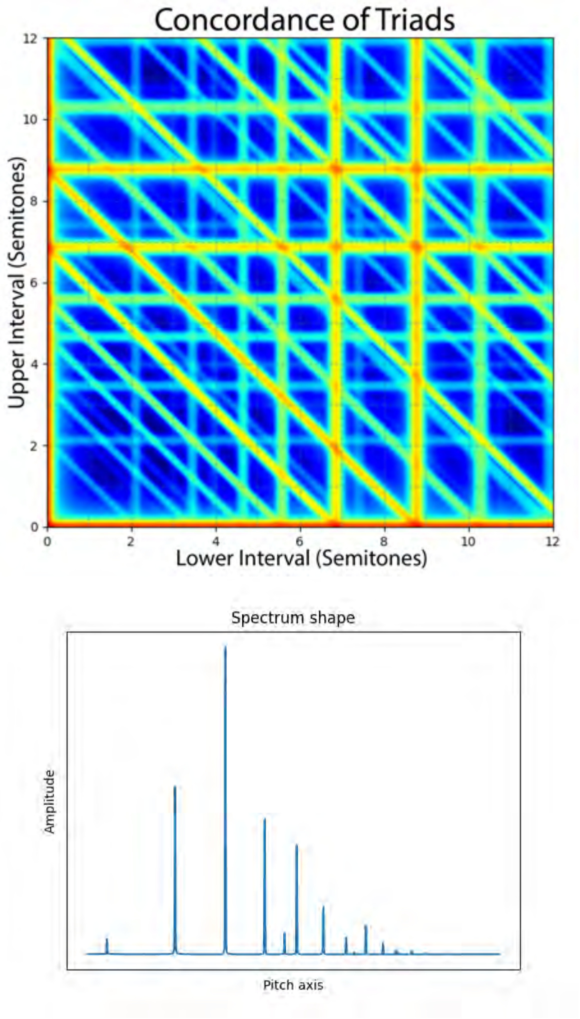


Figure 7. Concordance map for the tubular bell (Top), an inharmonic and non-sustained sound, and spectrum of this sound (Bottom).

6.4 Influence of Harmonicity

The spectra studied so far are harmonic spectra, but Harmonic Maps can also be constructed from inharmonic sounds, i.e. sounds whose partials are not multiples of the same fundamental frequency, such as from bells or the xylophone. Even in the case of very inharmonic sounds, for which it is complicated to perceive or define a fundamental frequency, the method remains applicable, since the spectra of the notes are obtained by the translation of the reference spectrum.

Figure 7 shows the concordance map of a tubular bell, an instrument used in the orchestra. The sound comes from UVI's Orchestral Suite library. The structure is very different from the previous structures. The absence of a segment marked on the edges of the map at the top and on the right shows that the instrument is not octaviant, in other words the spectrum does not include an octave interval on the first partials. The large over-diagonal shows that the octave equivalent for this sound is slightly larger than the 12 semitone interval. Likewise, the fifth is slightly lower than the interval of 7 semitones.

The sound exploration of this map by the Harmonic Maps application seems particularly relevant to guide the ear. The major chord on the first C-Eb-Ab inversion, with coordinates (3-5), is located on a blue part, that is to say a trough of concordance, and sounds very hard to the ears of authors, while the C-D-A chord, with coordinates (2-7), is close to a local maximum and has a softer sonority.

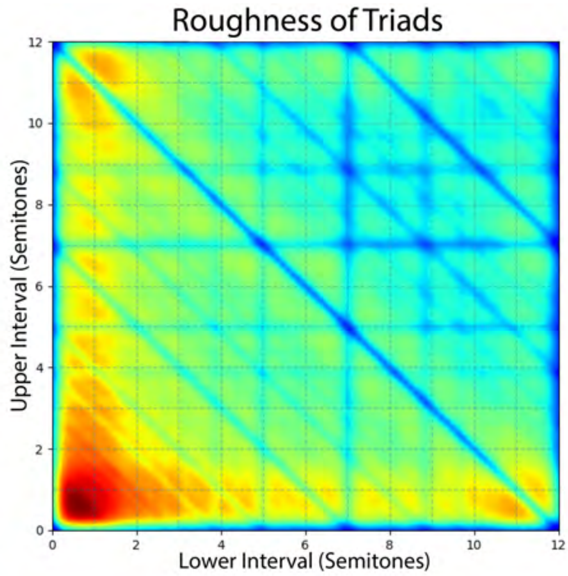


Figure 8. Roughness harmonic map for a synthesized sound with 11 partials.

6.5 Roughness

Roughness is a harmonic descriptor that reflect perceptive models, in contrast to concordance, a purely acoustic descriptor, which has not been the subject of perceptual studies.

Figure 8 shows the roughness map of a synthetic sound with 11 partials, with a very different structure from the concordance maps. The red zone on the triads formed by small intervals corresponds to a maximum of roughness. Roughness distinguishes major and minor chords in extended position C-G-E, with coordinates (7-9), and C-A-E, with coordinates (9-7), which correspond to local minima, but does not distinguish major and minor chords in tight voicings.

6.6 Third Order Concordance

Concordance and roughness are reductionist models, in the sense that they reduce the study of triads to the study of their constitutive intervals. The symmetry of concordance and roughness maps comes from this, since two symmetrical triads along the long diagonal have the same set of constituent intervals (the C-E-G and C-Eb-G triads are both made up of a fifth, a major third and a minor third). Third order concordance differs from concordance and roughness in that it is not reductionistic, as shown by the lack

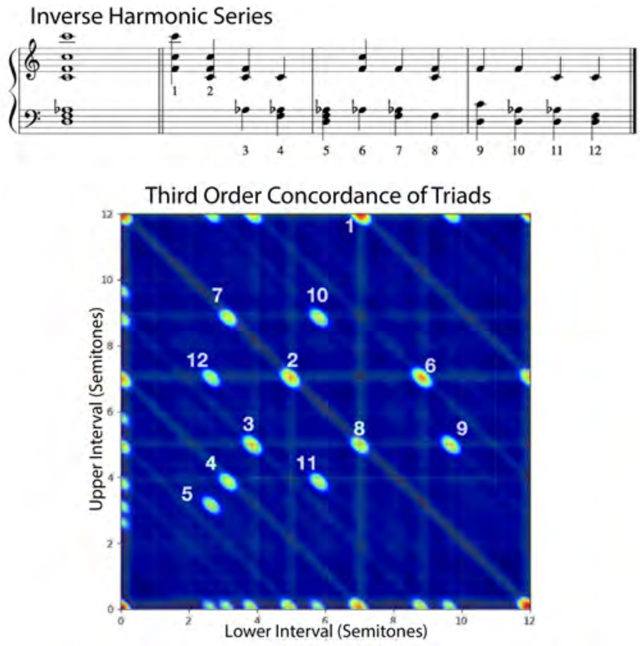


Figure 9. Example of a third order concordance map for a synthesized sound with 7 partials (Bottom) with a corresponding inverse harmonic series (Top). The numbered chords in the series are shown by their numbers in the map.

of such symmetry in the third order concordance map of a synthesized sound with 7 partials figure 9.

The local maxima of the third order concordance correspond to the triads for which there is a simultaneous coincidence of partials resulting from the three notes. The frequencies of the fundamental notes then divide this common frequency. In other wards, the triad is contained in the inverse harmonic series resulting from this common frequency, where the inverse harmonic series is defined as the series of frequencies which divide a given frequency. An inverse harmonic series, which does not have the physical reality of the harmonic series, is shown in figure 9, with the first 12 chords from this series. Reduced to the same lower note, they are projected on the map, and correspond exactly to the local maxima. Just as the harmonic series contains the major chords, the inverse harmonic series contains the minor chords, which are effectively identified by the third order concordance, while the minor chords are not.

7. CONCLUSIONS & FUTURE WORK

Harmonic Maps is a new way to visualize the spectral structure of triads in a continuous space, based on real sounds, which maintains the link with symbolic notions of notes and chords. Such a hybrid approach takes advantage of signal processing methods but makes their results more applicable for musicological purposes. Musicological subjects on which Harmonic Maps can shed light include the relationship between timbre and harmony, microtonality, and the study of non-harmonic sounds. We presented an interface for exploring the connection between Harmonic Map visualizations and the sounds they depict, allowing users to hear sounds by moving around on a map with a mouse or stylus, or by playing notes on a keyboard and seeing their trajectory on the map. Our visualization contributes to the literature of music theory and analysis, and the interactive interface makes a contribution for compositional and educational contexts.

In terms of future work, we plan to add a drag and drop feature, where new sound recordings can be uploaded to generate new maps without leaving the Max patch. A more ambitious future feature is the ability to generate maps based on real-time input sounds, which will involve optimizing our harmonic descriptor calculations to run in real-time. We can also imagine Harmonic Maps that can accomodate sounds played by different instruments.

Acknowledgments

We grateful to Sorbonne Université and IReMus for their support. Many thanks for Christophe D'Alessandro, Julio Estrada and Xavier Hascher for their feedback and encouragemet.

8. REFERENCES

[1] H. Riemann, *L'harmonie simplifiée ou Théorie des fonctions tonales des accords*. Augener, 1890.

[2] A. Forte, *The structure of atonal music*. Yale university press, 1973.

[3] G. Peeters, B. L. Giordano, P. Susini, N. Misdariis, and S. McAdams, "The timbre toolbox: Extracting audio descriptors from musical signals," vol. 130, no. 5, pp. 2902–2916, 2011.

[4] R. Plomp and W. J. M. Levelt, "Tonal consonance and critical bandwidth," vol. 38, no. 4, pp. 548–560, 1965.

[5] I. Lahdelma, J. Armitage, and T. Eerola, "Affective priming with musical chords is influenced by pitch numerosity," 2020.

[6] P. Harrison and M. T. Pearce, "Simultaneous consonance in music perception and composition." vol. 127, no. 2, p. 216, 2020.

[7] J. Tenney, "A history of consonance and dissonance," vol. 13, no. 3, p. 94, 1988.

[8] H. v. Helmholtz, *Théorie physiologique de la musique*. J. Gabay, 1863.

[9] P. Bailhache, A. Soulez, C. Vautrin, H. v. Helmholtz, E. Mach, and C. Dahlhaus, "Helmholtz, du son à la musique," 2011.

[10] N. D. Cook, "Harmony perception: Harmoniousness is more than the sum of interval consonance," vol. 27, no. 1, pp. 25–42, 2009.

[11] J.-M. Chouvel, *Esquisses pour une pensée musicale: les métamorphoses d'Orphée*, 1998.

[12] W. A. Sethares, *Tuning, Timbre, Spectrum, Scale*. Springer Science & Business Media, 2005.

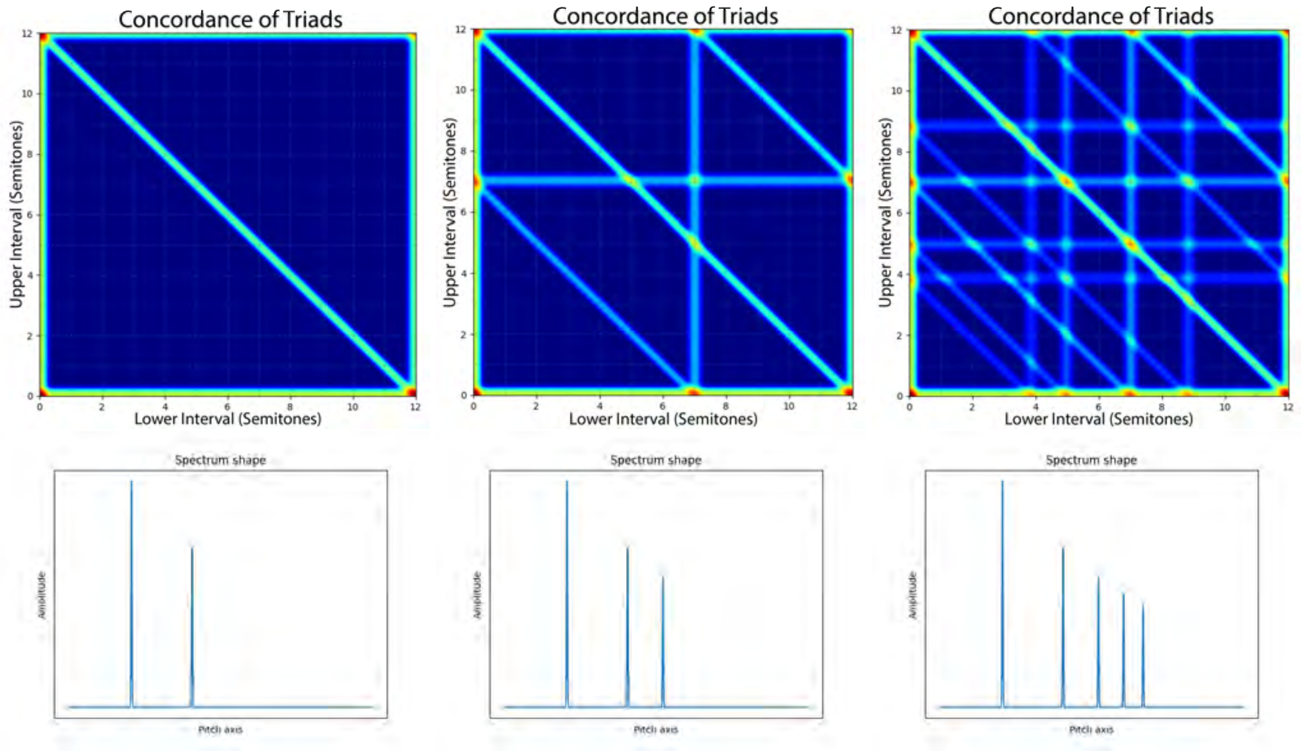


Figure 10. Influence of the number of partials. (Top) Concordance map for synthesized sounds with 2 (Left), 3 (Middle) and 5 (Right) partials. (Bottom) Corresponding spectra.

[13] R. Parncutt, *Harmony: A Psychoacoustical Approach*. Springer Science & Business Media, 1989.

[14] N. D. Cook and T. X. Fujisawa, “The psychophysics of harmony perception: Harmony is a three-tone phenomenon,” 2006.

[15] M. Gaulhiac, “Les descripteurs harmoniques: étude théorique et applications musicologiques,” Ph.D. dissertation, Sorbonne Université, 2021.

[16] ——, “Description acoustique des résolutions cadentielles sur le tonnetz,” in *Journées d’Informatique Musicale 2021*, 2021.

[17] C. Schörkhuber and A. Klapuri, “Constant-q transform toolbox for music processing,” in *7th Sound and Music Computing Conference, Barcelona, Spain*, 2010, pp. 3–64.

[18] G. Bonnet, “Considérations sur la représentation et l’analyse harmonique des signaux déterministes ou aléatoires,” vol. 23, no. 3, pp. 62–86, 1968.

[19] J. Driedger, M. Müller, and S. Disch, “Extending harmonic-percussive separation of audio signals.” in *ISMIR*, 2014, pp. 611–616.

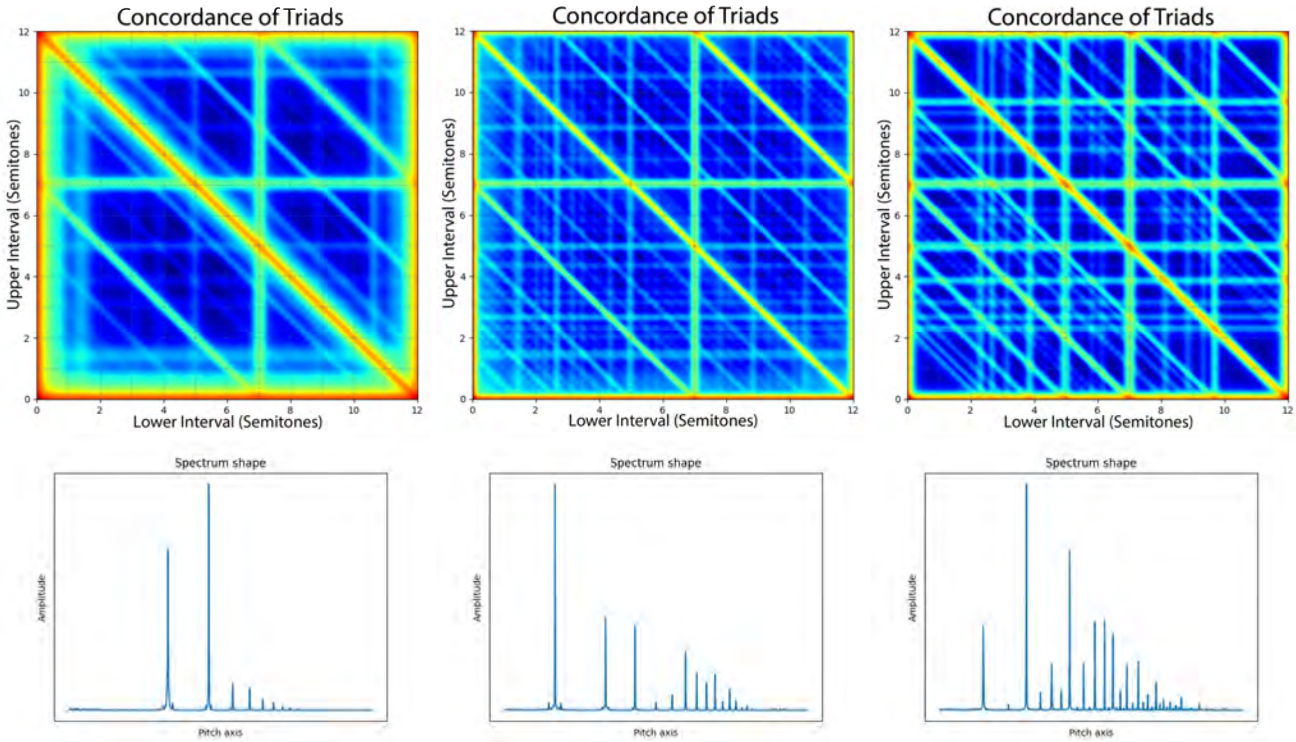


Figure 11. Influence of timbre. (Top) Concordance map for 3 different organ stops: unda maris (Left), vox humana (Middle) and tutti (Right). (Bottom) Corresponding spectra.

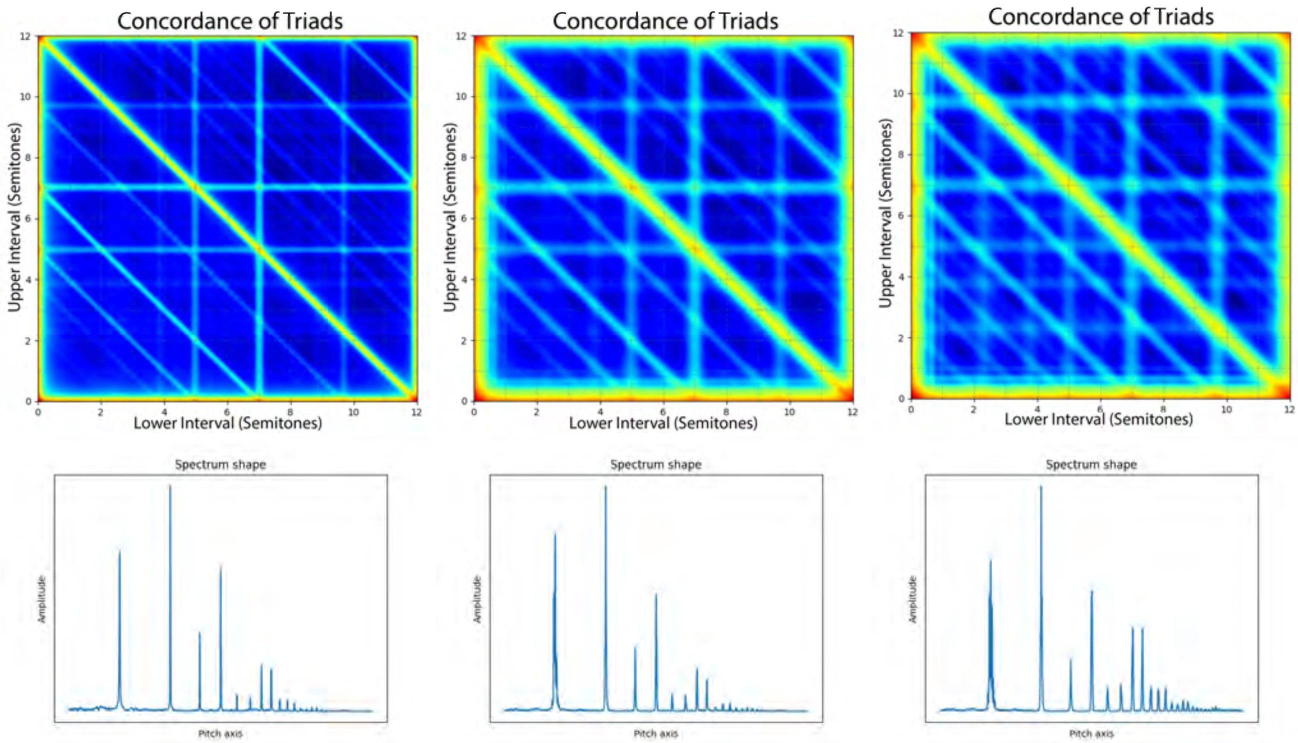


Figure 12. Influence of dynamics and playing technique. (Top) Concordance map for 3 different ways of playing on cello: nuance piano without any vibrato (Left), nuance piano with vibrato (Middle) and nuance forte with vibrato (Right). (Bottom) Corresponding spectra.

BEAM IT UP! — A CLASSIFICATION GRID FOR HISTORIC AND CONTEMPORARY PRACTICES OF BEAMING BY MATHEMATICAL RE-MODELLING

Markus Lepper
semantics gGmbH
Berlin, DE
post@markuslepper.eu

Baltasar Trancón y Widemann
Nordakademie
Elmshorn, DE
baltasar@trancon.de

ABSTRACT

In conventional Common Western Notation (CWN) there are different notations styles for flags and beams, evolved historically. We present a classification based on an algorithm which sets beams according to positions in a musical metric space. This algorithm contributes to more clarity also for the human discussion of the historic phenomena thanks to its stratified architecture: (a) assignment of canonical beaming to nodes of the metric tree, then (b) data transformations coming from pauses, dotted notations, etc., (c) breaking of beams according to further parameters like motifs, playing techniques, etc., and finally (d) transformations according to the needs of graphic appearance. For phases (a) and (b) an exact algorithm is presented; for (c) and (d) a semi-formal classification grid.

1. INTRODUCTION

In Common Western Notation (CWN, starting in the mid-dle of the seventeenth century) a beam is a means to no-tate the duration of a note event. It has basically the same meaning as the (much older) flag: The maximum number of beams from each side of a stem is taken for its flag count.

Early examples of beams in printing can be found in *Hans Neusidler: Ein Newgeordent Künstlich Lautenbuch* (Nuremberg, 1536), urn:nbn:de:bvb:12-bsb00041542-7. The duration symbols were borrowed from the white men-sural notation. Here prevails the "fusa", which looks like the modern eighth. And instead of drawing three flags each for four neighboring "thirty-second notes", it suggests it-self to simply join them into three continuous lines that link all four stems. See Figure 1, from page 99 of the cited work. In its foreword this is called "Leiterlein" = "little ladder". [1]

The idea of *mathematical re-modelling* is to mimic the syntactic operations and semantic outcomes of symbol sys-tems from historically evolved cultural techniques by a col-lection of mathematical definitions and algorithms. Its main

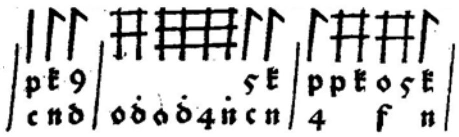


Figure 1. Early Example of Beams in Lute Tabulature

aims are (a) to define more precisely the terminology necessary for human discourse, (b) to lay the foundations for a transparent documentation of automated processing, and (c) to provide a collection of parameters which modify the behavior of the model and can thus be used as a grid for more precise classification and comparison of artefacts, methods and automated processing tools. [2] [3] [4]

Mathematical re-modelling is applied in the following to the problem of finding the adequate beaming configuration for given meter and rhythm, after the sequence of basic symbols (note heads, stems, and flag counts) has been found. The described algorithm has recently been added to our metricSplit Java implementation, found at `http://bandm.eu/downloads/DemoMetric.jar`. This also can be used as a library for own programming, see `http://bandm.eu/sig/doc/api`.

2. BEAMING RULES AS A TRANSFORMATION PIPELINE

Complex transformations which have evolved during centuries of historic practice, can best be analyzed and documented by modelling them as a transformation pipeline, a sequence of distinct transformation phases, each with well defined input and output interfaces and well defined inner behavior. In this paper we present a proposal as a basis for discussion, research and implementation. The first three phases have been implemented and thoroughly tested.

Figure 2 shows the chosen architecture: The first phase operates on the metric tree as such, the second gets the rhythmic information, and the third phase incorporates additional data like sung text, hand distribution, tempo, etc. The very last phase deals with the concrete rendering of the graphic appearance according to pitch heights. Every phase gets as input its specific main input data (= left column in the Figure), the output of the preceding phase (center column), and a collection of further parameter settings, to modify its way of operation (right column).

[1] According to [1, p. 88] beams appear not before 1690 in printed *staff* notation.

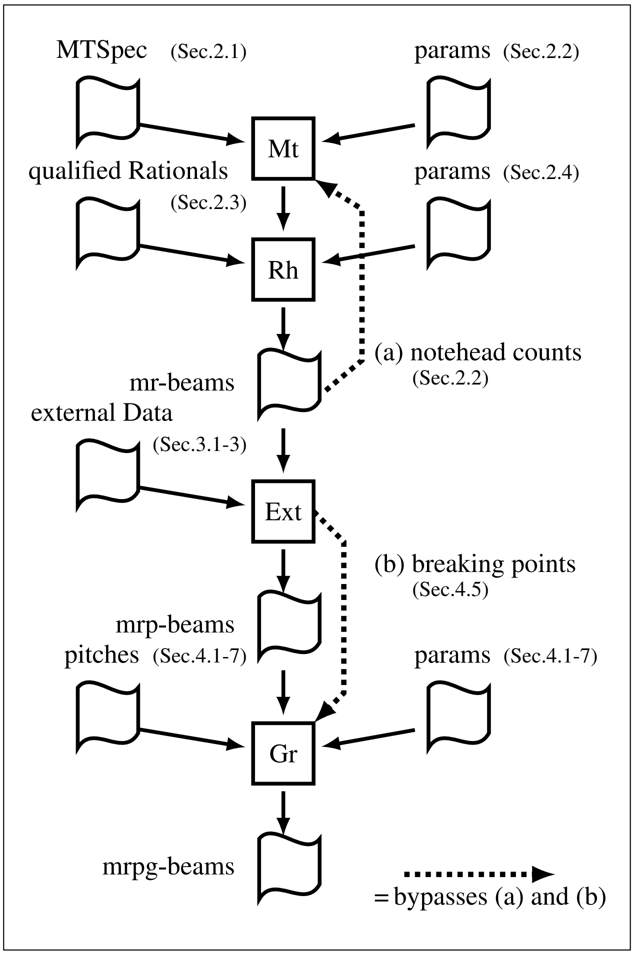


Figure 2. Beam Layout Processing Pipeline: Inputs, Result, Data Flow, and Bypasses

A natural classification grid for different notation styles, epochs, software systems, etc., is given by the selection of the applied transformations, together with their parameter values.

Dividing a complex transformation into distinct phases accomplishes a clean separation of concepts, data, and information flow. But it also helps to identify those critical aspects which can *not* be restricted to only one phase. We have found two such bypasses, shown in Figure 2 by the dotted lines and discussed in sections 2.2 and 4.5.

2.1 Foundation: Genuine Beams

The first phase deals only with the structure of the meter, not with a particular rhythm. A good starting point for any rendering (including beaming) in a metric context is a *metric tree*. Widely varying concepts of metric trees have been defined, esp. for the purpose of music recognition, automated transcription and automated interpretation ([5] [6] [7] [8] [9] [10] [11] [12] [13] [14] [15] [16]—for a survey see [17].) They all have in common:

- A musical meter is represented by a tree of nodes.
- The root of the tree represents the flow of time during a complete measure.
- Each node has an ordered sequence of child nodes; they represent consecutive and adjacent sub-intervals of the time interval represented by their parent node.

Many further attributes, like "metric weight", agogics, harmonic roles and rules, etc., can be attached to each node of such a graph. The following is based on the metric trees from metricSplit [16]. There the only additional requirement is that durations and start and end points are given as rational numbers and the measure starts at timepoint 0. [2] MetricSplit supports arbitrary complex metric tree specifications like $Mts = \mathtt{7/8} + \mathtt{8/7}$.

For notation of rhythms in general and for beaming in particular, it is ergonomically crucial that the notation expresses the relation of each single notation event (as a foreground structure) to a particular node of such a metric graph (as its middleground structure) in a direct and easily readable way. As a consequence, not every sequence of neighboring notes can be joined by a beam, but only those which represent tree nodes under the same parent. [18, p. 27 pp.][1, p. 91 pp.] [19, p. 43 pp.,47] [20, p. 80 pp.][21, p. 153 pp.]

This was already understood by Neusidler: the "sixteenths" in measure three in Figure 1 are only partially joined.

So the first and fundamental rules [3] are:

PROP. nota.trabes.trabesUtVexilia*: Add the numbers of beams and beamlets from both sides of the stem separately. The higher of these sums indicate the duration of the note in the same way as the same number of flags would do.*

PROP. nota.trabes.notaeUtNota*: Let there be a single note A, which directly corresponds to a node of a metric tree, and a group of notes B by which this note is replaced. Then the left(/right) side of the leftmost (/rightmost) note in B has the same beam appearance as the left(/right) side of A, resp.*

Both rules seem easy and trivial, and indeed they are widely applied in conventional engraving. Together they determine completely a beaming for each sequence of notes which corresponds to the complete list of child nodes of a particular metric tree node. These are called *genuine beams*, as shown in the top part of Figure 5 for a very regular "2/4" meter. For instance, the third and fourth thirty-second notes have the same beaming structure at their left and right end as the sixteenth note one line above, which they can replace. [4]

But two severe caveats arise: (1) A node in the metric tree may not correspond to one single note symbol, for instance a node with the duration $\frac{5}{16}$. (2) Historical practice often violates this principle, see the next section.

[2] In particular, the mathematically intricate problem whether the nodes represent open or half-open intervals does not require consideration.

[3] The LMN project [4] has collected about eight hundred properties to classify conventional usage of CWN, all identified by a hierarchical nomenclature of Latin terms. The property names in this article are intended to fit in neatly.

[4] All "one-and-a-half-dimensional" renderings in this article have been produced automatically by our implementation. Their graphics is a mere control instrument, with no claim for beauty. Esp. it employs only linear proportional space allocation, without any "psychological adjustment", which is for instance in musixTex executed by a dedicated external program run. [22]

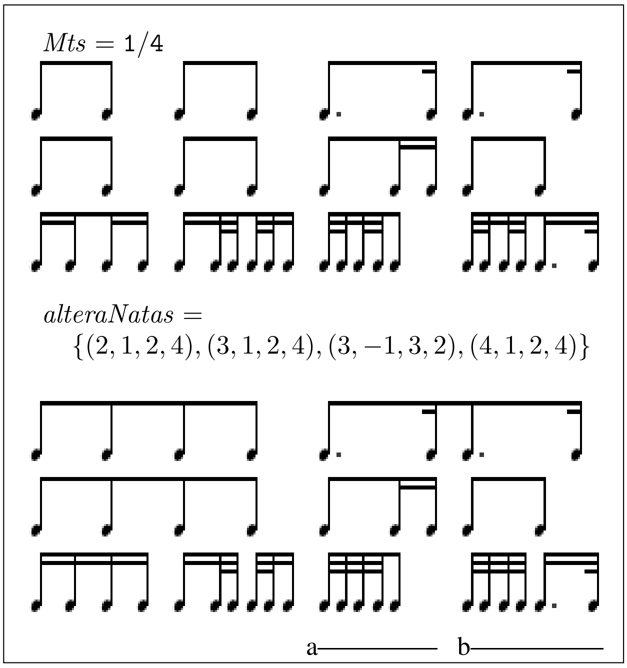


Figure 3. Modified Genuine Beams I

2.2 Modification of Genuine Beams

As mentioned, the historic practice has deviated from these simple rules above:

(A-a) Neusidler did not follow them: In the second measure of Figure 1, the first two "sixteenth" are *not* connected to the "thirty-seconds" by a single "eighth" beam.

(A-b) Nowadays, esp. in the context of traditional 4/4 meters, *all* four sixteenth notes of one quarter are joined by two beams, including the middle ones.

(A-c) Contrarily, when groups of four thirty-second notes follow, these are often also completely joined by three beams, but the single beam connecting the eighth nodes is dropped.

Modifications of these kinds can be modelled by the data type

PROP. nota.trabes.alteraNatas*:* $\mathbb{P}(\mathbb{N} \times \mathbb{Z} \times \mathbb{N} \times \mathbb{N}) \rightarrow$
Each contained tuple (a, b, c, d) *says that on the level* a *of the metric tree, (counted with positive numbers descending from the top node at level* 0*) the value* b *is added to the number of beams if and only if the conditions given by* c *and* d *are met.*

If b *is positive, then* c *gives a maximum number of note symbols which must not be exceeded by both halfs to be connected, and* d *is a maximum for the sum of these numbers. If* b *is negative, than* c *is a minimum which must be reached on either side, and* d *a minimum for both sides.*

These rules are a violation of the pipeline architecture, see the bypass line (a) in Figure 2: Conceptually the genuine beams are determined by the meter only, but the number of actual note heads on both sides of these formulas is not fixed until merging transformations (the "MX" in the next phase, see below) have been applied. When programming a concrete implementation, this twist causes real problems for documentation, testing, and maintenance.

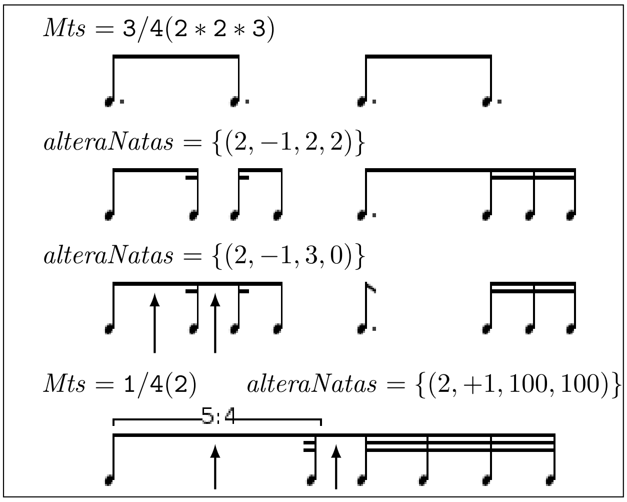


Figure 4. Modified Genuine Beams II

We found that these rather coarse and somehow arbitrary rules are sufficient to model most of the modified genuine beamings found in historic and contemporary engraving practice:

$\{(2, -1, 0, 0)(3, 1, 100, 100)\}$ produces the notation chosen by Neusidler in Figure 1, by removing the connecting beam between the eighths, when reading it (a-historically!) as modern notation.

The top of Figure 5 shows the genuine beams, and Figure 3 shows modern standard engraving of some rhythms in a $\frac{4}{4}$ meter: Code value $(2, 1, 2, 4)$ adds an additional beam between sibling quarters (= at level 2 of the metric tree), but only in case there are at most two notes on each side, see the second example line.

Similarly, $(3, 1, 2, 4)$ joins sibling eighths by two beams instead of one, of course only if possible w. r. t. the rhythmic values, compare lines three and six. [5]

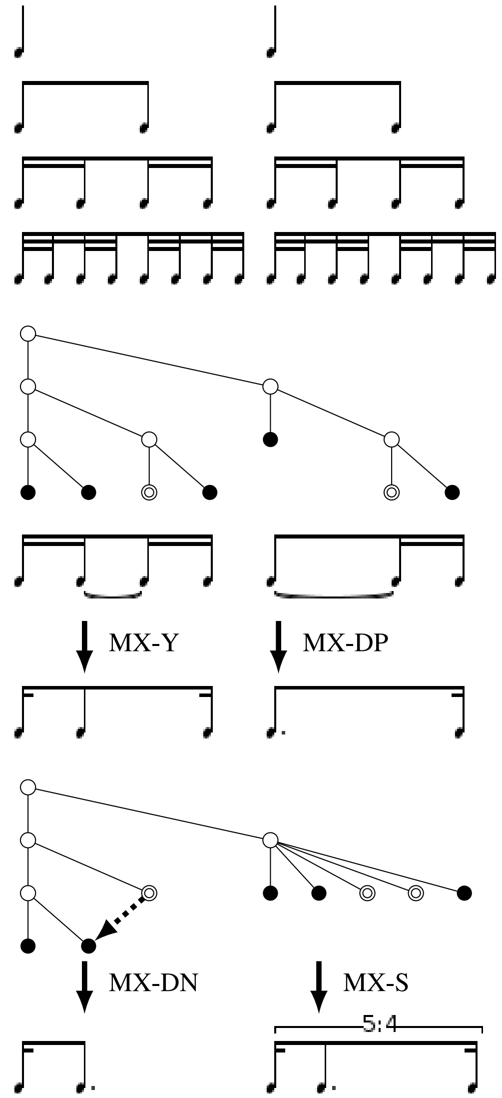
Contrarily, $(3, -1, 3, 2)$ suppresses the beam between two groups of notes which fill sibling eighths, if one of them is too crowded, for a better separation of these groups for ergonomic reasons. Comparing the groups (a) and (b) at the end of Figure 3 shows that $c \geq 3$ must be fulfilled by only one of their constituting eighths, but $d \geq 2$ by both.

Finally, $(4, 1, 2, 4)$ joins four thirty-second notes by an additional beam connecting the sibling sixteenths, but only if not divided finer.

Also counter-intuitive and confusing modifications can be applied. These are not prevented automatically by the current implementation. For instance, the widely spread textbook [18] shows on page 27 a table of examples for rhythmic notations, which notoriously contains patterns like

This is not wrong but somehow paradox: When it shall be stresses that not "two times two" thirty-seconds belong

[5] In the context of metricSplit [23], [16]. the same printing effect could also be realized by changing the definition to `4*1/16`, which means constructing four sibling nodes on the same level, under one common parent node. But then also the metric structure with all other roles and functions is a different one, not only the graphical rendering.



$\langle(1/16, \mathsf{true}), (3/16, \mathsf{true}), (1/20, \mathsf{true})$
$(3/20, \mathsf{true}), (1/20, \mathsf{true})\rangle$

Figure 5. Genuine Beams and Merging Transformations

together, but all four of these on equal terms, which thus cover the complete duration of an eighth, it is confusing to obfuscate these eighths by sixteenths' beams.

(B)

It is a common phenomenon in the history of notation, that a well-proven device is ab-used or re-used in a new context. So the prolongation dot has been advanced from part of the fixed topos "3:1" to a general sign for Mersenne numbers $2^n - 1$ as duration factors.

As a consequence, one and the same graphical beam level can become necessary on two different but neighboring levels of the metric tree, see the arrows in Figure 4. The first line shows an ergonomically sensible layout, but with finer divisions, cancellations of the higher beam become sensible, as also shown.

The same can happen with local divisions by non-Mersenne numbers, see the last line in the figure.

2.3 Beams for Rhythms

A rhythm to be rendered is given as a sequence of pairs of rational numbers (representing time points relative to the measure start, or durations to be added up) and one Boolean value each, qualifying the event as sound or pause, see the last lines in Figure 5. An *initial coverage* is the minimal front in the metric tree which has a node at every start point occurring in the rhythm, see again Figure 5. The initial coverage can always be rendered immediately, using the genuine beaming from the top of the Figure: When a sounding event is represented by more than one note symbol, these must be connected by a *tie symbol*.

Only the first nodes assigned to any event show up in the rendering necessarily. All others can possibly be *merged* with their predecessors. The rules when to apply a particular merging transformation (MX), and the different style parameters to re-model the different historical practices can become very complicated, especially when a *total* rendering function is to be described, which supports arbitrary meters and rhythms—for details see [16].

By each MX, a contiguous sequence of note symbols is cancelled from the notation and the immediate predecessor is prolongated. This may affect beaming. On the one hand, positive dottings (MX-DP) increase the duration, but they never reach the factor 2; therefore they never affect beaming. On the other hand, with negative dottings (MX-DN) the last of the merged note symbols appears at the rhythmic position of the first one and the beam structure must be transferred, see the dotted arrow in the figure.

Syncopes (MX-Y) and hemiolas (MX-H) print the parent node's note symbol at a child node's position. Thus the left(/right) side of this note must copy the left(/right) side of the left(/right) parent involved, resp.

The merging of equidistant siblings (MX-S) enlarges the duration of the first (=the only printed) note and thus reduces the beaming according to the resulting multiplication factor f. (More precisely: by the highest n such that $2^n \leq f$.)

Thus after all these transformations a subsequent *cutting down* step is needed: All beams which are foreseen as genuine but may only touch one of their two stems, because the other's note value has been prolongated by a merging transformation, are cut down to *beamlets*. [6]

2.4 Local Transformations of Beam Patterns

Concerning pauses, there are three alternative ways how to apply the resulting beams and beamlets to the note symbols [18, p. 49][19, p. 46][24, p. 15 p.][20, p. 88, 213]

PROP. nota.trabes.sopraPausam.perCaudulam: *Each pause is treated like a sounding note and gets a* stemlet.
...transiens: *Pauses are not connected to beams, but a beam may span over a pause.*
...separans: *A pause is never spanned by a beam, but cuts all beams pointing to it down to beamlets.*

The format *perCaudulam* is the most recent developed format, and also the the most canonical: Each pause gets a

[6] Beamlets can also occur as genuine beams, but only in advanced use cases, see the last example in section 2.2.

“stemlet” and is treated like a sounding event, see line ① in Figure 6. The representation is somehow redundant because the duration of the pause is encoded twice: in the beam (but without prolongation dots) and in the pause symbol. Nevertheless, from the systematic perspective this is the canonical form.

For the variant *transiens* these stemlets are deleted, all beams between a stemlet and a stem are cut down to beamlets, and all beams between two stemlets disappear completely, see line ②.

Line ③ shows the form *separans*: All beams which touch pauses are cut down to beamlets. This variant is dominant in conventional sheet music. But it gives up one particular role of beaming which goes beyond the mere replacement of flags:

PROP. nota.trabes.significantVocem: *Connected beams implicitly have the role of voice-leading indicators (German “Stimmweiser”).*

There are rare but relevant critical cases:

Here (BWV 543, ms. 16) the first version (Neue Bach Ausgabe = NBA) of engraving forbids to read a crossing of the voices allowed by the second (Breitkopf Sämtliche Orgelwerke Urtext), which would but be the correct resolution of the d”, the seventh of the dominant. [7] The collection of all beams and beamlets which are connected by at least one contiguous top-level beam are called a *beam aggregate* in the following.

