

Co-creative orchestration of Angeles with layer scores and orchestration plans

Francesco Maccarini, Maël Oudin, Mathieu Giraud, Florence Levé

To cite this version:

Francesco Maccarini, Maël Oudin, Mathieu Giraud, Florence Levé. Co-creative orchestration of Angeles with layer scores and orchestration plans. International Conference on Artificial Intelligence in Music, Sound, Art and Design (EvoMUSART 2024), 2024, Aberystwyth, United Kingdom. 10.1007/978-3-031-56992-0 15 hal-04388097

HAL Id: hal-04388097 <https://hal.science/hal-04388097v1>

Submitted on 22 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

[Distributed under a Creative Commons Attribution - ShareAlike 4.0 International License](http://creativecommons.org/licenses/by-sa/4.0/)

Co-creative orchestration of Angeles with layer scores and orchestration plans

Francesco Maccarini¹, Mael Oudin², Mathieu Giraud¹, and Florence Levé^{3,1}

¹ Univ. Lille, CNRS, Centrale Lille, UMR 9189 CRIStAL, F-59000 Lille, France

² Department of Music Research, Schulich School of Music of McGill University, CIRMMT, Montreal, Canada

³ MIS, Université de Picardie Jules Verne, F-58000 Amiens, France

Abstract. Orchestration is the process of creating music for a group of instruments, combining, blending, and contrasting their sounds to produce a unique orchestral texture. In this research/creation project, our team was commissioned to create an AI/human orchestration of two movements of Angeles, a piano composition by Gissel Velarde. The project turned out as a perfect case study for computational creativity in music and orchestration, where the role of the model is between AI as a colleague and AI as a tool. Our main contribution is a preliminary framework for computer assisted orchestration. By modeling a layer score and an orchestration plan in the orchestration process, we implement a simple Markov model that selects possible instrumentations for each score segment. Personalization of the AI and AI/human interaction occur through human segmentation of the score at two stages of the process (layer score, orchestral segments with loudness profile), through instrumentation presets, and finally through selection of the final orchestral plan and through the actual orchestration. We detail the research aspects of this co-creative project and analyze the roles of the actors involved in the creation of the final piece: the Music Information Retrieval (MIR) researchers, the orchestrators, and the algorithms.

Keywords: Orchestration, Computational Creativity, Co-creativity, Music Generation.

1 Introduction

1.1 Orchestration and Creativity

Orchestration Orchestration involves composing or arranging music for a large group of instruments, mixing or contrasting their sounds to create an orchestral texture. Orchestration goes well beyond distributing the voices among the

Author revised version, as published in

Maccarini, F., Oudin, M., Giraud, M., Levé, F. (2024). Co-creative Orchestration of Angeles with Layer Scores and Orchestration Plans. Artificial Intelligence in Music, Sound, Art and Design (EvoMUSART 2024), LNCS 14633, 228–245.

https://doi.org/10.1007/978-3-031-56992-0_15

<https://hal.science/hal-04388097>

instruments. Formalization and teaching of instrumentation and orchestration has improved over time, with various treatises covering topics such as the musical capabilities of each individual instrument, and their combination in order to shape the sound of the orchestra, and to render specific perceptual effects [8,23,57,37,55,47,2].

We can characterize orchestral music and orchestration through *musical tex*ture. In Western music, there are several commonly discussed types of texture, including monophony, polyphony, heterophony, and homophony [18]. In his definition of musical texture, Huron proposes to analyze the sound material according to the number of elements happening simultaneously (density) and their homogeneity or heterogeneity *(diversity)* [30]. Benward and Saker [7] describe texture by identifying different parts (or layers), which have different roles, such as melodies or (static, harmonic, or rhythmic) support. Complex textures can be created through the varying timbers of the large number of orchestral instruments and their combinations [47].

This richness of possible instrument combinations makes the analysis and generation of orchestral music particularly challenging for computational musicology [52,9,15,50,49,19]. Perceptual studies focused on how instruments combination affects timbral response, blending qualities, and on the use of timbral consonance and dissonance to create sounds [36,60,62]. Open corpora have been created to help research in orchestration techniques, automation, and perception. Crestel et al. published the Projective Orchestral Database (POD) linking piano and orchestral scores [15]. We proposed with Le a bar-to-bar textural analysis of 24 movements of symphonies [38]. The Orchestration Analysis & Research Database (OrchARD, <orchard.actor-project.org>, data not published) rather targets on specific auditory effects [44].

Computational Creativity and Music Generation The spread of systems for AIbased music generation studies has contributed to drive a growing interest in machine creativity [48]. Many questions emerge regarding the definition, evaluation, and uses of creative machines [34,33]. Esling and Devis suggest considering generative AI algorithms as creativity-enhancing tools [20]. This proposal finds its place in Lubart's classifications of modes of human-machine interaction in a co-creative process [41]. Among them, the computer as a colleague mode implies a direct involvement of algorithms in the creation of the final output, rather than a mere assisting role (algorithm as a tool) for the creator. Similar categories are found in Kantosalo and Jordanous's description of the roles of AI in creative processes. They divide them in the categories of co-creative colleague and creativity support tool [35].

Some studies focused on human creativity in music [56], and on the way it can be enhanced with Machine Learning (ML) models [43] through modes of interaction between the artist and the model [24,5]. Difficulties emerge in the evaluation of the creativity of such systems [4]. The impact of a good interface for steering AI on the ability of users to express musical ideas and "own" the resulting creation has also been highlighted [40]. Co-creative systems in music generation [21,27,10,11,31] can be divided into two categories: for live performance improvisation $[5,22,54,65]$, and for composition and production $[6,1,61,58]$. The AI song contest, a competition specifically focused on co-creation in songwriting, inspired different uses of AI and reflexion [28]. Some of us have been part of such a team, involving a composer the right from the beginning of the system design, resulting in personalized AI models [16].

