DYCI2 agents: merging the ”free”, ”reactive”, and ”scenario-based” music generation paradigms

Jérôme Nika, Ken Déguernel, Axel Chemla–Romeu-Santos, Emmanuel Vincent, Gérard Assayag

DYCI2 agents: merging the "free", "reactive", and "scenario-based" music generation paradigms

Jérôme Nika[1], **Ken Déguernel**[2,1], **Axel Chemla–Romeu-Santos**[1,3], **Emmanuel Vincent**[2], **Gérard Assayag**[1]
[1] IRCAM STMS Lab (CNRS, UPMC, Sorbonne Universités), France - `firstname.lastname@ircam.fr`,
[2] Inria, F-54600 Villers-lès-Nancy, France - `firstname.lastname@inria.fr`,
[3]LIM, Dipartimento di Informatica, Università Degli Studie di Milano, Italy

ABSTRACT

The collaborative research and development project DYCI2, Creative Dynamics of Improvised Interaction, focuses on conceiving, adapting, and bringing into play efficient models of artificial listening, learning, interaction, and generation of musical contents. It aims at developing creative and autonomous digital musical agents able to take part in various human projects in an interactive and artistically credible way; and, in the end, at contributing to the perceptive and communicational skills of embedded artificial intelligence. The concerned areas are live performance, production, pedagogy, and active listening. This paper gives an overview focusing on one of the three main research issues of this project: conceiving multi-agent architectures and models of knowledge and decision in order to explore scenarios of music co-improvisation involving human and digital agents. The objective is to merge the usually exclusive "free", "reactive", and "scenario-based" paradigms in interactive music generation to adapt to a wide range of musical contexts involving hybrid temporality and multimodal interactions.

1. INTRODUCTION

DYCI2 is a collaborative research and development project funded by the French National Research Agency (ANR). [1] This project explores the creative dynamics of improvised interactions between human and artificial agents, featuring an informed artificial listening scheme, a musical structure discovery and learning scheme, and a generalized interaction / knowledge / decision dynamics scheme. The project links together three main research issues: informed listening aiming at analyzing an auditory scene to extrapolate its musical structure by exploiting observed regularities and prior knowledge; adaptive learning of musical structures aiming at combining sequence modeling and probabilistic approaches to take better account of the complexity of the musical discourse from a limited amount of data; and dynamics of improvised interactions aiming at conceiving multi-agent architectures and models of knowledge and decision in order to bring into play different scenarios of co-improvisation involving human and digital agents. We focus here on this last point, and more particularly on the design of listening mechanisms, learning schemes and music generation processes of the creative agents. [2]

The creative agents we develop, as well as other related projects such as PyOracle [3] or Mimi4x [4], belong to a family of related research and implementations on machine improvisation. They share a sequential model, that we call here "memory", learned from live or offline music streams, and that is explored interactively at performance time with various types and degrees of constraints. They follow on the work on "style modeling" initiated in [5] and developed in [6], and its implementation in the real-time improvisation system Omax [7, 8] dedicated to non-idiomatic and non-pulsed improvisation.

Section 2 presents the first achievement of the project, that is introducing control and guidance of the improvisation generated by the agents with two different strategies: reactive listening and anticipatory behavior regarding a temporal specification (or "scenario"). Section 3 develops the current dynamics of the project focusing on the versatility of the agents and their capacity to adapt to varying musical contexts and types of prior knowledge. Finally, Section 4 outlines the fact that the development of DYCI2 models and prototypes is deeply rooted in numerous creative projects [3] with renowned musicians, and gives an overview of the artistic collaborations, residencies, and performances carried out since the beginning of the project.

2. INTRODUCING CONTROL, GUIDANCE, AND AUTHORING IN CREATIVE MUSICAL AGENTS

2.1 Guiding: "Follow my steps" / "Follow that way"

This section presents the music generation strategies associated to two models of creative agents navigating through an online or offline musical memory (*e.g.* audio or MIDI sequences) to generate music improvisations. They cope with different musical settings and prior knowledge on the

[1] ANR-14-CE24-0002-01
DYC2, Creative Dynamics of Improvised Interaction
See `http://repmus.ircam.fr/dyci2/home`.

[2] The reader interested in the other aspects of the project, such as *informed listening* mentioned in introduction, should refer to the related publications *e.g.* [1, 2].