All forms ① to ③ are a basis to which further *local transformations* are applied. These aim at eliminating the following properties:

PROP. nota.trabes.trabulaeContraIdem: *All beamlets appear on the same height on both sides of a particular stem.*

This pattern is caused by an isolated inner node from a group of more than two equidistant siblings, see the triplets in Figure 6 line ③. The beamlets remaining from the above-mentioned down-cuts appear on both sides of the stem with equally good reasons. It may seem desirable for an engraver (for whatever reason) to eliminate this property. This can be done by replacing them by beamlets only to one side. This transformation reduces the “sensible ergonomic information”, because the fact that the note is connected *in the same way* to its left and to its right neighbor is no longer expressed.

Line ⑦ in Figure 6 shows an example and Table 1 shows a mathematical model. Each stem end is represented by five natural numbers: number of flags, number of left beams and beamlets, number of right beams and beamlets. The properties and the eliminating transformations are specified on these 5-tuples, the former as pre-conditions for the latter. [8]

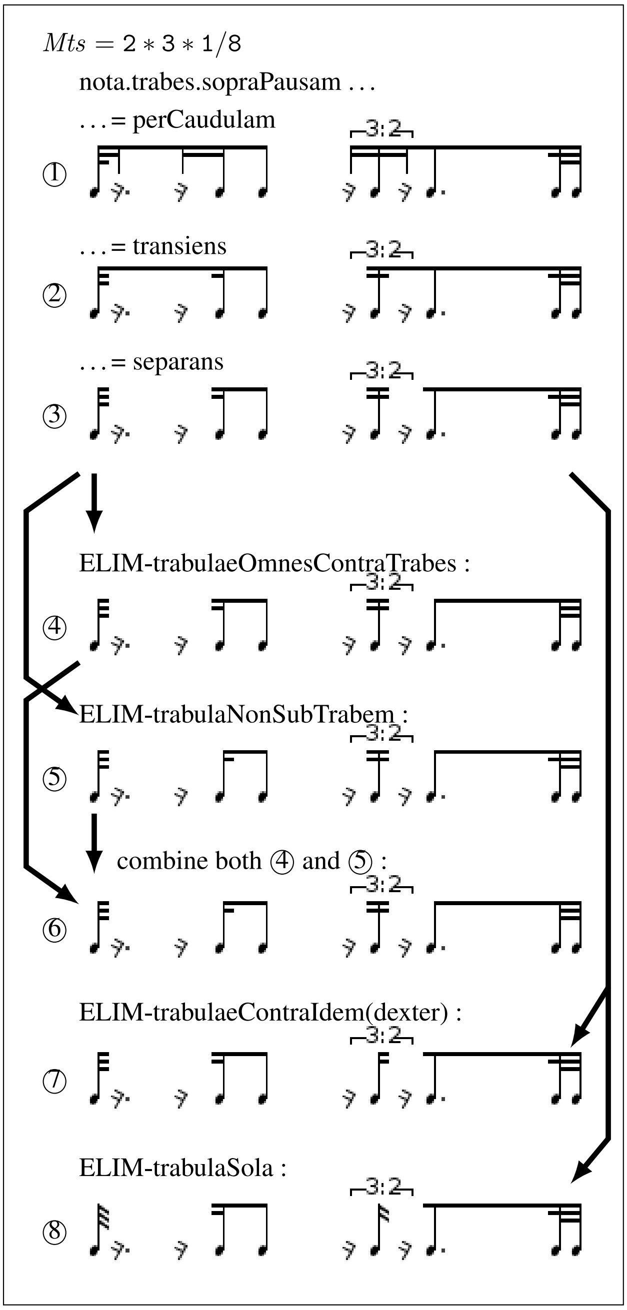
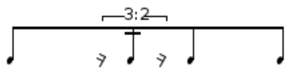
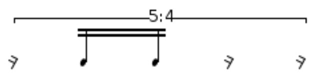


Figure 6. Local Transformations

The formula for *nota.trabes.trabulaeContraIdem* only requires the number of left and right long beams to be equal ($= a$). So it also holds for beamlets on both sides under (the same number of) beams on both sides, a case not depicted in Figure 6: 

A similar property can be defined to match cases like 

where a very similar distribution of beamlets is found at both sides not of a single

[7] Thus the voice identification algorithm in [25] in a preparatory step (p. 307) replaces beams (and slurs) by explicit voice-leading signs

[8] The prefix “ELIM-” works as on mere technical meta-level, not as part of the nomenclature. Then of course “elim-trabul*as*-contra-idem would be the correct case.

note but of a group of siblings. This case is *not supported* by the current model. It is left to future extensions of our work, because its eliminating transformation would not be a mere *local* one.

PROP. nota.trabes.trabulaeOmnesContraTrabes: *All beamlets one one side of the stem appear on the same height as a (long) beam on the other side.*

Due to *trabesUtVexilia*, each such beamlet is redundant for calculating the duration. It nevertheless indicates the relation of this stem's note and node to its neighbor in the indicated direction. In conventional engraving, this beamlet is removed—compare the last two beamlets in line ③ of Figure 6 to line ④.

While the eliminating transformations above reduce the information content, they do not introduce confusing contradictions to the genuine beams. This changes when trying to eliminate

PROP. nota.trabes.trabulaNonSubTrabem: *On one side of the stem one or more beamlets appear, but no single beam. On the other side there is at least one beam.*

Then the beamlets traditionally simply switch the side, see the second sounding note in Figure 6, the transformation's definition in Table 1 and the result in Line ⑤. Here the direction into which the sixteenth-beamlets point is *positively wrong* and *contradicts* the original genuine beams. Nevertheless this is a standard transformation in traditional engraving. This also eliminates

PROP. nota.trabes.trabulaSola: *There is a stem which carries beamlets but no beams.*

This is eliminated by the final transformation in Table 1 and Figure 6, namely by replacing all beamlets by the corresponding number of flags.

Conventional CWN engraving uses all these four transformations, but ELIM-trabulaeContraIdem is often superseded by ELIM-trabulaSola, compare lines ⑦ and ⑧ of Figure 6.

Not all possible cases are covered by these transformations, for instance $(0 \mid 0, a \mid 0, b) \wedge a \neq b \wedge a \neq 0 \wedge b \neq 0$. But in conventional usage of CWN it is transformed by ELIM-trabulaSola to $(max(a, b) \mid 0, 0 \mid 0, 0)$ anyhow.

3. ADDITIONAL EXTERNAL DATA

The beam aggregates constructed so far are determined solely by meter and rhythm. They are called *mr-beams* in the following. In each particular notation, they will be employed in contexts where *further parameters* influence their concrete appearance.

3.1 Indirect Influence by Stem Direction

First of all, the *stem direction* is of course relevant. Assume that every sequence of notes is part of a particular *notational voice*. Then there are basically three variants for determining the stem direction:

PROP. nota.cauda.significat...
... vocem: *the stem direction is needed to identify the voice among other voices in the same staff.*
... instrumentum/accentum/manum/modumAgitur/...: *the stem direction is employed to represent the named (binary valued) parameter.*
... nihil: *the stem direction carries no meaning at all, but is free to change according to graphic/ergonomic requirements.*

In the first two cases the stem direction is fixed and must be respected by all further transformations; it is a further input parameter to any beam layout algorithm. Normally it will change in the first case less frequently than in the second. Only in the last case it is free and thus is an output of the subsequent transformation processes.

For beam aggregates there are further sub-cases:

PROP. nota.cauda.significat.vocem...
... trabsSeparataCaudaMutata: *beams are broken when the stem direction changes.*
... trabsSeparataLineaMutata: *beams are broken when the voice moves into another (mostly: a neighboring) staff.*

In most conventional engravings of classical piano music the first property holds (to thin out the optical appearance for readability), the second does not (to make the transition of the voice even clearer, according to *trabes.significantVocem*).

Whenever such a breaking takes place, the situation is as if a pause had occurred in the mode *nota.trabes.sopraPausam.separans*: Subsequent application of *ELIM-trabulaNonSubTrabem* and *ELIM-trabulaSola* may or may not be applied.

A minimal case is a one-line percussion staff, with one notated pitch only and two stem directions for a parameter. Then a frequently found transformation is

PROP. nota.voces.unaUtDue_pausasPerdatas: *The voice is split into two voices, separating up and down stems, and both are beamed independently, except that the events in one voice stand in for the pause symbols in the other.*

This can also be applied in cases which are only slightly more complicated. Gould [21, p. 312] gives an example from piano notation, where the parameter encoded by stem direction is the hand selection = *cauda.significat.manum*:

With respect to the (invisible) pauses either *nota.trabes.-sopraPausam.transiens* or *... separans* can be applied to both resulting voices. The example shows the former, followed by *ELIM-trabulaeOmnesContraTrabes*.

$$(\mathit{flags} \mid \mathit{leftLong}, \mathit{leftShort} \mid \mathit{rightLong}, \mathit{rightShort}) : \mathbb{N} \times \mathbb{N} \times \mathbb{N} \times \mathbb{N} \times \mathbb{N}$$

$$\frac{b > 0}{\begin{array}{c}\mathsf{ELIM}\text{-}\mathit{trabulaeContraIdem}((0 \mid a, b \mid a, b), \mathsf{dexter}) = (0 \mid a, 0 \mid a, b) \\ \mathsf{ELIM}\text{-}\mathit{trabulaeContraIdem}((0 \mid a, b \mid a, b), \mathsf{sinister}) = (0 \mid a, b \mid a, 0)\end{array}}$$

$$\frac{b > 0 \qquad a + b \le c}{\begin{array}{c}\mathsf{ELIM}\text{-}\mathit{trabulaeOmnesContraTrabes}(0 \mid a, b \mid c, d) = (0 \mid a, 0 \mid c, d) \\ \mathsf{ELIM}\text{-}\mathit{trabulaeOmnesContraTrabes}(0 \mid c, d \mid a, b) = (0 \mid c, d \mid a, 0)\end{array}}$$

$$\frac{b > 0 \qquad c > 0 \qquad x = \mathsf{MAX}(d, b - c)}{\begin{array}{c}\mathsf{ELIM}\text{-}\mathit{trabulaNonSubTrabem}(0 \mid 0, b \mid c, d) = (0 \mid 0, 0 \mid c, x) \\ \mathsf{ELIM}\text{-}\mathit{trabulaNonSubTrabem}(0 \mid c, d \mid 0, b) = (0 \mid c, x \mid 0, 0)\end{array}}$$

$$\frac{b + d > 0}{\mathsf{ELIM}\text{-}\mathit{trabulaSola}(0 \mid 0, b \mid 0, d) = (\mathsf{MAX}(b, d) \mid 0, 0 \mid 0, 0)}$$

Table 1. Historically Defined Local Transformations

3.2 Direct Influence

The preceding section covered the case that an additional parameter breaks an mr-beam indirectly by determining the stem direction. But such a parameter can also affect the beamings directly. The result is called an *mrp-beaming*. Basically there are two cases:

PROP. nota.trabes.extera…
…separans: *a parameter value causes the breaking of a beam from the mr-beaming.*
…ligans: *a parameter value joins two groups of beams which are foreseen as separate by mr-beaming.*

The first case is much more frequent.

With *sung lyrics* one of the following properties may apply, which all are of type *extera.separans*:

PROP. nota.trabes.cumVerborum…
…syllabis: *beams may not extend further than the sung syllable.*
…nominibus: *beams may not extend further than the sung word.*
…lineis: *beams may not extend further than the sung text line.*

The first property had been standard in all engravings of classical and romantic music. The last property can sometimes be found in contemporary popular sheet music. For the second we have not found any evidence, but it logically closes a gap.

In contemporary advanced music neither is applied, but vocals are beamed as "any other" instrument. [1, p. 8] This can be called *NON-trabes.cumVerbis*.

PROP. nota.trabes.separatae.cumMelo/cumLigato: *Especially in piano music, but also in orchestral monodies, a sequence of adjacent motifs with the same rhythmic structure is clarified by breaking the beams between every onbeat and the following up-beat part.*

See chorale prelude op. 122, Nr. 3, ms. 7 f. by Johannes Brahms, where this technique is applied in the first measure (redundantly doubling the slurs), but not in the second:

Please note that this operation gives up *trabes.significantVocem*: The voice leading may become less clear.

Confusingly, two modifications with *contrary* meaning have the same optical result:

PROP. nota.trabes.ligataeContraNates: *Notes are connected contrarily to the grouping by the genuine beams.*

PROP. nota.metraMulta.perTrabem: *A polymetric situation is clarified by beam patterns in a particular subset of voices which are shifted against the beams in the other voices and/or the sequence of measure bars.*

These two are totally different techniques: The former only affects the way of writing (= the sphere of syntax) but still means the original pattern of stress, unchanged relations in agogics and motif, etc. (= the same semantics). But the second means that the meter shall indeed be shifted with all consequences in interpretation (= in the semantic sphere).

Often *ligataeContraNates* is applied to up-beats and their targets. For example, LilyPond [26], which claims to implement standard engraving conventions, renders the source text "`r8. d16 e r8.  r2`" as [9]

The appearance of this pattern in Chopin's sonata in b minor, [10] Largo, ms. 5–19, left hand, is a typical example for its combination with

PROP. nota.trabes.cumPositioneManus: *Change and identity of the hand position are expressed by breaks and continuity of the beams.*

[9] So done by LilyPond version 2.20.0.
[10] Engraving "Collection Litolff No 1087", `IMSLP638961--PMLP2364-ChopinSonataBMinor-KohlerEdition.pdf`.

This is one of the rare cases of the above-mentioned *nota.-trabes.extera.ligans*, where *additional* beams are caused by external data.

Both techniques can and often do lead to beams crossing bar lines and even line breaks, called *nota.metrum.secat.trabem* and *nota.linea.fracta.secat.trabem* by [4, p. 223, 229].

Examples of polymetrics in Romantic piano literature are often of a most simple type, namely a shift by a certain distance d with otherwise unchanged structure (= same duration, metric tree and tempo, labelled "$(d, =, =, =)$" by [4, p. 243 pp.]). But also constellation $(0, \neq, \neq, =)$ is possible and can easily expressed by beaming only, see again Brahms, Capriccio op. 67, Nr. 5 , m. 111 ff:

For more recent examples see [24, p. 116 pp.][21, p. 171, 175 pp.][20, p. 170 pp.]

3.3 Beams expressing Tempo – "Feathered" Beams

A more advanced concepts are "feather" or "fan" beams (German "Fächerbalken"). [19, p. 47] [20, p. 94] [24, p. 124, 141] [21, p. 158]

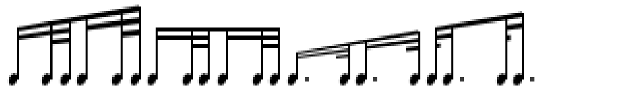
Primary, secondary, ternary beams etc. do not run in parallel but converge like a fan to a common point. The intuition is a ritardando when at the left end there are three beams (which would mean a thirty-second note) which collapse into only one beam at the right end (which stands for the much slower eighth note). Vice versa, starting at one point and running apart means an accelerando.

While in literature this principle is only defined for a "free" interpretation and nearly always written with three beams, indeed it is totally independent from the notated and played pattern and can (in sensible limits, for readability) be combined with any rhythm:

PROP. nota.trabes.accelerans*:* $\mathsf{Seq}(\mathbb{Q} \times \mathbb{N} \times \mathbb{Q})$.
= list of tempo changes which shall be expressed graphically by vertical beam dimensions.

Let this sequence be sorted by the first components ascending. Then each contained triple (a, b, c) says that at time point a the tempo b BPM shall rule, that this is expressed by the factor c applied to the widths and distances of all beams at this point, and that between these triples linear interpolation shall be applied.

So with *trabes.accelerans* $= \langle(0, 30, 1/1)(1/4, 90, 3/2), (1/4, 60, 1/1), (1/2, 60, 1/1), (1/2, 30, 1/2), (1/1, 120, 3/2)\rangle$ we get the rendering



which possibly is not too intuitive, but the canonical continuation of the principle of feathered beams.

4. TWO-DIMENSIONAL LAYOUT: VERTICAL POSITION AND PITCH HEIGHT

Up to here, all considerations have been related to a "one-and-a-half-dimensional" space: The x-axis is mapped to the flow of time, but the y-axis is made up by only the beam selections and the stem direction.

The most frequent context for beams is to be attached to noteheads which represent *pitches* by their vertical position relative to the staff lines. So now graphical criteria come into play, to produce a true two-dimensional arrangement of the beams. Here a beam aggregate from the mrp-beaming can even be broken again.

Any *layout algorithm*, whether applied manually or automatically, must always answer *four distinct questions*. These are very different in nature, but their solutions are tightly mutually dependent. [27, p. 153] How they are priorized or whether they are solved separately or in an intermangled way may differ.

The questions are:

PROP. nota.trabes.inclinatioSignificans*: How does the steepness of the top-most beam symbol indicate a tendency in the distribution of the pitches, or even a musical gesture?*
. . . ponuntCaudas*: Only in case that the stem directions are still free at this point of the processing pipeline: Does the fact that all notes shall be beamed together determine a preferred stem direction?*
. . . visio*: What are the graphical coordinates of the whole beam aggregate? Or those of its fragments, if a printable solution can only be found after breaking it?*
. . . inLineolas*: How does the graphic appearance of the beam symbols interact with the individual lines of the staff?*

Each of the relevant publications on historic engraving treats all or most of these questions, but only separately, see Table 3. There are *no complete and explicit* algorithms outside the heads of the engravers, who have done this job over centuries. On the other hand, nowadays digital note setting programs necessarily contain such an algorithm, but those are not published. Further research will thus include reverse engineering.

Table 2 shows a possible modelling of the algorithm's input data, and possible properties of its result. Again, restrictions on these can serve as further input parameters.

4.1 Ergonomic Significance of Beam Inclination

The first property, *inclinatioSignificans*, is discussed with very different results. [19, p. 42] [18, p. 46 p.] [27, p. 155, 168 pp.] [21, p. 22 pp., 169 pp.] A general consensus is that an overall tendency of the pitches shall be expressed by the inclination of the beam, which historically had been decided by the taste of the engraver (unless restricted by collisions, see below.) But the opinions in details, given only by examples, do differ widely. [19, p. 42] demands a maximum steepness of 30 degrees. Many authors on nineteenth to twentieth century engraving practice impose a *maximal lift*, not a steepness—required by the physics of copper plate engraving, see section 4.4 below. [11]

In general, this area appears to require *non-local* considerations, corresponding to **nota.trabes.priorInfluit/successorInfluit** from Table 2. For instance, to use horizontal beams for Alberti bass figures is motivated by the fact that these are immediately repeated, so that the distribution of the note head heights is stationary—not over a single beam aggregate, but over their sequence. [27, p. 154 p., 159 p.][19, p. 40 p.][21, p. 25] Here future work is required—our current formalization is restricted to *local* phenomena.

4.2 Stem Direction of Beam Aggregates

The question how the decision that a group of notes shall be beamed together influences the stem directions (of course only in cases when these are still free), is discussed by [1, p. 94 pp.] [20, p. 88] [21, p. 24 p.][27, p. 154] [19, p. 40 p.]

4.3 Graphical Placement of Beam Aggregates

The raison d'être of any layout algorithm is to find a final graphical position of the beam aggregate. Its input are the mrp-beaming computed so far, pitches of the note heads, and additional parameters. The properties listed in Table 2 can all be made input parameters by defining restrictions. Its output are the coordinates of one or more beam fragments, esp. their inclinations. [18, p. 45 pp.] [1, p. 97 pp.] [1, p. 115 pp.] [19, p. 42][21, p. 17 pp.][27, p. 155 pp.]

In most cases such an algorithm will be a *partial* function:

PROP. nota.trabes.conditionesConfligentes: *no (simple) solution can be found which fulfills all requirements stated by the input data (mrp-beaming, pitches, and parameters).*
. . . vocesConfligentes: *a conflict with the positions of the graphical representations of notes from* another voice *obstructs finding a solution.*

The second case is not formalized in our approach so far. A possible idea is to give a list of "blocked rectangles" as additional input parameters. Whenever no single beam aggregate can be found, the remedy of further splits can perhaps be applied, as described in sections 4.5 pp.

4.4 Fine Tuning against the Staff Lines

Historically, much attention had been payed to *nota.trabes.inLineolas*, the relation of the beam to the staff line, mainly to avoid too small areas of white paper, problematic with traditional mechanical engraving technologies. [18, p. 43 pp.] [1, p. 98 pp.] [19, p. 41 p.] [24, p. 9 pp.][21, p. 17 pp.][27, p. 25 p.,42, 161 pp.] But also nowadays, "wedges" and gaps should still be avoided for their bad psychological effects on legibility.

[11] See the sections "3.5.2.1 traditional steepness" v. "3.5.2.2 contemporary" in [27].

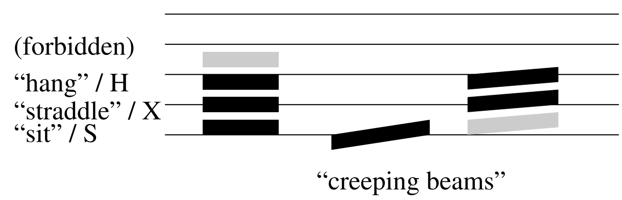


Figure 7. Beams versus Staff Lines

Most authors agree on

PROP. nota.trabes.tresInTresLineolis: *Let h be the height of a beam, d the distance between two beams, and s between two staff lines. Then it holds that $3*h+2*d = 2*s$.*

(Remarkably, all authors treat the width of the staff line as zero.) Most of the authors agree on the solution $4*d = 2*h = s$, but e.g. [24, p. 9] allows reducing h while sticking to the formula. Only [27, p. 42 p.] proposes the formula $4*h+3*d = 3*s$ (**...quatuorInQuatuorLineolis**, which is indeed much more flexible) and $h = 1.52*d$.

Under *trabes.tresInTresLineolis*, there are four positions of a horizontal beam relative to staff lines, one of which is forbidden. The others are called "sit" (S), "straddle" (X), and "hang" (H), and must follow in this order bottom-up for a horizontal 1/32 beam aggregate, see Figure 7. Continuing this rule, an aggregate of four or more parallel horizontal beams necessarily includes the forbidden position between the lines. [1, p. 125 p.] and [24, p. 11 p.] propose for this case to enlarge d to $s/2$, but only when the beams indeed fall into the staff. This implies

PROP. nota.trabes.sineLineolis: *Beams drawn in a staff and beams outside are treated differently.*

For *slanted* beams, the same relative positions S, X, and H apply for their starts and endings.

Most authors prefer

PROP. nota.trabes.subLineola: *beams which hold contact to one and the same staff line throughout.*

These are called "creeping beams" ("schleichende Balken") by Chlapik [19, p. 42], but he explicitly restricts their applicability, due to *inclinatioSignificans*.

A single creeping beam thus can have a maximum lifting of $s/2$, a double beam of $s/4$, a triple beam is not possible, see again Figure 7. (Creeping beams are the typical case where not the maximum steepness but the maximum lifting is fixed as an input parameter. Some authors allow the three slanted beams from the Figure with stems going up [="downstemmed" noteheads], because then no "wedge" appears. Else the beams must be shifted down by $s/4$.)

4.5 Resolving Conflicts by Breaking Beams

Giving the beam aggregates, the note head positions, and the additional parameters, an algorithm may fail to find a solution. The most simple and most frequently used

$$\mathsf{Heads} = \mathsf{seq}(\mathbb{Q} \times \mathbb{PQ})$$
$$H \in \mathsf{Heads} \wedge 1 \leq k_1 < k_2 \leq \#H \Longrightarrow \pi_1(H(k_1)) < \pi_1(H(k_2))$$
// = input data: sorted list of x-coordinates of stems and y-cooordinates of the attached note heads (possibly chords!)
// Sequence indexing starts with 1
$$\mathsf{Beams} = \mathsf{seq}(\mathbb{Q} \times \mathbb{Q} \times \mathbb{Q} \times \mathbb{Q})$$
$$B \in \mathsf{Beams} \wedge 1 \leq m \leq \#B \wedge 1 \leq n \leq \#B - 1 \Rightarrow \pi_1(B(m)) < \pi_3(B(m)) \wedge \pi_3(B(n)) \leq \pi_1(B(n+1))$$
// = output data: list of start and end point coordinates of the top-level beam or of its fragments
$$\mathsf{Layout} : \mathsf{Heads} \times \mathsf{Params} \nrightarrow \mathsf{Beams}$$
$$\mathsf{Layout}(H, P) = B$$

$$\mathsf{beamedBy} : \mathbb{N} \leftrightarrow \mathbb{N} \qquad \mathsf{beamHeight} : (\mathbb{N} \times \mathbb{N}) \nrightarrow \mathbb{Q} \qquad \mathsf{sides} : \mathbb{N} \rightarrow \mathbb{P}\{-1, +1\}$$
$$\mathsf{beamedBy}(m, n) \Longleftrightarrow H(m) = (x_m, \_) \wedge B(n) = (x_1, \_, x_2, \_) \wedge x_1 \leq x_n \leq x_2$$
$$\mathsf{beamHeight}(m, n) = y \Longleftrightarrow H(m) = (x_m, \_) \wedge B(n) = (x_1, y_1, x_2, y_2) \wedge y = (y_2 - y_1)/(x_2 - x_1) * (x_n - x_1) + y_1$$
$$x \in \mathsf{sides}(m) \Longleftrightarrow \exists n : \mathbb{N}, y : \mathbb{Q} \bullet y \in \pi_2(H(m)) \wedge x = \mathsf{sgn}(y - \mathsf{beamHeight}(m, n))$$
$$\mathsf{knees} = \{m : 2..\#H \mid \mathsf{sides}(m) \neq \mathsf{sides}(m-1)\}$$

Further sensible input parameters to a Layout Algorithm:

$nota.trabes.cauda.max = c_{\mathrm{A}} : \mathbb{Q}_{>0}$ the maximal length of the very first and last stems.
$nota.trabes.cauda.maxInterior = c_{\mathrm{B}} : \mathbb{Q}_{>0}$ the maximal length of all other stems in a beam aggregate.
$nota.trabes.cauda.min = c_{\mathrm{I}} : \mathbb{Q}_{>0}$ the minimal length of the very first and last stems.
$nota.trabes.cauda.minInterior = c_{\mathrm{J}} : \mathbb{Q}_{>0}$ the minimal length of all other stems in a beam aggregate.
($nota.trabes.cauda.solaMin/solaMax : \mathbb{Q}_{>0}$ the minimal/maximal length of a stem with no beams.)
$nota.trabes.maxInclinatio = d_{\mathrm{A}} : \mathbb{Q}_{\geq 0}$ the maximal steepness of a beam.
$nota.trabes.maxAltitudo = h_{\mathrm{A}} : \mathbb{Q}_{\geq 0}$ the maximal lifting of a beam (= absolute difference of first and last y pos).
nota.trabes.priorInfluit = the appearance of a beam aggregate changes with the contents of a preceding aggregate.
nota.trabes.successorInfluit = the appearance of a beam aggregate changes with the contents of a following aggregate.

Table 2. Graphical Properties of Beam Layout

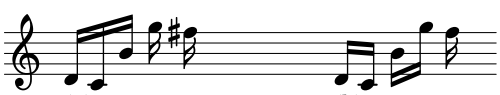
method for solving such a conflict is to break the beam into two parts, which are then dealt with separately. [21, p. 25] Since music is read from left to right, and the reader should be surprised as little as possible, this is naturally done at the rightmost position, i.e. as "late" as possible.

This break leads to a gap in the beaming; the resulting parts are then processed with the same layout algorithm recursively. The resulting aggregates are called *mrpg-beamings*. In Table 2 this is modelled by the result type Beams being a sequence of coordinate pairs. A gap after index position m of the input note sequence is characterized by $\mathsf{beamedBy}(m) \cap \mathsf{beamedBy}(m + 1) = \varnothing$.

Again it is worth mentioning that possibly *trabes.significantVocem* is given up and hence in multi-voice writing the voice leading may become less clear.

Applying just the break (in the narrow sense of this word) necessarily produces stemlets. These can subsequently be processed/removed in the same way as known from the pauses in the mode *nota.trabes.sopraPausam.separans*: by applying *ELIM-trabulaNonSubTrabem*, *ELIM-trabulaOmnesContraTrabes*, *ELIM-trabulaSola*, and (in rare cases) *ELIM-trabulaeContraIdem*, as already specified above in section 2.4 and in Figure 6.

The breaks inserted by this algorithm are in the syntactic sphere alone, they are introduced for mere graphical reasons. A straightforward solution is (a):



(a) (b)

Whenever the semantics shall influence these breaks, they must be applied in the preceding phase explicitly. Only (b) expresses correctly the jump of the hand position on a keyboard, according to *nota.trabes.cumPositioneManu*. Thus, for a flute (a) is correct, for a piano (b). Our strict pipeline architecture does not support these considerations, but could be enhanced by "predetermined breaking points" as known from the TEX typesetting system: the preceding semantic processing could assign "preferences and penalties" to breaking points which are used by the subsequent layout procedure only in case they are required. This is one of the bypasses we identified in our transformation pipeline, see the dotted line (b) in Figure 2 on page 2.

4.6 Resolving Conflicts by Knees

Another remedy with less loss of substance but more graphical overhead is the "extraordinary beam", called "knee" (from the German "Knie") in the following. It means a beam from which stems go *both up and down*. [1, p. 118 p.] [19, p. 43] [24, p. 12 pp.] [20, p. 86, 88 pp.] [21, p. 26] [27, p. 93, 135, 153] More formally the indexes after a knee are given by knees from Table 2.

PROP. nota.trabes.genus.inLineae*: a knee is contained totally in one staff.*
. . . interLineas*: a knee is contained in or crosses the space between two staves.*

The latter is a frequent case in two-staff writing. [12]

[12] [19, p. 43] says about this case even "[Die] Balkenverbindung von einer Zeile in die andere ist nicht als ≫Knie≪ anzusprechen." ("Beams from one staff into another shall not be called ≫knee≪.")

Significant properties are:

PROP. nota.trabes.numerusGenuum: *number of "knees" = points where the stems change the side* $= \#\mathsf{knees}$ *in Table 2.*
…numerusInGenui: *the number of beams which cross the knee.*
…numerusInterGenibus: *the number of consecutive note heads between two knees.*
…caudaVersusCaudam: *whether two stems to both sides can appear at the same x-coordinate. This corresponds to* $\{-1, 1\} \in \mathsf{ran(sides)}$ *in Table 2.*

Restrictions to these properties again can be used as further input parameters to an algorithm.

When multiple beams go "through" the knee (i.e. appear with a stem to one side immediately followed by a stem to the other side, or more complicated cases of change), they completely exchange their meanings. For instance in a group of three, the graphical symbol which is the "highest" beam on the one side, representing thus the "eighth" flag, becomes the "lowest" on the other, representing the thirty-second flag, et vice versa. [13] Anyway it should hold

PROP. nota.trabes.summaeInGenu: *The bundle of beams which cross the "point of change" at a knee are on both sides the top-most beams in the aggregate; all further beams are added to them "below" = towards the note heads.*

[24, p. 13 p.] exhaustively discussed this problem of beam addition, without finding a simple rule as our *summaeInGenu*. Herma by Xenakis violates *summaeInGenu*, but instead follows

PROP. nota.trabes.inGenuCumPluribus: *If there is a vast majority of one stem direction over the other, this defines the meaning of the beams.*

This alternative leads (in m. 87, as executed) to the left version, while our rule would produce the right one, clearly demonstrating the change of meaning:

The insertion of a knee may conflict with *nota.trabes.inclinatioSignificans*. In Figure 8 the cases for the sequence "up-stem followed by down-stem" are labelled by the comparison results of the note heads and the stem ends. (For the opposite sequence the comparison operators signs must be inverted.)

PROP. nota.trabes.genuParadoxum: *in the case "up-stem followed by down-stem", the knee cases* $(<,>)$ *and* $(<,>>)$ *are called* paradox *because an inclining sequence of note heads / pitches is graphically realized by a declining slope of the beam. (Vice versa in the case "down-stem to up-stem").*

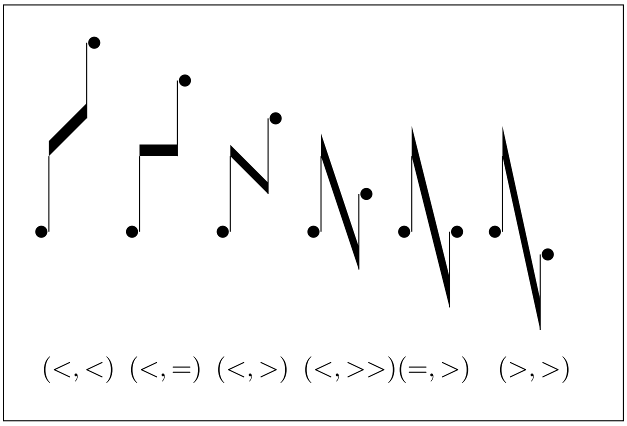


Figure 8. Categories of Knees

We leave it open if also the cases $(<,=)$ and $(=,>)$ shall be called "paradox". It holds:
a) If the lower note is stemmed down and the higher note is stemmed up, then no paradox shows up. (The tendency of the note heads is even *amplified* by the beam.)
b) Otherwise, if the absolute distance of the note heads is smaller than double the minimal stem length, then a paradox is unavoidable.

Paradox beams are explicitly forbidden by [19, p. 43].

While beams of the last category $(>,>)$ seem unpractical, they are indeed frequently required when more than two voices share a (piano) staff. See Brahms, Intermezzo op. 119 Nr. 1 m. 61: (Simrock, 1893, plate 10055):

4.7 Resolving Conflicts by Changing Height and/or Inclination

A rather modern means to resolve graphical beaming conflicts is to change the inclination or the height of the beam "underway", when crossing a particular stem S. Systematically this can be treated as breaking the beam into two fragments which overlap just at S. This is not one single break process, as in section 4.5, but two, namely immediately left and right of S. Then all beamlets resulting from these breaks are removed and both fragments are laid out independently. The last stem of the first and the first stem of the second resulting beaming aggregate must point into the same direction and will be unified in the final printed result. [14]

Such a result is represented in context of Table 2 by two elements of the sequence B according to the pattern

[13] While this indeed poses severe problems to formal definitions of syntax and semantics, it has no consequences in practice because only the *number* of beams is relevant, not their "undisturbed self-identity". This effect is called "non-deterministic determinedness" by [4, p. 298].

[14] These considerations do not cover to case of giving *two* stems (one in each direction) to the common notehead at S. This must be treated separately, because it possibly has stronger impact on the notated voice leading.

$\langle \ldots (x_0, y_0, X, Y_1), (X, Y_2, x_3, y_3), \ldots \rangle$. It can exhibit any combination of

PROP. nota.trabes.translataeInCaudis: *the distance of the top-level beam from the note head is different on both sides of the stem ($Y_1 \neq Y_2$).*
. . . angulusFractumInCaudae: *the top-level beams on both sides of the stem are not in 180° but in a different angle* $(Y_1 - y_0)/(X - x_0) \neq (y_3 - Y_2)/(x_3 - X)$.

See Figure 9 a) and b). The graphical syntax of these properties is modelled by *nota.graph.simplex.XIV* and *XV* in [4, p. 108]. The selection of allowed properties and combinations can be a further input parameter. The occurrence of neither would indicate inconsistent programming, because in this case the original rendering attempt (without any breaking) should have succeeded.

Both properties and their combination are rather frequent in classical avantgarde notation styles, see Stockhausen, Klavierstück X. But indeed *angulusFractumInCaudae* are found in much older prints and hand-writings. [1, p. 88 p.] gives two examples from 1690 (regrettably not specifying the source):

5. ASPECTS NOT COVERED

A mathematical re-modelling of inter-human symbol systems determined by history and culture is not and should not try to be exhaustive. But it should clearly circumscribe the areas of non-formalization. Whether extended formalization beyond these limits is sensible may be left to future work.

Some of the properties we have not (yet?) formalized:

PROP. nota.trabes.angulusFractumInterCaudis: *the top-level beam is bent at a point* between two *stems.*

This variant, see Figure 9 c), can be integrated into our model by breaking the aggregate between the stems, laying out the fragments independently and printing the resulting beams in a prolongated way, extending to their meeting point between the aggregates. Of course such a meeting point must exist, and research on its preconditions is necessary. This appears to be related to the problems of inter-aggregate dependencies and graphical harmonizations, see *nota.trabes.priorInfluit* etc. above.

PROP. nota.trabes.multaAdCaudam: *Multiple groups of beams are connected to the same side of the same stem.*

Different patterns of this kind can be found in advanced notation. A simple sub-kind is

PROP. nota.trabes.multaAdCaudam.perOrnamentum: *Additional beam groups extend only locally and indicate an ornament, i.e. a figure played in a sub-ordinated local organization of time.*

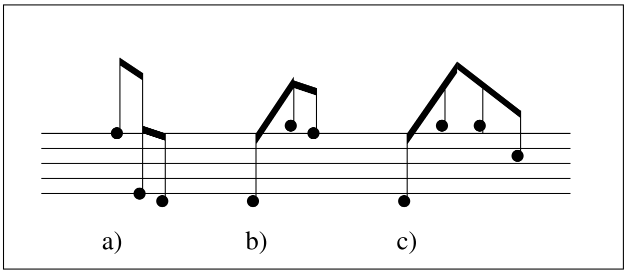
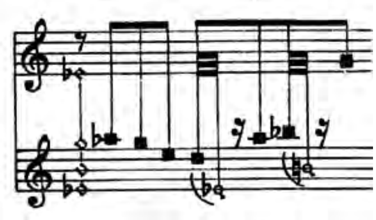


Figure 9. Displaced and Bent Beams

The following example (Lachenmann, Toccatina, p. 3, cited by [28] p. 24) is a such a simple case, where local "Nachschläge" are notated by a subordinate group of beams: [15]



PROP. nota.trabes.trabulaUtVexilium : *A beamlet is represented in the form of a flag, even when the stem has (long) beams.*

This is a mere graphical transformation. [20, p. 94] mentions Boulez, Le marteau sans maître, and rejects this kind of writing sharply.

6. CONCLUSION

We have presented a four-staged pipeline architecture for calculation and layout of a beam aggregate for a given meter, rhythm, and pitches. We have identified about sixty properties which influence the result. We have assigned statements from literature (which nearly exclusively is concerned with conventional CWN) to the different stages of processing.

The pipeline architecture helps to clarify input conditions, output specifications, tests of both, documentation, etc. considerably. Nevertheless we found two bypasses which do not fit cleanly into the sequential order, see Figure 2: (a) the number of note heads modifying the genuine beaming (an issue not treated in the literature so far), and (b) predetermined breaking points, prepared by semantic parameters to be used by the graphic layout process if required.

Table 3 in the Appendix shows where the cited standard works treat the different topics. Tables 4 and 5 list all properties found so far, referring to the corresponding section of this article.

Acknowledgments

Many thanks to Hartmuth Kinzler, Osnabrück, for a rich collection of beaming examples.

[15] Another important example is Stockhausen, Klavierstück X. But in its foreword special semantics of beams are explicitly defined anyhow.

7. REFERENCES

[1] T. Ross, *The Art of Music Engraving and Processing*. New York: Hansen Books, 1987.

[2] M. Lepper, M. Oehler, H. Kinzler, and B. Trancón y Widemann, "Diminuendo al bottom —- clarifying the semantics of music notation by re-modeling," *plosOne*, 2019. [Online]. Available: https://doi.org/10.1371/journal.pone.0224688

[3] M. Malt and B. Sluchin, "Geometric notation for time-bracket works, application and performance. the case of John Cage's Music for ___," in *Proceedings of the International Conference on Technologies for Music Notation and Representation – TENOR'20*, R. Gottfried, G. Hajdu, J. Sello, A. Anatrini, and J. MacCallum, Eds. Hamburg, Germany: Hamburg University for Music and Theater, 2020, pp. 114–119, http://tenor-conference.org/proceedings.html[20210102].

[4] M. Lepper, "De linguis musicam notare," Ph.D. dissertation, Universität Osnabrück, Osnabrück, 2021.

[5] D. A. Byrd, "Music Notation By Computer," Ph.D. dissertation, Indiana University, Bloomington, 1984.

[6] M. Gieseking, "Code-basierte Generierung interaktiver Notengraphik," Ph.D. dissertation, Universität Osnabrück, 2000.

[7] A. Fleischer, G. Mazzola, and T. Noll, "Computergestützte Musiktheorie: zur Konzeption der Software RUBATO für musikalische Analyse und Performance," *Musiktheorie*, vol. 15, pp. 314–325, 2000.

[8] G. W. Cooper and L. B. Meyer, *The Rhythmic Structure of Music*. Chicago: The University of Chicago, 01 1960.

[9] A. T. Cemgil, P. Desain, and B. Kappen, "Rhythm quantization for transcription," *Computer Music Journal*, vol. 24, no. 2, pp. 60–76, 2000.

[10] A. Cogliati, Z. Duan, and D. Temperley, "Transcribing human piano performances into music notation," in *Proc. of International Society for Music Information Retrieval Conference (ISMIR)*, J. Devaney, M. I. Mandel, D. Turnbull, and G. Tzanetakis, Eds., 2016, https://wp.nyu.edu/ismir2016/event/proceedings/.

[11] A. Ycart, F. Jacquemard, J. Bresson, and S. Staworko, "A supervised approach for rhythm transcription based on tree series enumeration," in *Proceedings of the International Computer Music Conference (ICMC 2016)*, H. Timmermans, Ed. Utrecht: Michigan Publishing, 2016, pp. 369–376.

[12] C. Agon, K. Haddad, and G. Assayag, "Representation and rendering of rhythmic structures," in *Second International Conference on WEB Delivering of Music*, C. Busch, M. Arnold, P. Nesi, and A. Schmuckler, Eds. Los Alamitos, CA, USA: IEEE Computer Society, 2002, pp. 109 – 113.

[13] M. Laurson and M. Kuuskankare, "From rtm-notation to enp-score-notation," in *Proceedings of Journées d'Informatique Musicale*, J. Clos, Ed. Montbeliard: L'ecole nationale de musique du pays te Montbeliard, 2003.

[14] A. Agostini and D. Ghisi, "A max library for musical notation and computer-aided composition," *Comput. Music J.*, vol. 39, no. 2, pp. 11–27, Jun. 2015. [Online]. Available: http://dx.doi.org/10.1162/COMJ_a_00296

[15] N. Didkovsky, "Java music specification language, v103 update," in *Proceedings of the International Computer Music Conference (ICMC 2004)*. Miami, Florida, USA: Michigan Publishing, 2004, pp. 742–745.

[16] M. Lepper and B. Trancón y Widemann, "metricsplit: An automated notation of rhythm aligned with metric structure," *Journal of Mathematics and Music*, vol. 13, no. 1, pp. 60–84, 2019.

[17] F. Gouyon and S. Dixon, "A review of automatic rhythm description systems," *Computer Music Journal*, vol. 29(1), pp. 34–54, 2005.

[18] A. C. Vinci, *Fundamentals of Traditional Musical Notation*. Kent, OH: Kent State University Press, 1985.

[19] H. Chlapik, *Die Praxis des Notengraphikers*. Wien: Doblinger, 1987.

[20] G. Read, *Music Notation: a Manual of Modern Practice*. New York: Taplinger, 1979.

[21] E. Gould, *Behind Bars: the Definitive Guide to Music Notation*. London: Faber, 2011.