Machine Learning, Creativity, and Orchestration In the realm of orchestral music, specific ML models were proposed for tasks related to arrangement [29,64], instrumentation [17], orchestration [53,26,14,12], and generation of orchestral music [39]. Beethoven X experiment was an AI-assisted composition of orchestral music that aimed to be a plausible Beethoven symphony [25], obtained by the collaboration between the composer Walter Werzowa, musicologists, and computational methods including generative ML models. No code nor data are publicly available, but their co-creative methodology is interesting, divided into continuation (expanding melodic lines and themes, from the original melodic material from Beethoven's own sketches), harmonization (composition of accompaniments parts for the melodic ideas, including, for example, homophony, counterpoint, and fugue), transition (orchestral/polyphonic inpainting, to connect different ideas), and orchestration (organizing across the available instruments and instrumental families of the orchestra).

1.2 Goal and Contents

We carried out this project to orchestrate two piano pieces by Gissel Velarde, focusing on co-creativity, i.e. putting human beings and computer models in the same loop to make music. In order to get an outcome of high quality scores, ready to be distributed to the orchestra, in a limited time frame, we focused on modeling the art of orchestration as a formal process, and on the possible ways of interaction between a human orchestrator and computational algorithms. We explicitly searched for a balance between having simple but high-level conceptual data that can be handled by AI and patterns that can be explained so that the musician can interact with them. Following Benward and Saker [7], we characterize orchestration in the "classical style" as an overlay of different layers, which have the roles of melodies or (harmonic/rhythmic) accompaniment, played by the different instruments of the orchestra and their combinations.

Our workflow of orchestration is the following (Figure 1):

- The first step is an analysis: the goal is to extract from the piano score a layer score, that is an abstract version of the piece analyzing musical texture (Figure 2 and Section 2). Voices are represented together with their role: melody, harmonic accompaniment, or rhythmic accompaniment. At this stage we do not provide any information about the instrumentation.
- $-$ The second step is to **build an orchestration plan**: We assign each layer to an instrument or a group of instruments, taking care of the balance between the different timbres. We have developed a Markov model for this stage, that uses probabilities of finding instruments together, of instrument sequences,

Piano si

nowledge base
Markov Mode models of
orchestratio Layer so

Fig. 1: The workflow of the co-creative orchestration builds on a layer score and on an orchestration plan.

and instrumental density, building from instrumentation presets as well as from a segmentation of the score (Section 3).

2. orchestration plar

orchestral scor

– The last step is to write the actual instrumental parts, following the orchestration plan and taking care of the peculiarities of range and dynamics of each instrument (Section 4).

For this experiment, the first and third step were done by human orchestrators, who are two of us: Mael Oudin (professional orchestrator, PO), and Mathieu Giraud (amateur orchestrator, AO). They both are also researchers for the project. The focus on human/AI co-creativity is thus here in the second step but also in the very decision of modeling the layer score and the orchestration plan to enable these three steps. We neither claim that this three-steps process fully models the art of orchestration, nor that it represents an optimal process, but rather that it is a plausible workflow to orchestrate a piece, which could be followed by human orchestrators alone, but that has the advantage to allow interaction between humans and AI.

The rest of the article details those three steps (Sections 2, 3, 4) and the results of this orchestration model, analyzing the roles of the actors – MIR researchers, orchestrators, and computational models. We conclude by discussing the challenges encountered, our positioning in computational creativity research, and perspectives (Section 5).

2 Modeling a Layer Score

2.1 Behind Orchestral Music: Analyses, Sketches

Any music, any score, may be seen as a rendering of high-level musical ideas, that may be intermediate steps when composing or improvising music, or serve analytical purposes. Describing and modeling these ideas is a challenge for (digital) musicology and music analysis.

Fig. 2: First eight measures of the first movement of Symphony $#9$ by L. van Beethoven, op. 125. (a) The orchestral score can be decomposed into three layers: one with an harmonic role (red), one with a rhythmic role (blue), and a third one with a *front* role, that is nevertheless difficult to categorize as a melody (yellow). (b) The layer score contains one part for each of these three layers, ℓ^{harmon} , ℓ^{rhythm} , and ℓ^{front} . (c) In the piano reduction, the harmonic and rhythmic layer are blended together and rendered with a pianistic texture, which is different from the orchestral version. The left hand alternates between that harmonicrhythmic layer, and the front layer.

(a) Full Orchestral Score 13 tracks/instruments (7 in this extract) Transcribed by ClassicMan on <musescore.com>

Textural labels from [38]. Front: Vl1, Vla, Cb; Rhythm: Vl2, Vc; Harmony: Cl, Hrn

(b) Layer Score 3 layers

(c) Piano reduction Arr. by Franz Liszt, ed. Breitkopf & Härtel

Analytical concepts. As usual in music analysis, an orchestral score can be studied at different levels. Taking an orchestral score as a "neutral level" [51], the analyst can examine orchestration techniques. Focusing on textural/instrumental aspects, music can be split into a number of parts, or layers, with a predominantly melodic, rhythmic, or harmonic role, or a mix of them [7]. In orchestral music functional layers are highlighted by combinations of instrumental timbres [38]. At a higher level, orchestral effects arise from auditory grouping processes [44].

Composer/orchestrator sketches. Were these layers and effects present in the composer's – or the orchestrator's – mind? Composers may have some sketches (in their mind and/or onto paper) of these parts and then complete those sketches into more elaborated music [59,66]. The study of Beethoven's sketches have revealed precise musical ideas such as patterns, themes, and sometimes orchestration sketches [32]. The options available to composers for orchestral music encourage them to explore *textural* spaces [63], but it is unclear to what extent they can model complex organization of layers in their mind.

Concurrent musical parts and layers can thus pre-exist the orchestral score, which can be heard as an expansion of these initial materials. Some MIR studies on orchestration already used this idea: Gotham et al. name as a short score a set of raw materials before orchestration [25]. Somehow, in jazz/pop styles, lead sheets with melodies and chords may also be seen as a condensed version of a music piece – or could correspond to a sketch of the final song. MIR and AI methods aiming at (co-)creating such music often generate such lead sheets at first, then proceed with accompaniment generation [28], even if the split is debatable and an end-to-end generation is sometimes preferred [3].

2.2 Defining and Modeling a Layer Score

Inspired by such hypothetical intermediary "sketch scores" and by analytical considerations, we define here the layer score as a score with a variable number of layers, each one with a given role, following the description of roles we gave in [38] for classical-romantic orchestral music, but without any indication on the instrument.