[3] Sound examples, videos of performances and work sessions with musicians can be found online:
`http://repmus.ircam.fr/dyci2/icmc2017`.

musical context: reactive listening (subsection 2.2) and anticipatory behavior regarding a temporal "scenario" (subsection 2.3). These orthogonal strategies respectively focus on the different meanings that "*guiding* music generation" can take in human-computer interaction: "follow my steps" or "follow that way". On the one hand, *guiding* can be seen as a purely reactive and step-by-step process [3, 9, 10, 11]. On the other hand, *guiding* can mean defining upstream temporal structures or descriptions driving the generation process of a whole music sequence. Some studies aim at generating new sequences favoring transitions or subsequences that are characteristic of a chosen corpus [12, 13, 14], at enforcing a formalized temporal structure in a music generation process [15, 16], or at running generative models using temporal queries [17, 18] (see [19] for an exhaustive taxonomy of the related work dedicated to guided human-computer music improvisation). The milestones of the project described in the two following subsections[4] serve as a basis for the ongoing research and development presented later in Section 3.

2.2 Somax: Guiding Music Generation with Reactive Listening

Local events or changes in musical situations occurring during performance have to be handled to address free or generative improvisation, or even to get closer to the real interactions between musicians in a more structured context. Therefore, the first model emphasizes reactivity to deal with human-computer co-improvisation in unknown musical contexts.

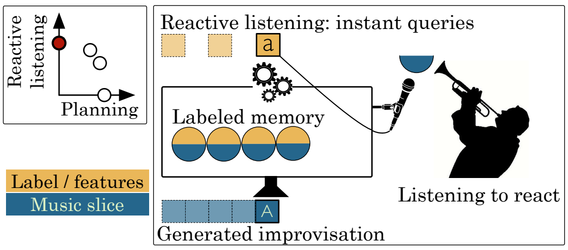

Figure 1. Somax: music generation guided by reactive listening.

Somax [20], an offspring to Omax developped by Laurent Bonnasse-Gahot as a post-doctoral study, extends the paradigm of automatic accompaniment by using purely reactive mechanisms without prior knowledge, and listens to its human or digital co-improvisers (Figure 1). Music generation is guided by an *activity profile*: a continuous score function representing the relevance of each state of the musical memory regarding the current musical environment, i.e. the similarity with regards to several musical dimensions such as melody and harmony. This activity profile evolves over time, and propagates with a forgetting behavior which introduces a mechanism of short-term cognitive remanence. A global score function is computed from the real-time analysis of the musical context by taking a weighted sum of the sub-scores corresponding to several *stream views* into accounts. This multimodal analysis of the live musical inputs drives the generation process and favors events of the memory sharing a common past with the current improvisation regarding the different musical dimensions while keeping pace with the evolving external sonic context by preferring high profil states whenever possible. The event-driven generation process is jointly steered by the reactive listening mechanisms and the internal logic of the musical memory. The continuity of the generated improvisation is insured by re-injecting the outputs of the system among the inputs processed by the selection routine (*self-listening*).

The first prototype is compatible with MIDI and audio data, and listens to pitch, harmony and rhythm in the musical environment in order to guide music generation. A new upcoming version extends this concept by allowing users to build custom architectures for listening and reacting, and adds time management routines based on hierarchical representations. It is modular, and open to extensions thanks to the genericity of its implementation as a Python library with OSC outputs, and offers easier definitions of new alphabets and musical dimensions, and the possibility to model the musical memory with different data structures.

2.3 ImproteK: Guiding Music Generation with a Temporal Scenario

Taking advantage of prior knowledge of the temporal structure is crucial if we want models that can address composed or idiomatic music [21] (jazz, rock, blues, pop, etc.). A chord *progression*, for example, is oriented *t*owards its future and carries more than the step-by-step conformity of a *succession*. The sequentiality of a progression must therefore be exploited to introduce *forward motion* at each step of the generation (trivial examples being harmonic anticipation, "going from the tonic to the dominant", "going from the tonic to the tonic", etc.).

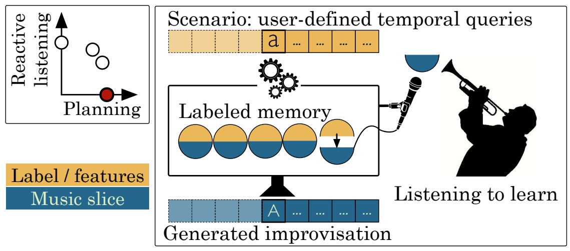

Figure 2. ImproteK: music generation guided by a temporal scenario.