[22] D. Taupin, R. Mitchell, and A. Egler, "MusixTEX – Using TEX to write polyphonic or instrumental music," 2002, www.texdoc.net/texmf-dist/doc/generic/musixtex/musixdoc.pdf[20200922].

[23] M. Lepper and B. Trancón y Widemann, "Metricsplit – Automated Notation of Rhythms, Adequate to a Metric Structure," Technische Universität Ilmenau, Ilmenau, Tech. Rep., Mar 2017. [Online]. Available: http://nbn-resolving.de/urn:nbn:de:gbv:ilm1-2017200259

[24] K. Stone, *Music Notation in the Twentieth Century: a Practical Guidebook*. New York: Norton, 1980.

[25] M. Lepper, M. Oehler, H. Kinzler, and B. Trancón y Widemann, "Crowded staves – rules for semantics and style of conventional multiple-voices-per-staff musical notation," in *Proceedings of the TENOR 2020 conference*, Hamburg, 2020, pp. 1–12, https://www.tenor-conference.org/proceedings.html.

[26] H. W. Nienhuys and J. Nieuwenhuizen, "Lilypond, a system for automated music engraving," in *Proceedings of the XIV Colloquium on Musical Informatics*, Firenze, 2003, pp. 167–172.

• Genuine beams:
[18, p. 27 pp.] [1, p. 91 pp.] [19, p. 43 p.] [19, p. 46 pp.] [20, p. 80 pp.] [21, p. 153 pp.]
• Modified genuine beams:
(no predecessors)
• Stem direction of beam aggregate:
[1, p. 94 pp.] [20, p. 88] [21, p. 24 p.] [27, p. 154]
• Beams crossing pauses:
[18, p. 49] [19, p. 46] [24, p. 15 p.] [20, p. 88, 213]
• Height of beam symbols:
[1, p. 94 pp.][1, p. 119 pp.][19, p. 41][24, p. 9] [20, p. 80][21, p. 17][27, p. 42 p.]
• Significance of the beam's steepness:
[19, p. 42] [18, p. 45 p.] [21, p. 22 pp., 169 pp.] [27, p. 155, 168 pp.]
• Value of the beam's steepness:
[18, p. 45 pp.] [1, p. 97 pp.] [1, p. 115 pp.] [19, p. 42][21, p. 17 pp.][27, p. 155 pp.]
• Knees:
[1, p. 126 p.] [19, p. 43 p.] [24, p. 12 pp.] [29, p. 56] [20, p. 86, 88 pp.] [21, p. 26] [27, p. 93, 135, 153]
• Relation to stafflines / micro positioning:
[18, p. 43 p., 47] [1, p. 98 pp.] [1, p. 119 pp.] [19, p. 41 p.] [24, p. 9 pp.][21, p. 17 pp.][27, p. 25 p.,42, 161 pp.]
• Beams for polymetrics:
[24, p. 116 pp.][21, p. 171, 175 pp.] [20, p. 170 pp.]
• Feathered Beams:
[19, p. 47] [20, p. 94] [24, p. 124, 141] [21, p. 158]

Table 3. Synopsis of the Literature w. r. t. the Discussed Issues

[27] H. Wanske, *Musiknotation – Von der Syntax der Notenstichs zum EDV-gesteuerten Notensatz*. Mainz: Schott, 1988.

[28] C. Dimpker, *Extended Notation – The depiction of the unconventional*. Berlin: Lit Verlag, 2013.

[29] K. Hader, *Aus der Werkstatt eines Notenstechers*. Wien: Waldheim-Eberle, 1948.

A. APPENDICES

Table 3 gives a synopsis of the discussion in standard literature. Tables 4 and 5 list all properties appearing in this article.

A.1 Polymetric Constellations Expressible by Beams

Following the classification grid in [4], a simple polymetric situation can be characterized by a quadruple indicating the relations between (1) the start points of the metric pattern, (2) the physical lengths, (3) the inner metric structure, and (4) the notated tempo, i.e. the relation from notated values to physical time.

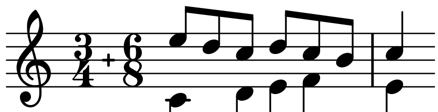
Many simple constellations can be expressed by beams alone. Examples without further comments:

$(\neq, =, =, =)$ (Petterson, Sinf 11)

$(=, \neq, \neq, =)$ (Petterson, Sinf 10)

$(=, =, \neq, \neq)$ (Beethoven op. 111)

$(=, =, \neq, =)$ (constructed example)



$(=, =, =, \neq)$ (Mozart Klaviersonate K 457)
nota.trabes.unificataIndicantRubatum[Kinzler]

• *nota.trabes.utVexilia*
= basic principle how beams express duration, see section 2.1.
• *nota.trabes.notaeUtNota*
= basic principle that a single note and a replacing group behave identically towards their neighbors, see section 2.1.
• *nota.trabes.alteraNatas* : $\mathbb{P}(\mathbb{N} \times \mathbb{Z} \times \mathbb{N} \times \mathbb{N})$
= collection of rules to modify the genuine beaming, see section 2.2
• *nota.trabes.sopraPausam* : $\{\mathsf{perCaudulam}, \mathsf{transiens}, \mathsf{separans}\}$
= way how beams cross pauses, see Section 2.4.
• *nota.trabes.significantVocem*
= a beam fulfills also the role to indicate voice identity, thus voice leading.
• *nota.trabes.trabulaeContraIdem*, *...trabulaeOmnesContraTrabes*, *...trabulaNonSubTrabem*, *...trabulaSola*
= local transformations of beamlets, see Section 2.4.
• *nota.cauda.significat.vocem*
= the stem direction indicates the notational voice.
• *...instrumentum/accentum/manum/modumAgitur*
= the stem direction indicates the employed instrument (with percussion) / the active hand / a certain way of sound production.
• *nota.cauda.significat.nihil*
= the stem direction is free and can be determined according to graphical needs.
• *nota.cauda.significat.vocem.trabsSeparataCaudaMutata*
= when the stem direction changes midways, the beam is broken.
• *nota.cauda.significat.vocem.trabsSeparataLineaMutata*
= when the voice changes the staff midways, the beam is broken.
• *nota.voces.unaUtDue_pausasPerdatas*
= one single voice with two kinds of events (e.g. by two hands) is notated as *two* voices, where the events in the one voice are read as pause symbols for the other.
• *nota.trabes.extera.separans*
= external parameters make a beam break.
• *nota.trabes.extera.ligans*
= external parameters make two separated beams join.
• *nota.trabes.cumVerborum.syllabis/nominibus/lineis*
= the beams are broken according to the syllables/words/lines of the sung text.
• *nota.trabes.separatae.cumMelo/cumLigato*
= the beams are broken according to motif structure/legato execution.
• *nota.trabes.ligataeContraNates*
= beams are connected against the genuine beams prescribed by the meter.
• *nota.metraMulta.perTrabem*
= the beams are used to clarify a multi meter situation (shifted or shifted meter in different voices=.
• *nota.trabes.cumPositioneManu*
= the beams follow the positional changes ("jumps") of a hand on a keyboard.
• *nota.trabes.accelerans* : $\mathsf{Seq}(\mathbb{Q} \times \mathbb{N} \times \mathbb{Q})$
= tempo curve for compressed and expanded beam heights and distances.
• *nota.trabes.inclinatioSignificans*
= the steepness of the beam is related to the distribution of the pitches, or even to the gesture of the motif.
• *nota.trabes.ponuntCaudas*
= the fact that the notes are beamed together influences the stem direction.
• *nota.trabes.visio*
= the final visual layout of the beams in its graphical context.
• *nota.trabes.inLineolas*
= the relation of the particular beams to the lines of the staff.
• *nota.trabes.cauda.max/maxInterior/min/minInterior* : $\mathbb{Q}$,
nota.trabes.maxInclinatio/maxAltitudo : $\mathbb{Q}$
= input parameters for a layout algorithm.
• *nota.trabes.conditionesConfligentes*
= the case that the input parameters conflict and prevent a (simple) solution.
• *nota.trabes.vocesConfligentes*
= the case that the spatial requirements of another voice (in the same staff) prevents a (simple) solution.

Table 4. All Properties Found for Beams and Their Transformations—Part I

• *nota.trabes.tresInTresLineolis*
= formula from copper engraving which allows to put three beams into two spaces of a staff.
• *nota.trabes.sineLineolis*
= the case that beams outside a staff are treated differently than inside.
• *nota.trabes.subLineola*
= beams which hold contact to one and the same staff line throughout. = "creeping beams".= "schleichende Balken".
• *nota.trabes.priorInfluit/successorInfluit*
= context dependency: a change in the input data for a preceding/following beam aggregate influences the layout.
• *nota.trabes.genus.inLineae*
A "knee" ("exceptional beams", German "Knie") is contained in the graphic area of a staff's lines.
• *nota.trabes.genus.interLineas*
A knee is contained in graphic space between two staves.
• *nota.trabes.numerusGenuum*
= the number of knees in one particular rendering.
• *nota.trabes.numerusInGenui*
= the number of beams which "go through" the knee.
• *nota.trabes.numerusInterGenibus*
= the number of stems between two knees/a knee and the end of the aggregate.
• *nota.trabes.caudaVersusCaudam*
= whether a stem in both directions appears at the same point of a beam.
• *nota.trabes.summaeInGenu*
= the rule that the top-most beams only cross a knee.
• *nota.trabes.inGenuCumPluribus*
= the rule that the majority side of stems decides the meaning of the beams across a knee.
• *nota.trabes.genuParadoxum*
= that the beam in a knee aggregate goes contrarily to the pitches.
• *nota.trabes.translataeInCaudis*
= that the beams on two sides of a stem start at different distances.
• *nota.trabes.angulusFractumInCaudae*
= that the beams on two sides of a stem start in different angles.
• *nota.trabes.angulusFractumInterCaudis*
= that a beam makes a "turn" between two stems.
• *nota.trabes.multaAdCaudam*
= multiple groups of beams at the same side of a stem.
• *nota.trabes.multaAdCaudam.perOrnamentum*
= multiple groups of beams at the same side of a stem, but only for sub-ordinated, local time with fast events.
• *nota.trabes.trabulaUtVexilium*
= beamlets are printed as flags, even under beams.

Table 5. All Properties Found for Beams and Their Transformations—Part II

SONIFICATION, MUSIFICATION AND DRAMAFICATION OF ASTRONOMICAL DATA IN THE MULTIMEDIA PRODUCTION "A SPACE JOURNEY"

Georg Hajdu
Hochschule für Musik und Theater
Hamburg, Germany
georg.hajdu@hfmt-hamburg.de

Xiao Fu
Hochschule für Musik und Theater
Hamburg, Germany
xiao.fu@hfmt-hamburg.de

ABSTRACT

In this paper, we will discuss, in the context of a recent multimedia production at the Hamburg University of Music and Drama, the notions and concepts surrounding the musical treatment of astronomical data. We will distinguish between the three categories of sonification, musification and dramafication, forming a continuum in terms of how accurately the underlying data are represented and give examples for their application in the multimedia production.

1. INTRODUCTION

A Space Journey – Perspectives of the Unknown is a multimedia project instigated by Georg Hajdu which premiered on October 21, 2022 at the large hall (Forum) of the Hamburg University of Music and Drama (HfMT). As a collaboration between the HfMT multimedia department, its theatre academy as well as the Hamburg University astrophysics department, namely Marcus Brüggen and his team, it involved approximately 80 persons ranging from composers (teachers and doctoral students), directors (bachelor students), scientists (teachers and doctoral students), musicians, stage designers, video artists, sound designers, to actors (including a ventriloquist) and technical personnel. The idea was born out of a discussion as to how to best conclude a research and translation project which, in 2019, had already sponsored the Symphony in the St. Pauli Elbe Tunnel, a large-scale networked music performance project with 144 musicians reading animated notation [1]. Marking the end of the 5-year project, this production was to use the items acquired and installed during this period, including a 146-speaker Meyer Sound Constellation system, 60 m2 LED video wall components, nearly 50 iPads used for animated notation and a novel Bohlen-Pierce contra clarinet. Eight teams, each consisting of a composer, a theater director and an astronomer were formed during the second quarter of 2021 under the direction of the dramaturg Elise Schobeß and the director Ron Zimmering. The aim was to create a diverse yet coherent dramaturgy consisting of a sequence of eight 10-minute scenes, each on a different topic related to fundamental questions of astronomy such as black holes, the big bang, the fate and structure of the universe, exo-planets, etc. (Table 1). We soon also decided on forming a large ensemble of eleven Bohlen-Pierce instruments (Table 2) and an eight-voice mixed choir (SSAATTBB). We decided on the Bohlen-Pierce scale for mainly two reasons: Firstly, it has been studied and put into practice at the HfMT since 2007. Various instruments had been built or repurposed to this end while software had been developed to meet the challenges posed by reading and playing music in this scale [2]. Secondly, due to its alienness to the unexperienced listener it functions well as a metaphor of the extra-terrestrial. Because of its derivation from the just twelfth instead of the octave (3:1 vs. 2:1) and a set of exclusively odd-integer-ratio intervals, the study of the scale can also be seen as a viable analogy to xenobiology and the exploration of alternative hereditary material.

	Composer	Title	Music notation and delivery
1	Georg Hajdu	Solaris I	Animated notation, audio score
2	Aigerim Seilova	hidden darkness	Animated notation
3	Todd Harrop	Quarks and Queries	Static parts
4	Goran Lazarević	Call of the Void	Graphic notation
5	Georg Hajdu	Solaris II	Animated notation, audio score
6	James Cheung	Message in a boomerang	Animated notation
7	Greg Beller	UME – Unidentified Missing Encounter	Static parts
8	Xiao Fu	100 Milliarden Sonnen	Animated notation, audio score
9	Benjamin Helmer	Kein Respekt den Sternen	Static part
10	Georg Hajdu	Solaris III	Static parts, animated notation, audio score

Table 1. The order of the compositions in A Space Journey

2. HISTORICAL CONTEXT OF MUSIC RELATED TO ASTRONOMY

Since antiquity, there has been a tendency to relate the (presumed) inner workings of the universe to music. Kepler coined the term Harmonice Mundi [3] upon demonstrating a mathematical relationship between planetary motion and of musical harmony. *Music of the spheres* remained a trope throughout modernity, both in a metaphorical and a concrete sense. Gustav Holst in his suite *The Planets* (1914-16) expanded on Wagnerian harmony to create a musical language denoting the extraterrestrial—a language that remains largely intact in Hollywood productions such as *Star Wars* (episodes created since 1977) with music by John Williams. The "apocalyptic" composers Scriabin (1871-1915) and Stockhausen (1928-2007) also had a strong affinity towards the cosmic, manifested by pieces such as *Vers la flamme* (1914), *Sternklang* (1969-71) and *Sirius* (1975-77). In contrast, composer and theorist Clarence Barlow correlates the gaps in the asteroid belt density between Mars and Jupiter to the harmonic intensity of orbital intervals obtained by his *harmonicity* formula without actually creating a musical illustration for this [4]. In 2007, Hajdu wrote the piece *Beyond the Horizon* for narrator, two Bohlen-Pierce clarinets and synthesizer on a text about the accelerated expansion of the universe due to dark energy by Scientific American authors and astronomists Lawrence M. Krauss and Robert J. Scherrer [5]. Sonification of astronomical data by NASA and others often serves a dual purpose: representing data non-visually and aestheticizing them at the same time. Prominent examples are the sounds obtained from the collision of two black holes and the sonification of the remnants of a supernova explosion, employed by Fu in her scene.

3. SONIFICATION VS. MUSIFICATION VS. DRAMAFICATION

I'm not like Stockhausen, I'm not creating music, it's already there ...

-- Morton Feldman

Sonification, musification and dramafication[1] can be conceived as approaches towards extra-musical data[2] located on a continuum but functionalizing the source materials in different ways. While sonification relies on an *auditory display* which needs to represent the original data as accurately as possible [6], musification permits adjustments and modifications of the data with the aim to represent the *structural principles* inherent in the data without necessarily having to spell them out accurately [7]. In contrast, dramafication permits the transformation and distortion of the source material subjecting it to a *musical and/or dramaturgical narrative*.

The term "sonification" was introduced in the 1990s to denote the "transformation of data relations into perceived relations in an acoustic signal for the purposes of facilitating communication or interpretation" [6]. Thus, data sonification aims to create "an appropriate mapping between data and sound features in a sonification display". Yet, due to its very nature, a sound whether to be used to map data or not, always carry an aesthetical dimension[3], with the propensity of interfering with the perception of the sonified data [9]. To reduce unwanted effects, special care needs to be taken when creating such mappings.

In contrast to sonification, musification pays homage to the aesthetics of sound and incorporates "data features that represent more traditional elements of a musical work such as melody, harmony, and rhythm" [10].

In dramafication, finally, the connection between the original data set and the musical outcome is blurred or overridden by the dramatic effect. It ends up having the similar significance as a musical cryptogram found in the works of Olivier Messiaen and other composers [11].

	Instrument	Notation
1	BP soprano clarinet	fingering notation
2	BP tenor clarinet	fingering notation
3	BP contra clarinet	fingering notation
4	trombone	eighth-tone notation
5	violin	scordatura with tablature
6	viola	scordatura with tablature
7	cello	scordatura with tablature
8	BP 9-string electric guitar	tablature
9	fretless bass	scordatura with tablature
10	keyboard	BP keyboard notation
11	percussion: • BP glockenspiel • BP metallophone	 R clef N clef

Table 2. Instrumentation of the Bohlen-Pierce orchestra

4. GEORG HAJDU: SOLARIS

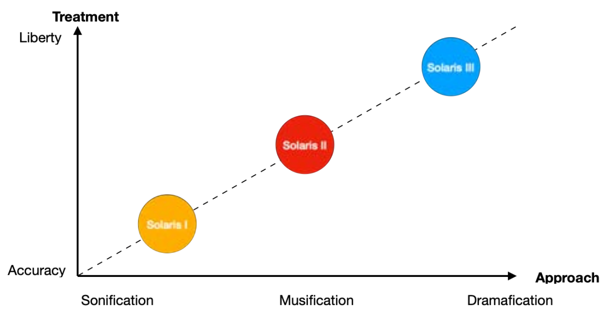


Figure 1. The three parts of Solaris treat the original astronomical data with a varying degree of artistic freedom.

[1] We are using term *dramafication* instead of *dramatisation* for two reasons: (a) to align the method linguistically with the terms sonification and musification, and (2) to distinguish it from *dramatization* which is defined as the process of adapting a novel or presenting a particular incident in a play or film while stressing that the data treatment is motivated by an overarching musical principle.

[2] The data don't need to be extra-musical: much of Lindsay Vickery's work is concerned with the (re)sonification of sonic materials.

[3] In *Doktor Faustus*, German novelist Thomas Mann ironically referred to this quality of music as "Kuhwärme" (bovine warmth).

In a team consisting of director Ron Zimmering, astronomist Kathrin Böckmann (later joined by Denis Wittor) and composer Georg Hajdu, it was decided to focus on turning the filament structure of the universe into music and use excerpts from Stanisław Lem's novel *Solaris* to be narrated alongside the music. Lem's popular work is about the failed mission to comprehend the intelligence of a large ocean covering the planet Solaris and the solipsistic nature of the human quest for alien intelligence in general.

The filament structure of the universe consisting of walls of gravitationally bound galaxy superclusters was discovered in the late 1980's. They are thought to represent the inhomogeneities of matter present during the first moments of the Big Bang. The filaments form the most complex structure in the universe, followed in complexity only by the human brain [12]. Both show a high degree of self-similarity and fractality, a quality that, to a lesser degree, can also be attributed to (tonal) music. We can therefore form a thematic triangle for which Lem's novel (excerpted by Zimmering) serves as a point of departure. Instead of a single 10-minute presentation, *Solaris* was split into three parts forming a prolog, an intermezzo and an epilog, thematically moving from the macrocosm of the filaments to mesocosm of Solaris' ocean-brain and the microcosm of neural activity, thus framing the seven other scenes.

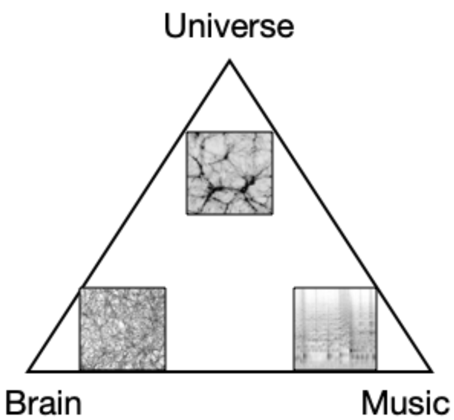


Figure 2. The universe, the brain and music are highly hierarchical and self-similar in nature and thus lend themselves to the creation of artistic analogies. The figure displays a simulation of the distribution of visible matter in the universe (top), neurons grown in a petri dish (left) and a sonogram of musical sound (right).

4.1 Workflow

The point of departure was a cross section of a simulation of the distribution of visible matter in the universe created by Wittor, consisting of a cube of 900x900x900 pixels. The x axis of the cross section was to represent ten minutes of music unfolding from left to right (Figure 3), while the y axis was mapped to a musically meaningful range of six octaves (C1 to C7).

Hajdu used software to

1. manually map the filaments to lines of varying lengths and widths and
2. isolate the largest discernible galaxy clusters and apply them to ovals with varying radii.

In a second step, the data obtained were mapped to frequency (y coordinates), loudness (width/radii) and time (x coordinates) and converted into SPEAR text format [13]. In SPEAR, the data files were transcoded to SDIF, the Description Interchange Format [14], and read into Macaque [15] allowing the conversion of the partial tracks into music notation. In the MaxScore Editor, the "approximate" feature of the n-TET Entry Tool was used to bend the pitches to the nearest Bohlen-Pierce notes [16].

```
par-text-partials-format
point-type time frequency amplitude
partials-count 373
partials-data
0 2 0. 2.
0. 2019.708 0.032 2. 2019.708 0.032
1 2 452.821 454.821
```

Table 3. SPEAR text format used to translate geometric data into partial tracks.

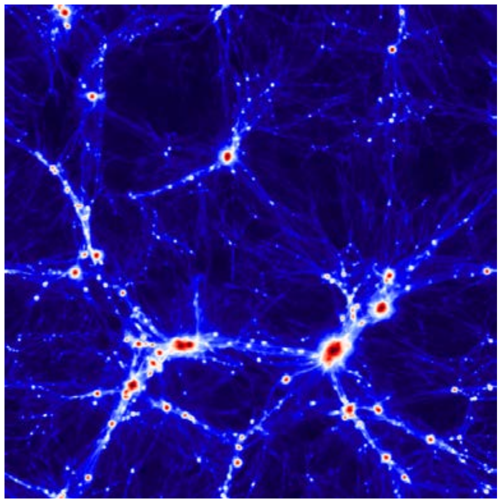


Figure 3. False-color representation of the distribution of the intra-cluster medium in the universe (image from a simulation by Denis Wittor).

Diagonal filaments were transcribed into portamenti whose rendering required a new *polybend* message to be implemented in the *MaxScore* object. To this end, a *pb* editor was added to *Picster Expressions* [15] permitting the generation of a multi-channel breakpoint functions which at runtime are translated by the *maxscore.makenote* abstraction into a sequence of *polybend* message (see 4.2.1).

4.2 The parts

4.2.1 Solaris I

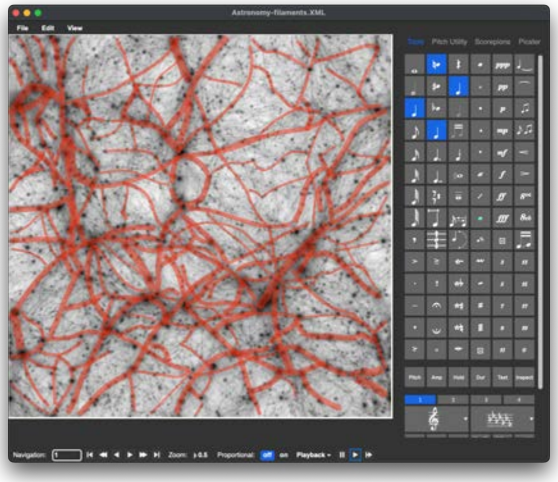


Figure 4. A map of the filaments, parsed manually in the MaxScore Editor.

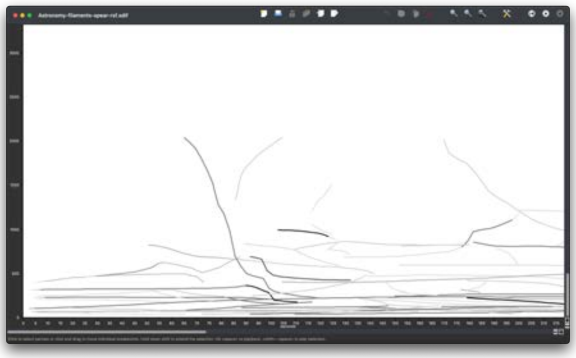


Figure 5. Translation of the filament structures shown in Figure **1** into partial tracks.

The first part, scored for narrator, eight singers and playback, has a duration of three minutes. By transcribing the filaments into partial tracks and ultimately into notation, Hajdu obtained 32 voices which were played back with glass harmonica sounds and recorded into 32 individual tracks in Ableton Live, to be spatialized on the HfMT's Meyer Sound Constellation system and played back in sync with the voices that fell into the singable range of D2 to A5 and sung by the choir.

As the score is in the Bohlen-Pierce scale, intonation posed a challenge which we met by creating a playback device in JavaScript formatting messages for drawsocket's tone.js oscillator method [17].

Figure 6. Portamenti resulting from the transcription of the filaments are rendered graphically and musically in MaxScore using Picster expressions.

For this, the maxscore.makenote object passes two types of event messages to the tonejs-osc JavaScript object (Figure 7):

- *note*, with the arguments pitch [MIDI cents], velocity, channel, e.g. "note 6025 89 29"
- *polybend*, with the arguments reference_pitch, pitch_deviation [cents], velocity deviation, channel, e.g. "polybend 6025 -100 -24 29"

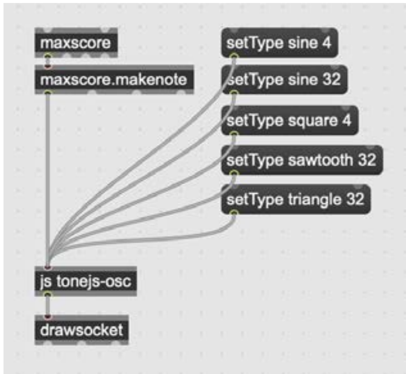


Figure 7. The js tonejs-osc object translates MaxScore event messages into dictionaries served by the drawsocket abstraction to connected browsers. The object also responds to a *setType* message setting the waveform for the synthesis performed on the browsers.

During the performance, the singers held iPads, wirelessly connected to the drawsocket server and received tone.js messages, the result of which they could via earbuds plugged into the iPads' audio jack. The score was rendered in metered proportional notation and fanned out to the performer (or groups thereof) by defining corresponding staffgroups. MaxScore now distinguishes between *metered* (barlines, flags and beams are shown) and *non-metered* (only noteheads are shown with lines extending from them to indicate the duration of the event) proportional notation.

To ensure that the singers heard their sounds well before the corresponding events hit the playhead, the score animation was offset by 500 msec ("pre-delay").

A flight through the cube (representing the universe) was prepared by the astrophysics student David Smolinski and shown on 20 individual LED screens hanging on a computer-controlled fly system. Being the first contribution to the show, the panels were slowly lowered one by one to create the sense of an introduction.

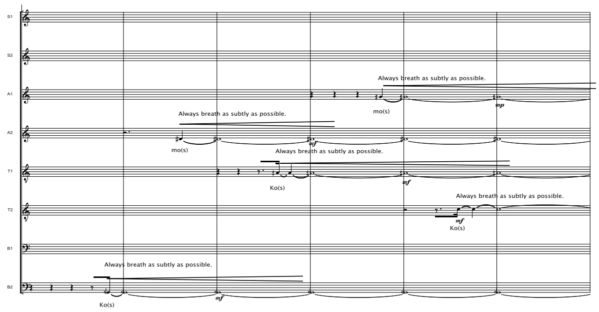


Figure 8. The opening of Solaris I in metered proportional notation. For dramaturgical reasons, the entrances of the voices were modified so they would enter one by one at nearly equal distances. A few of the pitches were adjusted to improve the resulting harmony between the voices.

Figure 9. A frame from the live stream of the premiere showing the performance of Solaris I. The eight singers holding their iPads are positioned in front of the LED video wall panels mounted on a computer-controlled fly system.

4.2.2 Solaris II

The second part lasts for 3"20" and is scored for narrator, choir and Bohlen-Pierce orchestra. In this part, Hajdu first catalogued the largest discernable objects (galaxy clusters) and translated them into a set of partial tracks with y-centers representing pitch (Figure 10). Similarly, the partial tracks were transcribed into notation and bent to the nearest Bohlen-Pierce scale tones. Using this approach, Hajdu's intention was to create a sense for the density distribution in the universe. Dense moments (frequent attacks) would alternate with moments of low density (prevalence of sustained notes) representing the voids between the filaments—creating a sonic landscape akin to Morton Feldman's later pieces (such as string quartet no. 2), yet not shying away from occasionally harsh wolf tones obtained by the mapping.

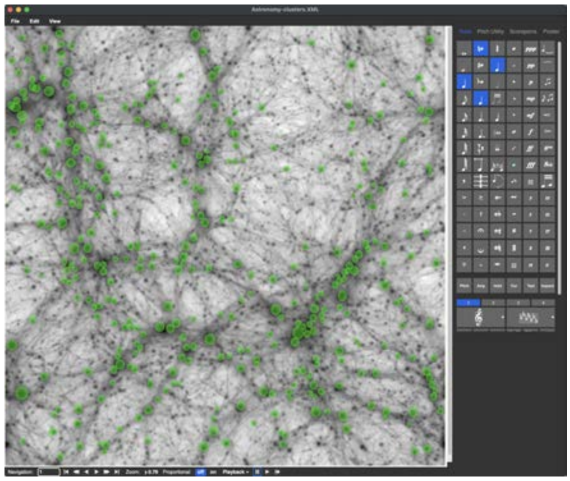


Figure 10. Major galaxy clusters were encircled and mapped to time and frequency.

Figure 11. The circles shown in Figure 10 were translated into short partial tracks and transcribed into notation (Figure 12).

Figure 12. Page 7 of the score of Solaris II.

Hajdu decided to treat dynamics and orchestration freely to highlight the inherent dramaturgy of single notes, dyads and chords resulting from the transcription. In this part, the music was displayed on the iPads of both the singers and the instrumentalists. As with the first part, the exact pitches were delivered to the singers via tone.js but this time with a pre-delay of 500 ms, just in time for their entrances.

The music was accompanied by (a) the narration about the mystery of Solaris' ocean, (b) a soundscape consisting of a multi-channel rendering of the sound of ocean waves enveloping the audience and (c) a video displaying simulations of ocean waves done by Nicolas Desmars, a doctoral student of hydrodynamics and ocean engineering at the Hamburg technical university. The flying system was programmed so that the panels individually performed smooth up and down motions, thus mimicking the behavior of waves.

Figure 13. A frame from the performance of Solaris II with the singers placed between the LED video wall panels displaying the ocean-wave simulation by Nicolas Desmars.

4.2.3 Solaris III

Solaris III, 3'40" in duration and scored for narrator, choir and Bohlen-Pierce orchestra, is made of three sections with the same proportions that the Solaris parts form with each other (i.e. 9:10:11). In the first section, the music returns to the original material of the filaments but replaces the sustained notes by fast arpeggios. Hajdu developed a variable arpeggiator which steps through the frames of the SDIF file and applies patterns which Hajdu derived from Clarence Barlow's indispensability function [18]. These patterns when applied to pitch (or in this case to an index number representing a pitch within a particular frame) create interesting and aesthetically pleasing progressions which Hajdu has used in several compositions of his.

While in the first section, the patterns are being played by the guitar, bass guitar and percussion (forming a sort of rhythm group), the second section adds chords with variable lengths performed by the choir and the other instruments. These chords are derived from SDIF frame and carefully voiced to avoid highly dissonant wolf tones occurring in the Bohlen-Pierce scale. After a climax in which the clarinets play ascending and descending scales, the rhythm group performs a strong rallentando and segues into the coda consisting of four sustained fff chords representing the filament structure at specific time points. They are contrasted by repeated notes played at heart rate by the glockenspiel which also concludes the piece and thus the entire show. The music is accompanied by abstract animations created by video artists Janina Luckow alluding to neural network chatter, finally transitioning to imagery that can be interpreted as a supernovae explosion.

As the synchronicity of parts when scrolling proportional scores with very fast pulses is problematic due to small differences in response time between the browser instances within the WiFi network, we decided to use a hybrid approach: The conductor was to listen to a click track and conduct the instrumentalists while the singers would read their scrolling parts—still depending on the tones delivered with the 500 ms pre-delay. We were positively surprised about how stable and reliable this approach turned out to be.

Figure 14. Section I from Solaris III is scored for BP 9-string guitar, fretless bass with scordatura, BP glockenspiel, BP metallophone and synthesizer. The guitarist requested encircled string numbers to be written above the notes for easier orientation. This feat was automated with the *Add String Index To Tablature* Scorepion operating on tablature notation (**Figure 15**).

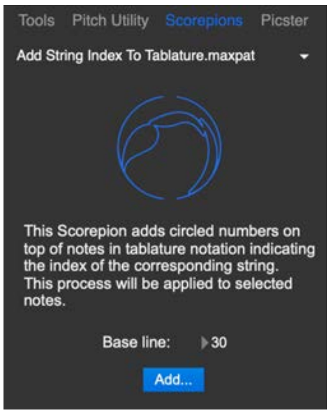


Figure 15. GUI of the *Add String Index To Tablature* Scorepion for the MaxScore Editor.

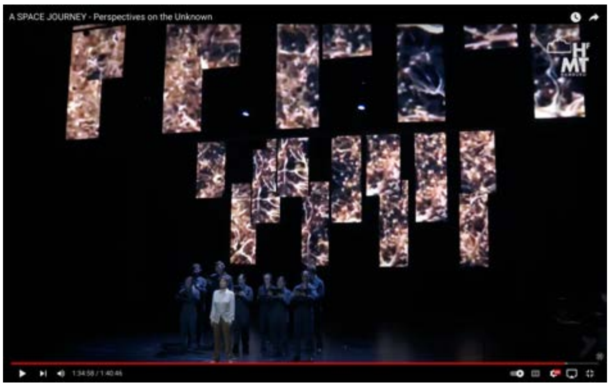


Figure 16. A moment from the dramatic ending of Solaris III with narrator and choir standing below the fly system.

5. XIAO FU: 100 MILLIARDEN SONNEN

The space - between self and the cosmos, birth and death - between each unique experience of life.
A star, a human being, a piece of sky, all live as expressions of creation, and all die as the same.
A star gathers into itself inner tension stirs transformation change explodes through its boundaries what once defined releases itself into the undefinable.

-- Sara Ezzell, choreographer

100 Milliarden Sonnen (*100 Billion Suns*) is composed in four parts: Birth, Life, Death and Sublimation and is performed by a women's choir, three BP clarinets, trombone, violin, viola, cello, percussion, tape and dancer (Sara Ezzell). It makes use of the results of the sonification project led by staff of NASA's Chandra X-ray Observatory and the Universe of Learning. She chose the sonification of Tycho's Supernova Remnant performed in the optical range (other wavelengths have also been subjected to sonification)[4].

Starting in 2012, when Fu commenced her collaboration with the architecture department of the Hamburg HafenCity University (HCU) on various space-sound installations, she developed a technique which takes a sound design as a starting point and derives the instrumental parts from it. The compositions are mostly very slow, but with a lot of subtle changes within the long tones. Stylistically, it borrows from pieces by Dutch composer Louis Andriessen: combining long tones and fast, small, moving figures. This results in an evolving harmonic movement and thus in the build-up of tension.

In *100 Milliarden Sonnen*, her point of departure was the Tycho's Supernova Remnant sonification as an *objet trouvé* which she dramafied by freely choosing and orchestrating the pitch material she obtained from it.

For this, she developed the following workflow:

1. Audio extraction
2. Audio to MIDI conversion in Ableton Live
3. MIDI import into MaxScore
4. Adjustment of the pitch material to the BP scale
5. Orchestration of select notes for the choir and BP ensemble

In the coda of her composition (approximately nine minutes into her piece), the original sonification is presented as a kind of afterthought and revelation of the source material.

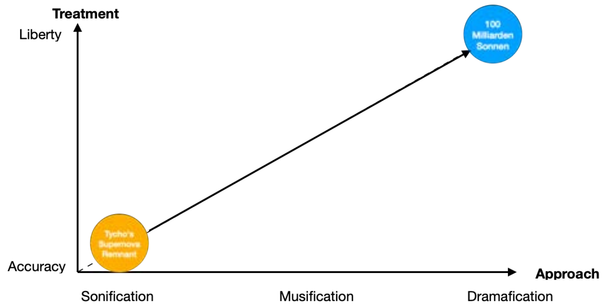


Figure 17. Path taken in *100 Milliarden Sonnen* by Xiao Fu dramafying an existing sonification obtained by NASA's Chandra X-ray observatory.

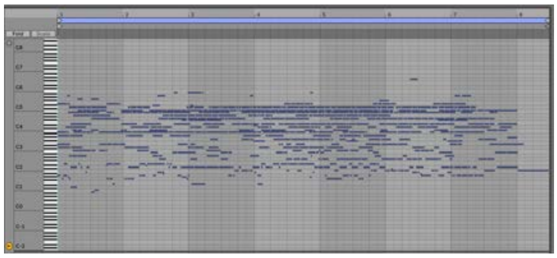


Figure 18. Display of the audio to MIDI conversion of the Tycho's Supernova Remnant (optical) sonification performed in Ableton Live.

[4] https://chandra.si.edu/photo/2021/sonify4/

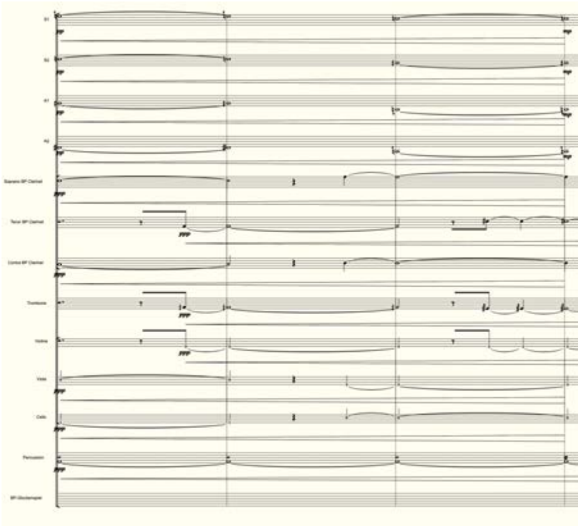


Figure 19. Score obtained from dramafication of an existing sonification.

Figure 20. Choreographer and dancer Sara Ezzell performing in front of the LED video wall.

Like Hajdu in Solaris II, Fu used the drawsocket server to distribute the parts to the performers and feed the audio score to the four singers via tone.js. The choreography was developed by Sara Ezzell while the video work was created by Janina Luckow.

6. CONCLUSION

This paper aims to present various approaches to the sonification of astronomical data. Whereas, originally, sonification aims to represent data on an auditory display as accurately as possible, we have presented several cases in which their musical transformation serves a dramatical purpose which takes them out of the original context and recontextualizes them within the framework of a large multimedia production. While the director Ron Zimmering praised the expressiveness of the sonified material, the audience response was overwhelmingly positive, which also convinced the astronomists who admitted that they had been quite skeptical at the outset of the project and now were suggesting a sequel to it.

Acknowledgments

We would like to acknowledge the Federal Ministry of Education and Research in Germany (BMBF) and the Hamburg Authorities for Science, Research, Equality and Districts (BWFGB) for their support of this research and thank the Forum team at the Hamburg HfMT for their continued help and support. We also gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for supporting this project by providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS Supercomputer JUWELS at Jülich Supercomputing Centre (JSC), under project no. hhh44.

7. REFERENCES

[1] G. Hajdu and R. Gottfried, "Networked music performance in the St. Pauli Elbe Tunnel," in *Proceedings of the International Conference on Technologies for Music Notation and Representation (TENOR'19)*, Melbourne, 2019.

[2] N.-L. Müller. *The Bohlen-Pierce Clarinet: Theoretical Aspects and Contemporary Applications.* Dissertation. BoD–Books on Demand, 2020.

[3] J. Kepler, M., Caspar, W. von Dyck, and F. Hammer, *Harmonice mundi*. CH Beck'sche Verlagsbuchhandlung, 1940.

[4] K. Barlow. *The Ratio Book: A Documentation of The Ratio Symposium, Royal Conservatory, The Hague, 14-16 December 1992*, vol. 43, Feedback Studio, 1999, pp. 14-15.

[5] L.M. Krauss and R.J. Scherrer, , "The end of cosmology?" *Scientific American*, *298* (3), 2008, pp. 46-53.

[6] G. Kramer, ed., "Auditory Display: Sonification, Audification, and Auditory Interfaces." In *Santa Fe Institute Studies in the Sciences of Complexity. Proceedings vol. XVIII.* Addison-Wesley, 1994.

[7] D. Williams, (2016) "Utility Versus Creativity in Biomedical Musification", In *Journal of Creative Music Systems.* 1(1).
doi: https://doi.org/10.5920/jcms.2016.02

[8] G. Kramer et al. Sonification Report: Status of the Field and Research Agenda, [Online] Available: https://www.icad.org/websiteV2.0/References/nsf.html

[9] P. Vickers, B. Hogg and D. Worrall, "Aesthetics of sonification: Taking the subject-position." In *Body, sound and space in music and beyond: multimodal explorations,* 2017, pp. 89-109. Routledge.

[10] A.D. Coop, "Sonification, musification, and synthesis of absolute program music." In *Proceedings of the 22nd International Conference on Auditory Display*, 2016.

[11] E. Sandzer-Bell. What is a Musical Cryptogram? [Online] Available: https://www.audiocipher.com/post/musical-cryptogram

[12] F. Vazza, “On the complexity and the information content of cosmic structures.” *Monthly Notices of the Royal Astronomical Society* vol. 465, 2017, pp. 4942-4955.

[13] M. Klingbeil, “Software for spectral analysis, editing, and synthesis.” In *ICMC: International Computer Music Conference*, 2005.

[14] D. Schwarz and M.C.Wright, “Extensions and applications of the SDIF sound description interchange format.” In *ICMC: International Computer Music Conference,* Berlin, 2000.

[15] G. Hajdu, “Macaque–a tool for spectral processing and transcription.” In *International Conference on Technologies for Music Notation and Representation. (TENOR'17)*, A Coruña, 2017, pp. 9-15.

[16] G. Hajdu and N. Didkovsky, “MaxScore: recent developments.” In *Proceedings of the International Conference on Technologies for Music Notation and Representation (TENOR'18*), Montreal, pp. 138-146.