On the opening of the Beethoven's 9th Symphony, we analyze three layers with different roles, ℓ^{front} , ℓ^{thythm} , and ℓ^{harmony} (Figure 2b). Those layers are distributed on several instruments in the full orchestral score. Here, one part in the layer score roughly corresponds to one layer in the orchestral score, but the actual music in the orchestral part could be more different from the content of the layer score, possibly depending on the capabilities of the instruments chosen to render a layer. For example, rhythmic motion could be rendered with a different density of notes if reproduced through a timpani roll, or through tremolo strings sul ponticello.

At the opposite, if we consider now a piano reduction of an orchestral score, the same layers will be blended together into the two staves of the piano reduction – a single pianist should be able to play it. In the layer ℓ^{r} hy^{thm} of the opening of the Beethoven's 9th Symphony, the lower strings repeat chords made of A and E, whereas the piano reduction by Liszt alternates between the same pitches (Figure 2c). Again, the actual music in the piano reduction could be different, in the rhythm organization for instance. In practice, the octaves often differ in piano reduction, the pianist hands not having the same ambitus as the orchestra. In some cases, the actual pitches may even differ between instrumentations, for example including patterns or scales in some of them.

The layer score therefore generally has fewer staves than a full orchestral score and more staves than a piano reduction, but could be (anachronistically) viewed as the common ancestor between the two – or, more generally, an ancestor of any other instrumentation of the music. The layer score includes the structural, melodic, harmonic content, and some of the textural content, but not the actual music rendering. This concept of layer score allows thus to decouple composition from instrumentation and orchestration (although this decoupling may be artificial) and to devote our attention mainly to the latter ones.

Back to the question of composer sketches, a layer score is not intended to reproduce an existing compositional practice. We do not claim either that writing such layer scores would be a desirable practice. Writing music for the orchestra (or for any other instrument) can be a non-linear process, with iterations between high-level ideas and actual music content. For example, the constraints given by the ambitus of each instrument influence the composition itself, forcing the composer to rethink some of the choices they have already made, and even to rewrite major sections of the composition.

Anyway, layer scores give new perspectives on topics related to orchestration. For example, piano reduction and orchestration from piano are symmetrical tasks. Orchestration from piano usually requires "de-pianotizing" the piano music, including voice separation [42] or texture analysis [13] – as piano reduction requires "de-orchestrating" the orchestral music. Modeling such a layer score thus increases the possibilities for a human intervention.

2.3 Results on Angeles

We illustrate our methodology through the first two measures of El Jardin Etero, the last movement of Angeles (Figure 3). In the layer score, we have separated into two different staves the melodic ℓ^{mel} and rhythmic ℓ^{rhyl} layers that emerge from the piano score. We also decided to stress the two last notes in a separate layer ℓ^{rhy2} and to add another rhythmic layer ℓ^{rhy3} . Such decisions in the analytical process have their part of subjectivity and contribute to the co-creativity, allowing to recreate new textures. The professional orchestrator states:

PO: My role was to make human musical choices during all stages of the AI-assisted orchestration process. As an orchestrator, identifying the texture is my first job. Analyzing the piano score allows me to deconstruct the music into different roles: the main melody, harmony, rhythm, res $onance... I$ prepare the addition of new parts. The piano is limited by its technique. As I orchestrate, I will add what is "absent but suggested",

Fig. 3: First two measures of movement 6 "El Jardin Etereo" from Angeles by Gissel Velarde, op. 7. (a) Original Piano score, provided by the composer. Clearly, the main melody is at the bass, and the right hand plays a rhythmic layer. (b) Layer Score elaborated as an intermediate step of the orchestration process. A melody in the low register ℓ^{mel} , and a rhythmic layer ℓ^{rhyl} have been directly identified by splitting the right and left hand of the piano score. Other layers $(\ell^{\text{rhy2}}, \ell^{\text{rhy3}})$ have been added to stress the importance of some notes and to have more possibilities of rhythms, departing from a typical pianistic texture. (c) Layer instrumentation, assigning instruments to each one of these three layers. The human orchestrator decided to have one rhythmic layer ℓ^{rhy1+3} , with added notes in the downbeats. (d) Orchestration of the piece by Mael Oudin. The instrumental parts have been written following the selected orchestration plan. Scores and rendered audio for selected extracts of the piece are available at <http://www.algomus.fr/data>.

resonance, missing registers, textures to be recreated. The first stage involved the creation of the "layer score" for this movement.

Creative choices, such as adding a layer, can already be made at this stage, going beyond a pure analysis of the original score. For example, a music pattern may have at the same time a melodic and an harmonic role. In that case, the layer score should define a mixed melodic/harmonic layer (see Section 5).

3 Modeling, Generating, Selecting an Orchestration Plan

3.1 Orchestration Plans

Once a score is split into layers, we have to map them to the available instruments, and instrument groups, in the ensemble, in order to describe the envisioned instrumentation for the piece. Given a set of layers such as $\mathcal{L} =$ $\{\ell^{mel}, \ell^{rhy1+3}, \ell^{rhy2}\},\$ a layer instrumentation is a list LI of instruments in each layer, each instrument occurring in at most one layer (see Figure 3c).

Each orchestration change, being it progressive, or a contrast, uses a new layer instrumentation. We thus define an *orchestration plan* as a set of layer instrumentations LI^1, LI^2, \ldots, LI^n , one for each *orchestration segment* $s \in \{1, 2, \ldots, n\}$ of the score. In the current model, it is the musician who segments the score at relevant points where they want orchestration changes. The syntax of the orchestration plan describes, for each segment, the layers using instruments in a predefined order and details the layers and roles (Figure 4c).

3.2 Personalized Orchestration Plan with Markov Models

Once they have set the orchestration segmentation, the artist could themselves write the layer instrumentation for every segment, and create an orchestration plan. Here, instead, we decided to have a simple knowledge-based algorithm to experiment with AI/human interaction. This model for semi-automated layer instrumentation follows two goals. The layers should include instruments that blend together, drawn from presets of possible instrumentations. Moreover, the "loudness" of the instrumentation at each segment should be close to the musician's desired outcome, for which they provide a loudness profile as input. These concepts are detailed in the following paragraphs.

Loudness profile and acoustic weights. To underline the form, the musician inputs a *loudness profile* as a list of targeted loudness values $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ for the n segments. The effective loudness depends on the dynamics, but also on the number and the qualities of each instrument⁴. A simple model is to consider that each instrument *i* has an *acoustic weight* w_i . We decided here to have higher coefficients for brasses and instrument of lower range, using the following values:

⁴ We call "instrument" an instrument group. Groups may include several people (Vl1).

(Fl:1, Ob:1, Cl:1, Fg:1.5, Hrn:1.5, Trp:2, Vln1:1, Vln2:1, Vla:1, Vc:1, Cb:1.5)

The loudness could be estimated as the sum of the weights of the instruments involved. However, selecting only the instruments according to such values would not realize a proper orchestration, as it would ignore blending qualities and orchestrator preferences.

Possible instrumentations. For each layer $\ell^{\alpha} \in \mathcal{L}$, the musician will thus define a set of *possible layer instrumentations* $pli^{\alpha} = \{pli^{\alpha}_1, pli^{\alpha}_2, \ldots\}$, each one being a weighted list of instruments. For instance, a rhythmic layer ℓ^{rhy1} could be associated to two distinct instrumentations, either on woodwinds, or brasses:

$$
p\ell i^{\text{thy1}} = \begin{cases} p\ell i^{\text{thy1}}_{wood} = (\text{Fl} : .3, \text{Ob} : .1, \text{CIBb} : .2, \text{Fg} : .15) & L(p\ell i^{\text{thy1}}_{wood}) = .825 \\ p\ell i^{\text{thy1}}_{brass} = (\text{HrnF} : .7, \text{TrpBb} : .2) & L(p\ell i^{\text{thy1}}_{brass}) = 1.45 \end{cases}
$$

Each component (i, p_i) tells that the instrument i should have a probability p_i of being used in this pli : The actual instruments that will be used will be a subset of that pli . Selecting a pli ensures that these instruments blend together for this particular layer. The sum $\sum p_i$ of the probabilities of a pli is the expected number of instruments in that pli . We rather use the *expected loudness* of the $p\ell i$, that is $L(p\ell i) = \sum w_i p_i$, weighting each probability by the acoustic weight of each instrument.

Selecting the pli then the instrumentation for each segment. Given a layer $\ell^{\alpha} \in \mathcal{L}$ and a segment $s \in \{1, 2, ..., n\}$, the $p\ell i^{\alpha,s}$ is selected in $p\ell i^{\alpha}$ according to a Markov model (Figure 5) that depends on the previous $p\ell i^{\alpha,s-1}$. In order to match the prescribed loudness λ_s , the model also tries to minimize $\delta^{\alpha} = |\lambda_s - \lambda_s|$ $L(p\ell i^{\alpha,s})$ by applying a further $e^{|\delta^{\alpha}\tau|}$ coefficient with $\tau = 2.0$.

For a given segment s, once all $p\ell i^{\alpha,s}$ are selected for all layers ℓ^{α} , instruments are assigned following the individual probabilities p_i . At the end of this step, it may happen that either a layer has no instrument assigned, or that an instrument is assigned to more than one layer. Such cases are resolved by further random assignations, based again on the p_i in the pli .

The personalization of the AI "to the style of the orchestrator" and the possibility to steer the AI are thus done both on the presets/ pli selections and on the loudness profile input. Moreover, the implemented method generated segments with three levels of relative loudness for each segment (Figure 4b), enabling the orchestrator to further select instrumentations at each segment, but still keeping the coherency of the pli .

3.3 Results on Angeles

PO: To develop the model that generates orchestration plans, I collaborated with the Algomus team to propose instrument combinations that I enjoy using, and to fine-tune the model.

Fig. 4: Creating the orchestration plan of Angeles, mvt 6. (a) The score is split by the musician into 16 instrumentation segments, each with a target loudness. (b) The model generates, for each segment, three *layer instrumentations* taking into account the expected segment loudness and another relative loudness coefficient $(0.5, 1.0, 4.0)$ (c) In the selected orchestration plan, for the segment $[p01]$ (first two measures), there are four layers instrumentations $\ell_{brass}^{\text{rhy1}}, \ell_{words}^{\text{rhy2}}, \ell_{brass}^{\text{rhy3}},$ and ℓ_{mel2}^{mel} . The layers are mapped to the instruments appearing in the order declared in InstList: For example, the "<2.2.|" bloc in the woodwinds refers to the layer instrumentation $\ell_{wood}^{\text{rhy2}}$, with here flutes (F1) and clarinets (C1Bb).

(a) Loudness profiles

LOUDNESS = $[0.20, 0.40, 0.20, 0.50, 0.30, 0.40, 0.20, 0.50,$ 0.30, 0.50, 0.30, 0.60, 0.40, 0.50, 0.60, 0.70]

(b) Generated orchestration plans, with several relative loudness

(c) Final orchestration plan

According to our orchestrator, there are several goals that a good orchestration should pursue: respecting and enhancing the piano composition, a good balance between layers (especially, the melody should not be muzzled by accompaniment), variation (contrasting moments in the piece should carry different orchestrations), and efficient dynamics (through loudness values and coefficients).

Respecting the piano composition means that the score brings constraints in register and dynamics that need to be reflected in the orchestration plan. For example, movement 6 was a Vivace, with the melody on the bass and an accompaniment more rhythmic than harmonic. The orchestration then had to address the character of the piece and abide by the register of each layer.

PO: The bass melody could only be performed by three instruments in the orchestra: the cellos, the contrabasses, and the bassoons. But not every choice would give a satisfying balance to the other layers played by the rest of the orchestra. The bassoons or contrabasses alone, for instance, would not be prominent enough so cellos were necessary here. Any generated orchestration plan that would not nominate cellos for that layer would be in practice almost unusable.