[17] R. Gottfried and G. Hajdu, “Drawsocket: A browser based system for networked score display.” In *Proceedings of the International Conference on Technologies for Music Notation and Representation (TENOR'19),* Melbourne, 2019, pp. 15-25.

[18] C. Barlow, “Two essays on theory.” *Computer Music Journal*, vol. 11, 1987, pp. 44-60.

COLLABORATIVE LIVE COMPOSITION WITH FRANKIE

Esa Onttonen
Sibelius Academy, University of the Arts
Helsinki, Finland
esa.onttonen@uniarts.fi

ABSTRACT

This paper describes a web application being developed for the purpose of collaborative live composition. In this concept, musical information such as pitch sets, chord progressions and textual parameters can be created and communicated in real-time by all members of the group. Primarily designed for tablets, the application aims for high usability values and has been tested and developed with professional jazz musicians. This paper presents the background for the concept of collaborative live composition, along with the main features, design principles and the development process of the application. Finally, the paper describes a concert performance and reflects on the affordances, opportunities, and challenges of the concept.

1. INTRODUCTION

Several methods to shape ensemble music in real-time have been introduced in the digital and non-digital domains. With their predefined signs and gestures, the conduction of Lawrence D. “Butch” Morris and the soundpainting of Walter Thompson are two better-known examples of non-digital methods, with Morris often crediting jazz drummer Charles Moffett as an important influence for his conduction technique [1, liner notes]. Having been described as group improvisation (conduction) and live composition (soundpainting) [2], they share some similarities with the rule-based game pieces by John Zorn, most notably *Cobra*, where the role of the prompter—usually played by Zorn himself [3]—is less about conducting and more about being a ”conduit of information” who responds to the requests of the musicians [4].

In the digital domain, various systems have been developed for real-time ensemble music making as well, such as John [5], Adaptive Markov Network for Free Improvisers [6], The Bucket System [7], live coding systems such as BEER [8], or more compositionally oriented systems such as ZScore [9] or Indra [10]. Unlike conduction, soundpainting, or Zorn’s game pieces, these systems usually do not have a conductor, unless the system itself or a generated score [5] is considered to be one. Additionally, the digital and non-digital examples presented have been mostly concerned with the music’s structure and the coordination of the structure, often seen as a challenge in ensemble improvisation [8, 11].

Although The Bucket System is implemented for QuNeo pad controllers without any visual notation, some of its fundamental ideas resonate with the project presented in this paper. They are both created as generic collaborative tools for professional musicians with the intention of allowing ”on-stage compositional/improvisation interaction”, and giving control to *all* musicians [7]. However, while The Bucket System selects randomly from the input given by the musicians, and John generates a random score based on constraints, Frankie does not employ randomization in the decision-making, but allows a single musician to be in control for any period of time until someone else takes the control, thus resembling alternating decision control of rotating leadership [12].

There are other differences to The Bucket System and the previously described digital improvisation systems as well. The improvisational aesthetics of the project presented in this paper are closer to Clifford or James than Earle Brown, often cited in the previous research [7, 5, 10], as well as by Zorn [3]. This has resulted in some features that are probably more useful for musicians coming from the jazz tradition. For example, improvisation in many subgenres of jazz is based on tonal and modal chord progressions and loops. Therefore, the application presented has been primarily developed for musicians who improvise on chord symbols and chord progressions, and pitch sets such as scales or melodic fragments, all available in the skill set of professional jazz musicians.

To define the concept of collaborative live composition, soundpainting provides a useful point of reference. In soundpainting, defined as *the* “universal multidisciplinary live composing sign language” [2], the composition is created by the soundpainter. In the project presented here, all members of the group can contribute to the directions the music will take, therefore it is defined here as collaborative live composition.

Furthermore, the live composition can be played *as such*, or it can provide a springboard for improvisation, as in jazz compositions. Although the term *comprovisation*—already employed by Butch Morris in the 1980s before settling on conduction [13]—is often used for compositional-improvisational practices especially outside jazz [14, 7, 5, 9], in this paper, the term composition is used with the jazz approach: a composition may consist only of a minimum amount of information—e.g. Carla Bley’s miniatures [15, p.16]—which is realized by the improvising musicians with a varying degree of connection to the source material.

The long-term development goal of Frankie is to create an easy-to-use tool that runs on a standard web browser on a typical mobile device and requires very little learning effort from musicians who are trained in Western common practice notation (CPN). Musicians can use their skills in sight-reading CPN which subsequently can provide triggers for their imagination [16]. As a communication tool for collaborative live composition, it does not come loaded with musical content or rules: it is up to the musicians to decide how to deal with the information communicated through the application. Additionally, the application does not try to replace any existing and perfectly valid cue mechanisms such as indicating a downbeat with an arm movement or a nod of the head; in fact, it does not currently provide any timing features. For its research part, the project aims to find out the musical affordances of this kind of practice by exploring the domain between composition and improvisation with the assistance of the application.

2. MAIN FEATURES

The following list describes the types of information that can be communicated with the current version of the application, including the descriptions of the interactions. All parameters, options, keywords, dynamics and roles can be freely configured in JSON (JavaScript Object Notation) files stored on the server.

- Short notated score fragments—simply called scores—of either durationless pitch sets (*set mode*) or chord progressions with automatic chord symbol recognition (*chord mode*). The score is entered with a MIDI input device and when complete, a swipe-up gesture on the score area sends the score to other musicians (alternatively, a button can be pushed). A short animated flash happens on the receiving devices to inform about the new score more visibly than simply changing the notation on the staff.
- Time signature, tempo and other custom parameters in drop-down menus that are displayed in the top section of the score area.
- Keywords as toggle buttons. Multiple keywords can be selected simultaneously, and they can be cleared with a dedicated button. These are displayed in the bottom section of the score area. The keywords are comparable with *karmas* of John [5] and to a lesser degree with the *signals* of The Bucket System [7].
- Dynamics which can be set individually or for all musicians at once from a drop-down menu. The dynamic mark is displayed in the lower left corner of the score area below the clef.
- Roles. Separate from the leader/musician division described later, each musician can have multiple simultaneous roles that are selected from a checkbox menu. For musicians their roles are displayed in the upper left corner of the score area above the clef. The roles can be used to define the musical function a musician should take. For example, a *bass* role indicates that a musician should treat the incoming material as a bass player, while *solo* could direct towards playing a solo. A combination of the previous roles, *bass* and *solo*, would suggest playing a solo in the bass register.
- Text messages that can be configured in advance or written during a performance.

By default, all interactions are transmitted immediately. For example, clicking a keyword sends it right away to the musicians. This works well when the objective is to change the music incrementally, one parameter at a time, for example from quiet to louder dynamics or from slow to a faster tempo. Sometimes it might be more suitable to build a more comprehensive content set where a score, parameters and keywords are all sent at once. This can be achieved with a hold-release functionality, where turning on a hold switch holds all communication until a release button is pressed. For a more fine-grained communication, it is possible to select the recipients and send different scores and parameters to different musicians.

In addition to the transmittable content described before, the application comes with some utility features. The musician who is the leader can turn on a blackout mode which dims all screens. This mode can be used for more traditional free improvisation, and it allows the musicians to take a break from following the screen. A score storage, another utility feature, allows the user to store up to 12 scores [1] with their parameters. Although this functionality is in some contradiction with the concept of *live* composition, it allows creating musical structures by revisiting previously played scores.

There are two user interfaces, leader and musician, depending on the status of the musician (see Figure 1). Although the design and the purpose of these interfaces are different, they share certain elements. The color scheme of the application is dark, so that the screen does not shine too brightly on dimly lit stages. In both interfaces, the main score area is displayed in the top area of the screen with most of the other elements placed below. While a tablet screen could contain even larger scores with more staves, the area has been deliberately kept compact to allow smaller devices such as mobile phones to be used for displaying the notation in reasonable size. Additionally, in the bottom bar there are some features that are common to the leader and musician interfaces (musician and session info, tutorial, reload, settings, full screen, clock), as well as some features that are available only in specific interfaces, such as the blackout mode in the leader interface or the 'Take Leadership' button of the musician interface.

2.1 Leader interface

The leader interface is used to compile a score and set of instructions for other musicians, and it is available for the musician who is currently in the leader role. The leader can

[1] The limitation of 12 scores is not technical but by design to allow using a MIDI input device to quickly send a score from the storage with any C note sending score 1, C# sending score 2 and so on.

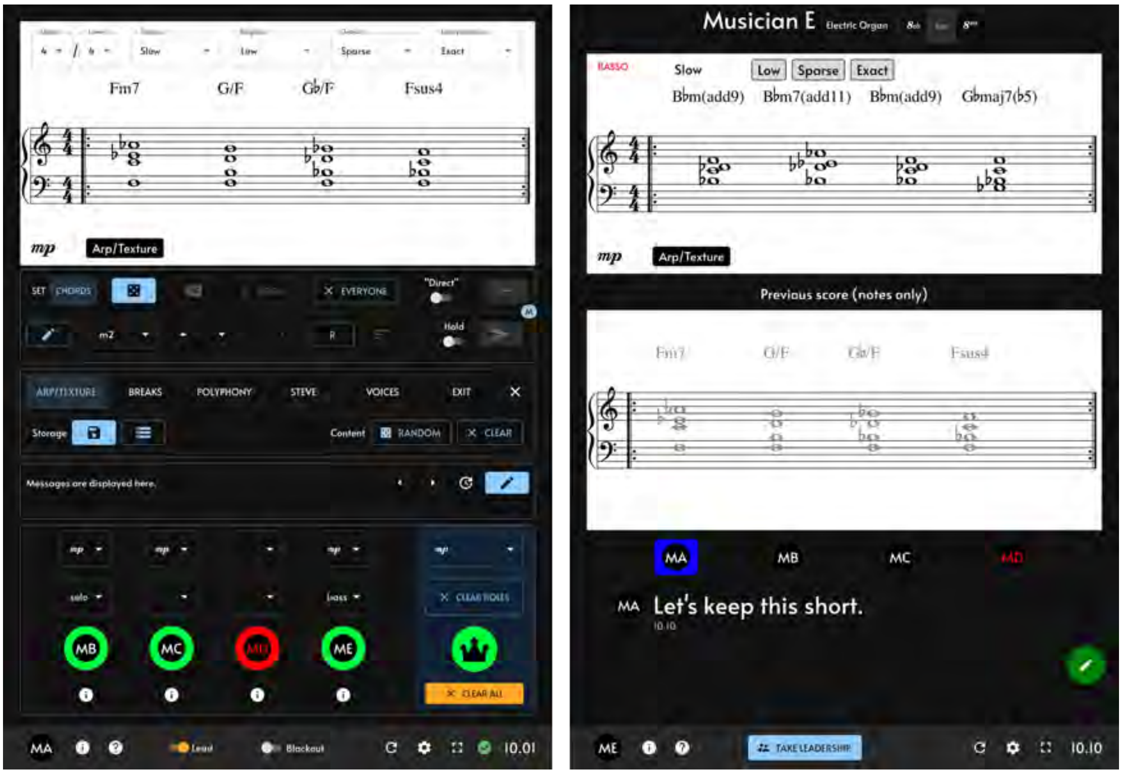


Figure 1. Two interfaces on a tablet: leader (left) and musician (right).

send a signal to indicate abandoning leadership, meaning that any another musician could jump to that role. However, in the current implementation the leadership can be taken by anyone at any moment, even during the blackout mode.

The screen is divided into five main sections. Below the score are various editing functions such as transposition by a selected interval, inversion and retrograde. With the switch labeled *direct* turned on, the score is sent automatically after a user-configurable period of inactivity, allowing to enter the score with a MIDI input device and distributing it for the musicians without touching the interface.

Below the editing functions are the keywords and buttons for the storage and content operations. A narrower section is reserved for text messages. Due to differences in available screen space, the text messages are displayed differently in the leader and musician interfaces. In the leader interface, a dialog window pops up when a new message arrives, requiring the leader to close the window before continuing with other tasks. In the musician interface, the messages are displayed in an auto-scrolling area similar to instant messaging applications, requiring no interaction from the musician.

The "mixer" at the bottom of the screen takes up most of the screen space, as it includes individual controls for all musicians that are displayed as avatar images or as initials if there is no image available. In its current design, the mixer can fit up to seven musicians (the leader is not displayed). The mixer area is used to set dynamics and roles for the musicians. Musicians can be also selected/deselected in order to send score or other parameters only to a selection of musicians. The master channel allows to change all dynamics or clear the roles for the selected musicians, as well as clearing all content with a single button.

2.2 Musician interface

The musician interface displays the score and other instructions received from the current leader, as well as text messages from any member of the group. Following the musician interface does not require any physical interaction with the device.

The header contains the musician's name and the instrument. For multi-instrumentalists a drop-down menu appears where it is possible to switch to another instrument with a correctly transposed part. A button allows to display the score either as transposed or in concert pitch, and an additional octave-up/-down transposition button is available to make it easier to read notation with many ledger lines.

The active score—the one to be played—is displayed below the header with the previous score placed below. A variation of a segmented score was selected (for an overview of various options, see [17, pp. 15–18]) for tablet use; on a smaller mobile phone screen there is enough space only for the active score, at least with instruments notated on a grand staff. It may feel counterintuitive to have the previous score displayed below the active one, but this arrangement was chosen to always have the active score in the same position in the leader and musician interfaces, and therefore provide continuity.

Similar to the leader interface, below the score are avatars for other musicians. The avatars can be clicked to see additional information about the musicians. The leader musician is framed with special colors which indicate if the leadership is active or inactive (“abandoned”). Finally, the lowest screen area is reserved for displaying text messages or entering the text message dialog window.

3. DESIGN PRINCIPLES

There are several principles that have affected the design of the application, most notably usability (learnability and efficiency) and cost-effectiveness. From the common definitions of usability (see for example [18, pp. 19–22]), learnability has been most the important design attribute for both the application and the whole concept. In this project, the importance of the learnability stems from the practicalities of working with professional musicians: they may not be able or willing to spend an inordinate amount of time learning complex applications, concepts and rules—or the signs and gestures of non-digital practices. Making the application easy to learn and use allows for rehearsal time to be spent more efficiently on making music with the application instead of learning to use it. Because the application can be used with a very limited initial configuration, learning can happen cumulatively by adding new parameters to the configuration as needed. For example, sending pitch sets or chord progressions entered via a MIDI input device could be the first step in the learning process. Alternatively and even more simply, dynamics could be the only parameter to be changed. After learning one or two features, more parameters could be incorporated based on the needs of the musicians (and the constraints of the application).

Efficiency is another usability attribute that has been taken into account in the design. First, the application is primarily designed for tablets, eliminating the need to move a pointer before clicking a user interface (UI) element. Second, within the limits of available screen space the UI elements have been optimized so that a minimum amount of interactions is necessary. This results in a somewhat crowded UI but affords a more efficient use of the application since most common operations can be launched directly instead of opening menus or dialogs.

Outside the realms of usability, cost-effectiveness has been another important factor, both for the development and for the user. The application uses only open source libraries and the development does not require any commercial software. For the user, the application works in a standard web browser (see, however Chapter 5). It also works on lower tier devices, although a more high-end device with a faster response does provide a better user experience.

4. DEVELOPMENT PROCESS

The aim of the development process has been to create a high-fidelity prototype [18, pp. 428–433] where the application would eventually evolve to a final product through continuous development, testing and evaluation. After the initial development stage based on the original ideas was largely complete, the application was taken to the first rehearsal session in September 2020. Since then, it has been tested and developed with feedback from the participating professional jazz musicians who are playing keyboards, guitars and drums. At this stage of the project, the instrumentation of mainly polyphonic instruments was selected so that all musicians except the drummer could take different roles such as bass, solo or chordal accompaniment.

All rehearsals produced log files and audio recordings that included both the played music and discussions in-between, providing immediate feedback about the system. For example, shortcomings of the application were collectively discussed right after they were found, as well as various strategies to deal with the information communicated with the application. Additional analysis was done after the sessions, for example, by scanning the log files for chords where the application had failed to generate a proper chord symbol, and fixing them for the next session. In addition, the music on the recordings was listened to with subjective aesthetic criteria: does the music work? If it does not, where is the problem?

Some development decisions were made by observing the musicians using the application. While fixing bugs is an ever-present feature of software development, optimizations were implemented after noticing that certain operations were requiring too many interactions. For example, editing the score required pressing an edit button before the editing functions were enabled. Keeping the editing functions always enabled eliminated the need for the edit button and subsequently made the application more efficient. Some features were implemented based on the suggestions made by the musicians, such as the score storage and the selection of roles (e.g. solo or bass), first quickly implemented through the use of keywords, but later refactored into a separate functionality.

5. TECHNICAL IMPLEMENTATION AND LIMITATIONS

Frankie consists of a web application and a server. The web client written in TypeScript uses React [2] and MUI [3] for common user interface components such as buttons, drop-down menus and so on. Other core components include Recoil [4] for state management, VexFlow [5] for notation rendering and JZZ [6] to handle the MIDI input happening via Web MIDI API. The communication between the server and the clients employs Apollo GraphQL [7] libraries. Excluding few experimental features, the Node.js [8] server does not currently modify the score or other parameters, and works mainly as a transmitter of data between the client devices. As an additional tool there is a log reader that converts the JSON log files into various formats for further analysis: PDF for easier

[2] https://react.dev/
[3] https://mui.com/
[4] https://recoiljs.org/
[5] https://vexflow.com/
[6] https://jazz-soft.net/doc/JZZ/
[7] https://www.apollographql.com/
[8] https://nodejs.org/

reading, CSV for importing markers into a DAW such as Reaper and SRT for creating subtitles to videos.

Since the WebKit browser engine does not implement Web MIDI API, [9] browsers based on WebKit cannot be used for MIDI input. These browsers include Safari and browsers for iOS and iPadOS devices where all vendors are required to use WebKit. However, WebKit browsers can be used if MIDI input is not necessary, for example with musicians who are not going to send staff notation. As of this writing, Google Chrome on an Android device is the recommended combination for MIDI use.

Due to the development focus on the usability of the user interface, the current server implementation can only run a single session at a time. For a proper user and session management, the server code needs a full rewrite and the addition of a database system. As a related limitation, authentication or authorization are not implemented, yet, so the system can be run most safely on a network that is not connected to the internet. As a result of these limitations and the general work-in-progress state of the whole system, it is not yet publicly available, either as source code or as a ready-to-use web service.

6. CASE: *LABRA*

The project was first presented publicly in *Labra*, [10] a concert held at the Black Box of Musiikkitalo in Helsinki on September 28, 2022. In this concert, a group of five musicians—organ player, keyboardist, drummer, electric guitarist and the author on another electric guitar and bass—were playing based on the material created and communicated in real-time with the application. In addition to their personal instruments, the musicians were equipped with various models of Android tablets and small two- or three-octave MIDI keyboards connected to the tablets either via Bluetooth MIDI or a USB cable (see Figure 2). Although operating a MIDI keyboard can possibly add an extra level of difficulty for non-keyboardists, a separate MIDI input device allows playing and simultaneously operating the application, for example, by sustaining a note or a chord on a guitar with a left hand while using the MIDI keyboard with a right hand.

The concert was prepared with five two-hour rehearsals that were scheduled so that some development of the application could be done between the rehearsals. These development periods were constrained only to bug fixes and optimizations, to avoid spending limited rehearsal time on learning new features. To try out the application outside the rehearsals, the musicians were given access to the demo version of the application deployed on the Heroku platform. [11]

6.1 General remarks

The concert performance lasted for 58 minutes without any breaks. Based on the JSON log file, there were 13 changes in the leadership and a total of 49 notated scores created and played during the performance. A single chord or a chord progression was the most popular score type; scores of the pitch set variety appeared only three times. Based on the rehearsals and musicians' preferences, the dominance of the chord progression score was to be expected. Eight of these chord scores were loops of 2–3 chords and the remaining chord scores were either single chords or pedal points sustained for a longer period. It had been found in the rehearsals that single chords and chord loops of 2–4 chords were short enough to quickly enter into the application during playing and had provided useful launching points for improvisation and further elaboration.

The parameters of time signature, tempo, density, register and interpretation were available in the drop-down menus, with 4–5 options for the text-based parameters and multiple options for the time signature. All parameters could be left undefined which was the default option. The parameter set had stayed fairly constant since the initial 2020 rehearsals that were held with partially different musicians. The interpretation parameter with the options [12] 'Completely free', 'Rather free', 'Rather exact' and 'Exact' was inspired by Butch Morris's conduction gesture *develop* which indicates the degree of development for an idea [13, pp. 178–179]. During the concert, tempo changes were communicated 6 times, density changes 8 times, register changes 7 times and interpretation changes 3 times. The only parameter not communicated with the application was the time signature, although it had been used in the rehearsals for various time signatures such as 3/4, 4/4, 7/4, 6/8, 9/8 and 12/8.

The keywords took their shape during the rehearsal period with 'Arp/Texture', 'Breaks', 'Polyphony', 'Steve', 'Voices' and 'Exit' ending up being used in the concert configuration. The meaning of these six keywords were discussed in the rehearsals. For example, 'Breaks' was inspired by the stop-time cues that Miles Davis used with his 1970s groups [19]. Two musicians used these keywords and except for 'Exit'—indicating a suggestion to exit the current musical idea—all keywords were used in the performance.

Free-form text messages were used by all musicians, 21 times in total. Compared to the keywords with predefined musical meanings, the text messages were sometimes more abstract such as 'Dry wood' or 'Backward jazz'. However, most messages were used for less ambiguous requests such as 'A single chord only' or 'Superfast beat with brushes shortly'. Although texting takes more time than using the predefined keywords or other more streamlined features, it was considered a very useful feature, with one musician comparing it to being as powerful as any other feature of the application.

While the screens were turned off with the blackout mode seven times by four different musicians, the score storage found more limited use, with a single content set saved to and sent from the storage during the performance. To summarize, most features of the application were used in the concert.

[9] https://webkit.org/status/#specification-web-midi

[10] *Labra* is a Finnish equivalent for the word *lab*, a laboratory.

[11] On November 28, 2022, Heroku discontinued their free plans and the demo version of the application is not hosted on the platform anymore.

[12] All predefined textual parameters used in the concert were in Finnish. For this paper, these have been translated into English.

Figure 2. Various placements of tablets and MIDI keyboards in *Labra*. The tablet screens have been turned off before the concert performance to save battery.

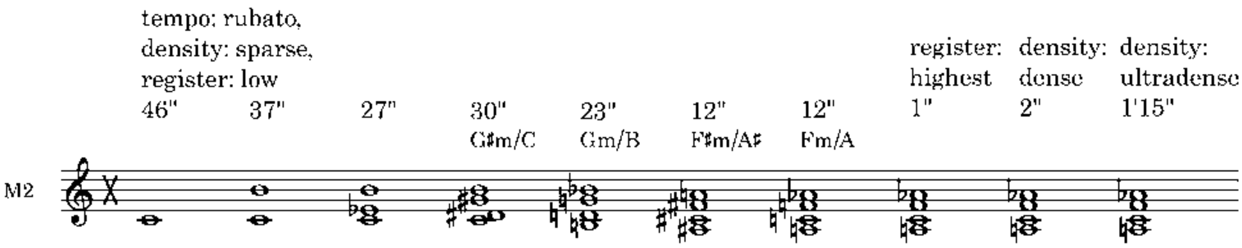


Figure 3. The first ten chords and parameters sent by musician 2 (M2) in *Labra*.

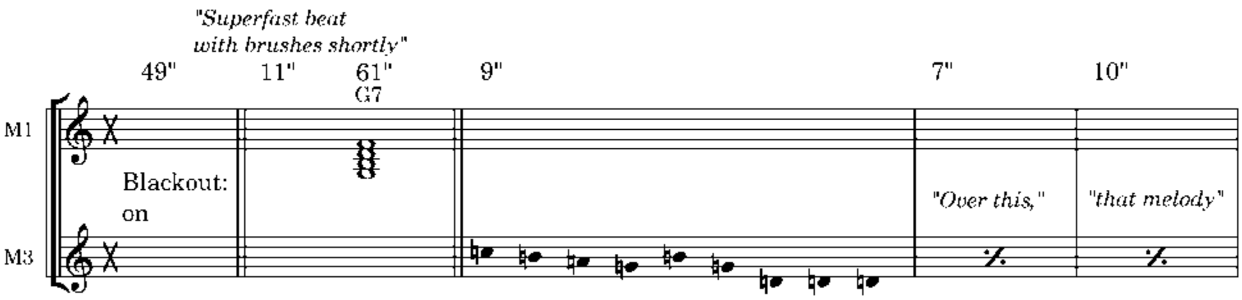


Figure 4. Changing leadership between musicians 1 (M1) and 3 (M3) in *Labra*.

6.2 Two examples

Figures 3 and 4 display two sections [13] from the log file as engraved with the Dorico notation software. The durations of the events are displayed above the staff in seconds. The first example (see Figure 3) contains the first ten chords and parameters sent with the application by musician 2 (M2) during the first ca. 4 minutes of the performance, after a short collective free improvisation had been played by the other musicians. The first event contains a single pitch (a middle C) and the parameters of tempo (*rubato*), density (*sparse*) and register (*low*). This slowly builds to a G#m/C chord which then begins to move downwards chromatically until reaching Fm/A. During that chord the parameters of register and density are changed again, resulting in *ultradense* Fm/A chord in the *highest* register of each instrument. This harmonic idea live-composed by M2 during the first minutes of the performance—a minor triad with a major third as the lowest note—resurfaced ca. 10 minutes later when musician 3 (M3) brought it back with the message 'A single chord only'. [14]

In the second example (see Figure 4) musician 3 (M3) has been leading the group for 4 minutes before abandoning leadership and turning on the blackout mode (the first event in the example). During the blackout mode musician 1 (M1) has taken the leadership and sent the message 'Superfast beat with brushes shortly' before giving the G7 chord as a harmonic guideline. After M1's short leadership, M3 retakes the control and sends a melodic fragment ('that melody') to be played over the current background ('over this'). While M1's leadership lasted for less than 90 seconds and contained only a single instruction for the drummer and a harmonic idea for the others, the music took a different direction both harmonically and rhythmically, and provided a platform for M3's subsequent melodic idea.

7. CONCLUSIONS AND FUTURE WORK

Based on the feedback collected from the participating musicians, [15] in its current state, the application has been relatively easy to learn and use, thus at least partly accomplishing one of the development goals. The application has enabled to shape music with established musical parameters without the need to have extensive training. Musicians have found it fun to make music with the method.

Since the method introduces a new modality for the musicians, this requires dividing attention between playing, listening and operating or following the application, a phenomenon recognized in previous research as well [7]. Musicians found it challenging to combine these modalities, although simply playing more gigs—not only rehearsals—was offered as a possible solution to becoming more fluent with the concept. To avoid becoming completely overwhelmed with the additional information it was found useful to abstain from playing while operating the application (see Chapter 6.2). Similar best practices still need to be found through more public performances.

Musicians also noticed a content gap between the musical ideas they communicated with the application and the realization that followed. Although the music did not always sound as they had imagined, this was not necessarily considered to be a negative thing. For example, one of the musicians explained having a "let's see what happens" type of attitude towards the material while another one reflected that a vague and unfinished idea may turn out to be the best moment of a concert, and that a better outcome would not be guaranteed by more "masterful" information.

Some audience members were missing the ability to follow the interactions, for example through the projection of the score or the application interface, a common practice in live coding performances [8]. The reasons for the decision to not offer score visualization for the audience are not unique for this project [8, 5] and are too manifold to be exhaustively described within this paper. However, when asked about the lack of projection, one of the musicians stated that it would have impacted the way the musicians approached the music making, for example by generating unwanted reactions in the audience. Nevertheless, the performative aspect still needs further consideration to keep audiences less puzzled.

As the application is still a work-in-progress, the future work includes typical programming tasks such as bug fixes, optimization and refactoring. The user interface is under constant revision to provide a better and more inspiring user experience, as well as responsive designs for different screen sizes and orientations. There are areas such as enharmonic spelling which are still suboptimal and in need of further development. As mentioned earlier (see Chapter 5), the server is being rewritten to allow easier access and more spontaneous use of the application. One of the most requested features by the musicians is the addition of durations, especially to the pitch sets. Designing an efficient method for live use is one of the next steps in the development of the application. Additionally, there have been a couple of experiments with server-side processors that modify the content, such as distorting the input pitch sets to give cluster-like results. More work will be done in this area later in the project.

Acknowledgments

The beginning of the project in 2020 was funded by a 12-month working grant from Kone Foundation. Without this grant the development would have not been possible. Later, University of the Arts Helsinki provided funding to organize the concert. Finally, I would like to thank Frankie the cat for standing in my sight when I needed a title for the application.

8. REFERENCES

[1] L. D. B. Morris, "Testament: A Conduction Collection," New York, NY, 1995.

[2] B. Faria, "Exercising musicianship anew through soundpainting: Speaking music through sound ges-

[13] The audio clips for Figures 3 and 4 can be listened to at `https://www.researchcatalogue.net/view/2008271/2008272`

[14] The reentry of the idea is not included in the notated example.

[15] The majority of the musicians were individually interviewed in semi-structured interviews which were transcribed, and analyzed especially on the themes of the method, authorship and the usability of the application.

tures," Ph.D. dissertation, Lund University, Lund, 2016.

[3] J. Zorn, "The Game Pieces," in *Audio Culture: Readings in Modern Music*, C. Cox and D. Warner, Eds. New York, NY: Continuum International Publishing Group, 2005.

[4] J. Brackett, "Some Notes on John Zorn's Cobra," *American Music*, vol. 28, no. 1, pp. 44–75, 2010.

[5] V. Goudard, "John, The Semi-Conductor: A Tool for Comprovisation," in *Proceedings of the 4th International Conference on Technologies for Music Notation and Representation*, S. Bhagwati and J. Bresson, Eds., Montréal, 2018, pp. 43–49.

[6] S. Kalonaris, "Adaptive Specialisation and Music Games on Networks," in *Proc. of the 13th International Symposium on CMMR, Matosinhos, Portugal, Sept. 25-28, 2017*, Matosinhos, 2017.

[7] P. Dahlstedt, P. A. Nilsson, and G. Robair, "The Bucket System - a computer mediated signaling system for group improvisation," in *Proceedings of the International Conference on New Interfaces for Musical Expression, Baton Rouge, LA, USA, May 31-June 3, 2015*, Baton Rouge, 2015.

[8] S. Wilson, N. Lorway, R. Coull, K. Vasilakos, and T. Moyers, "Free as in BEER: Some Explorations into Structured Improvisation Using Networked Live-Coding Systems," *Computer Music Journal*, vol. 38, no. 1, pp. 54–64, 2014.

[9] S. Zagorac and M. Zbyszynski, "Networked Comprovisation Strategies with ZScore," in *Proceedings of the International Conference on Technologies for Music Notation and Representation - TENOR2020*, R. Gottfried, G. Hajdu, J. Sello, A. Anatrini, and J. MacCallum, Eds. Hamburg: Hamburg University for Music and Theater, 2020, pp. 133–140.

[10] D. Andersen, "Indra: A Virtual Score Platform for Networked Musical Performance," in *Proceedings of the International Conference on Technologies for Music Notation and Representation - TENOR2020*, R. Gottfried, G. Hajdu, J. Sello, A. Anatrini, and J. MacCallum, Eds. Hamburg: Hamburg University for Music and Theater, 2020, pp. 227–233.

[11] J. Burrows and C. G. Reed, "Free Improvisation as a Path-dependent Process," in *The Oxford Handbook of Critical Improvisation Studies, Volume 1*, G. E. Lewis and B. Piekut, Eds. New York: Oxford University Press, 2016.

[12] J. P. Davis and K. M. Eisenhardt, "Rotating Leadership and Collaborative Innovation: Recombination Processes in Symbiotic Relationships," *Administrative Science Quarterly*, vol. 56, no. 2, pp. 159–201, 2011.

[13] T. T. Stanley, "Butch Morris and the Art of Conduction," Ph.D. dissertation, University of Maryland, 2009.

[14] S. Bhagwati, "Comprovisation – Concepts and Techniques," in *(Re)Thinking Improvisation: Artistic Explorations and Conceptual Writing*, H. Frisk and S. Östersjö, Eds. Lund, Sweden: Lund University, 2013, pp. 99–104.

[15] A. C. Beal, *Carla Bley*, ser. American Composers. Urbana: University of Illinois Press, Nov. 2011.

[16] S. Hayden and M. Kanno, "NEXUS: Live Notation as a Hybrid Composition and Performance Tool," in *Proceedings of the International Conference on Technologies for Music Notation and Representation - TENOR2020*, R. Gottfried, G. Hajdu, J. Sello, A. Anatrini, and J. MacCallum, Eds. Hamburg: Hamburg University for Music and Theater, 2020, pp. 13–23.

[17] S. Shafer, "Recent Approaches to Real Time Notation," Ph.D. dissertation, University of North Texas, Denton, May 2017.

[18] H. Sharp, Y. Rogers, and J. Preece, *Interaction Design: Beyond Human-Computer Interaction*, 5th ed. Indianapolis, IN: John Wiley and Sons, 2019.

[19] D. Liebman, "Miles and Me," https://davidliebman.com/home/ed_articles/miles-and-me/, 2009.

Non-linear volumetric music composition in a VR context: project “Omega”

Anna Shvets
Fablab by Inetum
Paris, France
anna.shvets@inetum.com

Samer Darkazanli
iMSA
Montauban, France
samer.darkazanli@msa.imsa.org

ABSTRACT

Technological progress in development of XR creative spaces influences our perception of the piece of art. The use of gestural control enables interaction with the spatially distributed elements and creates conditions, where the user participates in a co-creation of a virtual experience. When applied to music art, the spatialization along with interactivity promotes non-linearity of musical structure, resulting in a new instance of music, out of the initial set of possibilities. In this article, we present a concept of a volumetric music composition, which maps music form to a 3D VR space, with a possibility of a variance at each mapping point. We implement a concept in a volumetric VR music composition “Omega”, which allows the user to choose between two alternative paths at each stage of the form development. The user’s action affects not only the changes in music, but triggers the colorization of the surrounding space, color being the reflection of the chosen musical path. Along with the music form spatialization and surroundings colorization, we integrate the paths of gestural movement, which user may choose to follow, adding visual aesthetics to the act of music creation. The synesthetic nature of the composition tightly binds music, color, 3D space and movement into one act of expression in a co-creation of a musical form. The article describes the context, philosophy and details of realisation of the concept’s practical implementation.

1. INTRODUCTION

The graphical musical form representation has a long history, starting in XIX century with Eric Satie, who was the first to create a music content, which graphically represented shapes of a non-musical form in a two-dimensional space. The use of an additional dimension started with the geometric transcription of Bach’s music compositions done by Yannis Xenakis, followed by the appearance of three- and four- dimensional representations of harmonic structure in the XXI century [1, 2]. The technical possibilities brought with the virtual reality (VR) headsets, allowed distributed 3D compositional representation of painting, being experienced in a VR space [3, 4]. The combination of a distributed 3D compositional representation along with the graphical musical form evolution, leads us to the creation of a volumetric non-linear interactive music composition, experienced and co-created in a VR space.

The rising interest to the volumography in music videos production has been mentioned in scientific literature [5, 6], which reports the audience expectation of interactivity with the 3D objects [6]. The same direction follows the definition of Musical XR, which is characterized by the conditions, in which musical elements are controlled through manipulations of their properties, such as pitch, timbre or dynamics [7]. We propose the music architectonic manipulation through stages of the form development, mapped to a visual representation of 3D objects in a VR space, which extends the current state of the art. The interaction with these elements with a haptic feedback (vibration of the hand triggers) and implication of the proprioceptive function through different positions of the mapping points (low and high), engage all sensory modalities currently offered in VR systems.

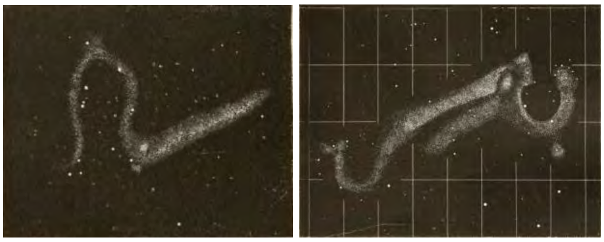
The nonlinearity of music composition appeared in the XX century and was closely related to the interactivity, since it gave the score performer a possibility to choose music events to follow during the performance. The nonlinearity of music form implicitly inherent to the postmodern music is rather seen as a non-centered music matter development, achieved by a small variation of gradually added rhythmo-melodic elements, as in minimal music by Steve Reich, or in a gradual addition of intervallic structures embraced by strophic forms, inherent to compositions by Arvo Pärt [8]. The absence of movement towards culmination results in a musical form-being with non-hierarchical local micro-development, which correlates with the non-hierarchical structure of rhizome introduced by a postmodern philosopher Gilles Deleuze [9]. In our project, both interpretations of non-linearity are employed, giving the user a choice of music events to follow, which are embraced in a strophic form with a static non-centered development. Thus, the project “Omega” embodies postmodern form-being in an interactive way through a new medium.

Given the above, we define of the concept “volumetric composition” as a piece of music, the constructive elements of which are distributed in a 3D space – either physical or virtual, allowing the user to interact with each of these elements in a linear or non-linear manner. The volumetric music composition should be distinguished from VR sound navigation systems [10, 11] where sound sources visualized in a form of 3D objects bare no initial composer intention.

2. VISUAL STRUCTURE MAPPING

The inspiration of the compositional structure shape in a 3D space comes from a stellar composition of the Messier 17 nebula, also referred as Swan Nebula or Omega Nebula, following John Herschel's drawing made in 1833 and 1837 respectively (shown in Figure 1). The points representing important music form development stages were placed in two parallel rows, following visual configuration of the reversed shape of the 1833 drawing [12], with reasonable modifications, to match the dual choice principle of the music composition. Thus, each of the parallel rows receives its proper name and color palette, which will be used for surrounding space colorization. The drawing of the nebula gives a clear path for points placements, while the scientific colorization of the image of the nebula made by Hubble (see Figure 2) inspired a color palette, summarized to a combination of blue and red spectra.

The Swan path is represented with the shades of a blue color and the Omega path uses the shades of red. The intensity of the color shade corresponds to the events ordering, where lighter shades are used in the beginning of the piece and the color intensity increase reflects the stages of the music matter development. The VR mapping of both paths is presented in Figure 3.



a) b)

Figure 1. John Herschel's drawings of Messier 17 nebula a) the drawing made in 1833 and referred by Herschel as "Swan Nebula"; b) the drawing from 1837 and referred as "Omega Nebula". Pictures adapted from [12].

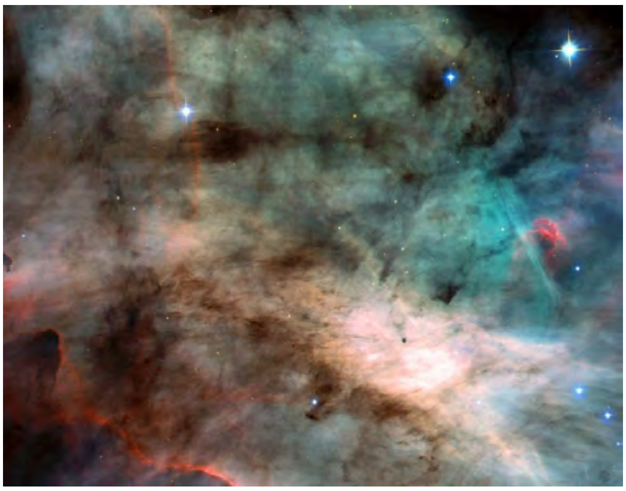


Figure 2. Hubble image of Messier 17 nebula. Pictures adapted from [13].

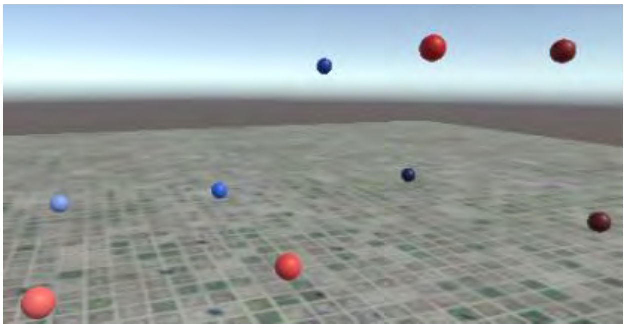


Figure 3. VR structure mapping in two parallel paths – Swan path (in the shades of blue) and Omega path (in the shades of red).

The media materials of the current project are available online [14].

3. COMPOSITION PHILOSOPHY

The semantics of chosen colors reflects the compositional philosophy, expressing a duality of existence through the opposite beginnings: utopian transcendence, metaphorically related to the image of a beautiful celestial bird swan, and dystopian fatality, related to the omega symbol - the last letter of a Greek alphabet, commonly associated with apocalyptic meanings.

The duality of existence is realized not only by visual means – the change of a fundamental tone, axed to the common tone of a third, expresses the same idea by means of music matter development. The change of the fundamental represents the main harmonic signature for both pieces, which makes possible the mutual exchange of the music form development stages.

The similar change of the fundamental, axed to the common tone of a third, was used by John Tavener in his *Funeral Canticle* to express the transcendence from the terrestrial world, embodied in a duo, performing a religious routine melody in F-dur, to the celestial world, with the choir singing much complex music matter, in terms of harmonic organization, marked by the tonality shift to fis-moll – a tonality of the common third with F-dur. The use of natural dominants and seventh chords in a choral part, along with a perpetual micro change of harmonic patterns, creates non-centered music matter development, perceived as endless, eternal state [15]. Similar to John Tavener's composition, a constant change of harmonic patterns, embraced by the common harmonic signature, allows creating form-state perception of both pieces.

4. MUSIC STRUCTURE

The harmonic language of the project is complex and although the pieces contain triadic structures, there is no central tonality to be defined. The music matter follows a multi-center tonal organization, when instead of one tonality, a cluster of tonalities appears and the relation between two neighboring chords defines the tonal center at a given point of time. The multi-center tonal organization may be considered as a natural path of harmonic development after late romantism with Wagnerian enharmonic and elliptic chains and tonality dissemination, which was pushed to the limits with atonal theory. However, since postmodernism seeks to restore the constructive elements of the past within a new context [16], a cluster of tonalities with preserved triadic structures seems to be a viable alternative to the tonality annihilation.

The structure of both pieces heavily relies on harmonic development. Both pieces are starting with the interval of a minor third, while the relation of primary tonalities fundamentals is of a major third (*c* for Swan and *e* for Omega). The minor-major duality of the third drives the development of both pieces in many configurations such as:

1) Intervallic succession of minor and major thirds in Omega path;
2) Major and minor third in the structure of the triad with the same fundamental, which adds dystopian dramatism in both paths;
3) Common third, which plays the role of either minor or major third, depending on a changing fundamental – a harmonic equivalence pattern of both pieces, which expresses the exaltation of the universal beauty with the unescapable finality of existence.

The form of each piece may be defined as a strophic form with variations and can be represented alphabetically as follows:

Swan: $A - A^1 - A^2\ B - A^3$
Omega: $A - A^1 - A^2\ B - A^3\ B^1 - A^4$

The primary tonality of each strophe reacts to the introduction of a *B* element by a semi-tone descent:

Swan: *c – c – h – c*
Omega: *e – e – es – es – e*

This tonal structure resemblance creates another level of connection between two pieces. Let's now analyze the harmonic development of each piece separately.