Fig. 5: Extract of the transition table of the Markov model modeling the evolution of $p\ell i^{\text{mel}} = \{p\ell i_{\text{H1}}^{\text{mel}}, p\ell i_{\text{H2}}^{\text{mel}}, p\ell i_{\text{B}}^{\text{mel}}, \ldots\}$ for movement 6. The transition table was created through iterations between the MIR researchers and the orchestrator. These coefficients are further adjusted by a loudness factor, then normalized.

$$
\begin{array}{rcl}\np\ell i_{\mathsf{H1}}^{\mathsf{mel},\mathsf{s}-1} \longrightarrow p\ell i_{\mathsf{H1}}^{\mathsf{mel},\mathsf{s}}: .8 & p\ell i_{\mathsf{H2}}^{\mathsf{mel},\mathsf{s}-1} \longrightarrow p\ell i_{\mathsf{H1}}^{\mathsf{mel},\mathsf{s}-1} \longrightarrow p\ell i_{\mathsf{H1}}^{\mathsf{mel},\mathsf{s}-1} \longrightarrow p\ell i_{\mathsf{H1}}^{\mathsf{mel},\mathsf{s}}: .3\\ \longrightarrow p\ell i_{\mathsf{H2}}^{\mathsf{mel},\mathsf{s}}: .1 & \longrightarrow p\ell i_{\mathsf{B2}}^{\mathsf{mel},\mathsf{s}}: .1 & \longrightarrow p\ell i_{\mathsf{H2}}^{\mathsf{mel},\mathsf{s}}: .3\\ \longrightarrow p\ell i_{\mathsf{B3}}^{\mathsf{mel},\mathsf{s}}: .1 & \longrightarrow p\ell i_{\mathsf{B4}}^{\mathsf{mel},\mathsf{s}}: .8\n\end{array}
$$

Balance in an orchestration is also reached through the separation of the orchestra into different groups (namely, the strings, the woodwinds, the brass, and the percussion). Harmonic blending is best achieved when all the notes of a chord are performed by instruments in the same group. This was a constraint to our model if we wanted to avoid too much disparity in the instrument combinations proposed in the orchestration plans. Movement 6 had a continuous 3-voice rhythmic layer and we wanted these three voices to be performed by instruments from the same group. This wasn't always the case in the orchestration plans generated so this was also a criteria in the selection of the best outputs.

The ensemble of these parameters creates many constraints on orchestration possibilities, so one of the first human tasks when analysing the output generated by the model is to remove what seems impossible, for reasons as varied as register limitations, number of voices to be played, and poor blending or contrast. Perspectives include (semi-)automatizing some of these tasks. At the same time, the challenge was to foresee the potential of each generated combination when formalized into a musical score at the next stage of the process. Some combinations, such as opening the rhythmic layer with brass only, were unexpected by the orchestrators but rather "proposed" by the model (Figure 4c).

PO: I worked on the outer sections of "El Jardin Eterno" using several dozens of orchestration plans generated by the model. It was my responsibility to sort through them and select the most convincing ones according to my taste (while also being open to surprises).

4 Writing and Performing the Orchestral Score

Once the orchestration plan is decided, the orchestrator has a large space of possibilities related to the range, dynamics, and playing techniques of each instrument that can still be creatively explored. The orchestration plan only suggests the instruments to be used in every portion of the piece, but many decisions still need to be taken to get a playable score, in particular to have idiomatic patterns for each instrument of the orchestra.

In the final orchestral score, on the same first two measures (Figure 3d), the choice has been made to fill the rest of the first beat of the rhythmic texture to provide a more efficient and easier line to the brass instruments at this fast tempo. The rhythmic layer is then rendered differently from the original pianistic texture, but it preserves the intention. Likewise, the choice of writing pizzicati for the contrabass part, to lighten the orchestral texture and express the mezzopiano dynamic, was taken at that stage of the process.

Our orchestrator shared some reflections on the artistic side of this final step:

PO: The distribution of instruments is suggested by the model, but there is still considerable freedom in the choice of notes and registers. It is also up to me to choose and indicate nuances, phrasing, playing modes and expressive indications. Some of the model's choices wouldn't have been what I would have done, like starting directly with the trumpets at the beginning of movement 6. But it's stimulating!

Movement 2 "Inexorable" and the outer sections of movement 6 "El Jardin Eterno" have been orchestrated with this procedure whereas the middle section of movement 6 has been orchestrated with a "traditional" method. The other movements have been commissioned to orchestrators outside of our team. The whole suite has been performed by the Orquesta Kronos conducted by Andrés Guzmán-Valdez at Nuna Theatre in La Paz, Bolivia, on 19th July 2023.

The experience of working with different methods showed that the AI-assisted method starts to offer a gain in productivity once the orchestration plan is reliable. It is also a tool for creative thinking:

AO: Like any creative work, an orchestrator may face the anxiety of the blank page. Especially as an amateur orchestrator, I enjoyed having such suggestions. Even when they were inappropriate, they stimulated creativity through reinforcement, contrast, or opposition.

5 Discussion and Perspectives

In [16] the use of *low-tech AI* is advocated, to ease the communication between the composer and the researchers, and to obtain tailor-made models with scarce data. We adopted a similar approach here, to focus on modeling the process of orchestration by identifying possible steps, and adapting it to the co-creation environment. The computational model used to generate orchestration plans has been conceived and designed with a continuous back and forth between the artist and the research team. It is meant to be the simplest possible, so that it can be more easily modified to experiment with different inputs and controls. Concepts like loudness profiles and coefficients have been added to the model to respond to the ideas and the necessities of the orchestrator.

We believe that we have succeeded in individuating three well-separated stages of orchestration (Fig. 1), which could be performed by three different (human or algorithmic) actors. Each step included a self-refining feedback loop. For example, the human task at the final stage of the process (3. writing orchestral score) can be described in two phases, the second of which is usual for "traditional orchestration": (3.1) interpret and adapt the orchestration plan, (3.2) write notes for the instruments. In the first stage, the orchestrator will read

the output made by the machine and mentally link it to the score to find the best strategy to transform these outputs into music notes (this phase is fundamental to selecting the best outputs from the machine). The second phase consists of writing the notes idiomatically for the instruments, but also the dynamics and the phrasing and expressive instructions.