Swan. The primary tonality, which creates a global structure, is followed by multiple changes within the cluster of tonalities for a given strophe and the relation between two neighboring tonalities may be expressed alphabetically using the following alphabetical designation and a color scheme:

1. ***A***: minor tonic – Dorian minor VI relation (ex.: c-moll – a-moll);
2. ***B***: natural dominant – minor tonic relation (ex.: a-moll – d-moll);
3. ***C***: chords with a common third (ex.: d-moll – Des-dur);
4. ***D***: chords with a common fundamental (ex.: c-moll – C-dur);
5. ***E***: minor tonic – double dominant relation (ex.: c-moll – D-dur);
6. ***F***: minor tonic – minor VI relation (ex.: e-moll – c-moll);
7. ***G***: major tonic – flat major III relation (ex.: C-dur – Es-dur);
8. ***H***: minor tonic – Neapolitan minor II relation (ex.: c-moll – cis-moll);
9. ***I***: parallel tonalities (ex.: C-dur – a-moll).

The colors chosen to mark the chord relations describe a degree of a major's "heat": the relations between two minor triads are marked by a cold palette of blue and green with their combination (***A***, ***B***, ***F***, ***H***), while the presence of major triads adds red shades to the palette (***C***, ***D***, ***E***, ***G***, ***I***). The color allows to create a correspondence between two pieces, while alphabetical representation reflects the order of the relation appearance in a specific piece. Using this relations scheme, let us build the harmonic structure for the Swan piece:

1st strophe:	***A – B – C – C***
2nd strophe:	***D – C – E – D – F – C – C – C***
3rd strophe:	***D – C – D – G – C – C – D – H – C – D***
4th strophe:	***A – B – C – I – F – B***

The hybrid visualization of the harmonic development inherent to the Swan piece reveals the logic of the new harmonic relations addition. After the exposition of the three first relations in the 1st strophe (***A***, ***B***, ***C***), which creates the arch with the 4th strophe, new relations are introduced via alternation with the main harmonic signature (***C***) at first in a singular mode (***D***, ***E***) in the 2nd strophe; then, being coupled with the common fundamental relation (***D***) in a middle of the 2nd strophe (***D – F***)) and the 3rd strophe (***D – G***, ***D – H***). We can see that the common fundamental relation (***D***) expands starting from its appearance in the 2nd strophe, invading the harmonic space of the 3rd strophe. The harmonic development of the 3rd strophe is structurally close to the harmonious concentric circles form (***D – C*** .. ***C – D***) with the main harmonic signature relation in the center (***C – C***), if new relations ***G*** and ***H*** are considered as variations of ***D***, which precedes their appearance. Finally, the parallel tonalities relation (***I***) is introduced in the last 4th strophe.

We also may observe the increase of the harmonic intensity trough a passage from a Dorian minor VI (***A***) and natural dominant (***B***) relations to the less related chords with a common third (***C***) already in the 1st strophe. The 2nd strophe, along with close double dominant relation (***E***) and less close relation of chords with a common fundamental (***D***), contains very distant intense harmonic relation of the minor VI (***F***). In the 3rd strophe, the harmonic intensity increases with the flat major III relation appearance (***G***): the only major-major relation of the piece, which is further neutralised by a minor tonic – Neapolitan minor II relation (***H***). Finally, the 4th strophe brings the harmonic intensity discharge with very close parallel tonalities (***I***) and natural dominant (***B***) relations. A certain level of intensity remains with the Dorian minor VI (***A***), chords with a common third (***C***) and minor VI (***F***) relations to finish the piece.

Exploiting the same color scheme, let's visualize the list of tonalities, which appeared within described relations:

1sts.: *c-a*|*a-d*|*d-Des*|*Des-d*
2nds.: *c-C*|*C-cis*|*des-Es*|*Es-es*|*dis-h*|*h-B*|*B-h*|*h-B*

3rds.: *h*-*H*|*Ces*-*c*|*c*-*C*|*C*-*Es*|*Es*-*e*|*e*-*Es*|*Es*-*es*|*es*-*e*|*e*-*Es*|*Es*-*es*
4ths.: *c*-*a*|*a*-*d*|*d*-*Des*|*Des*-*b*|*b*-*d*|*d*-*a*

The migration of the tonalities within common third relation (***C***) reveals minor second movements in downward (*d*-*cis*) and upward (*h*-*c*) directions, being in the range of the ancient *passus duriusculus* melodic pattern, which expresses intense dramatism. The migration is finished by a major second downwards movement (*e*-*d*), returning to the initial point (*d*), which correlates with the existential finality. The common fundamental relation (***D***), which augments its presence in the 2nd and 3rd strophes, also tends to expand its region of tonalities coverage (*c*-*es*-*h*-*c*-*es*). The Dorian minor VI (***A***) and natural dominant (***B***) relations remain tightly bound to their tonalities (*c*-*a*, *a*-*d*, *d*-*a*), adding stability to the constantly changing harmonic patterns.

Along with the tonal relations, the means of harmonic expression comprise the structure of chords with additional tones and their alterations. Thus, a sharp fourth regularly appears within the common third relation on a feeble time, adding to the melodic development, which is mostly drawn by a harmonic flow. Major seventh within a minor seventh chord augments a dramatic effect of existential finality, while major/minor third inside the same chord structure expresses the duality of existence.

Omega. Let's define an alphabetical order of the relations appearance in the Omega piece, using a common color palette with the colors choice logic described earlier:

1. ***A***: chords with a common fundamental (ex.: c-moll – C-dur);
2. ***B***: parallel tonalities (ex.: C-dur – a-moll).
3. ***C***: tritone relation between 2 minor triads (ex.: fis-moll – c-moll);
4. ***D***: chords with a common third (ex.: d-moll – Des-dur);
5. ***E***: minor tonic – Dorian major VI relation (ex.: es-moll – C-dur);
6. ***F***: tritone relation between major and minor triads (ex.: Fis-moll – c-moll);
7. ***G***: minor tonic – minor VI relation (ex.: e-moll – c-moll);
8. ***H***: minor second relation between two minor chords (ex.: cis-moll – c-moll);
9. ***I***: major tonic – flat VI minor relation (ex.: C-dur – as-moll);
10. ***J***: major tonic – flat VI major relation (ex.: C-dur – As-dur).
11. ***K***: minor tonic – Dorian minor VI relation (ex.: c-moll – a-moll).

Let's build the harmonic scheme of Omega piece development using described formalism:

1st strophe:	***A – A – B – C – D***
2nd strophe:	***A – A – E – A – F – D – B – D – A – G***
3rd strophe:	***A – A – B – H – A – I – D – D – D***
4th strophe:	***A – A – E – A – J – I – K – D – D – D***
5th strophe:	***A – A – A – H***

The common fundamental relation (***A***) stabilizes the dynamism of the relations change, returning at the beginning of each strophe. It also tends to precede the introduction of a new relation: ***A – B*** in the 1st strophe, ***A – E*** along with ***A – F*** and ***A – G*** in the 2nd strophe, ***A – I*** in the 3rd strophe, and ***A – J*** in the 4th strophe. Very distant relations of minor triads – of a tritone (***C***) and of a minor second (***H***), are preceded by a close relation of parallel tonalities (***B***) in the 1st and the 3rd strophes respectively, creating relaxation of a harmonic density just before its immediate concentration. The common third relation (***D***) gradually increases its presence through the piece, being replaced by a very distant and harmonically dense minor second relation (***H***) in the last 5th strophe. It is quite remarkable that the last new relation introduced in the Omega piece (***K***) is the very first relation that appeared in the Swan piece; it is also appears close to the golden ratio zone. This event is preceded by the only major-major relation of the Omega piece (***J***), representing a flat major relation, similar to the flat major relation – also a unique major-major relation in the harmonic structure of the Swan piece. The difference lies in the third's direction – the Swan piece contains an upward flat relation (flat major III), while Omega contains a downward flat relation (flat major VI), related to a major tonic. This creates a complementary relation between two pieces, as a reflection of the utopian (upwards) and dystopian (downwards) beginnings.

Let's now visualize the list of tonalities, which appeared in the Omega piece within described relations:

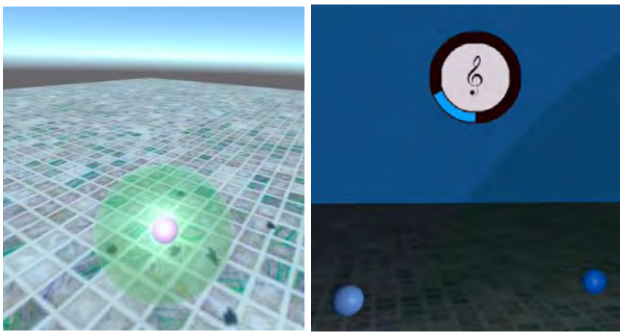
1sts.: *e*-*E*|*e*-*E*|*E*-*cis*|*cis*-*g*|*g*-*Ges*
2nds.: *e*-*E*|*e*-*E*|*E*-*g*|*g*-*G*|*G*-*cis*|*cis*-*C*|*C*-*e*|*e*-*Es*|*Es*-*es*|*es*-*g*|
3rds.: *es*-*Es*|*es*-*Es*|*Es*-*g*|*g*-*ges*|*ges*-*Ges*|*Fis*-*d*|*d*-*Des*|*Des*-*d*|*d*-*Des*|
4ths.: *es*-*Es*|*es*-*Es*|*Es*-*ges*|*ges*-*Ges*|*Ges*-*D*|*D*-*b*|*b*-*cis*|*cis*-*C*|*C*-*cis*|*cis*-*C*|
5ths.: *e*-*E*|*e*-*E*|*E*-*e*|*e*-*es*|

The tonalities movement within the common third relation is started with a downward jump to a tritone (*g*-*cis*), then upward pass to a minor third (*cis*-*e*) and gradual descent to the tritone's base tone (*e*-*d*-*cis*). The tonalities movement within this relation is much more dramatic, than in the Swan piece, where the movements on small intervals within a range of a perfect fourth (h-e) took place. A dissonant tritone interval, suddenly taken by a downward movement, the attempt to raise on the interval of a minor third and a fatal descent to the tritone jump's hole creates dystopian associations. The tonalities coverage within the chords with a common fundamental is similar to the tonalities movement within the same relation in the Swan piece and do not surpass the interval of a major third (*e*-*g*-*ges*-*es*-*ges*-*e*). However, instead of raising above the starting tonality, the tonal movement in the Omega piece returns to its initial point.

The means of harmonic expression in the Omega piece also comprise the use of additional chord tones with alterations. The sharp fourth along with the major/minor third within the same chord structure are also used in the Omega piece to express the similar meaning as in the Swan piece. Unlike the Swan piece, the alternation of the minor and major seventh within the major seventh chord is used to add harmonic density and the dynamism of a dominant sounding. It is also should be noted, that either resolved directly or via ellipsis dominants add dramatism and instability to the whole piece, compared to a more static Swan piece. Finally, the relation of stability within two pieces is supported by a rhythmic organization. The same metric structure was used for both pieces with twice more agitation in Omega (100 bpm) compared to Swan (50 bpm).

5. INTERACTION AND GESTURAL CONTROL

The interactive aspect in the context of the processual art requires a clear indication of the elements that may be triggered at a given moment of time. In our project we wrap the potential points to trigger inside semi-transparent spheres, which gradually shrink until the exact moment of interaction, requesting the user to decide which audio event will be activated (see Figure 4). Since the sphere is bound to the point and it might go offset the user's field of view, we integrate an additional indicator synchronized with the sphere movement, which stays in the user's field of view at all times. This solution not only helps to maintain the order of interaction, but also guaranties the continuity of performance and integrity of the music piece.



a) b)

Figure 4. Semi-transparent spheres, indicating the next element to trigger (a); an additional indicator of the timing (b).

The surrounding space elements colorization is affected by the sound events choice and the surrounding canvas is gradually colored in shades of colors inherent to the chosen points. The surrounding canvas is divided into regions and each region may be equally affected either by Swan path's color palette or by the color palette of the Omega path. An example is given in Figure 5.

The gestural control makes an integral part of the user's experience and is implemented via gestural movement paths. The movement paths are represented as semi-transparent traces of the dancer's hand movements, recorded from touch controller's coordinates during the dancer performance in a VR space. The dancer's guiding movement paths are unrolled in time, engaging the user to follow. The user's hand movements are leaving their own traces, which allows to assess how close the user movement was to the guiding dancer's movement. To prevent the saturation of lines, both guiding and the user's hand traces begin to fade just after their appearance. Since gestural movements are tightly bound to the points being triggered, we provide 5 pre-recorded combinations of interaction possibilities with both paths, an example of such variability is given in Figure 6. There is also a possibility to turn off the pre-recorded movements, leaving the user a complete freedom in a piece stages combination and gestural movements.

Figure 5. Canvas regions after interaction with one Swan element (shades of blue) and two Omega elements (shades of red).

The choreography of the guiding movements represents modern dance style, including *chaine tour* (chain turns), *glissade* (gliding steps) that predicate *jetes* (jumps) or *tour de basque* (leaps), along with front-to-back and side-to-side steps. The dance choreography is interpretative and conveys the main artistic message, related to the compositional philosophy, expressing the fragility of human existence and the eternal beauty of the universe – the duality of physical and metaphysical presence, embodied in a transcendent artistic experience.

Figure 6. Guiding traces interacting with the Omega and Swan points.

6. CONCLUSIONS

The article presented a concept of a volumetric music composition along with its practical implementation in a VR application. The compositional philosophy and its realization in a music matter has been shown via hybrid visualization and detailed analysis of harmonic relations and tonal movements. The principles of the interactive points with the related surrounding regions colorization were shown and the models of interaction with the project's visual interface via guiding traces have been discussed. The project presents a great potential for co-creation hybrid VR experiences of a music art.

7. REFERENCES

[1] D. Tymoczko, *The Geometry of Music*. Science, 2003.

[2] G. Baroin, *Applications de la théorie des graphes à des objets musicaux*. *Modélisations, visualisations en hyperspace*. Thesis (PhD). Université de Toulouse, 2011.

[3] S. Belloli, S. Porro, V.S. Virk, R. Etzi, A. Gallace, M. Bordegoni, M. Carulli, "The Kandinsky experience: A multisensory augmented reality application for cultural heritage", *Computer-Aided Design & Applications*, 18(4), 2021, pp. 799–814.

[4] A. Zhilyaeva. Live painting performance at the Louvre museum Paris. [Online] Available: https://twitter.com/AnnaDreamBrush/status/1601230437841502210

[5] A. Smolic, K Amplianitis, M. Moynihan, N. O'Dwyer, J. Ondrej, R. Pagés, G. W. Young, E. Zerman, "Volumetric Video Content Creation for Immersive XR Experiences", *London Imaging Meeting*, 2022, pp 54 - 59, https://doi.org/10.2352/lim.2022.1.1.13

[6] G. W. Young, N. O'Dwyer, M. Moynihan, A. Smolic, "Audience experiences of a volumetric virtual reality music video", *IEEE Conference on Virtual Reality and 3D User Interfaces*, 2022 pp. 775–781.

[7] L. Turchet, R. Hamilton, A. Çamci, "Music in Extended Realities", *IEEE Access,* 9, 2021 pp. 15810–15832.

[8] A. Shvets, "Mathematical bases of the form construction in Arvo Pärt's music", *Lietuvos muzikologija*, vol. 15, 2014, pp. 88–101.

[9] G. Deleuze, *Empirisme et Subjectivité*. Paris: Presses Universitaires de France, 1973.

[10] L. Turchet, M. Carraro, M. Tomasetti, "FreesoundVR: soundscape composition in virtual reality using online sound repositories", *Virtual Reality*, 2022 pp. 1–13.

[11] V. Zappi, D. Mazzanti, A. Brogni, D. Caldwell, "Concatenative synthesis unit navigation and dynamic rearrangement in vrgrains", *Proceedings of the sound & music computing conference,* 2012 pp. 403–408.

[12] E. S. Holden, "The Horseshoe Nebula in Sagittarius", *Popular Science*, vol. 8, January 1876, pp. 269–281.

[13] Constellation Guide. Omega Nebula – Messier 17: https://www.constellation-guide.com/omega-nebula-messier-17/ (retrieved 02 March 2023).

[14] Volumetric Music Composition in the VR context: https://shvetscomposer.wordpress.com/2023/03/14/volumetric-music-composition-in-the-vr-context/ (retrieved 02 March 2023).

[15] A. Shvets, "Contemporary church music and postmodernism: a compositional perspective", [Conference presentation]. *Between faith and art*. *Art in sacred space*, 31 May – 2 June 2012, Lublin, Poland.

[16] A. Shvets, 2013. "Modern model of narration: about relation between cyberculture and postmodernism", in L. STANSBIE, A. BOURLESCU (Eds.), *Reflections on Narrative: Interdisciplinary storytelling*, Oxford: Inter-Disciplinary Press, pp. 77–88.

MAPS AS SCORES: “TIMBRE SPACE” REPRESENTATIONS IN CORPUS-BASED CONCATENATIVE SYNTHESIS

Jonathan Bell
Aix-Marseille Université, PRISM-CNRS, XR2C2-UCA
Marseille, France
belljonathan50@gmail.com

ABSTRACT

The present study investigates ways in which the “timbre space” metaphor may be used in creative ways for instrumental composition. Numerous tools for concatenative sound synthesis share today the ability to represent in an n-dimensional space large quantities of sound, thus displaying on a map data which originally unfolded in time. If the potential of such systems for creating interactive instruments is an evidence, their affordance as musical scores needs further assessment, for control over time becomes unknown territory. When porting to VR such representation of sonic data, the *score* becomes a 3-dimensional map (or world) in which the user typically navigates freely. Experimentation through composition, instrument design and improvisation have shown a potential of simulation of plausibly automatised acoustical instrument, using machine learning techniques to model virtual instruments out of relatively small quantities of data (e.g. 20 minutes of audio to model a clarinet). The method offers promising avenues for the exploration of instrumental fragments clustered by timbre, register, dynamic, instrumental techniques. Whether or not such maps identify as musical scores, they contribute to addressing a problem formulated by Lev Manovitch: “how to merge database and narrative into a new form”.

1. INTRODUCTION

1.1 Are scores maps?

The question raised in this article is in no small part inspired by a study by Daniel Miller, *Are scores maps? A cartographic response to Goodman*[1], in which a dialectical tension between two concepts (score and map) leads to interesting questions on the role and function of musical notation. Miller proposes that, in spite of surface-level conventions, the underlying structure of scores closely relates to maps: “*notational components of scores are better understood as contingent surface-level features leveraged by an underlying map-like representational structure. On this account, scores are seen to be highly conventionalized maps, and the notational symbols of scores constitute just one of multiple modes of representation and depiction harnessed by this framework*”. Scores may therefore be conceived as a mere subset of maps. The rapprochement between both notions is best illustrated with a historical example: famous works of the post-war avant-garde (by M. Feldman, J. Cage, E. Brown, P. Boulez, K. Stockhausen and A. Boucourechliev to cite a few) took the form of graphical scores which the performer could freely navigate. Boucourechliev named many of his compositions “Archipels” (archipelago) which evokes the same metaphor. Those scores typically presented common notation fragments spread out all over the page, so as to emancipate the work from the linearity imposed by traditional notation, an idea thoroughly discussed in U. Eco’s *Open Work* [2]. In Boulez’s *second sonata* for piano for instance, this co-existence of snippets of conventional notation with a map-like layout on the page illustrates how maps allow for compelling hybrid cases, in which not every symbol on the page functions as a map. Similarly, in more conventional/linear scores, Miller underlines that only some features are isomorphic (or maps-like): “*Scores are maps that are isomorphic with the spatial and temporal structures of the musical works they represent, while other graphical features may be purely contingent or incidental. This highlights an interesting property of maps: they need only be isomorphic with regard to a subset of the properties of the space they represent*”. Maps, in a manner reminiscent of scores, aim to guide a user or prompt a performer for action, and this goal needs not obey strict, systematic, one to one mapping relationships. For Miller indeed, this mixture of isomorphism and contingency responds to Goodman’s famous attack against John Cage, formulated as follows: “*Without stipulation of minimal significant units of angle and distance, p. 53 from John Cage’s Concert for Piano and Orchestra from 1960 is not syntactically differentiated*”. The philosopher’s observation led him to vitriolic criticism of John Cage’s approach to graphic notation: “*Under the proposed system there are no disjoint and differentiated characters or compliance classes, no notation, no language, no score*”[3], a point of view which, in return, has proven widely unpopular.

Reflexions on maps and paper scores may seem here overly theoretical. The digital age, however, urges composers to think about scores in new ways, in which hybridization between interactive systems and notation leaves more room to our notion of map as score. In “The digital score, musicianship, creativity and innovation”[4], Craig Vear proposes that: "*The core purpose of any digital score is to communicate ideas between musicians using digital technology.*", thus placing technology at the center contemporary notation, and expanding at the same time the range of systems hitherto coined as "scores". Vear also proposes that "*Some digital scores might feel like computer games where the performer makes decisions about what happens next.*" The image of the computer game, just as the one of a map which the performer can browse, evokes freedom in the first place, but also an important shift of focus when considered through the perspective of a traditional composer. His craft then becomes closer to the one of an instrument designer, which Vear still considers belonging to the realm of digital scores. In a chapter entitled "The nature of digital scores: expanding the core signatures", he states: "*the score might be embedded within the design of an instrument, instrument might be a system-score of electronics controlled by generative software*". Under such circumstances, the frontier between score and digital instrument design blurs, and diverse forms of interactive system querying a database will be considered a score.

In some early accounts of artistic uses of CataRT [5], Diemo Schwartz described his 2d representations of a sound corpus as "navigable score" or "score instruments": "*The [piece's] subject was "navigable scores" or score-instruments, in which different kinds of users would play the sound or music, cruising through the score.*"[6] Schwartz's discovery then presents this paradox that it proposes a score that does not represent time.

1.2 Maps do not represent time

Musical scores are used to convey information about musical material function of time.

Maps, on the other hand, are static representations of an area. They do not represent time or changes over time, and are typically used to convey information about geographical features and spatial relationships, and are also a powerful tool for visualisation in data science, as will be developed in Chapter 5.

"*Many new media objects do not tell stories; they don't have a beginning or end*" [7]. This quote by Manovitch helps us here introducing a challenging idea for a traditional composer: ignoring the temporal dimension of a score may lead to innovative approaches to composition. We will now expose how such data plot, or more globaly databases can also be considered as an art form. Finally, the absence of time inherent to these concept encourages to use in Section 1.2.2 aesthetic arguments to understand what might stay creatively interesting within a loose control over time in musical composition.

1.2.1 Databases as an art form

Databases and data mining in music can be thought of as an art form in the sense that they involve using creativity and analytical skills to extract insights and knowledge from large sound datasets. This requires the ability to inform musical creativity by identifying patterns, trends, and relationships in data that may not be otherwise apparent.

Novels and film, are, as music, dance, or theatre, temporal art forms. Manovitch's insight on database art form suggests that new media isn't subordinate to time, or narrative, in the same manner; this consideration may help a composer alter some preconception on his approach to form, articulation, or narrative: "*After the novel, and subsequently cinema, privileged narrative as the key form of cultural expression of the modern age, the computer age introduces its correlate - database. Many new media objects do not tell stories; they don't have a beginning or end; in fact, they don't have any development, thematically, formally or otherwise which would organise their elements into a sequence.*"[7] This absence of narrative in new media art, which Manovitch relates to post-modernity, finds an interesting echo in the thought of Morton Feldman, who saw in the European Avant-Garde an excessive desire for control over time.

1.2.2 Morton Feldman and the European clock makers

In his collected writings [8], Morton Feldman recalls a discussion he once had with Karlheinz Stockhausen: "*He was convinced that he was demonstrating reality to me. That the beat, and the possible placement of sounds in relation to it, was the only thing the composer could realistically hold on to. [. . .] Frankly, this approach to time bores me. I am not a clockmaker. I am interested in getting Time in its unstructured existence. That is, I am interested in how this wild beast lives in the jungle - not in the zoo*". Feldman, often with humour, insisted on that idea : "*Let the sounds alone, Karlheinz, don't push them — not even a little bit?*". This almost passive approach to composition echoes the zen-inspireed thought of John Cage, and also importantly took his inspiration from painters: "*A painter will perhaps agree that a color insists on bieng a certain size, regardless of his wishes [...] He can simply allow it to 'be'. In recent years we realize that sound has a predilection for suggesting its own proportions [...] Any desire for differentiation must be abandoned*"

Feldman's music explored this idea of surface in a number of ways, such as through the use of long, slow and static musical lines, the repetition of simple motifs, absence of contrast (which he called differenciation), and the use of unconventional temporal dimensions (some works such as the second string quartet last over four hours). This unconventional approach to form relates more broadly to an opposition between European and American avant-garde in the 50-60s.

The tools described in Chapter 5 inclined the author to think of musical form as long static *time canvases*, as would M. Feldman call them: " "*My obsession with surface is the subject of my music. In that sense, my compositions are really not "compositions" at all. One might call them time canvases in which I more or less prime with an overall hue of the music*". Listening to almost any of his works, one realises that the entire piece most often shares the same atmosphere from its beginning until its end. A piece with a recognisable instrumental combination such as *why patterns* for flute, piano and celesta or *Clarinet and Percussion* have a strong acoustical footprint and illustrate how machine machine listening could be used on such materials.

2. CORPUS-BASED CONCATENATIVE SOUND SYNTHESIS (CBCS) TODAY

Corpus-Based Concatenative Sound Synthesis (CBCS) is a technique used in computer music that involves constructing a sound or music piece by concatenating (joining together) smaller units of sound, such as phonemes in speech synthesis or musical phrases in music synthesis. It can be used to model an improvising instrumental musician by creating a database of recorded musical phrases or segments that can be combined and rearranged in real-time to create a musical performance that sounds like it is being improvised.

Today nearly 20 years old if one refers to the first CataRT publications [5], CBCS today enjoys an increasing popularity. Various apps today are based on similar principals (AudioStellar, Audioguide, LjudMAP or XO). The democratisation of audio analysis and machine learning tools such as the FluCoMa package (for Max, SuperCollider and Pure Data) encourages computer music practitioners to engage in this field at the crux between music creation and data science/machine learning.

2.1 Timbre Space

In spite of promising advances in the domain of deep learning applied to sound synthesis [9] [10], CBCS tools may earn their popularity from a metaphor which leads back to the early days of computer music: the notion of timbre space, developed by Wessel [11] and Grey [12], according to which the multi-dimentional qualities of *timbre* may be better understood using spatial metaphors (e.g. the timbre of the English horn being closer to basson than is it of trumpet).

The heritage of the Timbre Space metaphor can also be found in various iterations of the Orchidea [13] project, and continues to inspire generations of composers and music technologists. [1]

Pioneers in the perception of timbre studies such as Grey [12], J.C. Risset, D. Wessel, [16] or Stephen McAdams [17] [18] most often define timbre by underline what it is not. Risset and Wessel, for instance, define it as follow: *It is the perceptual attribute that enables us to distinguish among orchestral instruments that are playing the same pitch and are equally loud.*

The co-variance such parameters (pitch, loudness and timbre), however, leads Schwarz to distinguish timbre space and CBCS notions: '*Note that this concept is similar but not equivalent to that of the timbre space put forward by Wessel and Grey [7, 24], since timbre is defined as those characteristics that serve to distinguish one sound from another, that remain after removing differences in loudness and pitch. Our sound space explicitly includes those differences that are very important to musical expression.*" [19]

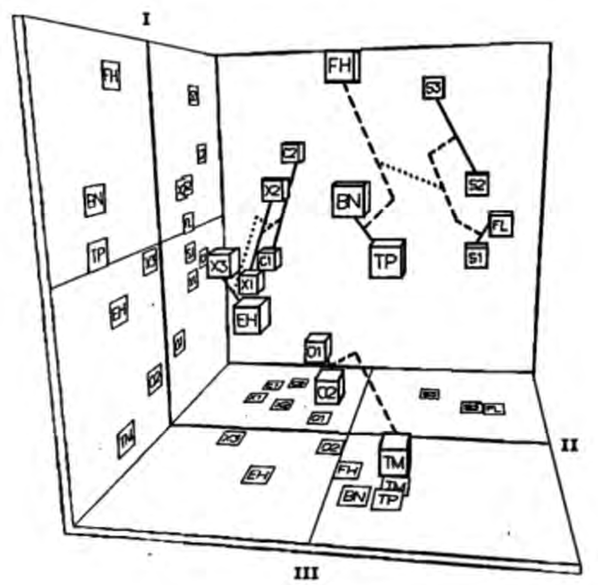


Figure 1. Multidimensional perceptual scaling of musical timbres (John M. Grey[12]). Sounds available at: https://muwiserver.synology.me/timbrespaces/grey.htm.

The workflow described in Chapter 5 gave in practice strong evidence of inter-dependance between register, timbre and dynamics, particularly when the analysis run over a single instrument sound file (e.g. 30 minutes of solo flute), and chopped in short samples. The system will then precisely be able to find similarity between instrumental passages played in the same register, same dynamic, and same playing technique (e.g. a flute playing fast trills mezzo forte, in mid-low register, with air).

2.2 Corpus-Based Concatenative Synthesis - State of the art

A wide array of technologies today can be called corpus-based concatenative synthesis, in the sense that they allow, through segmentation and analysis, to explore large quantities of sound. Some of them are presented as "ready-made" solutions, such as the recent *Audiostellar* [20], or SCMIR [2] for SuperCollider. Hackbarth's AudioGuide [21] offers a slightly different focus because it uses the morphology/timeline of a soundfile to produce a concatenated output. Within the Max world finally, two environments appear as highly customizable: IRCAM's MuBu [22] and the more recent EU funded FluCoMa [23] project. CataRT is now fully integrated in MuBu, whose purpose encompasses multimodal audio analysis as well as machine for movement and gesture recognition [24]. This makes MuBu extremely general purpose, but also difficult to grasp. The data processing tools in MuBu are mostly exposed in the pipo plugin framework [25], which can compute for in-

[1] Daniele Ghisi, co-author of the *bach* [14] package for Max, occupies a role role here as he's worked both on the *Orchidea* and *FluCoMa* projects. Some objects of his later *dada* library [15] also show an influence of CataRT (the dada.catart object was later renamed dada.cartesian). The dada.base based, finally, might have been a source of inspiration for the manipulation of databases in Max. An extract of one of his presentation is available here: https://youtu.be/LD0ivjyuqMA?t=3032

[2] A demo is available at: https://youtu.be/jxo4StjV0Cg

stance mfcc analysis on a given audio buffer [3] by embedding the pipo.mfcc plugin inside the mubu.process object.

FluComa also aims to be general purpose, but seems particularly suited to perform two popular specific tasks. With only limited knowledge of the framework nor of theory laying behind the algorithms it uses (such as those dimentionality reduction, mfcc analysis, or neural network training), the framework allows: 1/ to segment, analyse and represent/playback a sound corpus 2/ to train a neural network to control a synthesizer, in a manner reminiscent of Fiebrink's Wekinator [26].

Only the tools for segmentation, analysis representation and playback (described in detail in Chapter 5) were used here, for they precisely fit the needs of corpus-based synthesis.

3. FIRST ATTEMPTS

Some of the early versions of the FluCoMa package already provided efficient onset detection algorithms (onsetslice [4]), which encouraged the author to further dig into their environment.

At that time, all the compositional material generated with the package used the fluid.bufstat [5] object to run statical pitch analysis (and pitch confidence analysis) on each slice of a given pre-existing sound file. As can be heard in this accompanying example [6], slices were classified by pitch (the lower the index, the lower the register), although some inaccuracy could occur due to the simplicity of the analysis (section 5 will describe more elaborate strategies).

Onset detection combined with statistical pitch analysis were first used on a large scale in a piece (Chef 2.0 [7]) in which all the instrumental parts had been generated with comparable rudimentory Music Information Retrieval techniques.

Some experiments of various forms of VR control over such analysis tools already seemed promissing at that time [8]. The PatchXR software will be discussed in section 5.3.

4. MOTIVATIONS

One of the goals of the present study is to take advantage of the numerous capacities of visualisation, interaction and motion gesture data available within a VR environment. Porting to VR analysis made in Max/FluCoMa has its challenges but I am currently exploring various possible ways of interacting with a corpus-based analysis in VR. After a few trials in which the x y z coordinates of a world directly represented audio descriptors such as loudness pitch and centroid [9], I more systematically used mfcc analysis and dimensionality reduction, as described in the following section.

The use of machine learning (dimensionality reduction) in the latter case renders a world in which the absolute coordinates of each point has no more link to the descriptor space (the high sounds cannot be mapped to the y axis for instance), but offers compelling results for clustering information relating to the different playing styles of the instrument that is being analysed: as an example, in this extract [10] based on flute sounds, the opening shows a clear opposition between two types of gestures: 1/ staccato notes and 2/ legato scale-like type of material. This contrast in sound is made explicit by a movement of the avatar which jumps from a cluster of buttons to another.

5. WORKFLOW

After numerous attempts at listening to and playing with this system [11], I am now investigating how to diversify in VR metaphors for exciting the corpus based synthesis engine, as well as the different ways in which the synthesis may be rendered on an orchestra of RaspberryPi-equipped loudspeakers (see Chapter 6).

5.1 Corpus Selection

My experiments have focussed on musical instrument corpora almost exclusively. The tools presented here can efficiently generate plausible virtuosic instrumental music (as was sometimes the case in the piece Chef 2.0 discussed earlier), but recent uses found more satisfying results in slower, quieter, "Feldman-like" types of textures. Various limitations on the playback side (either in VR, or on a Pure Data sampler for RaspberryPi described in Chapter 6) have imposed restrictions in the first stages on the amount of data it could handle (less than 5minutes in AIFF in VR) or the number of slice the sample could be chunked into (256 because of limitation of lists in Max). Both limitations were later overcome (use of the ogg format in VR, increase of internal buffer size in fluid.buf2list), thus allowing for far more convincing models.

5.2 Analysis in FluCoMa

Using concatenative synthesis to model an improvising instrumental musician typically involves several steps:

1. Segmentation of a large soundfile: This involves dividing a large audio recording of the musician's performance into smaller units or segments.

[3] MFCC stands for Mel-Frequency Cepstral Coefficients. It is a type of feature extraction method that is commonly used in speech and speaker recognition systems. MFCCs are used to represent the spectral characteristics of a sound in a compact form that is easier to analyze and process than the raw waveform. They are calculated by applying a series of transformations to the power spectrum of a sound signal, including a Mel-scale warping of the frequency axis, taking the logarithm of the power spectrum, and applying a discrete cosine transform (DCT) to the resulting coefficients. The resulting coefficients, which are called MFCCs, capture the spectral characteristics of the sound and are commonly used as features for training machine learning models for tasks such as speech recognition and speaker identification.

[4] https://learn.flucoma.org/reference/onsetslice/

[5] https://learn.flucoma.org/reference/bufstats/

[6] https://youtu.be/UNj7_TI8SVs

[7] Simulation: https://www.youtube.com/watch?v=MkMKVm3G3W8 Result: https://youtu.be/Sc4Ye_rnSO8?t=9893. Although not relevant to the discussion here, the arm movement of the conductor was mapped to the speed of the cursor on the performers' screens with the help of INScore [27] and Gesture Follower [28] in MuBu for Max.

[8] https://youtu.be/DC_BL_HGPLA

[9] https://youtu.be/1LHcbYh2KCI?t=19

[10] https://youtu.be/777fqIIJCY4

[11] For cello: https://youtu.be/L-MiKmsIzjM For various instruments: https://www.youtube.com/playlist?list=PLc_WX6wY4JtnNqu4Lwe2YzEUq9S1IMvUk For flute: https://www.youtube.com/playlist?list=PLcWX6wY4JtlbjLuLHDZhlx78sTDm

2. Analysis: These segments are then organised in a database according to various descriptor data (mfcc in our case).
3. Scaling/pre-processing: scaling is applied for better visualisation.
4. Dimension reduction: Based on mfcc descriptors, the dimensionality of the data is reduced in order to make it more manageable and easier to work with. This can be done using techniques such as principal component analysis (PCA) singular value decomposition (SVD), or Uniform Manifold Approximation and Projection (UMAP, preferred in our case).
5. Near neighbours sequencing: Once the segments have been organised and analysed, the software selects and combines them in real-time based on certain input parameters or rules to create a simulated musical performance that sounds like it is being improvised by the musician. We use here a near neighbours algorithm, which selects segments that are similar in some way (e.g., in terms of pitch, loudness, or timbre - thanks to similarities revealed by umap on mfccs in our case) to the current segment being played.

We will now describe these steps in further detail:

5.2.1 Slicing

We saw in Chapter 3 how slicing musically trigger possibilities. In MuBu onset detection is done with pipo.onseg or pipo.gate. FluCoMa expose five different onset detection algorithms:

1. fluid.ampslice: Amplitude-based detrending slicer
2. fluid.ampgate: Gate detection on a signal
3. fluid.onsetslice: Spectral difference-based audio buffer slicer
4. fluid.noveltyslice: Based on self-similarity matrix (SSM)
5. fluid.transcientslice: Implements a de-clicking algorithm

Onsetslice only was extensively tested. The only tweaked parameters were a straigh-forward "threshold" as well as a "minslicelength" argument, determining the shortest slice allowed (or minimum duration of a slice) in hopSize. This introduce a common limitation in CBCS: the system strongly biaises the user to choose short samples for better analysis results, and more interactivity, when controlling the database with a gesture follower. Aaron Einbond remarks in the use of CataRT how short samples most suited his intention: "*Short samples containing rapid, dry attacks, such as close-miked key- clicks, were especially suitable for a convincing impression of motion of the single WFS source. The effect is that of a virtual instrument moving through the concert hall in tandem with changes in its timbral content, realizing Wessel's initial proposal.*"[29]

A related limitation of concatenative synthesis lies in the fact that short samples will demonstrate the efficiency of the algorithm [12], but at the same time moves away from the "plausible simulation" sought in the present study. A balance therefore must be found between the freedom imposed by large samples, and the refined control one can obtain with short samples.

A direct concatenation of slices clicks in most cases on the edit point, which can be avoided through the use of ramps. The second most noticeable glitch on concatenation concerns the interruption of low register resonances, which even a large reverb fails making sound plausible. Having a low threshold and large "minslicelenght" results in equidistant slices, all of identical durations, as would do the pipo.onseg object in MuBu.

Because we listen to sound in time, this parameter responsible for the *duration of samples* is of prior importance.

5.2.2 mfcc on each slice - across one whole slice/segment

Multidimensional MFCC analysis: MFCC (Mel-Frequency Cepstral Coefficient) analysis is a technique used to extract features from audio signals that are relevant for speech and music recognition. It involves calculating a set of coefficients that represent the spectral envelope of the audio signal, and can be used to capture the spectral characteristics of the musician's playing style.

5.2.3 statical analysis over each slice

BufStats is used to calculate statistical measures on data stored in a buffer channel. A buffer here is a type of data structure that holds time-series information, audio descriptor data in this case. BufStats calculates seven statistics on the data in the buffer channel: mean, standard deviation, skewness, kurtosis, low, middle, and high values. These statistics provide information about the central tendency of the data and how it is distributed around that tendency. In addition to calculating statistics on the original buffer channel, BufStats can also calculate statistics on up to two derivatives of the original data, apply weights to the data using a weights buffer, and identify and remove outlier frames. These statistical measures can be useful for comparing different time-series data, even if the original data is different lengths, and may provide better distinction between data points when used in training or analysis. The output of BufStats is a buffer with the same number of channels as the original data, with each channel containing the statistics for its corresponding data in the original buffer.

5.2.4 Normalization

The FluCoMa package proposes several scaling/preprocessing tools, amongst which normalization and standardization were used. Standardization and normalization are techniques used to transform variables so that they can be compared or combined in statistical analyses. Both techniques are used to make data more comparable, but they work in slightly different ways.

Standardization involves scaling a variable so that it has a mean of 0 and a standard deviation of 1. This is done by

[12] e.g. https://youtu.be/LD0ivjyuqMA?t=3032

subtracting the mean of the variable from each data point and then dividing by the standard deviation. Standardization is often used when the variables being compared are on different scales or have different units of measurement. It allows for comparison of variables that would otherwise be difficult to compare directly.

Normalization involves scaling a variable so that it has a minimum value of 0 and a maximum value of 1. This is done by subtracting the minimum value of the variable from each data point and then dividing by the range (i.e., the difference between the maximum and minimum values). Normalization is often used when the variables being compared have a skewed distribution, or when the variables are not normally distributed. It allows for comparison of variables that would otherwise be difficult to compare directly due to the skewness of their distribution.

Standardization scales a variable to have a mean of 0 and a standard deviation of 1, while normalization scales a variable to have a minimum value of 0 and a maximum value of 1. Normalization ccaling was found easier to use both in 2-D (in FluCoMa, the fluid.plotter object), as well as in the VR 3D world in which the origin corresponds to a corner of the world. The fluid.nomalize object features an "@max" attribute (1 by default), which then maps directly to the dimensions of the VR world.

5.2.5 Dimensionality Reduction

Dimensionality reduction is a technique used in machine learning to reduce the number of features (dimensions) in a dataset. The goal of dimensionality reduction is to simplify the data without losing too much information. Various dimensionality reduction algorithms are presented in an early FluCoMa study[30], with interestingly no mention of UMAP, later favoured.

SOM in one of the most popular algorithms for dimensionality reduction. It is implemented in the ml.star[31] library for Max, a simple hands-on library for machine learning, just one amongst the vast ammount of frameworks and machine learning algorithm famous across the NIME community [32] [33] [34] [35].

SOM (Self-Organizing Map) and UMAP (Uniform Manifold Approximation and Projection) are both techniques for dimensionality reduction. SOM is a type of neural network that is trained using unsupervised learning. It consists of a grid of neurons, each of which is associated with a set of weights. The SOM is trained by presenting it with input data and adjusting the weights of the neurons so that similar input patterns are mapped to nearby neurons on the grid. The resulting map is a low-dimensional representation of the input data that preserves the topological structure of the original data. UMAP, on the other hand, is a non-linear dimensionality reduction technique that is based on the principles of topological data analysis. It uses a combination of techniques such as k-nearest neighbours, weighted graph construction, and low-dimensional embedding to produce a low-dimensional representation of the input data. Unlike SOM, which is limited to a fixed grid structure, UMAP can produce a continuous, flexible representation of the data. Both SOM and UMAP can be useful for visualising high-dimensional data and for discovering patterns and relationships in the data. However, UMAP has some advantages over SOM, including the ability to handle large datasets more efficiently and the ability to produce more interpretable results.

UMAP (Uniform Manifold Approximation and Projection) can be used to visualize high-dimensional data in a lower-dimensional space. When applied to sound data analysed with MFCC (Mel-Frequency Cepstral Coefficients), UMAP reduces the dimensionality of the data and creates a visual representation of the sound in a 2- or 3-dimensional space. MFCCs, again, are a feature extraction technique commonly used in speech and audio processing. They involves decomposing a sound signal into a set of frequency bands and representing the power spectrum of each band with a set of coefficients. The resulting MFCC coefficients capture important spectral characteristics of the sound signal (albeit hardly interpretable by the novice user), such as the frequency and magnitude of the spectral peaks. By applying UMAP to the MFCC coefficients of a sound signal, it is possible to create a visual representation of the sound that preserves the relationships between the different MFCC coefficients (see Fig. 2). This can be useful for tasks such as exploring the structure of a sound dataset, identifying patterns or trends in the data, and comparing different sounds.

Figure 2. Dimensionality reduction of MFCCs help revealing spectral similarities. UMAP outputs coordinates in 2d or 3d.

UMAP is therefore used for its clustering abilities in the first place, helping for classification purposes. It helps identifying patterns or trends that may not be evident from the raw data. Most importantly, the non-linear dimensions proposed by UMAP (whether in 2d in Max or in 3 dimensions in PatchXR, and when compared to linear analyses in which, for instance, x, y and z correspond to pitch, loudness and centroid) gave far more "intelligent" clustering than more conventional parameter-consistent types of representations.

5.2.6 Neighbourhood queries

The neighbourhood retrieval function in a slightly different way each time, but is based in FluCoMa on K-d trees and and the knn algorythm. In MuBu , the mubu.knn object, as well as the ml.kdtree object ml.star, give very comparable result than those achievable with fluid.kdtree.

K-d trees (short for "k-dimensional trees") and k-nearest neighbours (k-NN) are two algorithms that are related to each other, but serve different purposes.

A k-d tree is a data structure that is used to efficiently store and query a set of points in a k-dimensional space. It works by partitioning the points into a binary tree, with each node in the tree representing a hyperplane that splits the space into two halves. The points are recursively partitioned into the left and right subtrees based on which side of the hyperplane they fall on. By organising the points in this way, it is possible to quickly find the nearest neighbours of a given point by searching only a subset of the tree rather than the entire set of points.

On the other hand, the k-NN algorithm is a machine learning algorithm that is used for classification or regression. Given a set of labeled points and a new, unlabelled point, the k-NN algorithm determines the k points in the set that are nearest to the new point, and then uses the labels of those points to predict the label of the new point. The value of k is a hyper-parameter that is chosen by the user, and it determines the number of neighbours that are considered when making the prediction.

In summary, a k-d tree is a data structure that is used to store and efficiently query a set of points in a k-dimensional space, while the k-NN algorithm is a machine learning algorithm that is used for classification or regression. Both algorithms are often used in applications such as pattern recognition, image classification, and data mining.

While CataRT or Audiostellar are typically used for generation of electronic textures/sound design, I have most often used FluCoMa to generate monophonic instruments (one performer plays one instrument at a time), in which the avatar reproduces what knn would do with an automated instrument: he will privilege in his choice the sample he can reach at hand, rather than jump large distance between 2 items (see Fig. 3) .

5.3 PatchXR

PatchXR [36] is a playful digital audio workstation for making music in VR. It's core metaphor corresponds to what the FluCoMa team calls CCE (creative coding environments) insofar as it functions in many ways like Max or Pure Data.

One reason for using VR to explore a 3D dataset is that it allows users to interact with the data in a more natural and immersive way, using it as a tool for data visualisation and analysis. Users can move around and explore the data from different angles, which can help them to better understand the relationships between different data points and identify patterns. Users get a more intuitive sense of the data and better understand how it is structured and how the different data points relate to one another.

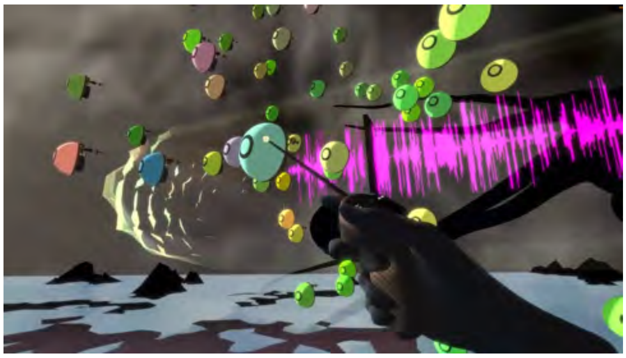


Figure 3. A VR interface in which each button in the world corresponds to a slice of the sound file. Machine learning helps bringing closer sounds that share common spectral characteristics.

The structure of a .patch file (a patchXR world) follows the syntax of a .maxpat (for Max) or .pd file (for pure data) in the sense that it first declares the objects used, and then the connexions between them. This structure made it relatively trivial to generate a javascript routine taking as input a dictionary (json file) with each segment's 3d coordinates, and as output a new .patch file (a world accessible in VR, see general workflow on Fig. 4).

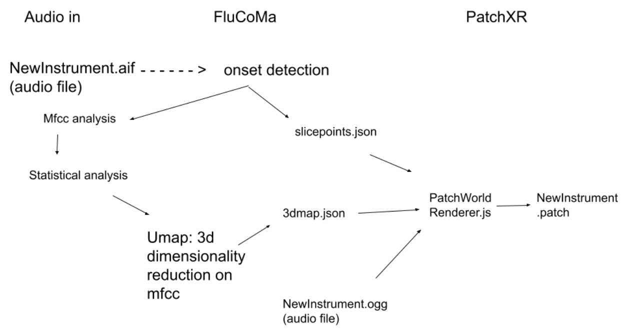


Figure 4. General Workflow: from an input audio file to its .patch 3d representation in PatchXR.

5.3.1 Interaction and OSC communication

PatchXR exposes a wide range of blocks (a block corresponds to an object in Max or Pure Data) making it simple to access gesture data such as:

- The position/distance between hands/controllers and a reference.
- The rotation angles (x y z) of both hands' controllers
- 2-d touchscreen like controllers, where the user moves the xy position of a selector across a plane by manually grabbing it.
- 2-d lazer-like controllers, where the user moves the xy position of a selector remotely, as if using a lazer pointer towards a remote screen

- 2-d pads, which allow to access the velocity at which the pad is hit
- 3-d theremine like controllers, where the user moves the xyz position of a selector across a plane by manually grabbing it.
- 1-d sliders, knobs, buttons...

One of the current challenges consists in diversifying the ways in which the corpus is queried. One to one mapping of UMAP results such as those described in Chapter 5.3 favour simulations for soloists, or duo in a multiplayer mode, in which the button interface buttons are facing each other, in order to prompt the players to face each other (see https://youtu.be/LP1g79BdIpY).

A simulation for more instruments, particularly when playing alone, encourages to use higher level type of control over automata, most importantly the simple ability to automatically concatenate: play the next sample as soon as the previous one has stopped.

6. FUTURE WORKS: THE RASPBERRY PI ORCHESTRA

In the frame of an artist residency at UCA (Université côte d'azur), the investigation questions how the tools presented above (those concern with the domain of Music Information Retrieval - MIR) may serve the control an immersive platform made of an orchestra of 64 *Pré* modules. [37]

At the time of writing this report, most satisfying results were achieved by sending messages to each RaspberriPi independently, according to its specific (static) IP address, with a simple syntax of a 2-integer list corresponding to: 1/which buffer to lookup 2/ which slice in this buffer to play. Pursuing on elaborations of timbre-space illustrations, the *Pré* modules, with the different acoustics its mobility allows, will encourage contrasted density of events according to the acoustics of the space in which the listening experience is happening.

7. CONCLUSIONS

After a discussion on the valuation of scores as maps, we've proposed a workflow for corpus-based concatenative synthesis CBCS, arguing that machine learning tools for data visualisation offer revealing and exploitable information about the timbral quality of the material that is being analysed. From a composer's perspective, the disappearance of the x time axis prompts to envisage composition not in a narrative sense (understood according reflections on new media and the notion of data-based art developed by [7]), but rather, as "time canvasses", as understood by Morton Feldman.

The discussed tools for "machine listening" (FluCoMa, MuBu) help building intelligent instruments with relatively small amounts of data, the duration of samples appear crucial in CBCS. A balance must be found between 1/ short duration sample analysis which are easier to process and categorise and 2/ long samples which sound more natural in the context instrument-based simulations.

Acknowledgments

I am grateful for the support of UCA/CTEL, whose artist residency research program has allowed to hold these experiments, and for the support of PRISM-CNRS.

8. REFERENCES

[1] D. Miller, "Are scores maps? a cartographic response to goodman," in *Proceedings of the International Conference on Technologies for Music Notation and Representation – TENOR'17*, H. L. Palma, M. Solomon, E. Tucci, and C. Lage, Eds. A Coru na, Spain: Universidade da Coru na, 2017, pp. 57–67.

[2] U. Eco, P. Eco, A. Cancogni, and D. Robey, *The Open Work*. Harvard University Press, 1989. [Online]. Available: https://books.google.fr/books?id=7jroM0M8TuwC

[3] N. Goodman, *Languages of Art: An Approach to a Theory of Symbols*. Hackett, 1976. [Online]. Available: https://books.google.fr/books?id=e4a5-ItuU1oC

[4] C. Vear, *The Digital Score: Musicianship, Creativity and Innovation*. Routledge, 2019. [Online]. Available: https://books.google.fr/books?id=oSblwQEACAAJ

[5] D. Schwarz, G. Beller, B. Verbrugghe, and S. Britton, "Real-Time Corpus-Based Concatenative Synthesis with CataRT," in *9th International Conference on Digital Audio Effects (DAFx)*, Montreal, Canada, Sep. 2006, pp. 279–282, cote interne IRCAM: Schwarz06c. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01161358

[6] D. Schwarz, R. Cahen, and S. Britton, "Principles and Applications of Interactive Corpus-Based Concatenative Synthesis," in *Journées d'Informatique Musicale (JIM)*, Albi, France, Mar. 2008, pp. 1–1, cote interne IRCAM: Schwarz08a. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01161401

[7] L. Manovich, "Database as symbolic form," *Convergence: The International Journal of Research into New Media Technologies*, vol. 5, pp. 80 – 99, 1999.

[8] M. Feldman, B. Friedman, and F. O'Hara, *Give My Regards to Eighth Street: Collected Writings of Morton Feldman*, ser. Exact Change. Exact Change, 2000. [Online]. Available: https://books.google.fr/books?id=hfgHAQAAMAAJ

[9] J.-P. Briot, G. Hadjeres, and F.-D. Pachet, *Deep Learning Techniques for Music Generation – A Survey*, Aug. 2019. [Online]. Available: https://hal.sorbonne-universite.fr/hal-01660772

[10] P. Esling, A. Chemla-Romeu-Santos, and A. Bitton, "Generative timbre spaces with variational audio synthesis," *CoRR*, vol. abs/1805.08501, 2018. [Online]. Available: http://arxiv.org/abs/1805.08501

[11] D. L. Wessel, "Timbre space as a musical control structure," *Computer Music Journal*, vol. 3, no. 2, pp. 45–52, 1979. [Online]. Available: http://www.jstor.org/stable/3680283

[12] K. Fitz, M. Burk, and M. McKinney, "Multidimensional perceptual scaling of musical timbre by hearing-impaired listeners," *The Journal of the Acoustical Society of America*, vol. 125, p. 2633, 05 2009.

[13] C.-E. Cella, "Orchidea: a comprehensive framework for target-based computer-assisted dynamic orchestration," *Journal of New Music Research*, vol. 0, no. 0, pp. 1–29, 2022. [Online]. Available: https://doi.org/10.1080/09298215.2022.2150650

[14] A. Agostini and D. Ghisi, "A Max Library for Musical Notation and Computer-Aided Composition," *Computer Music Journal*, vol. 39, no. 2, pp. 11–27, 06 2015. [Online]. Available: https://doi.org/10.1162/COMJ_a_00296

[15] D. Ghisi and C. Agon, "dada: Non-standard user interfaces for computer-aided composition in max," in *Proceedings of the International Conference on Technologies for Music Notation and Representation – TENOR'18*, S. Bhagwati and J. Bresson, Eds. Montreal, Canada: Concordia University, 2018, pp. 147–156.

[16] J.-C. Risset and D. Wessel, "Exploration of timbre by analysis and synthesis," *Psychology of Music*, pp. 113–169, 1999.

[17] S. Mcadams, S. Winsberg, S. Donnadieu, G. De Soete, and J. Krimphoff, "Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes," *Psychological research*, vol. 58, pp. 177–92, 02 1995.

[18] A. Caclin, S. Mcadams, B. Smith, and S. Winsberg, "Acoustic correlates of timbre space dimensions: A confirmatory study using synthetic tones," *The Journal of the Acoustical Society of America*, vol. 118, pp. 471–82, 08 2005.

[19] D. Schwarz, "The Sound Space as Musical Instrument: Playing Corpus-Based Concatenative Synthesis," in *New Interfaces for Musical Expression (NIME)*, Ann Arbour, United States, May 2012, pp. 250–253, cote interne IRCAM: Schwarz12a. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01161442

[20] L. Garber, T. Ciccola, and J. C. Amusategui, "Audiostellar, an open source corpus-based musical instrument for latent sound structure discovery and sonic experimentation," 12 2020.

[21] B. Hackbarth, N. Schnell, P. Esling, and D. Schwarz, "Composing Morphology: Concatenative Synthesis as an Intuitive Medium for Prescribing Sound in Time," *Contemporary Music Review*, vol. 32, no. 1, pp. 49–59, 2013. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01577895

[22] N. Schnell, A. Roebel, D. Schwarz, G. Peeters, and R. Borghesi, "Mubu and friends -assembling tools for content based real-time interactive audio processing in max/msp," *Proceedings of the International Computer Music Conference (ICMC 2009)*, 01 2009.

[23] P. A. Tremblay, G. Roma, and O. Green, "Enabling Programmatic Data Mining as Musicking: The Fluid Corpus Manipulation Toolkit," *Computer Music Journal*, vol. 45, no. 2, pp. 9–23, 06 2021. [Online]. Available: https://doi.org/10.1162/comj_a_00600

[24] F. Bevilacqua and R. Müller, "A gesture follower for performing arts," 05 2005.

[25] N. Schnell, D. Schwarz, J. Larralde, and R. Borghesi, "Pipo, a plugin interface for afferent data stream processing operators," in *International Society for Music Information Retrieval Conference*, 2017.

[26] R. Fiebrink and P. Cook, "The wekinator: A system for real-time, interactive machine learning in music," *Proceedings of The Eleventh International Society for Music Information Retrieval Conference (ISMIR 2010)*, 01 2010.

[27] D. Fober, Y. Orlarey, and S. Letz, "INScore - An Environment for the Design of Live Music Scores," in *Linux Audio Conference*, Stanford, United States, 2012, pp. 47–54. [Online]. Available: https://hal.archives-ouvertes.fr/hal-02158817

[28] F. Bevilacqua, B. Zamborlin, A. Sypniewski, N. Schnell, F. Guédy, and N. Rasamimanana, "Continuous realtime gesture following and recognition," vol. 5394, 02 2009, pp. 73–84.

[29] A. Einbond and D. Schwarz, "Spatializing timbre with corpus-based concatenative synthesis," 06 2010.

[30] G. Roma, O. Green, and P. A. Tremblay, "Adaptive mapping of sound collections for data-driven musical interfaces," in *New Interfaces for Musical Expression*, 2019.

[31] B. D. Smith and G. E. Garnett, "Unsupervised play: Machine learning toolkit for max," in *New Interfaces for Musical Expression*, 2012.

[32] M. M. Wanderley and M. Battier, "Trends in gestural control of music," 2000.

[33] F. Bevilacqua, R. Müller, and N. Schnell, "MnM: a Max/MSP mapping toolbox," in *New Interfaces for Musical Expression*, Vancouver, France, May 2005, pp. 85–88, cote interne IRCAM: Bevilacqua05a. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01161330

[34] R. Fiebrink and B. Caramiaux, "The machine learning algorithm as creative musical tool," *ArXiv*, vol. abs/1611.00379, 2016.

[35] B. Caramiaux and A. Tanaka, “Machine learning of musical gestures,” in *New Interfaces for Musical Expression*, 2013.

[36] V. Bauer and T. Bouchara, “First steps towards augmented reality interactive electronic music production,” in *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, 2021, pp. 90–93.

[37] *≪ PrÉ ≫ : connected polyphonic immersion*. Zenodo, Jul. 2022. [Online]. Available: https://doi.org/10.5281/zenodo.6806324

Interpreting notated works using the *Terpsichora* Pressure-Sensitive Floors

Iran Sanadzadeh
Monash University
Melbourne, Australia
Iran.sanadzadeh@monash.edu

Cat Hope
Monash University
Melbourne, Australia
Cat.hope@monash.edu

ABSTRACT

The *Terpsichora* Pressure-Sensitive Floors are a new digital musical instrument which uses whole-body motion to control electronic music. The instrument continues the development of early models for pioneering dancer Philippa Cullen (1972), expanding its use as an expressive and versatile musical instrument. Two works by Australian composer Cat Hope were adapted for performance with this new instrument. *Delay Taints* (2018), for dancer, cellist and subtone, is an animated graphic score that provided an opportunity to freely assign sonic choices to the instrument, and read notated body movement to control those choices. This adaptation contrasts with that of *Majority of One* (2016), for four sustaining instruments and room feedback, where two of the notated parts were interpreted on the instrument. Methods to produce continuously controlled sound using limited movements of the body were developed to replace the instruments featured in the original performances of this work. This work explores the difference in the embodied connection of gesture to sound between acoustic and electronic instruments and explores the idiosyncrasies in the navigation of time elements in music for the Floors. In addition, methods of performing with the Floors produces a new form of communicating electronic performance to audiences using full body gesture. Interpreting these two compositions by Hope using the *Terpsichora* Pressure-Sensitive Floors contributes new strategies for adapting animated scores for electronics using direct body movement.

1. INTRODUCTION

1.1 The Terpsichora Floors and Cullen's originals

In the past three decades, the development of gestural controllers and digital musical instruments interfacing with digital audio workstations has rapidly expanded to bridge the connectivity gap in performance of electronic music with computers [1, 2, 3]. Philippa Cullen, a prolific Australian choreographer and dance artist working with sound, created a range of instruments in the early 1970s through collaborations with designers and technologists [4]. Cullen aimed to gain a new level of control of sound as a dancer, freeing her movements from sounds created by a separate instrument or performer. Cullen's *Vernus* Pressure-Sensitive Floors, built in 1972, remain a distinct application of interaction design, as the movement of the entire body is required to control a single sensor underneath each of her four Pressure-Sensitive Floors, as seen in Figure 1. The Floors are made of individual platforms that fit together and can be used in anu number or combination. Most synthesiser controllers use many small sensors and control surfaces to allow for control of many parameters; in contrast, Cullen's floors use a large interactive surface for fine control of few sonic parameters. On this instrument, data is not directly observed from the body, and the platforms do not require specific physical movements such as tapping or a particular gesture of a limb to be activated. Rather, they respond to changes in pressure, applied via body weight, on each of the different floor platforms. The new instrument follows the original design principles: the *Terpsichora* Pressure-Sensitive Floors are a set of wooden platforms, sending control voltages, and respond to movement using a load cell; each of the floors have one sensor underneath them [5]. This more recent version of the Pressure-Sensitive Floors (hereon referred to as 'the Floors'), built in 2015, works as a MIDI controller by converting the control voltage signal, interfacing with Max/MSP and Ableton Live software. The design and implementation in a DAW is combined with the development of a gestural vocabulary for performance on the new Floors. This vocabulary is used to increase the expressivity of performance, toward the Floors operating primarily as an expressive instrument, rather than a performance interface [6].

Figure 1 Left: One of Cullen's four identical *Vernus* Pressure-Sensitive Floors, built in 1972. Right: The *Terpsichora* Pressure-Sensitive Floors, built in 2015. 7 platforms forming a tangram shape, each with one sensor underneath, responding to movement

With a new access to expressivity, the Floors have recently entered the space of navigating composition. The gestural vocabulary developed for the instrument and the instrument's performance affordances are still in an early stage of gestation. For their development, new compositions for the instrument, adapting existing works, and improvisation are all necessary components. Much of the

Sanadzadeh's performance on the Floors is not notated, and pieces which are notated are often done in a shorthand style specific to the performer. Thus, adapting existing scores for movement and sound was considered simultaneously a diagnostic tool about the instrument, a creative endeavour and a method for growing the performance and design of the instrument.

2. ADAPTATION OF CAT HOPE SCORES

Two works by Cat Hope were adapted to be performed with the Pressure-Sensitive Floors and acoustic instruments. The first, *Majority of One,* is an animated, scrolling timeline-based score with a playhead indicating the point of performance. Originally conceived for sine tone oscillators alongside acoustic instruments [7], a 2021 recording of the work performed by Hope's ensemble Decibel [8] features four acoustic instruments and feedback. The performance featuring the Floors on two of the four 'instrumental' parts, featured bass flute and double bass on the other two parts [9]. *Majority of One* notates sonic directions using different coloured lines that sweep up and down between 'highest and lowest point' markers. The accompanying instructions for the piece explain how the lines are read by different instruments, noting that the performances should "Follow your colour for the whole piece… the movement mut always be seamless and without interruption. Bend your note by whatever mechanism you choose, but it must be without steps [7]. Adapting two parts for the Floors required simultaneous interpretation of the differing pitch trajectories of each part, presenting a unique challenge for performance on the instrument. A way of compartmentalising the Floors for the purposes of this investigation was examining it as an electronic instrument and separately considering human motion in performance and notation. Rather than having a new piece composed for the instrument, which would be designed to work with its existing performance style, adapting two pieces written for an electronic oscillator (a fundamental element of electronic music) and a score for movement, would offer unique findings. In a performance examining these divided parts of using the Floors, the affordances of the instrument could be better contextualised. First, in comparison with a sine wave oscillator in *Majority of One*, the adjustments required to perform a fundamental element of synthesis directly were found; the freedom and sonic possibilities in interpreting notation for human movement, outside of the medium of dance, were examined in adapting *Delay Taints*.

Delay Taints, initially composed for dancer, cello and subtone is also an animated, scrolling score with the subtone embedded in the digital score [10]. Both works are presented in the Decibel ScorePlayer software on iPad [11]. In this performance of *Delay Taints,* the Floors read the 'dancer' notation, and a double bass read the cello part. The interpretation of the dance part for this instrument provided an avenue to explore the shape and limitations of the connection between movement and sound when performing on the Floors.

In the adaptation of the dance part for the Floors, two factors were considered: first, the appropriate gestural vocabulary with which to adapt the action directions of the score; second, a responsive sound design to enable musical synergy to be achieved between the double bass and Floors part. The double bass actions can be more directly interpreted by the player, yet complex action shapes in the score present challenges to interpretation for the Floors part.

3. INTERPRETAITON OF *DELAY TAINTS*, AN ACTION SCORE

Delay Taints has two intersecting performance parts that use slow gestures, subdued dynamics, and small movements. On the score, the height of each line is an indicator of pitch for the cellist and height of the dancer's body in relation to the ground. Lines, circles and crosses are directed in the score instructions as "Crosses for a gentle, Bartok pizzicato, or hard clap/slap/hit; smooth round dots are gentle pizzicato or finger click/softer sound. Dynamics are soft unless the line is thicker" [10]. To interpret this action score, dancer movements were first interpreted in terms of their efficacy to create sound on the Floors. This adaptation involved considering the change in pressure exerted on the Floors by changing the positioning and height of the player. The sounds that were chosen for the performance mapping were intended to stay outside of the double bass' timbres, but allow for moments of unity where the two lines intersect on the score (see Figure 2). The score provides instruction for movement, and the Floors sonify that movement in unique ways.

Figure 2. A screenshot of the *Delay Taints* score in the Decibel ScorePlayer, showing intersection of the instrument in blue (double bass) with the dancer (the Floors) in red lines.

3.1 Adaptation of Dancer Part to Floors

As the Floors respond to movement, a dancer's body is the closest analogy to a performing body for the player of the Floors. However, whilst for dancers, the torso, arms, and the head can move expressively and independently from each other, for the player of the Floors, their movement affects the subtle shift of the entire body's weight, thus impacting the sound as it responds to changes in pressure. Whilst independent movements of the arms and neck are inconsequential to the sound, movements of the torso strongly affect the sonic outcome; much of pressure distribution of the body is determined by the relative positioning of the hips and shoulders. The placement of feet on each platform, the relative weight on each foot, and the shifting of this weight across the instrument are the focal points of performance [12]. This anchoring of the body in the feet means the movements of the Floors player are more restricted than that of the dancer. Figure 3 shows stills from the movement of dancer Laura Boynes performing in the premiere performance of *Delay Taints* and that of the Floors performer in the 2022 adaptation.

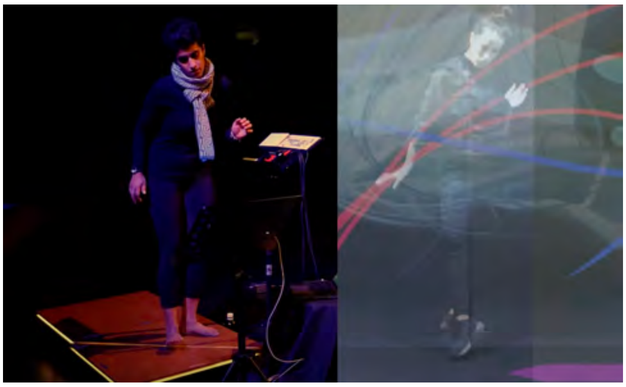


Figure 3. Comparison in movement adaptation and similarities between movement in Floors adaptation (left) and Dancer in the original adaptation (right)

Expressivity and the role of the performer are distinct between the dancer and the Floors performer. Watching a performance of electronic music with the Floors is more dance-like, as the entire body of the performer is the main mode of communication between audience and player [13]. This mode creates a point of connection between the electronically-generated sound and the gesture, thus bringing a new type of liveness to the performance [14, 15]. However, unlike a dancer, the individual shape of movement is not a priority for the Floors player. The sonic result is the focus of the performer, even when performing an action score. In performance with the Floors, the gesture is in service of creating sound, and the action score is interpreted as one intended to create musical meaning. This difference in priority changes how the score is interpreted.

Increasing height for the performer of the Floors does not always aid an accurate realisation of a score, as it can limit the control of individual placements of weight accessible for the performer. In this way, the original score's direction of height for the dancer needed to be adapted [10]. This height was translated to changing the amount of pressure on each platform. Much of the control of the Floors in performance is done with small gestures and minute shifting of the weight of the body. The larger gestures in the score for *Delay Taints* were thus interpreted as larger musical gestures, rather than larger body gestures. In the performance, the author tried to access new sounds by using larger musical gestures which often involved complex small movements across the Floors. Figure 4 shows the performer on multiple Floors for complex gesture and on toes with smaller contact for interpreting height.

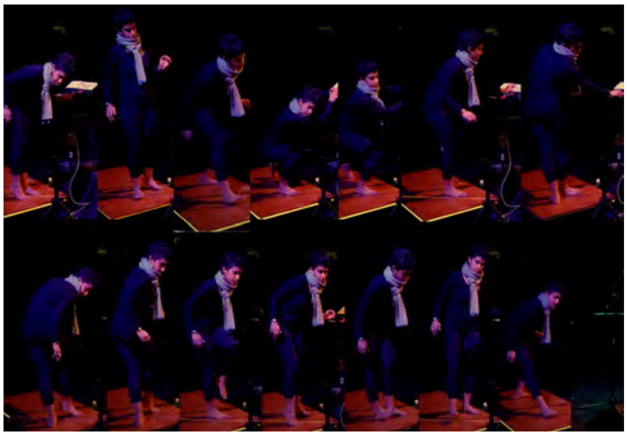


Figure 4. Compilation of gestures from Floors performer to interpret complex subtle movements in the score. Movement on the floors is combined with changes of sound on the computer

3.2 Interpretation of Intersecting lines

Delay Taints features varying number of lines for performance that appear, disappear and intersect. They also 'fade' in and out (see Figure 2). These movements are interpreted by the dancer in the original performance as individual and independent layers of action. For performance on the Floors, this interpretation needed adjustments as individual movements of the limbs of the performer are inconsequential to the sound. Two actions were taken for this interpretation

1. Taking the number of lines to direct the number of simultaneous layers of sound
2. Interpretation of individual occurrences of unique shape combinations as one sonic gesture or sonic motif, which recurs.

In performances with the Floors, two types of sounds are used: those triggered by the Floors, and looping sounds that are shaped by the movement. Triggered sounds, used for staccato crosses, are used with each floor. Thus, layers of individual staccato sounds are performed separately. The small number of used Floors in this interpretation (three out of an available seven) means that beyond 3 layers of staccato sound, the interpretation must fall on adding simultaneous layers. Many layers of sound can be added which can be triggered using the same action. However, as they are triggered simultaneously, using the same small number of sensors, they cannot then be independently controlled.

Within the design of the Floors, each looping sound is enabled using the keyboard and its parameters controlled using all the available platforms. Whilst performance with more Floors could enable separating layers of sound, the limitation of the human body, in restricted distribution of weight across limbs and being able to have limbs available per floor, remains at 4-5 Floors (using the feet, hands and knee of the performer).

3.3 Sound Design and Mapping of Movement

In the original form of *Delay Taints*, the dancer's expression is a silent one, thus the action score is translated to a visual and kinetic medium. Adapting it to a sonic interpretation meant that the shape of the dancer's gestures would need to be translated into timbres and placement of the sounds in time. The score part for the double bass was taken as an interpretive guide for this adaptation. Samples of 'bubble-like' sounds, created using the granular synthesiser in SuperCollider [16], as well as small crackle and bell sounds, were used to interpret in the short circle and dot notations, with longer oscillators and looping bell samples as those used for the sustaining line notations. The lines in the score were interpreted using individual channels of looping sounds to follow the movement of the lines in the Decibel ScorePlayer software [11]. Individually triggered sounds correspond to specific floors, but looped sounds are triggered across the entire instrument, with each the parameter of a looped sound (duration, harmonics, effects, ADSR envelope, etc.) controlled by an individual floor. Whilst performing a gesture, multiple parameters of all enabled looping sounds are changing, whilst triggered sounds are added with the same gestures. Thus, the performance of all simultaneous lines is inextricably linked and cannot be actioned individually.

The larger, linear vertical notation in the score were challenging to interpret sonically, since time is represented horizontally and as a whole gesture, meaning a vertical (or steep) line would be 'instantaneous'. To do this, a combination of activating multiple loops briefly and combining them with short triggered sounds was used. In performance of *Delay Taints*, triggering and reshaping of these lines limited the Floors performer's ability to move freely.

3.4 Comparing The Floors with Acoustic Instruments

Three distinct differences between The Floors and acoustic instruments shaped the interpretation of the score: the Floors lack of immediacy or instantaneous sound, the complex nature of the Floors mapping and the way they obscure cause and effect, as well as the layering of sounds [5].

First, the Floors have a smaller range of available 'immediacy' in performance. This is due to the mapping design chosen for the Floors, which dictates a complex divergent mapping of musical parameters, tied to their nature as continuous controllers, which send data as a stream, detecting change, rather than MIDI triggers, which would respond to individual actions on them.

Secondly, the Floors' mapping is complex due to multiple parameters being controlled with a small number of sensors. Thus, whilst the body movement appears to communicate the score to the audience in these works, the connection between action and sound remains hidden. By contrast, the physical movement of the performers body with the acoustic instruments can be observed readily by the audience, by watching the player's fingers landing directly on different notes, or moving across a string instrument's fingerboard.

Finally, the Floors can layer many sounds simultaneously. Unlike acoustic instruments, the dynamic and timbral range immediately available at any point is broad on the Floors, enabling a wider interpretation of action scores. In combination with the lack of immediacy, this facet enables a wider sonic range in duo performances with the Floors player but simultaneously obscures the method of sonification [14].

4. ACTIONING SONIC DIRECTIONS IN *MAJORITY OF ONE*

Unlike *Delay Taints,* the adaptation of *Majority of One* was sonically clear. Two notated parts on the score were performed on the Floors. Each part is an expression of pitch at a relatively low and unchanging amplitude throughout, and with rests interpreted as amplitude of 0. The pitches of each oscillator were placed on two of the Floors. This meant that to raise the pitch of one oscillator, more pressure needed to be applied to the corresponding floor. Increasing the pressure applied to a floor is done by either moving to a more flexible part of the platform or by increasing the amount of weight the performer puts on the floor [5]. The performer has a finite amount of weight; in performance, adding weight to one floor involves removing weight from another. Since the parts are independent, at points in the score one oscillator is changing in pitch whilst the other remains still. To accommodate for performing these independent lines, a new element was devised in the gestural vocabulary for the Floors. If an oscillator's pitch is changing whilst the other oscillator was remaining steady, a counter level of 'pushing down' on the steady floor was used whilst transferring weight to the floor corresponding to the changing oscillator (see Figure 5).

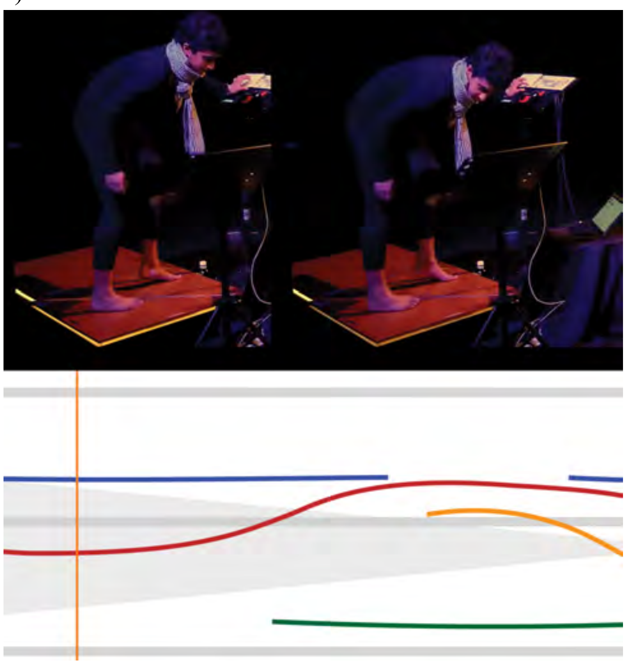


Figure 5 A screenshot (detail) of the score for *Majority of One*. The Floors performed the blue and red parts. Photograph shows the Floors performer in the corresponding moment pushing on left foot to compensate for weight distribution changing on right foot, which corresponds to red part. Movement of y-axis on trackpad facilitates silence in blue part.

Whilst *Delay Taints* allows for free movement of the performer, the sine-wave parts in *Majority of One* are precise in pitch, thus invite minute movements.

The amplitude of each sine wave part was controlled by the x and y position of the trackpad. This mapping strategy allowed for independently controlling the amplitude of each sine wave part. Executing the two parts thus became inextricably linked as four gestures of up/down for controlling the first sine wave part, left/right for controlling the second, and diagonally across in both directions facilitated amplitude changes to be executed. This mapping integrated with the movement of the timeline on the score in the slower sections of the piece. Yet in the latter parts, where oscillator entries and exits are rapid, this gesture did not facilitate the sharper entries and exits. This amplitude mapping inevitably gave a slower attack to each oscillator, which shaped the performance of these two parts.

5. OVERARCHING CONSIDERATIONS

Performing *Delay Taints* and *Majority of One* using The Floors revealed contrasts in the relationship between score, gesture, sound, and timbral range, as they present between the Floors and acoustic instruments.

In performance with the Floors, the relationship between action, sound and time, is different from acoustic instruments and from standard electronic instruments. Whilst the action can be quick, the Floors send a continuous signal, which broken up to create a trigger, means that an increased anticipation is required by the performer to 'ramp up' to the particular trigger, akin to sounding a large drum or bell. Simultaneously, the action of performance uses the shifting of weight, rather than a moment of contact with force, so unlike a mallet or clapper, the sounding action is akin to a heavy bow. This anticipation is distinct from latency since the instrument is sending out a continuous signal that is changed by the performer rather than individual triggers that are each enacted by the action of the performer. As the sounds are electronically generated instead of being activations of a resonant body, increased speed or force does not immediately translate to increased amplitude. These factors mean that the action and force from the Floors performer do not visibly reflect the sonic outcome. In addition, to change sounds, the use of the keyboard is required, which further creates a gap in the gesture; it also makes it harder for the audience to understand the relationship between action and sound.

Although the double bass has a longer anticipation time and attack for bowed sounds than other string acoustic instruments, it still retains a direct relationship between action of the player, sounding of the note, and resonance through the body. This inherent acoustical nature makes it a more immediate sounding body than that of the Floors, where parameters are shifted by moving weight. Mappings are complex and gradual to unfold, and there is no resonant body to sonify an action immediately. These differences meant that the faster reaction of the double bass part in *Delay Taints* did not immediately lock in with the sound of the Floors.

The nature of triggering sounds by shifting weight means that precise rhythms of movement and fast triggering of individual short sounds presents a challenge, as it requires a rapid movement of the entire body that is unlike the rapid movement of an arm and bow for the double bass; the bass player can instantaneously replicate individual small sounds (dots and cross notations), whilst the Floors performer must create a gesture that encapsulates a group of small triggered sounds together, as a cluster. Whilst the scores are linked, the action interpretation for the Floors player is temporally disjointed from the bass player where precise triggered moments are required [10].

The timbral range available to the Floors is far broader than acoustical instruments, as any sample can be programmed into the instrument. Enabling and disabling ranges of sound within *Delay Taints* thus considered two factors: time and cohesion. The enabling of individual types of sound for the Floors is done via a computer keyboard, affecting the movement by the Floors performer at the moment of enabling. To create the right sound when desired, timbral changes need to be enabled prior to action on the floor. This pre-selection is similar to the change in organ stops pulled for a timbral shift. Whilst this pre-selection is helpful to performance, its utility in rapid moments of shifting within the score or in response to improvised elements within a scrolling score prove challenging, and its preparation within a scrolling score is more difficult to program.

Similarly, in pitch, exactitudes are harder to achieve with the Floors. The range of the pitch change available needs to be calibrated to the needs of the performance and a very controlled lack of movement is required to avoid the pitch from changing. This consideration affects how the Floors interlink with the acoustic instruments in *Majority of One.* In the final section of this piece, where the oscillator pitches remain steady for some minutes, the corresponding floor values were 'frozen' to facilitate the requisite steadiness and allow the performer to breathe without affecting the pitch.

In the adaptation of *Majority of One*, the inclusion of two musical parts for the one Floors performer resulted in one complex interlinked part. One Floor platform was used for each notated part, connected to its own oscillator. The performer was linked to one or both parts simultaneously; natural required movements of the body had to be limited as they would affect the pitch of one of the two oscillators. In future adaptations of these two parts for one Floors performer, a new way to control amplitudes with two controllers can assist with the separation of the parts and freeing the body of the performer by creating a third contact point can help achieve more accuracy with the sine wave parts. For further accuracy of performance, amplitude mapping would need to be reconsidered in terms of facilitating faster changes with a new gesture for volume control and a way to centre the body of the performer.

The Floors provided the basis for the development of a unique technique for the gestural performance of these scores. Given the nature of the instrument and its response to weight and pressure, this technique is tied to the unique body of the performer. Spatz considers technique as that which is "not tied to specific bodies and local contexts" [17]. The findings of this research, whist facilitating a specific performer's action, have found techniques that are not unique to the performer, but translate broadly as adaptation strategies for other electronic gestural instruments. The type of movements considered by the performer and the embodied understanding of weight and motion provide a template for the use of other gestural controllers as well as an embodied understanding of electronic sound control. The effect of this has been observed in the use of other gestural controllers by the author and the adaptations of her movement technique in other players' gestural vocabulary of movement. The movement techniques in relation to the Floors themselves is useful for building a language of performance on the instrument, allowing for adaptation of other pieces for it and for its use by other performers. Here, the transferrable skills of new techniques in training other players appear as existent issues in performance with electronic gestural controllers [18]. An instrumental facility is developed on a bespoke controller, often for its techniques to remain locked to the single user-developer-performer figure. Through adaptation of existing action scores, and discussion of the developed movement language, it is the aim of this research to illuminate embodied issues in electronic performance that can aid performers of other gestural instruments and bring additional connection between players of different bespoke controllers.

6. CONCLUSIONS

The adaptation of *Delay Taints* and *Majority of One* for performance with the Floors provided a new pathways and considerations for the interpretation of digitally animated, action scores. Through interpreting the two pieces, new minute body movements were found to enable a change in one floor with minimised effect on the other, and new forms of interpretation were developed for action scores. In this process, sound design choices, mapping of movement parameters, and gestural control were examined to create a cohesion between the individual musical parts, to enable a coherent final work. The scrolling timeline nature of the two works allowed for a stronger connection between immediate gesture and sound elements using the Floors, yet there remains room for interpretation and ability to perform longer phrases with more expressivity.

In addition, the Floors provided a more expressive gestural performance of the sustaining instrument part. Issues of accuracy in timing and achieving expressive cohesion between the parts were observed along with differences between instruments in the connection between sound and gesture. The movement of the Floors player is visible unless subtle gradations of sound are achieved by minute shifting of weight. Using the Floors requires an indirect control of sonic parameters which further changes the relationship between action and sound. On this instrument, subtleties of electronic sound control in these pieces have become accessible in new ways whilst enriching the performance strategies of this instrument and informing its design in response to animated scores.

Acknowledgments

The authors would like to thank the New North concert series team for the curation of the program, Sebastian Collen for the hardware work on the Floors, and Monash University colleagues for their support.

7. REFERENCES

[1] S. Sapir "Gestural Control of Digital Audio Environments" *Journal of New Music Research,* vol. 31, no. 2, 2002, pp. 119-129.

[2] D. Wessel and M. Wright "Problems and Prospects of Intimate Musical Control of Computers" *Computer Music Journal* vol. 26, no. 3, 2002, pp. 11-22.

[3] E. Miranda, and M. Wanderley, *New Digital Musical Instruments: Control and Interaction beyond the Keyboards,* A-R Editions, 2006.

[4] S. Jones, "Philippa Cullen, Dancing The Music" *Leonardo Music Journal,* vol. 14, 2004, pp. 65-73.

[5] I. Sanadzadeh. *The Terpsichora Pressure-Sensitive Floors* [Online] Available https://www.youtube.com/watch?v=EbN7H2tqE6k

[6] A. Tanaka. "Musical Performance Practice on Sensor-based Instruments", in M. Wanderley, E. Miranda (Eds.), *Trends in Gestural Control of Music,* IRCAM, 2000.

[7] C. Hope. *Majority of One*. Monash University. Composition. 2019. https://doi.org/10.26180/8135354.v4

[8] Decibel. *Decibel - Cat Hope.* CD. 2022. ezz-thetics: LC 91771

[9] New North Concert Series. *New North Concert 5: Cat Hope 'Majority of One'* [Online] https://www.youtube.com/watch?v=A1Jxvt9ArJU

[10] C. Hope. *Delay Taints,* for cello, dancer and subtone, 2018. [DOI: 10.26180/5e2ac844a6413]

[11] C. Hope, A. K. Wyatt, and L. Vickery, "The decibel scoreplayer – a digital tool for reading graphic no- tation," in *Proceedings of TENOR, First International Conference on Technologies for Music Notation and Representation, 2015*, Institut de Recherche en Musi- cologie, 2015, p. 59.