The current model has limitations. The process is not always linear: it was sometimes in phase (3.2) that choices made in phase (3.1) retrospectively appeared to be pitfalls, and the entire process had to be made again (when, for example, a particular arrangement was not compatible with the plan selected for the subsequent section). Other points could also be improved, as for example the formation of mixed layers. More generally, a challenge is to better model largescale orchestral thinking. Orchestral contrast is typically achieved when the same instrument (or combination of instruments) is not used in the same way in two successive contrasting parts. This may lead orchestrators to "reserve" an instrument on purpose for a specific moment in the piece. Somehow, the *pli* presets combined to the acoustic weights and the targeted loudness profiles help such a large-scale homogeneity and steerability, but these models could be refined.

Looking back to the categories proposed by Kantosalo and Jordanous [35], we have experimented with a process rooted in the interaction between the computational models and the artist, in which the role of the model is in between AI as a colleague and AI as a tool. Dividing the orchestration process into steps has facilitated the introduction of computational models. In this way, the role of the AI is to act on a well defined and specific task, making the model an essential tool in the overall process. The algorithm is acting on the product of human actions (the layer score), and is enabling further human processing with its output (writing the final orchestral score). At the same time, the model is able to "suggest" unforeseen ideas to the humans, acting more as a co-creative colleague, who can inspire and enhance the inspiration of the human artist.

Any co-creative project confronts us with questions related to the authenticity and the ownership of the such art [46]. When dealing with AI-generated art, ethical, legal, and moral concerns emerge, questioning the status of the product itself as having artistic qualities [45]. In this research and creation project, the development of the algorithm and the artistic creation of the orchestrated score were intertwined processes. For this reason, the score have been signed with Orchestrated by Mael Oudin and the Algomus team, recognizing authorship to all members of the project, musicians and computer scientists.

In summary, we proposed a framework for AI assisted orchestration, in which the orchestrators craft their art in collaboration with AI algorithms. We divided the process into three steps, modeling layer scores and orchestration plans as intermediate objects. The code for generating orchestration plans is available under an open-source licence at <algomus.fr/code>. Through this preliminary project, this approach has proven to be effective in formalizing the art of orchestration, enabling the involvement of both machine and human actors, each contributing at different moments. The possibilities for the employment of AI in the process are not limited to the ones selected for the scope of this project. Perspectives include modeling other tasks in the process with Deep Learning AI, both for texture analysis tasks related to the creation of the layer score, and for constrained notes generation, in the creation of the final score. Co-creative interactions would be allowed through model parameters, and through creative modifications of the outputs at several stages: when writing the layer score, the orchestration plan, and the final rendering of the notes. All these steps can be accomplished partly by the machine and partly by the human being, with a fruitful continuous exchange of information.

Acknowledgements

We deeply thank Gissel Velarde, Andrés Guzmán-Valdez and the Orquesta Kronos, and the Nuna Theatre in La Paz, Bolivia. We thank the valuable input and comments provided by Dinh-Viet Toan Le, the Algomus team, and the anonymous reviewers. We also extend our gratitude to the ACTOR project for facilitating collaboration among the authors of this paper and inspiring enriching discussions.