[12] New North Concert Series. *New North Concert 5: Cat Hope 'Delay Taints'* [Online] Available https://www.youtube.com/watch?v=gTOM87Yw0-w&list=PLeuq3buh9Ey5VtDpPDg3Ya0Nc4AdvOuIJ&index=3

[13] I. Sanadzadeh. *Axes 2022* [Online] Available https://www.youtube.com/watch?v=nX4xboeNVKo

[14] J. Kirby. "Defining and evaluating the performance of electronic music", in R. Hepworth-Sawyer, J. Hodgson, J. Paterson, R. Toulson (Eds.) *Innovations in Music*, Routledge, 2019.

[15] T. West, B. Caramiaux, S. Huot and M. Wanderley "Making Mappings: Design Criteria for Live Performance" *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME 2021),* Online and Shanghai, 2021. https://doi.org/10.21428/92fbeb44.04f0fc35

[16] J. McCartney. 1996. SuperCollider. https://supercollider.github.io/

[17] B. Spatz, *What a Body Can Do: Technique as knowledge, Practice as Research.* London: Routledge, 2015.

[18] D. Levitin, S. McAdams and R. Adams "Control Parameters for Musical Instruments: a Foundation for New Mappings of Gesture to Sound," *Organised Sound* vol. 7 no. 2, 2002, pp. 171-189.

The Digital Score Project: Review of Ongoing Research into Digital Score Creativity and Digital Musicianship (2022-23)

Craig Vear
University of Nottingham Nottingham, UK
Craig.Vear@nottingham.ac.uk

Solomiya Moroz
University of Nottingham Nottingham, UK
Solomiya.Moroz@nottingham.ac.uk

ABSTRACT

This paper presents an overview of *DigiScore* research ac-tivities in the first 18 months of the project. *Digiscore* is an ERC-funded project studying the technological transfor-mation of the music score through technology. The paper is split into three sections each dealing with a major re-search theme. The first section presents design considera-tions for digital scores that have emerged out of the case studies, including findings from TENOR 2022, Marseilles, and other projects that incorporated technologies such as Unity gaming engines, machine learning, robotics and EEG readers. The second section focuses on two case stud-ies looking at the impact of this research upon inclusive and accessible music-making through digital scores. Spe-cifically, *Digital Syzygies* – a digital score as a shared cre-ative platform for d/Deaf and autistic musicians, and *Jess+*
- an AI/Robot Digital Score that operates as a creative in-terface between an ensemble of musicians of mixed abili-ties. The final section reports on insights from a digital mu-sicianship investigation that seeks new knowledge in the skills, perceptions, contexts, cultures, awareness and knowledges of digital musicians engaging with digital scores.

1. INTRODUCTION

The Digital Score: Technological Transformations of the Music Score (*DigiScore*) is a European Research Council funded research project operational between 2021-2026. The core aims of this frontier-research project are to: (1) determine scientific knowledge of how digital scores stimulate new creative opportunities and experiences within a range of music practices, (2) develop a theoretical framework for digital scores as an important transdisciplinary area of research, (3) build a scientific study of inclusive digital musicianship through the transformative potential of the digital score.

A digital score is defined by the project as *a communications interface of musical ideas between musicians utilizing the creative potential of digital technology*. The meaning of this in practice is being investigated through a series of practice-based case studies that place the experts at the centre of their practice. The purpose of this is to offer the musicians experiences with which to report back to the researchers of *DigiScore* by contributing to a comprehensive dataset, who in-turn synthesize the results into a developing theory.

A central theoretical construct in studying digital score music-making is that of Christopher Small's notion of *Musicking* [1]. In this book, he states that 'to music is to take part', and that taking part can happen 'in any capacity, in a musical performance, whether by performing, by listening, by rehearsing or practicing, by providing material for performance (what we call composing)'. Critically, Small stresses that 'the act of musicking establishes in the place where it is happening a set of relationships, and it is in those relationships that the meaning of the act lies'. Simon Emmerson clarified Small's principle of 'meaning' to infer the 'what you mean to me' [2], (this subtle shift circumvents the significant issues of value and who is doing the evaluation of meaning). Therefore, meaning (or the what-you-mean-to-me) is to be found in the relationships formed between the new creative acts of musicking and the technologies and media of the communications interface: the digital score.

This paper presents an overview of DigiScore research insights in the first 18 months of the project.

2. DESIGN CONSIDERATIONS

Over the past 18 months the DigiScore team has analysed and evaluated a total of 16 case studies. A case study is defined as an ecology of relationships between all agents involved (personnel, machines, media, music etc) therefore:

- the creation of a digital score = 1 case study ("composer's "perspective)
- the performance of a digital score = 1 case study ("performer's" perspective)
- a repeat performance of a digital score with new musicians e.g. at another centre = 1 case study (performer's perspective)

The purpose of a case study is to

1. innovate DS creativity through the 7 project themes of artificial intelligence, machine learning, internet

networking, robotics, virtual and augmented reality, gaming and physical computing

2. Analyse in-vivo experience/ practitioner-knowledge using the theoretical framework (i.e. the ecology of relationships inside musicking)

3. Evaluate and enhance the theoretical framework of Digital Score

4. Build a wider understanding of the scope of digital musicianship through the creative practice associated with Digital Scores

The types of digital scores that we have studied so far have ranged from fixed media, animated scrolling scores that use a combination of standard/ Western European Art Music (WEAM) notation and graphic symbols. (Type 2 in the *DigiScore* typology [3], to Unity engine platforms supporting interactive scores that reconceptualise notation and musical direction through Non Player Characters, glyphs and graphic material (type 8), to AI and robotic based scores that co-create in real-time guided by belief systems [4] and embedded musical/ compositional behaviours (type 10).

From the analysis process introduced as the methodology of the DigiScore project at SYSMUS'22 [5], we are able to start to see patterns common across this range and present here some of the more critical insights as a set of design considerations [6]:

- Digital Scores are fun and engaging as evidenced by interaction with music students on our recent research tour (please see section 4.1)
- They allow for a level of complexity where the temporal stress of accuracy is handed over the to the digital media, enabling the musicians to focus on the music-making
- Animated scores engaged the audience in a media space leading to expectation and macro involvement [7]
- The digital score offers existing composition and performance students the opportunity to move away from the 'culture' of contemporary composition
- Digital media supports inclusivity and accessible music-making enabling musicians from outside the WEAM culture to engage creativity with score and communicating musical ideas using a language that supports what they wish to say (rather than forcing what can be said upon them).
- Musicians (old and young) are excited about digital scores and are willing to engage openly and creativity with them (some caveats below)
- Unity engines and python platforms acutely support the reconceptualization of a music score and the communicative parameters / language of notation (e.g. dynamics, tempo, feel, world-space)
- They encourage new forms of collaboration between, say, coders and visual artists with musicians, and blur the boundaries of definitions of "composer", "performer", "audience".

But …

- A lot of digital scores do not seem to consider the needs of the musicians that are rehearsing and then realising them, for example is it possible to rehearse a small section of a fixed-media animated score? What happens when the ensemble wishes to workshop the middle minute? How do they identify "sections" when they are reflecting, discussing and trying to improve upon your score?
- The audience is within the world of the music, but, generally does not wish to see the score – perhaps there needs to be a second visual output that is made specifically for them?
- New media and technology can be unstable, with crashing and missing adapters factored into the rehearsal process. Do you have a plan B?
- Some of the notation and conceptualisations of a digital score require a lot of explanation, and some degree of involvement with the musicians prior to engagement should be factored in.

3. INCLUSIVE AND ACCESSIBLE MUSICKING

We are currently involved in two case studies that deal directly with inclusive and accessible musicking through the digital score. Or more accurately, how the digital score concept and technologies can support inclusive and accessible musicking.

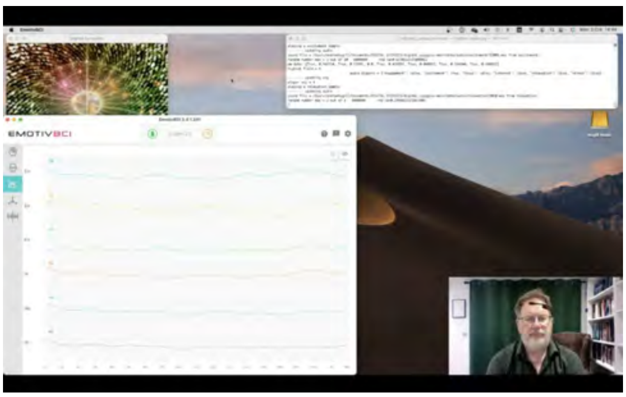


Figure 1. Andrew Hugill performing the digital score *Dygital Syzygies*

3.1 Digital Syzygies (2022)

The first is *Digital Syzygies.* Here we developed a Digital Score platform for shared creativity with d/Deaf and autistic musicians, that unites creativity and collective involvement in the development of a musical idea. This Digital Score used EEG headsets to transmit musicians' brainwaves into a set of neural-controlled musical compositions and performances (see Fig 1.).[1]

The Digital Score project worked in collaboration with Prof Andrew Hugill (University of Leicester) to build a practice-based case study (*Digital Syzygies*) that addressed the central concerns introduced above. *DigiScore's* principal investigator Prof Craig Vear's involvement was 2-fold: 1) providing the theoretical and conceptual guidance offered to Hugill, which supported his creativity and framing the conceptual development of *Digital Syzygies*. 2), Vear designed, developed and deployed a technical solution that facilitates the translation of brainwaves into music and supported shared creativity for the *Digital Syzygies* team, while embedding the definitions and concepts of *DigiScore.*

Digital Syzygies used the Emotiv Insight EEG Headset Brain Interface to connect four musicians who were remotely located (in Sweden, Sri Lanka/London, Brighton and Leicester). This technology was defined by Vear and Hugill as a novel solution with which to bind the quartet who are either autistic or are d/Deaf (Hugill is both), and as an inclusive mechanism with which to support a shared creative pursuit.

This project resulted in some transformational revelations which emerged on several levels: technical, musical and personal.

Vear's Brainwave-to-Music app supported and framed a digital score that built connections or "syzygies" between people with neurodivergent brains and hearing differences. This app was embedded with the core principles of his *DigiScore* project ensuring that it supported the *Digital Syzygies* team's development of the project. It was also designed to be a creative sandbox with which co-located musicians could share ideas.

This app supported Hugill to design a case study using the digital score format that enabled musicians to connect and see similarities in the way their neurodivergent brains react and interact with sounds. It also enabled those who are d/Deaf to relate to each other's sounds and hearing peculiarities. Overall, musicking with the help of this digital score opened many new ways to making and listening to music for the musicians of *Digital Syzygies*. On a technical level, it united d/Deaf or neurodivergent musicians in a creative space and way of working by allowing them to generate, develop and share musical ideas packaged in a novel digital and technical solution (the core definition of a digital score [1]).

For Hugill, the digital score made him unlearn everything he has previously learned in making music as an autistic musician with hearing limitations. Such things as listening to one's instrument, listening precisely and accurately to what you're doing, and being aware of others around you are not needed from him when he makes music using this digital score. Using the digital score also allowed him to not have to listen to sounds for long periods to make and perform music since he usually experiences pain in doing so. The digital score accessed via the brainwave reader does all the performing when it receives the sounds, Hugill can also just wear earplugs and let the brain do all the work: "I'm seeing this self-actualising brain that is making interesting music regardless of me, so it's kind of the opposite of conventional music making" (Hugill 2022). This is inevitably a new musical skill for him which he hopes to use more in his music-making from now on.

3.2 *Jess+* (2023)

Jess+ is an intelligent digital score system for shared creativity with a mixed ensemble of able-bodied and disabled musicians. The digital score uses AI and a robotic arm to enhance the real-time dynamic engagement of the disabled musician with the able-bodied musicians allowing her to thrive in a realtime communication beyond what has so far been possible (see fig 2.). This has flattened any hierarchy through movement and enhanced the sense of togetherness and inclusivity in musicking.

This digital score project is working in collaboration with Sinfonia Viva, Orchestras Live and Digit Music. It is involving three talented and highly experienced musicians: the disabled musician has reduced mobility and uses a powered wheelchair, her normal music engagement technology is CMPSR a modified chair joystick developed by Digit Music, and she usually engages with music by making individual pitch or chords in Ableton using the CMPSR to navigate around a Live environment. The other two musicians identify themselves as being able-bodied professional orchestral musicians who also specialise in SEN workshops.

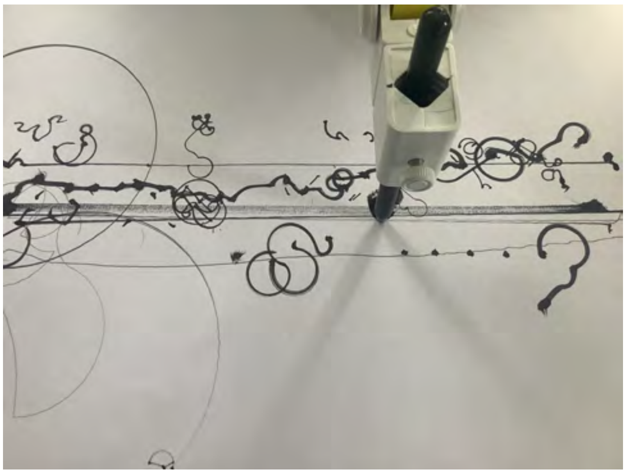


Figure 2. Image of proof-of-concept version of *Jess+*, illustrating the notational language inspired by Cardew's work *Treatise*.

AI and robotic technologies are used to extend the creativity of disabled musicians. In a realtime music feedback loop, a disabled musician (A) is "plugged into" a creative

[1] Full details can be found here http://www.digitalsyzygies.org.uk/

AI robot arm that draws graphic music notation which is performed live by other musicians (B). The creative AI and the disabled musician work as an extended system, with each feeding the other: the AI reads the real-time brainwaves and arousal data of the human (A), and (A) feels the embodied behaviour of the robot arm. The feedback loop is closed as the live musicians (B) make sound, which in turn is heard by both the AI and humans (A+B). In this system, the AI and the robot arm is not an assistive tool, but rather co-creative other.

The point with this digital score is that the inked notation is only part of the score: how the arm moves, when it moves, its velocity and acceleration; how each of the marks relate to the ongoing flow, and legacy of thinking inspired by its previous marks on this paper; how the presence of the AI is continuously reading the human musicians are all elements that communicate the idea and nature of the digital score as a whole. The robot arm in this sense should not be considered an assistive tool, but a co-creative agent, as the musicians and the AI work together in a synergetic relationship to create the music.

About the AI/ robot: one musician found the AI "surprisingly communicative". Another said that she "FELT interactive from the immediate observation that the robot's responses weren't purely imitating us." [original emphasis]. She went on to say that "It seemed to 'listen' and reflect (for example: use a shape or stroke again later when we hadn't repeated the music that invoked this) so this extended the 'conversation' between us. As a result, the creativity of the robot started to match that of the musician."

Another musician felt that she was "in control but very collaborative", and yet at times she wished for it to be "more demonstrative and make larger gestures to initiate a response". On this topic, and with relation to it 'joining in' with the ensemble, one of the musicians wrote that "Once I got my head around not imitating what the robot was doing, I found it was rewarding and engaging to play with. The fact that it "listened" and then responded (or not!) made it easier to accept as a member of a chamber group."

Upon reflection of their collective experiences, the disabled musician described how after a few days she is still "feeling really relaxed" because of the experience, and that it had "opened my eyes to what can be possible in ways, I hadn't even thought about it's like I'd met a new friend in a way." The other musicians echoed this and wrote that it was "all very exciting", and "mind-bogglingly inspiring and interesting". Collectively they were looking forward to making music again in the following sessions, but crucially it is "the process that means so much and the resulting drawing an evocative memory of what happened, which is also evidence of the interaction."

4. DIGITAL MUSICIANSHIP FINDINGS FROM THE DIGISCORE ROADSHOW

A large task of the DigiScore research project is the analysis of case studies, and an equally large task is the "Roadshow". The primary aim of this task is to evaluate higher education music students' wants and needs from digital musicianship education across the globe through engagement with digital scores. Throughout 2021-22 we visited various institutions across Europe, and at the start of 2023, we embarked upon our first Partner roadshow of North-East America/ Canada, in association with our Canadian partner Concordia University, Montreal. Further roadshows are being planned with our other project partners for West-Coast USA (with UCSB), Australasia (with Monash), and China/ Far East (with Central Conservatory of Music, Beijing), as well as continuing our visits to Europe, and potentially India and Central and South America.

Throughout the 4 weeks, we visited public and private universities with music departments that ranged from large internationally renowned centres for innovating electronic music (such as Columbia and Carnegie-Mellon), or music performance and composition (such as Illinois at Urbana-Champaign, Northeastern and New England Conservatory), to small departments with a focus on pop music production and music business (such as New Haven). We spoke with undergraduate, and postgraduate students and PhD researchers, post-docs and faculty.

Our study of digital musicianship was understood from the perspective of 'a person's ability to perceive, understand and create sonic experiences', through the broad range of musicking activities revolving around digital score creativity. We established the following areas as points of interest:

- **Skills**: what are the skills needed to articulate and interpret features and affects of digital score musicking?
- **Contexts**, **Cultures & Literacy**: what contextual, cultural literatures and insights are required to inspire creative thought and support musicking ideas
- **Musical Identity and Creative Practice**: what are the new modes and possibilities of creative practice?
- **Perception and awareness of (digital) music**: how do musicians actively analyse digital score music, and what interpretations are they generating when making music?

We conducted two types of polls, the first was embedded Mentimeter slides that allowed the students to engage and contribute to the discussion through the 90-minute lecture. These questions were associated with the above points of interest at appropriate points in the DigiScore lecture. This served three key purposes: 1) to gather immediate data from the students, 2) to have a higher yield of lower depth data, and to start their process of deeper thinking about the nature of their digital musicianship which we hoped would be developed in the second poll as an online questionnaire. Having started the roadshow series with only the online questionnaire we realised that yield was very poor, so the Mentimeter interactive polls, provided us with a safety net.

Overall, 60+ students engaged with the in-lecture Mentimeter polls, from which 50 completed the online form. At the time of writing the analysis of results has yet to be completed, there are some provisional critical insights that are worth highlighting here:

4.1 Mentimeter Insights

The first poll asked the question "How would you define yourself as a musician? (multiple answers possible)" (see fig 3.). This aligned with the 3rd point of interest "Musical Identity and Creative Practice" and was chosen to be the first to draw their sense-of self directly into the discussion. The list of possible types is listed in figure 3, and this distribution was almost identically replicated across each university department. Of interest was how many students identified themselves as BOTH composer and performer, adding nuance to that identity with the remaining types (e.g. hacker). When we discussed this distribution and choice, it was revealed that any division between "composition" and "performance" pathways within a department was generally a negative divisive split. Whilst there were some monotypes who only chose "composition" or "performance", this was generally isolated to PhD researchers, who would naturally be focused on their field of study.

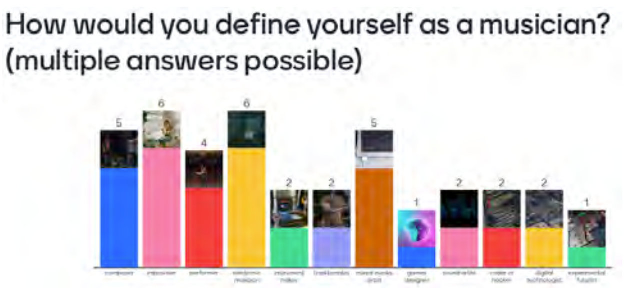


Figure 3: 1st slide from the DigiScore roadshow questioning strategy.

The second poll followed a series of slides that outlined the continuum of types of digital scores, from type 1 "The Digitised Score", to type 4 "The Animated Score", to type 6 "The Creative System", to type 10 "The Living Score". This question asks them "What skills do you have that could help you with digital score creativity?", and was immediately followed by a slide that asked, "What digital music skills would you like to have?". A major theme in this poll was the repeated focus on AI, data and coding skills. When questioned whether this was because the final digital score example focused on AI and robotics had skewed their answers, the majority of students responded that it was more to do with critical skills they are being told are crucial for the workplace in general. Our observation here is that music is a great laboratory with which to learn about AI, data and coding in composition, performance and musicological studies. If implemented in the curriculum could be a very popular programme.[2]

The third poll focused on surveying the underlying cultures and contexts that inspire these musicians to make music. This was a broad litmus test asking them to choose one or more of the following:

- Types of music/ musicians
- Books/ theories
- Art forms (art, theatre, dance etc)
- Media (film, gaming, video etc)
- Nature and the environment (this one was added half way through the roadshow and was inspired by Seth Cluett at Columbia)

Overall, music/ musicians were the dominant choice, but not by much. Generally, only peeking over the top of others by 1 or 2 counts. With the others generally coming out equal. One conclusion is that, as the world gets smaller due to the internet and knowledge becomes readily available due to the internet, and other musics and cultures are within grasp, because of the internet, we have a pluralistic society that is taking inspiration from many more sources than merely music studies.

The final poll was the weakest in terms of the core question that it asked, but the responses were always surprising, so it wasn't replaced. The question asked, "What do you value most in your music-making" was designed to address the fourth point of perception and awareness. The expected responses were to do with core skills of music engagement, but the reality was that they focused almost exclusively on themes of "musical soul", "human connection". "emotion", "experience", "multi-disciplinarity", "my creative voice" or "honesty".

4.1 Online Questionnaire insights

There were two online survey questionnaires made for the roadshow tour, one addressing the students who would attend the DigiScore lecture and another to collect responses from the students participating in the workshop with digital scores. Since we presented lectures to more students than to workshop participants, we received more responses from the Digital Musicianship Questionnaire.

The Digital Musicianship Questionnaire followed a similar logic of questioning strategy as the one targeted through the Mentimeter polls. While the questions were similar, they allowed for a more in-depth description of each participant's digital/musical skills, the context in which they are creating, their digital music identity/creativity as well as knowledge and awareness of their music practice. Besides more traditional music skills, it was noticed that the majority of students already have some digital music skills such as DAW, mixing/producing, MIDI and Max/MSP/Pure Data skills. A lesser number of students expressed proficiency in coding, however, some expressed a desire for acquiring those skills. This aligns well with some of the Mentimeter poll answers where students generally expressed a desire for more coding, AI and

[2] As is evident in the University of Illinois's Computer Science and Music degree https://cs.illinois.edu/academics/undergraduate/degree-program-options/cs-x-degree-programs

machine learning skills in higher education. Another, correspondence to the Mentimeter poll that was answered in more depth in the surveys was to do with students' digital music identity/creativity. A lot of the students see themselves not just as instrumentalists/composers but as music producers and creative technologists. Thus, the use of digital technology greatly augments and changes how these musicians see themselves. Based on our classroom discussions and survey responses, it seems like most of the students are aware of how the tools of digital technology shape their musicianship, citing "a learning tool for shaping own's musicianship" and allowing for "flexibility in one's music-making". Similar to the Mentimeter poll but in more depth, when evaluating what one finds important in one's music, most students answered personally citing, "fun", "authenticity", "originality", "passion" and "self-expression".

When it came to the Digital Musicianship Workshop Questionnaire, we had fewer responses, but they were valuable in assessing how the digital scores presented to the musicians engaged them in the process of music-making. This is also an important part of *DigiScore* research, gathering evidence on the connections, experiences of flow and the transformative effect digital scores have on musicians. Thus, the first two questions addressed musicians' *creativity*, looking at the kinds of relationships musicians formed with the materials of the digital scores and how the digital score engaged the musicians in musicking. Part three asked them to reflect on any new musical experiences they might have had with the digital scores that could be thought of as transformational, mentioning a digital score in particular that had that effect. In addition to these questions, we were also interested to know what musicians might need from higher education, in general, to facilitate their engagement with digital scores in the future.

Overall, the above questions were answered in greater depth and some trends that emerged from them reflect our impressions from the workshops: musicians preferred scores with clear instructions, which facilitated a fun and engaging experience for them. In addition, they found video games' format engaging where some of the behaviours of video games (such as pursuing a goal, or 'winning') add to the excitement of taking part in playing the digital score.

5. CONCLUSIONS

Although this project is still in the relatively early stages, it is possible to conclude the following. However, we stress the need to be open about these insights and to move away from them if required as new insights emerge over the coming years if they present challenging or contradictory evidence. From a creativity perspective, digital scores are proving themselves to be engaging and versatile. They are supporting new expressive ideas and new cultural voices. The case studies we have thus far conducted are enhancing musicians' creativity and supporting interdisciplinary collaboration. Creating digital scores through innovative use of digital media and technology does, however, come with a new set of concerns and responsibilities for those constructing them (what we have called composers) and those realising them (performers). We are starting to see a trend where this divide of "composer" and "performer" is divisive and unhelpful with digital score creativity, and this is leading to false expectations and negative enaction. The solution we are seeing emerge is more interactive and collaborative discussion and involvement between the musicians so that the needs of the realising musician ("performer") are embedded into the core construction of the digital score. Additionally, as a lot of digital scores deal with rich media and audience-friendly spectacle, should their experience be factored into the construction? For example, should they simply watch the digital score as shown to the "performers" or could there be a parallel version that heightens their involvement in the experience of the music?

The *DigiScore* project has focused on inclusivity and accessibility, and this is something that will grow over the coming years. In the 2 case studies presented in this paper, it is evident that the digital score format and approach are enhancing musical engagement for those involved. It is flattening any hierarchy of creativity and expression by neutralizing the specialism and training needed to share and communicate musical ideas (the core definition of a music score). Additionally, binding assistive and digital technology into the fabric of a digital score system, perhaps with AI, is amplifying the communicative interactions between musicians regardless of ability.

Finally, by evaluating digital musicianship through digital score creativity we can start to see patterns that suggest it is shifting across 4 realms: skills, contexts, identity, and awareness. However, we wish to stress that these four realms are immutably interconnected and should not be isolated to the point of exclusivity of the influence of others. For example, skills that are enacted in real-time are done so because of the individual's context and education, which has informed how they perceive and what they are focusing their awareness on.

There are more insights to be found in these three sections of the project. Our approach will be to challenge these provisional conclusions in the hope that they are strengthened or transformed. Ultimately they will be used to determine scientific knowledge of how digital scores stimulate new creative opportunities and experiences within a range of music practices, and to build a scientific study of inclusive digital musicianship through the transformative potential of the digital score.

Acknowledgments

The Digital Score project (DigiScore) is funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. ERC-2020-COG – 101002086).

REFERENCES

[1] C. Small, *Musicking*. Middleton, CT: Wesleyan University Press, 1989

[2] S. Emmerson *Living Electronic Music* London: Routledge 2007

[3] C. Vear, *The Digital Score*. Routlegde, 2019

[4] C. Vear "Embodied AI and musicking robots" in C. Vear, & F. Poltronieri (eds) *The Language of Creative AI*. Springer: Cambridge, 2022

[5] S. Moroz and C. Vear, "Methodology of the Digital Score Project through an Analysis of Nautilus Case Study, an Immersive Game Environment Digital Score," in *Proceedings of the 15th Conference of Students of Systematic Musicology (SysMus22),* Ghent, 2022, p. 45.

[6] Digiscore, Website https://digiscore.github.io 2023 (accessed 22 April 2023)

[7] G. Calleja. *In Game*. MIT Press: Cambridge Massachusetts. 2014

JUDGING A SCORE BY ITS COVER: THE ROLE OF VISUAL DESIGN IN INTERPRETING COLOUR SCORES

Frankie Dyson Reilly
Queensland Conservatorium Griffith University
Brisbane, Australia
frankie.dysonreilly@griffithuni.edu.au

ABSTRACT

Music notation is typically viewed to be an interface for the transfer of musical information, with a performer's individual interpretation of a score determining the aural outcome. Performers rely on learned symbols and context clues to interpret a score, supplemented by semantic information inferred by the style and font of the score. When scores contain novel graphic elements that have no standardised framework for interpretation, such as colour and shape, the semantic information contained in the visual presentation of the graphics becomes integral in influencing a player's unique interpretation. Though marketing and graphic design literature demonstrate the clear importance of visual design in mediating the relationship between viewer and media, examination of this phenomena remains largely absent from most academic score analyses. In this paper, I use colour as the primary lens through which to explore the role of visual design in mediating a performer's response to a score. I present three original and visually distinct compositions as case studies, each uniquely demonstrating the role of colour and peripheral extra-musical content (such as font, shape, size, and layout) as mediators of interpretation. I centre verbal and written responses from performers to explore how the interaction of and with these visual characteristics shapes interpretation of the score and resulting music.

1. INTRODUCTION

The score is at the centre of many Western classical and contemporary music traditions, and the role of notation in mediating the exchange of musical ideas continues to foster boundless discussion and innovation. In my own experience of reading and interpreting notation, the physicality of the score has always felt vitally important in shaping my engagement with the work. I recall sitting in the university library as an undergraduate composer and sifting through large scores, feeling the yellowed grain of the paper, learning to recognise the unique style of each composer or engraver. The huge size and aged quality of these scores seemed to suggest the enormity and richness of the musical material within. It was a vastly different experience to what I had previously experienced in picking up thin, white A4 scores with the same stock Finale or Sibelius fonts, and evoked a lasting regard for the ability of a physical score to seemingly reflect the music material held within. This sentiment is echoed in discussion with colleagues and enforced in composition courses and texts which emphasise the importance of good score engraving**.** However, though music notation is historically an explicitly graphic medium, academic analyses of scores typically focus on *what* musical material is being communicated through notation, without touching on *how* that notational material is visually designed and represented (such as font, format, size, medium, colour, and shape) and how it mediates a player's interpretation. Building on the importance of visual design established in fields like advertising, I formed my hypothesis that the visual design of notational material (both in terms of aesthetics and functionality) might create an intrinsic difference in how that notation is interacted with and interpreted by performers.

Within the vast field of novel and/or graphic notations, I limited my scope to the use of "colour-notation" as a lens through which to examine the role of visual presentation in shaping players' interaction with a score. In this paper, I present three case studies of original compositions which I have workshopped and discussed with performers in an iterative process. Using an artistic research framework, I highlight the lived experience of performers through thematic analysis of their verbal and written responses from collaborative workshops and surveys, and then compare this to my own reflections on the scores I created. Each case study uniquely demonstrates the value of colour as a visual design tool which mediates a performer's interpretation and realisation of the score.

2. VISUAL DESIGN AND INTERPRETATION

Schuiling's definition of musical notations as "interfaces for imagining virtual musical relations" [1, pp. 432] understands the score as a site for interaction and interpretation, with meaning that is flexible depending on the interpreter. In Western classical scores, some musical characteristics are prescribed and proscribed, and some are left open to the judgement of the performer, who typically draws on historical and contextual clues to seek out and embody the composer's emotional and narrative intentions implicit in those technical actions [2]. Silvermann describes two approaches to interpretation: formalist, in which the performer is tightly bound to and protecting the composer's perceived intention, or subjective, allowing a performer more 'poetic license' to experiment [3]. For notated compositions, this process of analysis and interpretation is essential to aurally realising a work.

This relationship becomes complicated by non-standardised notations, such as graphic and/or coloured scores. In [4], Sobakina likens the process of interpreting such scores to that of experiencing synaesthesia, as without an existing frame of reference provided by Western staff notation, the performer must personally perceive intrinsic relationships between the graphics and a musical outcome. Here, musicians often rely on semiotic associations triggered by colour, shape, size, and visual interaction of elements. Therefore, musical responses and interpretation become closely tied to the performer's individual response to the visual presentation of the score, including graphic components, format, style, and aesthetic value. Including open elements evokes Silvermann's *subjective* approach to interpretation, centring the musical voice of the performer.

For both standardised and non-standardised notations, visual design of the page plays a role in influencing a performer's interpretative process. In staff-notated compositions, familiar elements such as font, typesetting, and format may subtly guide a player's approach to a score. Different fonts possess distinct emotional connotations, described as 'semantic signatures' by Kulahcioglu and de Melo [5], and can influence value judgements of a text [6]. In scores, the most common example is the use of 'classical' sans-serif font versus 'jazz' script-style font, which separate perception of each genre by drawing on connotations of 'seriousness' and 'looseness'/'fun' that Shaikh and Chaparro [7] attribute to serif/sans-serif vs. script fonts respectively. In a study on font and brand choice, Doyle and Bottomley [8] suggest that a font's perceived appropriateness to its context can determine its appeal; in a score context, appropriate typesetting may indicate both professionalism and serve to reinforce a narrative for a player. Similarly, the appearance of staff notated scores carries extra-musical connotations. Experienced performers can differentiate baroque, classical, romantic, and "new" music simply by looking at how information is distributed on the page; their own personal preferences for or against certain eras (and learned understanding of how they "should" be played) will immediately influence their attitude and approach to the learning process. The dense appearance of musical content in "new complexity" scores such as Brian Fernyhough denotes complexity and technical difficulty, and thus intellectual-musical value [9]. In contrast, sloppily engraved scores may be viewed as the work of an inexperienced or inattentive composer, and therefore are given less care in the interpretative process.

Manipulation of otherwise-familiar content may also serve to illustrate the narrative of a work. George Crumb's famously circular and spiral-shaped piano scores such as *Makrokosmos* could have been notated using a traditional staff format, but McKay suggests that the unique calligraphic format perhaps adds an intangible narrative quality which "changes something fundamental about what the work communicates to a performer" [10, pp. 13].

Colour is a particularly versatile component of graphic scores, as demonstrated by the wealth of unique colour-scores and notation systems curated by sources such as Theresa Sauer's *Notations 21* [11], Read Garner's *Source Book of Music Notation Reforms* [12], and Cat Hope's curated site *Drawn From Sound* [13]. Because there is no standardised symbolic meaning to colour as a notational construct, colour can be assigned open, ambiguous, or highly specific meanings depending on context. When colour is used as an open parameter, performers may draw upon associations between music and colour mediated by emotion, personal experience, and cultural metaphor [14]. For example, the association of red with excitement, anger and strength translates musically to speed, agitation and loudness, while cooler hues are associated with softer, passive sounds [15] [16] [17] [18]. John Cage's *Aria* (Figure 1) is a notable example of an open colour-score which invites a performer to harness personal colour associations to create a variety of vocal timbres [19].

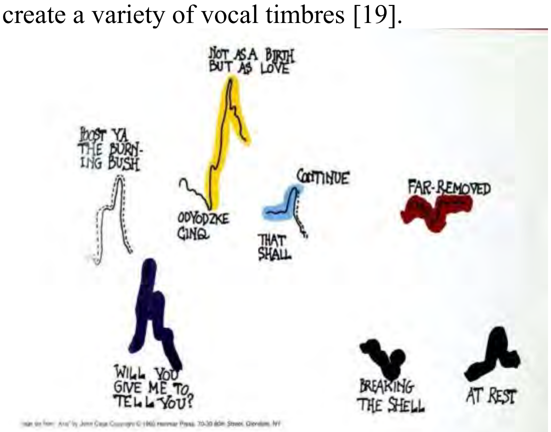


Figure 1. Excerpt of *Aria* by John Cage (1958)

Composers – such as Michael Poast, Deborah Pritchard, Anthony Braxton, and Cilla McQueen – join this tradition either by using coloured shapes, lines and spatial relationships alongside some traditional markings, or emancipating themselves from traditional notation styles entirely in experimental, abstract scores [20] [21] [22] [11]. By applying graphic design principles in scores such as these, a carefully selected colour palette can communicate mood, style and concept, while also drawing the eye to or from specific elements [23]. Recognisable colour combinations can be used to illustrate narrative in a score, guiding an interpreter. The design (or in-/ex-clusion) of elements such as staff notation, text, colour, and imagery might suggest a genre or interpretative approach to a performer, e.g. more non-standardised characteristics might connote open interpretation.

Colour-scores can also be explored through a variety of visual mediums, from pencil and paint to digital, with each medium carrying different connotations for a viewer. In the 21st century, digital scores and score-sharing present a particularly compelling and adaptable platform with which to experiment with colour and visual format, as the physical constraints of colour printing and paper no longer exist [24]. The Decibel ScorePlayer is a prominent example, hosting scrolling, animated scores that frequently feature line, shape and colour-based graphics, most notably by composer-researchers and developers such as Cat Hope, Aaron Wyatt, and Lindsay Vickery [25]. Cat Hope's *Human Cathedral* (see Figure 2) is an example of a left-to-right scrolling score supported by the ScorePlayer [26].

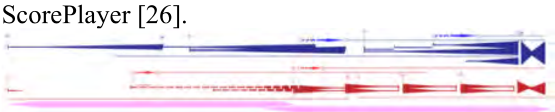


Figure 2. Excerpt of *Human Cathedral* (2019) by Cat Hope in the Decibel ScorePlayer

The movement of the score allows colour to be used as a linear narrative tool in itself, unfolding as the piece progresses. This exciting format offers an opportunity for visual immersion to performers and even audiences. Its applications for colour-scores are explored as one case study in this paper.

3. CASE STUDIES

In my recent work, I've become increasingly aware of the limitations of standardised classical music notations and performer-composer relationships enforced therein. The use of colour in my scores provided a gateway for me to begin exploring alternative notational spaces which gently disrupt those hierarchies. To centre the voices of performers in realising my work, I primarily explore the use of colour as an open, expressive device which invites players to draw deeply on their own creativity.

What follows are three case studies, each analysed through an artistic research framework in order to illustrate the role of score design in mediating performer-score interactions. Each work uses a different medium, format, ensemble, purpose, and intended role of colour as an expressive tool, with the goal of examining how the visual qualities of a score can mediate and transform the musical response of performers. These works are part of an ongoing and iterative creative immersion in colour notation, during which I created and workshopped "colour-scores" over a period of years. In order to conduct an analysis of the effectiveness of each notation system, I sought feedback from performers using questionnaires and a semi-structured interview process.

Using an action research spiral (O'Leary, 2004), each composition was developed and workshopped through a *workshop – reflect on feedback – re-compose – workshop* cycle in order to hone in on the effectiveness of colour notation. To gather feedback, performer-participants completed a detailed survey before and after workshopping the piece, and I guided and transcribed in-rehearsal discussions regarding players' evolving impressions of the score. By triangulating these data with my own reflections and recordings of each work, I can include the perspective of both composer *and* performer(s) in evaluating the role that presentation and aesthetic value played in interpretation of each work. In each of the case studies that follow, I synthesise these two perspectives to discuss how colour is uniquely utilised and presented, consider the role this played in performers' interpretations, and evaluate its function as a visual notation tool. As the case study scores all contain open and semi-open notations – wherein one performer's response might differ from another's – each performer's lived experience is offered equally as a unique perspective on the potential of the different scores, and my discussion of each case studies suggests the implications and questions that arise from my findings rather than a decisive conclusion. Due to the surprising lack of writing regarding this topic within music literature, the scope of this paper is limited to my analysis of this phenomena in the three case studies, alongside synthesis of my observations with literature from non-musical fields to offer an explanation as to how and why the performers in this study engaged with the scores as they did.

4. FIREFLIES (2021)

Fireflies (2021) is a piece for piano and open-instrumentation ensemble of 6 or more players. Composed with the intention of creating a minimalist soundscape with a progressive narrative driven by player decisions, the six players act as "fireflies", gradually modifying an individually assigned three-note motif which is underlaid with a different colour (Figure 3). Performers are instructed to create a "personality" for their motif by manipulating the dynamics, articulation, speed, register and presence of the motif according to their personal interpretation of the given colour.

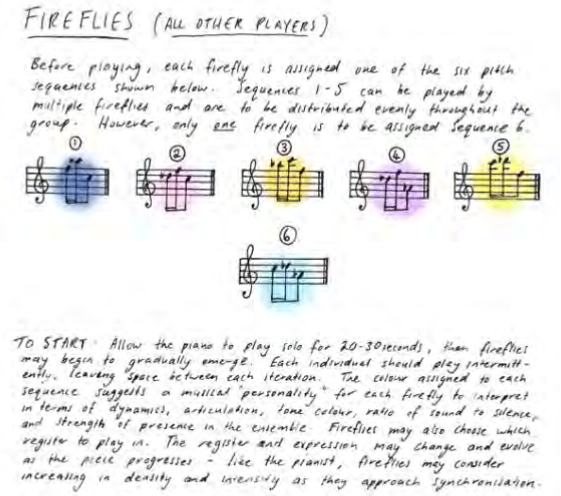


Figure 3. Excerpt from *Fireflies* (2021) displaying coloured 'firefly' motifs

Two ensembles workshopped this piece in September 2021 and August 2022; this analysis draws from the responses of both groups. Instrumentation for group 1 included piano, vibraphone, flute, two violins, marimba, recorder, and banjo, and read from paper booklet-style scores. Group 2 included a range of pitched percussion instruments and mostly read from digital tablets. The score was originally hand-drawn, with coloured graphics later added digitally to the scanned score. Following suggestions made during the 2021 workshop, I re-drew the score in 2022 using Procreate for iPad and altered the format slightly. Figures 4 displays an excerpt from the 2022 fully-digital score.

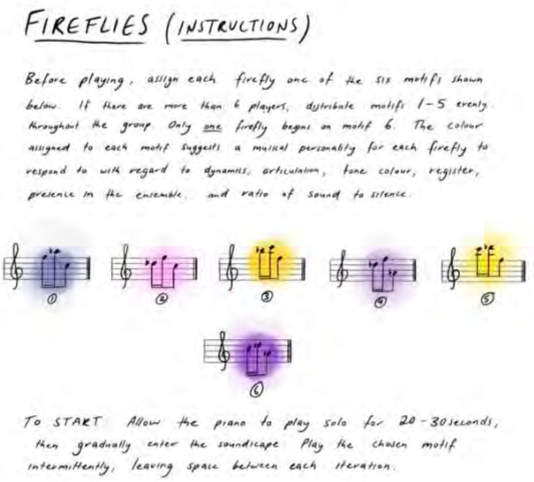


Figure 4. Excerpt from *Fireflies* (2021, rev. 2022) displaying coloured 'firefly' motifs

Within my survey data and rehearsal transcriptions, I identified several key themes which pointed to the role of both the style of the score and its aesthetic appeal in shaping the way that participants approached their performance.

4.1 Creative immersion and engagement

First, the "homespun" presentation (as one player put it), including watercolour graphics and handwritten text instructions in lieu of bars and systems, engaged players differently than a black-and-white score. Several players believed that, although traditional notation could have been used to achieve a similar result, they might have had a more "narrow-minded view… just counting rests and playing at the right time". The existing format provided "more freedom… consciously thinking about how I could and should make the motif sound and where I should place it". Another player suggested that if the score did not include colour, they "would have interpreted it more as [they] would a Reich score" and that they were "not so much worried about executing the correct notes as much as [they were] committed to evoking the colour of my firefly". These responses echo McKay's suggestion that Crumb's *Makrokosmos* may have lost some intangible immersive quality if traditional notation were used [10].

Without the score containing any pre-defined expressive approach to their motifs, players reported having to "dive into [their] imagination" and incorporate "more personality and individual interpretation" as they 'interpreted' colour. Reflecting Sobakina [4] and Palmer *et al* [14], their vivid descriptions of their colour interpretation drew on natural and emotional metaphor, including "a sombre, rainy blue"; "cherry blossom gardens"; "sadness and an anguish with a hint of tenderness"; "like the warmest hug". Phrases like "embody", "evoking", and "see[ing] the music" suggest a deeply immersive interpretative process for each player. Here, the coloured score provided a visual and musical framework that invited players to feel comfortable contributing their personal musical ideas, and experimenting creatively within that framework. When presented with a work that visually deconstructed their expectations of what a score should look like, the group understood the need to subvert their own interpretative assumptions.

4.2 Narrative subtext

Another key role of colour and visual design in *Fireflies* is to reflect and enforce the narrative and emotional atmosphere of the work. Players repeatedly mentioned the visual appeal of the score, calling it "beautiful" and "a joy to look at", with one player stating that "the hand-written, drawn and painted elements created a kind of organic homespun quality that worked for the piece." Ren *et al* [27] found that when used in marketing, hand-written font styles project qualities of warmth and personality; the response from performers suggests these qualities are also present in a score context, perhaps allowing players to feel more personally connected with the score. Further, the handwritten style and "beautiful" score design is indexical of the personal care, time, and attention present in the making of the score, which perhaps inspired players to engage with a similar degree of care.

Through each player's understanding of their motif colour, the piece takes shape and progresses organically; the unpolished, freely handwritten style of the score may have subtly reflected the creative freedom of the music itself. The warm, ethereal musical atmosphere was further enforced by the colour palette for the motifs: I intentionally selected colours which, to me, evoked either the warmth of nightlights or the colours of dusk. A player commented that "the colour palette… allowed for the score to evoke the delicacy and vibrancy of fireflies". A stronger, darker colour palette may have evoked a different atmosphere. In this sense, visual design served to immerse players in both their creative roles and in the narrative of the work; as put by a player, "it is quintessentially a firefly piece in aesthetic!".

4.3 Reflection

Overall, the unique design of the score mediated a positive relationship with the performers. The open notation elements, soft colours, and care implicit in the visually engaging hand-drawn score each contribute to a greater level of musical immersion and understanding of its narrative. Participants offered positive comments mentioning how "fun" and "engaging" they found the workshop, and when invited to add suggestions for future colour scores, even players who had previously expressed apprehension toward open scores were enthusiastically offering ideas. These included suggestions for shading and colouring the motifs, and a percussionist proposed adapting the work for pairs with a gradual colour gradient representing changes in sound. In future developments of *Fireflies,* I plan to re-create the score digitally with a black background and white text to take a more literal approach to visually illustrating the firefly narrative.

A continued direction of inquiry for colour-scores is summarised by a player's suggestion that "small injections of colour 'de-colonise' the score". The work mostly uses text instructions, which contain prompts for verbal group discussion, alongside colour, which invites players to include their personal narrative. The reduction of staff notation in favour of alternative notational tools means that a player needs only basic knowledge of western notation to successfully perform the piece. By using tab or letter-names in noteheads, the score could easily be adapted to require *no* staff notation experience whatsoever. With such adaptations, the visual design of *Fireflies* suggests some possibilities for making tonal, contemporary art music more accessible to ear-trained and/or non-classically trained players.

5. VENUS FLYTRAP (2022)

Venus Fly Trap was commissioned in 2021 by the Night Window Trio, including piano, Bb clarinet, and viola. It is a programmatic work which follows an unlucky fly's demise in the jaws of a Venus flytrap. *Venus Fly Trap* is an experiment in combining colour with an otherwise very conventional staff-notated format (Figure 5).

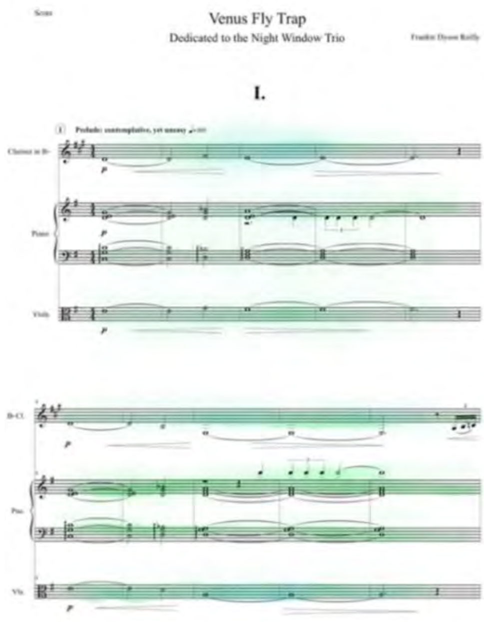


Figure 5. Excerpt from *Venus Fly Trap* (2022)

I used Finale to engrave the staff notation – using mostly default font settings to keep the style as familiar as possible – and edited in the colour using Procreate. Performers were not given any instruction as to how colour should be interpreted, other than that they should draw on their own intuition to connect it to an expressive musical response. My intention for this work was to understand how colour might complement (rather than replace) the traditional score, and how presenting it might inform or disrupt the perceived meaning of the score.

5.1 Expressive cues

A major benefit of combining colour and staff notation is that it provided an intuitive visual framework within which players could shape their expressive performance. This was present first at a micro level, as players used the colours provided as expressive referents to shape each individual phrase of the work. Throughout the work, the underlaid colour was "easy to read" and blended visually with the staff. It allowed players to absorb all information simultaneously, perceiving nuances in colour and their relationship to pitch and time. Positioning the colour within/behind the staff ensured that it was seen as a welcome enhancement – rather than a replacement – of the familiar staff-notated content. Players responded to this addition by using colour as a tool to find the expressive meanings they had sometimes found staff notation to be lacking in. They noted that colour "evoked more of [their] emotions while playing" and "made [them] consider things like timbre and emotional intent". This corroborated their pre-workshop predictions that colour would "enhance" and add "more direction and meaning" to traditional notation. Players mentioned the creative challenges arising from combining expressive markings and colour, such as when their personal interpretation of the colour information did not line up with assigned dynamics. A suggested solution was to combine perceived expressive qualities for more depth. As one player put it, "A red *piano*, may be produced with a slower, weightier bow with an intensely fast vibrato. This creates an intensity without the instrument being loud."

5. 2 "Highlighting" structure

Colour's role in expressing narrative was also evident at a macro level, as colour contrast provided immediate visual reference to the overarching programmatic structure of the work. Flipping through the score, players can immediately differentiate green, orange/red, and purple as distinct thematic material, with the return to green at the end signalling a return to the same musical material (Figures 6 – 7).

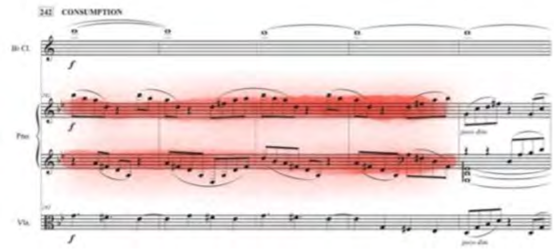


Figure 6. Red excerpt from *Venus Fly Trap*

Figure 7. Purple excerpt from *Venus Fly Trap*

Like in *Fireflies,* the group's interpretation of each colour emphasised the accompanying narrative content, though in this case it was likely complemented by expression markings. Green sections (marked as '*a summer's day, an innocent fly*') were described as having a "light alive green sound"; the red/orange sections (a fly being consumed by the flytrap) "made for a darker and more intensive timbre"; and the purple (alongside the marking *'a brief memorial'*) added "a sadness". Though expression markings might have been sufficient, the inclusion of colour "played an important role in defining character and storytelling aspects of the music" and "definitely enriched my response to the emotion and character of the different sections." Within the comfortable boundaries of the familiar A4, white-paper, engraved staff notation, performers were able to explore colour's expressive potential without becoming totally untethered from the familiar.

5.3 Rehearsal cue

The inclusion of colour played a major role in opening a dialogue around tone-colour and expression among the players and helping to shape rehearsals. Colour was seen as a useful prompt for "meaningful ensemble conversation about musical intentions and sonority"; this view reflected my own experience of the rehearsals, which were full of discussion surrounding colour and its relation to expression. Though one player expressed concern in the pre-workshop survey that the ensemble wouldn't interpret colour consistently, the same player amended this opinion in the follow-up survey: "I believe [colour notation] adds to group unity in terms of phrasing and musical intention… and could even help to make the rehearsal process more streamlined and efficient".

5.4 Reflection

In *Venus Fly Trap,* colour as a visual tool provided clarity in three ways: first, in providing an engaging enhancement to expressive information in staff notation; second, in visually illustrating the overarching narrative and structure of the work; and third, in prompting productive and creative rehearsal discussion. The trio's experience (and the musical outcome) was positive, yet the design still has potential to be improved upon. In writing the piece, I used colour sparingly, concerned that including colour for the entire score might be overwhelming for players. In contrast, players suggested that sections without colour "definitely felt a little uninteresting when compared to the coloured sections". Dark colours can sometimes obscure notes within the staff; in works since, I have experimented with colour above and below the staff. Both of these considerations exemplify the importance of *how* colour is styled (as opposed to simply *whether or not* it is included) within a piece in determining a player's interpretation.

6. STICKS, TWIGS, BRANCHES (2022)

Sticks, Twigs, Branches (2022) is a three-movement scrolling score compatible with the Decibel Score Player. It was composed for prepared piano with extended techniques, three violins, and electric guitar with the intention of exploring the varied sounds created by different branches of the string family. Drawing on the style of scrolling scores by composers like Lindsay Vickery and Kate Milligan, in which artistic graphics are presented alongside some standardised notation, *Sticks, Twigs, Branches* combines elements of staff notation and open graphics on a colourful bed of shapes and lines (Figures 8-10).

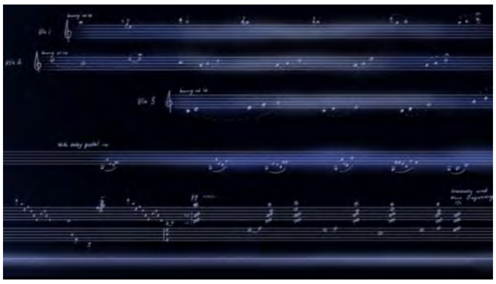


Figure 8. Excerpt from *Sticks* (2022)

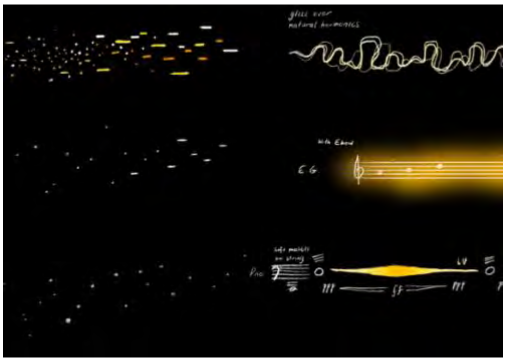


Figure 9. Excerpt from *Twigs* (2022)

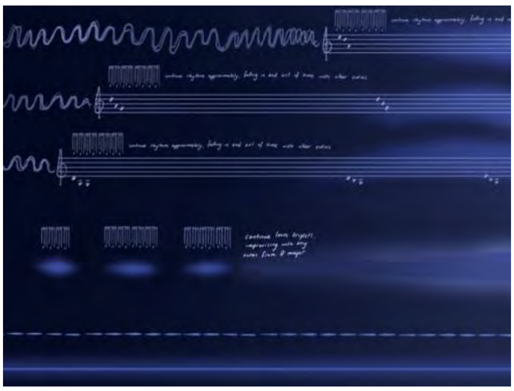


Figure 10. Excerpt from *Branches* (2022)

The work was created entirely using Procreate for iPad. During the three workshops of this piece, I played the piano part with the goal of being a part of the interpretive discussion alongside the other performers. Performer quotes in this section are from the three violinists and guitarist.

Unlike in *Fireflies* and *Venus Flytrap*, some explicit parameters were given to guide the interpretation of graphic elements; as a much greater proportion of this work included non-standard notations, establishing this framework provided coherency throughout the work. Performers were instructed to approach blue hues as passive, warm, and legato, and yellow hues as more active, bright, and angular, and relative size of shape was mapped to volume. Some instructions were informed by an early workshop in which the group freely improvised to simple graphic scores incorporating yellow and blue colour schemes, allowing me to embed the performer's intuitive responses to colour within the score. Players are reminded in the performance notes that the instructions are guidelines only and intuitive exploration of sound and colour is encouraged.

6.1 Immersive visual design

A primary role of visual design in *Sticks, Twigs, Branches* was to provide an immersive platform for performers to explore both intuitive and learned responses to visual characteristics. In the pre-workshop surveys, players were apprehensive about whether colour might impede their ability to "correctly" perform the work according to the composer's (my) intentions. After playing the work, players' focus had shifted away from this initial rigidity. Phrases like "I was thinking much more about sound in general than 'am I getting it right'", "it felt more fun and freer", and "[the piece was] engaging to play, anxiety free" show the significant change in attitude from start to finish. The change in attitude reflects a shift from the traditional 'formalist' interpretative approach described by Silvermann [3] to a more personal, performer-led 'subjective' approach. By presenting colour and shape as having negotiable sound outcomes within some mapped parameters, performers were given safe boundaries within which to exercise their own musical voice.

6.2 Font and colour contrast

Colour also provided an opportunity to harness semiotic associations and visually signpost the musical style. Throughout the movements, I purposefully paired different colours and styles with different sounds. The movements *Sticks* and *Branches* were designed in a clean, digital style with shades of blue, whereas *Twigs* had a rough, sketchy style in shades of yellow and orange. My stylistic and font choices were intended to enforce and enhance the ensemble's intuitive free-improvised responses to graphic drafts in an early workshop. The sketchy, unpolished *Twigs* (Figure 11) was played "a little more chaotic and energetic", "more disruptive”, and for the guitarist had “a more energetic feel and coarser texture that I would respond to in distorted sound and dissonance”.

Figure 11. Excerpt of ‘rough, sketchy’ style from *Twigs* (2022)

The roughness of the style denoted a similarly rough musical approach. The yellow-against-black colour contrast created by the dark background enforced the electricity intensity of the movement, as “higher contrast” equated to “more [musical] presence” with stronger dynamics and attack.

In comparison, one violinist was “more attentive to the timing of other players in *Sticks* and *Branches* because of the "cleaner" style”, as it complemented the “calmer/smoother” response to blue. The lack of contrast between background and foreground (Figure 12 and 13) may have further emphasised the soft, warm musical outcome.

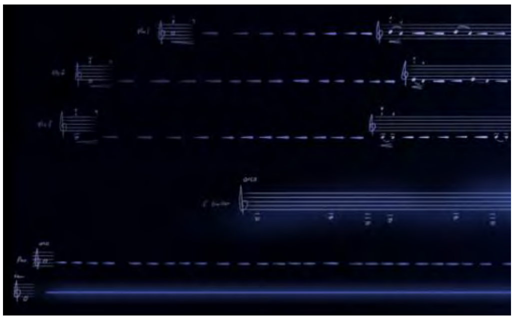


Figure 12. Dark blue against black in excerpt from *Branches* (2022)

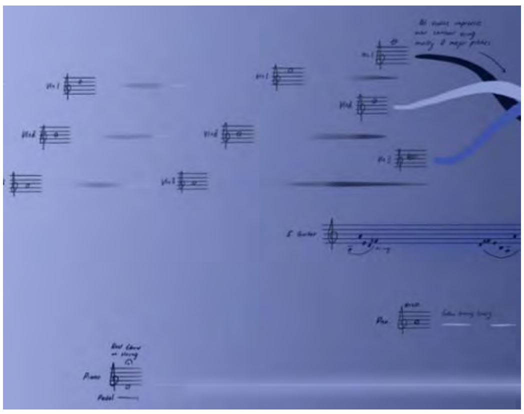


Figure 13. Light blue against blue in excerpt from *Sticks* (2022)

In both cases, the visual style complemented the explicitly notated musical content in each movement: *Twigs* includes more fully improvised gestures and distorted/dissonant sounds, while *Sticks* and *Branches* are more expansive, tonal, and include more pitch notation. This strategy, in which the visual design semiotically reflects the musical style, allowed players to immediately perceive (and become immersed in) the intended expressive approach upon looking at the score.

6.3 Reflection

The overall response to the work was very positive, particularly regarding the democratic process of determining the interpretation. During workshops, players discussed drawing from cultural metaphor, associations between colour and natural phenomena, and personal synaesthetic associations. Allowing musicians to play, discuss, and reconvene created an opportunity for free negotiation of the “meaning” of the colour amongst players. The process was described as “collaborative”, “fun”, and “anxiety free”, with colour providing a visually engaging framework within which the musicians in the ensemble could play together and make individual creative choices in response to score directions. Indeed, as this work was shaped around initial improvisational responses to coloured graphics, it felt more akin to a collaboration than other works. Though I developed and notated the score, each performer’s creative essence is present in the final work. In this sense, colour and peripheral visual characteristics provided a gateway to gathering collaborative musical ideas and disrupting the expected hierarchy of creative interaction between ‘composer’ and ‘performers’. In *Sticks, Twigs, Branches,* all of us existed somewhere in between those traditionally rigid roles.

4. CONCLUSION

When a score is flooded with colour and visual gesture, it can present a deeply immersive musical and visual experience for performers, offering a different way to engage with and “see” music. Throughout this paper, I have observed how manipulating the visual design of colour-notation can transform the interpretative process in both historical examples and in the context of my own three distinct case studies. Though performers’ *interpretation* of (or, musical responses to) coloured stimuli are not always consistent from one performer to another, my findings demonstrate the relatively consistent influence of score design in transforming their *approach* to interpretation. Throughout the case studies, three key themes emerge which frame performers’ interactions with score design, even when the specifics of those interactions differ from player to player.

1. Introducing a novel visual parameter (colour) to a score causes musicians to

engage differently with the score than they would have with a familiar format.

2. Colour, font, style, shape, format, and interaction of colour notation with standardised notation all distinctly contribute to musical response and shape the perceived narrative of a work, though the musical response may be variable from performer to performer.
3. Colour, particularly when framed as a semi-open notation, provides a framework for collaboration between musicians and for democratic creative interactions.

Primarily, these themes highlight how manipulating the appearance of notation can alter the way in which musicians engage with the score in a *functional* sense (i.e. how the score influences their audible musical response). Beyond the functional roles that score design can play in mediating performer's responses, comments from participants describing the scores as "beautiful" bring me to consider the role of the score as a visual accompaniment to a performance, or even as a stand-alone artwork. In other words, how might the score mediate an *audience's* interaction with a musical performance? In the context of screen-scores, Hope and Vickery suggest that a score projected alongside a performance could be a deeply immersive experience for audiences, but could also become a distraction from the musical performance, depending on the work [28]. If the score is intended to be able to stand alone without music – like in the case of Cilla McQueen's *Picnic* [11], which consists of sweeping coloured shapes labelled with instruments – can it still be considered a score? Further research into score design and score aesthetics may continue to bring light to these definitions and differences.

Within the scope of the score 'experiments' I have created thus far, the responses I have received from performers have impressed on me the immense value of learning new notation and design techniques as a composer. Colour and shape opened up new channels of communication between myself and the performers in this project, presenting a new way to suggest musical narrative and expression and, when the context called for it, welcoming performers into a more creative role. Performers often reported 'interpreting' colour by finding personal connections between colour and nature, symbolism, metaphor, and culture, and then triangulating these with aural associations – a fascinating and personal process which warrants further investigation to understand. As questions continue to emerge, the findings presented in this paper open a doorway for myself and others to deepen our exploration of the role of colour and score design in extending notation practices. In my own research and composition practice, I will continue experimenting with creative uses of colour, imagery, and design in written score formats, moving toward a better understanding of how score, performers, *and* audience might interact.

8. REFERENCES

[1] F. Schuiling, "Notation Cultures: Towards an Ethnomusicology of Notation," *Journal of the Royal Music Association,* vol. 144, no. 2, 2019, pp. 429-458.

[2] P. Thom, *The Musician as Interpreter.* The Pennsylvania State University Press, 2007.

[3] M. Silvermann, "Musical Interpretation: Philosophical and Practical Issues," *International Journal of Music Education,* vol. 25, no. 2, 2007, pp. 101-117.

[4] O. Sobakina, "The Musical Graphics in Connection With Synaesthesia in the Context of the Polish Avant-garde," in Proceedings of the 3rd International Conference on Art Studies: Science, Experience, Education. Moscow 2019, pp.146-149.

[5] T. Kulahcioglu, and G. de Melo, "Predicting Semantic Signatures of Fonts," in *Proceedings of 12th IEEE International Conference on Semantic Computing,* Laguna City, 2018, pp. 115-152.

[6] S. Juni, and J.S. Gross, "Emotional and Persuasive Perception of Fonts," *Perceptual and Motor Skills,* vol. 106, no. 1, 2008, pp. 35-42.

[7] D. Shaikh, and B. Chaparro, "Perception of Fonts: Perceived Personality Traits and Appropriate Uses", in M. Dyson, C.Y. Suen (Eds.), *Digital Fonts and Reading,* World Scientific, 2016.

[8] J.R. Doyle, and P.A. Bottomley, "Font Appropriateness and Brand Choice," *Journal of Business Research,* vol. 57, no. 8, 2004, pp. 873-880.

[9] S.P. Duncan, "Re-Complexifying the Function(s) of Notation in the Music of Brian Ferneyhough and the "New Complexity"," *Perspectives of New Music,* vol. 48, no. 1, 2010, pp. 136-172.

[10] T. McKay, A Semiotic Approach to Open Notations: Ambiguity as Opportunity. Cambridge University Press, 2021.

[11] T. Sauer, *Notations 21.* Mark Batty Publisher, 2009.

[12] R. Gardner, *Source Book of Proposed Music Notation Reforms.* Greenwood Press, 1987.

[13] C. Hope. Drawn From Sound. [Online]. Available: www.drawnfromsound.com

[14] S.E. Palmer, K.B. Schloss, Z. Xu, and L.R. Prado-León, "Music-Colour Associations are Mediated by Emotion," *Proceedings of the National Academy of Sciences,* vol. 110, no. 22, 2013, pp. 8836-8841.

[15] F. Adams, and C. Osgood, "A Cross-Cultural Study of the Affective Meanings of Colour," *Journal of Cross-Cultural Psychology,* vol. 4, no. 2, 1973, pg. 135-156.

[16] A.J. Elliot, and M.A. Maier, "Colour-in-Context Theory," *Advances in Experimental Social Psychology,* vol. 45, no. 1, 2012, pp. 61-125.

[17] R. Hupka, Z. Zaleski, J. Otto, L. Reidl, and N. Tarabrina, "The Colors of Anger, Envy, Fear, and Jealousy: A Cross-Cultural Study," *Journal of Cross-Cultural Psychology,* vol. 28, no. 2, 1997, pp. 156-171.

[18] H-J. Suk, and H. Irtel, "Emotional Response to Color Across Media," *Color Research and Application,* vol. 35, no. 1, 2009, pp. 64-77.

[19] J. Cage, *Aria* [score]. Peters Edition EP 6701, 1958.

[20] M. Poast, "Color Music: Visual Color Notation for Musical Expression," *Leonardo,* vol. 33, no. 3, 2000, pp. 215-221.

[21] D. Pritchard. 'Colour Circle' by Deborah Pritchard. [Online] Available: https://www.youtube.com/watch?v=aaxB6uLK93A

[22] P. Steinbeck, "Improvisation and Collaboration in Anthony Braxton's 'Composition 76'," *Journal of Music Theory,* vol. 62, no. 2, 2018, pp. 249-278.

[23] A. Jahanian, S. Keshvari, S.V.N. Vishwanathan, and J.P. Allebach, "Colors – Messengers of Concepts: Visual Design Mining for Learning Color Semantics," *ACM Transactions on Computer-Human Interaction,* vol. 24, no. 1, 2017, pp. 1-39.

[24] L. Vickery, "Some Approaches to Representing Sound With Colour and Shape,", in Proceedings of the 4th International Conference on Technologies for Music Notation and Representation, Montreal, 2018, pp. 165-173.

[25] C. Hope, and L. Vickery, "Decibel ScorePlayer – A Digital Tool for Reading Graphic Notation," in *Proceedings of the International Conference on Technologies for Music Notation and Representation (TENOR'15),* Paris, 2015, pp. 59-70.

[26] C, Hope. Compositions [Online]. Available: https://www.cathope.com/compositions

[27] X. Ren, L. Xia, and J. Du, "Delivering Warmth by Hand: Customer Responses to Different Formats of Written Communication," *Journal of Services Marketing,* vol. 32, no. 2, 2018, pp. 223-234.

[28] C. Hope, and L. Vickery, "The Aesthetics of the Score-Score," in *Proceedings of CreateWorld,* Brisbane, 2010, pp. 48-57.

声易经 Shēng Yì Jīng (Book of Sound Changes):

A Comparison of Spectromorphology and 声易学 Shēng Yì Xué (Sound Changes Study)

Ken Fields
Central Conservatory of Music
Beijing, China
Kenfields@gmail.com

ABSTRACT

This paper posits that there is a mapping between the Yijing (YJ) symbolic domain of change and the spectromorphological (SM) domain of sonic forms. Such a demonstration can substantiate the ontological terrain and primary categories of sonic states and their transformations. Spectromorphology emerged out of an empirical/phenomenological method as bound to sonic material, while being shaped by categorical, symbolic and historical considerations. The Yijing arose from a much longer relationship with a method of describing the abstract states and processes of change itself, independent of material or contextual application - determining its object or context at the time of reading or consultation. In this paper, I touch on the basic concepts that situate this study and how it might be applied in a prescriptive musical context (notation).

1. INTRODUCTION

声 (sheng) means sound, 易 (yi) means change, 经 (jing) means book, 学 (xue) means study. 声易学 (sheng yi xue) 'sound change study,' could directly translate to spectro-morph-ology. Morph means form or shape. Morphology refers to the study of forms; spectro-morph-ology: the study of sound forms. Meta-morphosis or morpho-genesis means change or generation of form, respectively.

To orientate the reading, in the following image, we see one example of a possible correlation between the Yijing hexagram Cui, *Clustering* and the spectromorphological graph of *Bulging*. A catalogue of such mappings aims towards a 声易经 (Sheng Yi Jing) or Book of Sound Changes.

Figure 1. Mapping YJ hexagram Cui to the SM symbol for bulging.

Form and change are related such that without form there can be no change and conversely, without change there can be no form. It would be a neat characterization to say that morphology is focused more on 'transformation of forms,' while the yijing system is focused more on 'forms of transformation;' but this has to be investigated. The premise of this project is that a comparative study of the Yijing as applied to the domain of spectromorphological concepts will unfold clear correspondences between the two domains. The YJ is thousands of years old, while SM is a relatively young field, yet this paper aims to show that the comparison can lead to mutual insights.

2. PROBILITY SPACE OF CHANGE

One could look at the field of all possible sound change as a probability space.

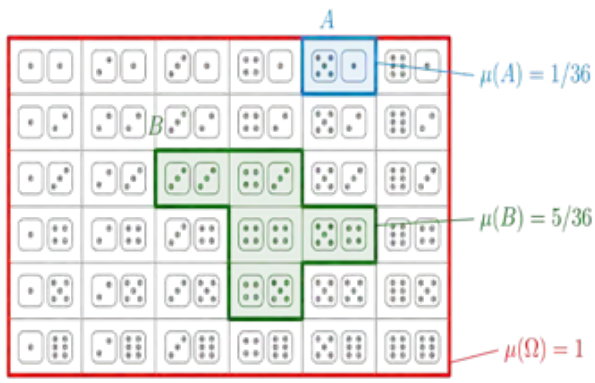


Figure 2. Probability space from https://en.wikipedia.org/wiki/Probability_space.

Let's look at the rules that formalize a generic probability space:

- All possible outcomes are called the 'sample space.'
- An 'event' is one particular outcome.
- Divide the number of event(s) by all possible events which will equal the 'probability measure'

In a similar way, Let the 64 Hexagrams of the Yijing to represent the *probability space* of sound changes.

Figure 3. Yijing hexagrams as a probability space.

Considering the hexagrams to represent sonic transformations, we can call the probability of all sound changes the '*sample space*,' that which represents any one particular abstract change schema an '*event*' and the probability for that change to occur, the '*probability measure*.' Further on, we can call a larger set of changes a 'performance event.'

One can look at the ancient text of the Yijing (易经) in terms of the Chinese equivalent of probability theory.

3. BACKGROUND

This section looks very generally at some underlying dynamics of the system of the Yijing as it might relate to a scientific understanding in general or field and wave phenomenon in particular.

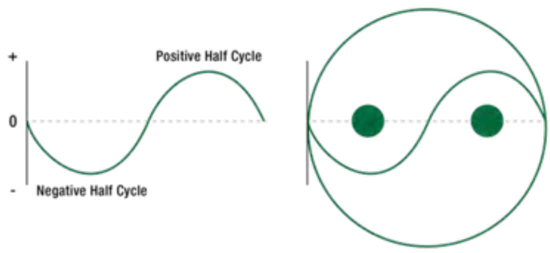


Figure 4. Sine wave compared to the Taiji symbol.

It is said in the Yizhuan[1] "Conjunction and alteration of yin and yang (阴阳) is called the dao (道)" [1, p. 517]. The interaction of yin and yang is the way things are formed and transformed; such is the essence of change or yi (易). Compare this to a famously attributed (dubious) quote to Einstein that, "everything in life is vibration." In any case, this is a fundamental premise in physics.

There is a striking correspondence between Chinese metaphysics and quantum theory. In physics there is the vacuum, or ether, or dark energy, which could correspond to the Chinese concept of *Tàixū (*太虚) – or ultimate void – a complex term originating with Zhang Zai (1020-1070) as a refutation of Laozi's conception of nothingness (wu). "Taixu contains qi… taixu is in both the dispersion and the coalescence of qi… one thing but two states (yi wu liang ti - 一物兩體)" [2, p. 47]. It is then straightforward to extend this concept to the fundamental energetic particles/fields (two states) of quantum theory. Qi is therefore a complex configuration of energies.

When considering the structuration of physical or biological matter by fields of energy (morphogenesis), we have the concept of *Lǐ* (理) in Chinese - *pattern*: holding that "there is no energy that is not patterned" or that "qi functions as emergent organizations of energy (氣之理 qi zhi li)" [3, p. 82]. Wang Fuzhi writes (400 years ago) following Zhang Zai (1000 years ago): "Patterning moves within qi. There is nothing that is not constituted by Qì, so there is also nothing that does not have pattern" [Ibid, p. 82]. This patterning is created out of the polarity of Yīnyáng (阴阳) and continually changes/evolves as subject to the nature of Gǎnyìng (感应 resonance).

This process of patterning as represented in the symbolism of the Yijing, distinguishes three kinds of change in general: no change, cyclic change and sequential change. Movement and rest have their definite laws according to these firm and yielding lines and are differentiated and motivated by energetic polarities [4, pp. 833-34].

4. SPECTROMORPHOLOGY

The philosophy of form and transformation in the West has ancient roots as well (Platonic forms and a persistent cultural/philosophical polarity of a different sort: ideal-real). Not by chance, this dialectic of form plays out well in the early history of electroacoustic music in the

[1] The Yizhuan or Dazhuan, translated as the *Great Commentary* of the Yijing. Attributed to Confucius or disciples.

theoretical approaches of musique concrete and spectromorphology [5, 6]. The concrete/discrete approach of Schaeffer focuses on the physical-phenomenological characteristics of sound objects (the real), while Smalley's SM leans more towards the abstract/reduced (ideal) by identifying sonic structure through the analytic bifurcation of their spectral and energetic (temporal) components. Both methods tend ultimately toward listening and categorization which this project hopes to cross-culturally extend further.

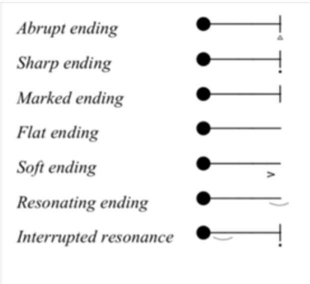


Figure 5. Offset *genres*. Image from auralsonology.com [Online]: Accessed January 2023.

As a broad sociological project spanning decades, the spectromorphological symbolic system has become hybrid and extended, coalescing into an organized language for describing sound shapes, spectra, and transformations. It classifies the abstract properties of amplitude, duration and spectra; onsets, continuants, terminations; motion and growth; harmonicity or density, and more [7]. The graphs can be deployed in analysis or composition, or are, for taxonomists, interesting in themselves.

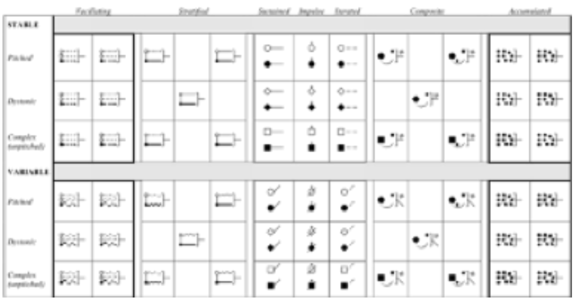


Figure 6. Spectromorphological categories. Image from auralsonology.com. [Online] Accessed January 2023.

The 21st century brings with it new frameworks to challenge spectromorphological approaches, experimenting with embodied, acoustemological, ecstatic, indigenous or even phonosophical approaches [8]. Also, a better understanding of the latest scientific perspectives can be, paradoxically, an opening to ancient modes and approaches that are non-mechanistic and less sonic materialistic.

5. THE YIJING

In comparison to SM, the Yijing utilizes a system of hexagramatic graphs (comprised of six binary yin/yang lines) with associated symbols and poetic imagery: the chinese characters and evocative scenes (treading on a tiger's tail). The symbols represent "all possible configurations of change in nature and in human life." [9, p. 1]. The situational images are abstract - to be applied in specific frameworks, contexts or situations. Here, they are applied to organized sounds.

Figure 7. Hexagram sequence of Fuxi, the mythic creator of the Yijing [9, p.27].

A 3000 year history of 'Guān' (观) or 'comprehensive observation' has led to the formation of the Guà 卦 (tri-/hexagramatic symbolic system), and their many layers of interpretation. When you look at any particular hexagram, you are looking into a virtual space, or house, that has been visited by generations of scholars over the centuries. On the walls of each space hangs their analytical poems, paintings, or commentaries written in thoughtful moments or in consideration of dire situations. The Yijing, while built on a binary-based arithmetical scaffolding, is lush and overgrown with the vines of deep organic/historical experience.

6. HEXAGRAMS

In considering this virtual probability space of sound changes (声易 sheng yi), the hexagrams are reoriented horizontally rather than vertically, in order to consider them from an aspect of resonance, left-right symmetry and from a perspective of mutual relationality.

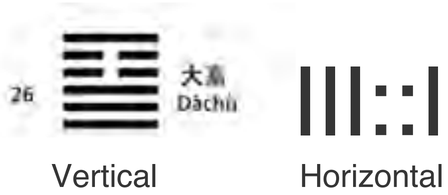
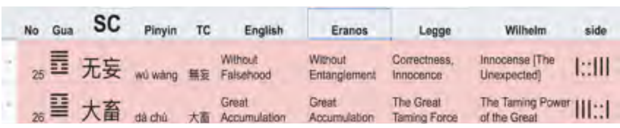


Figure 8. Horizontal orientation of hexagrams allows for emphasizing symmetry rather than hierarchy.

In the horizontal aspect, hexagrams can be weighed as a symmetry, rather than heaven-above and earth-below, or superior and inferior positions. Secondly, we find that 56 of the hexagrams are actually mirrored by their adjacent hexagram - according to their order of listing in the Book of Changes. In other words, they come in pairs and as such form conceptual continua. This is similar to what is known as hexagram fluctuation, or Guà biàn 卦变 as noted by the philosopher Zhu Xi. For example, hexagram pair 25 (Wuwang 无妄, Without Entanglement) and 26 (Dachu 大畜, Great Accumulation) reveal a mirror symmetry:

|::||| <–> |||::|



No	Gua	SC	Pinyin	TC	English	Eranos	Legge	Wilhelm	side
25	䷘	无妄	wú wàng	無妄	Without Falsehood	Without Entanglement	Correctness, Innocence	Innocence [The Unexpected]	|::|||
26	䷙	大畜	dà chù	大畜	Great Accumulation	Great Accumulation	The Great Taming Force	The Taming Power of the Great	|||::|

Figure 8. Mirror hexagrams 25 and 26, Wuwang and Da Chu, represents a continuum of entanglement.

To illustrate the above, the paired sonic archetype as represented by hexagrams 25/26, represents the idea of non-entanglement to full entanglement.

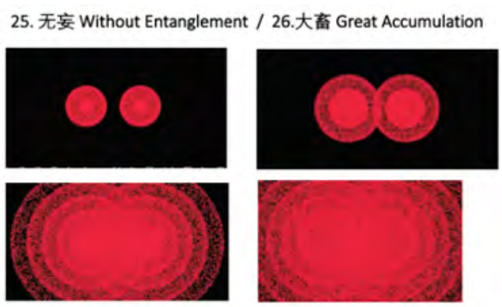


Figure 9. Illustration of the continuum of entanglement.

In total, 56 mirrored adjacent hexagrams are paired based on left-right reversal. The other 8 hexagrams are self-symmetrical or mirrors of themselves (such as ||::|| or :||||:). As pairs, 56 hexagrams are reduced to 28 archetypal concepts plus eight self-symmetrical hexagrams derive 36 sonic archetypes. Interestingly and perhaps non-incidentally, Schaeffer's TARTYP, which is "intended to present a presumably all-encompassing typology of sound objects" breaks down into 28 categories! [7, p. 3).

7. PUTTING SYSTEM TO SCORE

The guqin is an ancient Chinese instrument. An antiquated notation system that exists for the quqin is based on a system of 'reduced characters' (减字谱 jianzi pu). The system assembles a complex music gesture by indicating a string number, fret position, right hand plucking technique and left hand movement. Immediately, we are struck by the similarity of the idea of Schaeffer's "Typological Formulae" to indicate the "interdependence of a group-object and its constituent parts" [5, p. 372]. Six sections of the reduced character are filled in by simplified radicals as indicated on the following diagram [10].

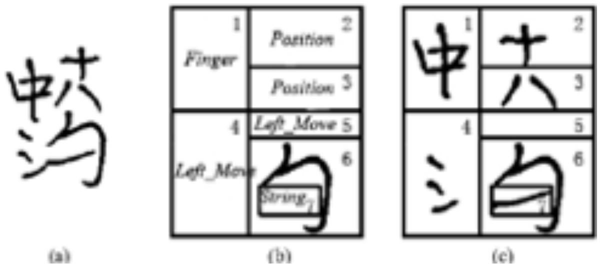


Figure 10. Schema of a reduced character for ancient guqin notation.

While the Sheng Yijing aims to be a general theoretical framework, the reduced character may be a natural solution to translate a hexagram into a six-part formulary to communicate more detail to a performer. But the underlying computer representation would indeed be a more mathematized specification as illustrated in Schaeffer's TOM [5, p. 373]. The Chinese radicals that comprise the displayed character would not indicate guqin gestures necessarily, but rather would represent more broadly sheng yixue / spectromorphological gestures - so that any instrument(s) could perform the score.

For example, while slot 6 (see diagram above) indicates a string number and pluck technique for a guqin, for our system, we could insert the Yijing character here which corresponds to a symbolic sound function (i.e. accumulation). Slots 1-3 might indicate the onset of a pitched or dystonic event and a transformation (pitch envelope of the center frequency of an EQ filter). Slots 4 and 5 might refer to gait and offset/release. This part of the system is yet to be determined.

Currently we are working with brainwave sensors to trigger the symbolic events - hexagram choice. We have started working with mappings - from brainwave event, to Yijing probability space and finally to reduced character that indicates the sonic gesture for the performer. Of course, a brainwave controller is just one way (much more complex than necessary perhaps) to trigger sonic events. That said however, the brainwave sensor output is a multichannel system of neuronal energy flux, which has its own potentialized change space within which we can look for further correspondences between mind, brain, change states and symbolic sonic event.

8. CONCLUSION

This paper has introduced a Chinese approach to sonic description based on its own internal philosophical tradition, the Yijing. Quoting Wang Fuzhi: "The function of nature (tian 天) to enact transformations is qi. Its transformations achieved in the way (dao 道) are patterns (li 理)" [3, p. 82).

Spectromorphology is a system for describing electroacoustic sound shapes and epitomizes the traditional Western approach to sonic description through abstraction and categorization.

The YJ is a process-based system encased within an arithmetical scaffolding. We analogized the hexagramatic space to a probability space of sound changes. We found possible mappings between the Yijing and spectromorphological categories. An important concept in Chinese philosophy is resonance or Ganying 感应 (resonance) which creates flux and movement in a polarized configuration of yin and yang. We begin to explore the resonance between the YJ and SM.

Work has begun on a score system to communicate the shengyi (sound changes) meanings to performers. The reduced character technique is one possible solution for enhancing or reinforcing the abstract nature of the shengyijing system. The simplified radicals in their appropriate slots act as a sonic gestural alphabet, similar to Schaeffer's formulae, to translate the words of the shengyijing images.

Finally, a toss of the coins finds that the space of dispersion-restraint (涣 huan and 节 jie) leads to the perilous pit of darkness. So at this early stage of this project, one would be advised not to venture a definitive conclusion. In general, spectromorphology is deployed in the mode of description more often than it is applied in prescription (notation); more as an analytical apparatus. The Yijing has traditionally been consulted as prognosticative of events and actions; more as a prescriptive apparatus. Thus the Sheng Yijing (Book of Sound Changes) might prove a better system for sound notation for performers, rather than as a precise tool for sound analysis.

Figure 11. Yi (Sound Change). New composite Character designed by Zhang Ruibo.

9. REFERENCES

[1] C.Y. Cheng, "Paradigm of Change (Yi) in Classical Chinese Philosophy: Part I," Journal of Chinese Philosophy 36:4, (December 2009) 516-530.

[2] R.R. Wang and W.X. Ding, "Zhang Zai's Theory of Vital Energy," in J. Makeham (ed) Dao Companion to Neo-Confucian Philosophy, Springer Science + Business Media, 2010.

[3] N. Brasovan, "Aesthetics of Qi: Building on the Internalist-Essentialist Philosophy of Art," in Springer Science + Business Media, (January 9) 2015.

[4] R. Wilhelm, C.F. Baynes (Tr), *The I Ching: or Book of Changes*, Bollingen Series XIX, Princeton University Press, 1967.

[5] P. Schaeffer, North and Dack (Tr), *Treatise on Musical Objects*, University of California Press, 2017.

[6] D. Smalley, "Spectro-morphology and Structuring Processes," in S. Emmerson (Ed), *The Language of Electroacoustic Music*, MacMillan Press, 1986.

[7] L. Thoresen and A. Hedman, "Spectromorphological Analysis of Sound Objects: An adaptation of Pierre Schaeffer's Typomorphology," Organised Sound, vol. 12, no. 2, 2007, pp. 129-141.

[8] K. Fields, "On Phonosophy," Organised Sound, vol. *25*, no. 3, 2020, 274–281.

[9] X. Zhu (Zhu Xi), J. Adler (Ed/Tr), *The Original Meaning of the Yijing: Commentary on the Scripture of Change*, Columbia University Press, 2020.

[10] E. Ni, M. Jiang, X. Ding, C. Zhou, "Handwriting Input System of Chinese Guqin Notation," ACM Journal on Computing and Cultural Heritage, vol. *3* no. 3, March 2011.

Host Institutions:

Sponsored by:

Northeastern University

9 780648 159278