References

- 1. Adkins, S., Sarmento, P., Barthet, M.: LooperGP: A Loopable Sequence Model for Live Coding Performance Using GuitarPro Tablature. In: Johnson, C., Rodríguez-Fernández, N., Rebelo, S.M. (eds.) Artificial Intelligence in Music, Sound, Art and Design. pp. 3–19. Lecture Notes in Computer Science, Springer Nature Switzerland, Cham (2023). [https://doi.org/10.1007/978-3-031-29956-8_](https://doi.org/10.1007/978-3-031-29956-8_1) [1](https://doi.org/10.1007/978-3-031-29956-8_1)
- 2. Adler, S.: The Study of Orchestration. Norton (1982 (first ed))
- 3. Agostinelli, A., Denk, T.I., Borsos, Z., Engel, J., Verzetti, M., Caillon, A., Huang, Q., Jansen, A., Roberts, A., Tagliasacchi, M., Sharifi, M., Zeghidour, N., Frank, C.: Musiclm: Generating music from text (2023)
- 4. Agres, K., Forth, J., Wiggins, G.A.: Evaluation of musical creativity and musical metacreation systems. Computers in Entertainment (CIE) 14(3), 1–33 (2016)
- 5. Assayag, G.: Creative symbolic interaction. In: Sound and Music Computing Conference (SMC 2014) (2014)
- 6. Ben-Tal, O., Harris, M.T., Sturm, B.L.: How music ai is useful: Engagements with composers, performers and audiences. Leonardo 54(5), 510–516 (2021)
- 7. Benward, B., Saker, M.: Music in Theory and Practice (Volume 1). McGraw-Hill Professional (2008 (8th ed))
- 8. Berlioz, H.: Grand Traité d'instrumentation et d'orchestration Modernes. Novello (1844 (first ed))
- 9. Bosch, J.J., Marxer, R., Gómez, E.: Evaluation and combination of pitch estimation methods for melody extraction in symphonic classical music. Journal of New Music Research 45(2), 101–117 (2016)
- 10. Briot, J.P., Hadjeres, G., Pachet, F.D.: Deep learning techniques for music generation. Springer (2019)
- 11. Briot, J.P., Pachet, F.: Deep learning for music generation: Challenges and directions. Neural Computing and Applications $32(4)$, $981-993$ (2020-02-01). [https:](https://doi.org/10.1007/s00521-018-3813-6) [//doi.org/10.1007/s00521-018-3813-6](https://doi.org/10.1007/s00521-018-3813-6)
- 16 F. Maccarini et al.
- 12. Cella, C.E.: Orchidea: A comprehensive framework for target-based computerassisted dynamic orchestration. Journal of New Music Research (2022). [https:](https://doi.org/10.1080/09298215.2022.2150650) [//doi.org/10.1080/09298215.2022.2150650](https://doi.org/10.1080/09298215.2022.2150650)
- 13. Couturier, L., Bigo, L., Levé, F.: Annotating symbolic texture in piano music: A formal syntax. In: Sound and Music Computing Conference (SMC 2022) (2022)
- 14. Crestel, L., Esling, P.: Live Orchestral Piano, a system for real-time orchestral music generation. In: Sound and Music Computing Conference (SMC 2017). p. 434 (2017), <https://hal.archives-ouvertes.fr/hal-01577463>
- 15. Crestel, L., Esling, P., Heng, L., McAdams, S.: A database linking piano and orchestral midi scores with application to automatic projective orchestration. In: International Society for Music Information Retrieval Conference (ISMIR 2017) (2017)
- 16. Déguernel, K., Giraud, M., Groult, R., Gulluni, S.: Personalizing AI for Co-Creative Music Composition from Melody to Structure. In: Sound and Music Computing (SMC 2022). pp. 314–321. Sound and Music Computing (SMC 2022), Saint-Étienne, France (2022). <https://doi.org/10.5281/zenodo.6573287>
- 17. Dong, H.W., Donahue, C., Berg-Kirkpatrick, T., McAuley, J.: Towards Automatic Instrumentation by Learning to Separate Parts in Symbolic Multitrack Music. In: International Society for Music Information Retrieval Conference (ISMIR 2021) (2021). <https://doi.org/doi.org/10.5281/zenodo.5624447>
- 18. Dunsby, J.: Considerations of texture. Music & Letters 70(1), 46–57 (1989)
- 19. Esling, P., Carpentier, G., Agon, C.: Dynamic musical orchestration using genetic algorithms and a spectro-temporal description of musical instruments. In: European Conference on the Applications of Evolutionary Computation (EvoCOP 2010). pp. 371–380 (2010)
- 20. Esling, P., Devis, N.: Creativity in the era of artificial intelligence. arXiv:2008.05959 (2020)
- 21. Fernández, J.D., Vico, F.: AI methods in algorithmic composition: A comprehensive survey. J. Artificial Intell. Res. $48(1)$, 513–582 (Oct 2013)
- 22. Fernández, J.M., Köppel, T., Lorieux, G., Vert, A., Spiesser, P.: GeKiPe, a gesturebased interface for audiovisual performance. In: New Interfaces for Musical Expression Conference (NIME 2017). pp. 450–455 (2017)
- 23. Forsyth, C.: Orchestration. Courier Corporation (1914 (first ed), 1935)
- 24. Ghisi, D.: Music across music: towards a corpus-based, interactive computer-aided composition. Ph.D. thesis, Paris 6 (2017)
- 25. Gotham, M.R.H., Song, K., Böhlefeld, N., Elgammal, A.: Beethoven X: Es könnte sein! (It could be!). In: Conference on AI Music Creativity (AIMC 2022) (2022). <https://doi.org/10.5281/zenodo.7088335>
- 26. Handelman, E., Sigler, A., Donna, D.: Automatic orchestration for automatic composition. In: Artificial Intelligence and Interactive Digital Entertainment Conference (AIIDE 2012) (2012)
- 27. Herremans, D., Chuan, C.H., Chew, E.: A Functional Taxonomy of Music Generation Systems. ACM Computing Surveys $50(5)$, $69:1-69:30$ (2017-09-26). [https:](https://doi.org/10.1145/3108242) [//doi.org/10.1145/3108242](https://doi.org/10.1145/3108242)
- 28. Huang, C.Z.A., Koops, H.V., Newton-Rex, E., Dinculescu, M., Cai, C.J.: AI song contest: Human-AI co-creation in songwriting. In: International Society for Music Information Retrieval Conference (ISMIR 2020) (2020)
- 29. Huang, J.L., Chiu, S.C., Shan, M.K.: Towards an automatic music arrangement framework using score reduction. ACM Transactions on Multimedia Computing, Communications, and Applications $8(1)$, 8:1-8:23 (2012). [https://doi.org/10.](https://doi.org/10.1145/2071396.2071404) [1145/2071396.2071404](https://doi.org/10.1145/2071396.2071404)
- 30. Huron, D.: Characterizing musical textures. In: International Computer Music Conference (ICMC 1989). pp. 131–134 (1989)
- 31. Ji, S., Luo, J., Yang, X.: A comprehensive survey on deep music generation: Multi-level representations, algorithms, evaluations, and future directions. arXiv:2011.06801 (2020)
- 32. Johnson, D.P., Tyson, A., Winter, R.: The Beethoven sketchbooks: History, reconstruction, inventory. Univ of California Press (1985)
- 33. Jordanous, A.: A standardised procedure for evaluating creative systems: Computational creativity evaluation based on what it is to be creative. Cognitive Computation 4(3), 246–279 (2012)
- 34. Jordanous, A.: Has computational creativity successfully made it "beyond the fence" in musical theatre? Connection Science 29, 350–386 (10 2017). [https:](https://doi.org/10.1080/09540091.2017.1345857) [//doi.org/10.1080/09540091.2017.1345857](https://doi.org/10.1080/09540091.2017.1345857)
- 35. Kantosalo, A., Jordanous, A.: Role-based perceptions of computer participants in human-computer co-creativity. In: AISB Symposium of Computational Creativity (CC@AISB 2020) (2020)
- 36. Kendall, R.A., Carterette, E.C.: Identification and blend of timbres as a basis for orchestration. Contemporary Music Review 9(1-2), 51–67 (1993)
- 37. Koechlin, C.: Traité de l'orchestration. Max Eschig (1941 (completed), 1954-1959 (posthumous ed))
- 38. Le, D.V.T., Giraud, M., Levé, F., Maccarini, F.: A corpus describing orchestral texture in first movements of classical and early-romantic symphonies. In: Digital Libraries for Musicology (DLfM 2022). pp. 22–35 (2022)
- 39. Liu, J., Dong, Y., Cheng, Z., Zhang, X., Li, X., Yu, F., Sun, M.: Symphony generation with permutation invariant language model. In: International Society for Music Information Retrieval Conference (ISMIR 2022) (2022). [https:](https://doi.org/10.48550/arXiv.2205.05448) [//doi.org/10.48550/arXiv.2205.05448](https://doi.org/10.48550/arXiv.2205.05448)
- 40. Louie, R., Coenen, A., Huang, C.Z., Terry, M., Cai, C.J.: Novice-AI music cocreation via AI-steering tools for deep generative models. In: Conference on Human Factors in Computing Systems (CHI 2020). pp. 1–13 (2020)
- 41. Lubart, T.: How can computers be partners in the creative process: classification and commentary on the special issue. International Journal of Human-Computer Studies 63(4-5), 365–369 (2005)
- 42. Makris, D., Karydis, I., Cambouropoulos, E.: VISA3: Refining the voice intergration/segregation algorithm. In: Sound and music computing conference (SMC 2016) (2016)
- 43. Mateja, D., Heinzl, A.: Towards machine learning as an enabler of computational creativity. IEEE Transactions on Artificial Intelligence 2(6), 460–475 (2021). <https://doi.org/10.1109/TAI.2021.3100456>
- 44. McAdams, S., Goodchild, M., Soden, K.: A taxonomy of orchestral grouping effects derived from principles of auditory perception. Music Theory Online 28(3) (2022)
- 45. McCormack, J., Cruz Gambardella, C., Rajcic, N., Krol, S.J., Llano, M.T., Yang, M.: Is Writing Prompts Really Making Art? In: Johnson, C., Rodríguez-Fernández, N., Rebelo, S.M. (eds.) Artificial Intelligence in Music, Sound, Art and Design. pp. 196–211. Lecture Notes in Computer Science, Springer Nature Switzerland, Cham (2023). https://doi.org/10.1007/978-3-031-29956-8_13
- 46. McCormack, J., Gifford, T., Hutchings, P.: Autonomy, authenticity, authorship and intention in computer generated art. In: International Conference on Computational Intelligence in Music, Sound, Art and Design (EvoMUSART 2019). pp. 35–50 (2019)
- 18 F. Maccarini et al.
- 47. McKay, G.F.: Creative Orchestration. Allyn and Bacon (1963)
- 48. Miller, A.I.: The Artist in the Machine: The World of AI-Powered Creativity. The MIT Press (10 2019). <https://doi.org/10.7551/mitpress/11585.001.0001>
- 49. Miron, M., Carabias-Orti, J.J., Bosch, J.J., Gómez, E., Janer, J.: Score-informed source separation for multichannel orchestral recordings. Journal of Electrical and Computer Engineering (2016)
- 50. Miron, M., Carabias-Orti, J.J., Janer, J.: Audio-to-score alignment at the note level for orchestral recordings. In: International Society for Music Information Retrieval Conference (ISMIR 2014). pp. 125–130 (2014). [https://doi.org/10.5281/](https://doi.org/10.5281/zenodo.1416150) [zenodo.1416150](https://doi.org/10.5281/zenodo.1416150)
- 51. Nattiez, J.J.: Fondements d'une Sémiologie de La Musique. Dufrenne (1975)
- 52. Nordgren, Q.R.: A Measure of Textural Patterns and Strengths. Journal of Music Theory 4(1), 19–31 (1960). <https://doi.org/10.2307/843045>
- 53. Pachet, F.: A joyful ode to automatic orchestration. ACM Transactions on Intelligent Systems and Technology $8(2)$, 18:1-18:13 (2016-10-03). [https://doi.org/](https://doi.org/10.1145/2897738) [10.1145/2897738](https://doi.org/10.1145/2897738)
- 54. Parmentier, A., Déguernel, K., Frei, C.: A modular tool for automatic soundpainting query recognition and music generation in Max/MSP. In: Sound and Music Computing Conference (SMC 2021) (2021)
- 55. Piston, W.: Orchestration. Norton (1955)
- 56. Reybrouck, M.M.: Musical creativity between symbolic modelling and perceptual constraints: The role of adaptive behaviour and epistemic autonomy. In: Musical Creativity, pp. 58–76. Psychology Press (2006)
- 57. Rimsky-Korsakov, N.: Principles of Orchestration: With Musical Examples Drawn from His Own Works, vol. 1. Édition russe de musique (1873 (begun), 1912 (posthumous ed))
- 58. Rosselló, L.B., Bersini, H.: Music Generation with Multiple Ant Colonies Interacting on Multilayer Graphs. In: Johnson, C., Rodríguez-Fernández, N., Rebelo, S.M. (eds.) Artificial Intelligence in Music, Sound, Art and Design. pp. 34–49. Lecture Notes in Computer Science, Springer Nature Switzerland, Cham (2023). https://doi.org/10.1007/978-3-031-29956-8_3
- 59. Sallis, F.: Music sketches. Cambridge University Press (2015)
- 60. Sandell, G.J.: Roles for spectral centroid and other factors in determining "blended" instrument pairings in orchestration. Music Perception 13(2), 209–246 (1995)
- 61. Sarmento, P., Kumar, A., Chen, Y.H., Carr, CJ., Zukowski, Z., Barthet, M.: GTR-CTRL: Instrument and Genre Conditioning for Guitar-Focused Music Generation with Transformers. In: Johnson, C., Rodríguez-Fernández, N., Rebelo, S.M. (eds.) Artificial Intelligence in Music, Sound, Art and Design. pp. 260–275. Lecture Notes in Computer Science, Springer Nature Switzerland, Cham (2023). [https://doi.](https://doi.org/10.1007/978-3-031-29956-8_17) [org/10.1007/978-3-031-29956-8_17](https://doi.org/10.1007/978-3-031-29956-8_17)
- 62. Schnittke, A.: Timbral relationships and their functional use. Orchestration: An anthology of writings pp. 162–175 (2006)
- 63. de Sousa, D.M.: Textural design: A Compositional Theory for the Organization of Musical Texture. Ph.D. thesis, Universidade Federal do Rio de Janeiro (2019)
- 64. Takamori, H., Sato, H., Nakatsuka, T., Morishima, S.: Automatic arranging musical score for piano using important musical elements. In: Sound and Music Computing Conference (SMC 2017). pp. 35–41 (2017)
- 65. Vechtomova, O., Sahu, G.: LyricJam Sonic: A Generative System for Real-Time Composition and Musical Improvisation. In: Johnson, C., Rodríguez-Fernández, N., Rebelo, S.M. (eds.) Artificial Intelligence in Music, Sound, Art and Design. pp.

292–307. Lecture Notes in Computer Science, Springer Nature Switzerland, Cham (2023). https://doi.org/10.1007/978-3-031-29956-8_19

66. Zembylas, T., Niederauer, M.: Composing processes and artistic agency: Tacit knowledge in composing. Routledge (2017)