ImproteK [22] addresses structured or composed improvisation. This project was initiated by Marc Chemillier in order to combine the "style modeling" approaches mentioned in introduction and the ability to cope with a beat and long-term constraints. The software introduces a temporal specification, or *scenario*, in the music generation process (Figure 2). Depending on the musical context, this temporal structure can for example represent the harmonic progression of a jazz standard, a profile of audio features, or any formalized structure defined on a user-defined alphabet. In a first approach, the consistency between the scenario and the memory ensures the conformity of the successive elements retrieved from the memory regarding the required temporal evolution. In a second approach, the scenario gives access to prior knowledge of the

[4] Both prototypes can be found online:
`https://forge.ircam.fr/p/Dyci2/`.

temporal structure of the improvisation which is exploited to introduce anticipatory behavior in the generation process. This way, the future of the scenario is taken into account when generating the current time of the improvisation. The scenario / memory approach is independent of the chosen musical alphabet, the musicians can therefore be involved in a meta-level of composition by designing scenarios as well as alphabets and their properties.

Recent developments include the ability to launch several instances of the system simultaneously, learning from different musical sources. In addition, the system was used as an application case to design the new scheduling engine of the OpenMusic environment [23] and to implement generative agents that can be integrated into a "meta-score" (new version of the OpenMusic "maquette" [24]). [5]

3. INTRODUCING HYBRID GUIDANCE, COMPOSITIONAL FORESIGHT, AND MULTIDIMENSIONALITY

This section develops the motivations and the ongoing research focusing on the versatility of the digital musical agents and their capacity to adapt to different musical settings, with or without prior knowledge about the musical context. The main goal is to merge the usually exclusive "free", "reactive", and "scenario-based" paradigms in interactive music generation to adapt to a wide range of contexts of human-computer music improvisation, capitalizing on the models described in the previous section. This requires to design a range of strategies going from a floating musical coordination on different musical dimensions (*e.g.* energy, harmony, melody, timbre) with live musicians and / or other digital agents, to the conformity to a predefined structure, through the discovery and inference of underlying structures dynamically feeding generation models.

3.1 Hybrid and Adaptive Temporal Guidance of Generative Models

The dialectic between planning and reaction raised by the previous section has been approached from a wide range of disciplines, including psychology, cognitive science, and neuroscience. Among them, action planning [25] studies the interaction between perceived events and planned actions when a gesture is prepared according to a given execution plan and has to be conciliated with new data from the environment. These interactions observed in a system of two agents (called joint actions in psychology) have been applied to music, and in particular regarding synchronization [26]. The forward models or efference copy models, coming from the field of motor control and now widely used in the field of speech production [27], are closer to our issue: an internal strategy produces an anticipation of the output of the system (efference copy) and is compared then to the actual output. A reaction is then seen as a revision of the strategy or of the output of the system.

In a collective improvisation, the conformity and / or forward motion in regard to a scenario are preserved and exploited when facing "unpredictable" events. To illustrate this point, Bonnerave [28] gives the example of a be-bop bass player who plays over an indicative chord progression in a different way every time, taking into account the rhythmic calls of the drummers, the accents of the pianist, etc. Studying the combination of planning and reactivity in music improvisation through the notion of "pré-voyance" introduced by Bourdieu [29], Bonnerave considers that a soloist or accompanist improviser senses interaction in its immediacy in order to bring out possibilities of prolongations which are present in a latent state. This concept of future seen as possible continuations existing in present time and selected depending on interaction underlines the link between reaction and anticipation in improvisation. This idea is closely related to that of Lortat-Jacob [30], for whom improvisers are involved in a permanent process of anticipation. Drawing insights from various music theorists to study improvisation in the social and political life, Citton [31] even considers that anticipation is an integral part of improvisation. Indeed, he explains that improvisation in our daily actions results from a certain mix of *on-the-spot improvisation* "thanks to which I can integrate the unavoidable novelty encountered in each situation", and *compositional foresight* "thanks to which I attempt to anticipate the future effects of my current behavior according to my rational understanding of the laws of nature".

We aim at designing models and architectures that can take part in collective improvisations. The agents have therefore to be capable of mid-term planning as well as reactivity to the environment, such as live players inputs. In this context, reactivity has to take advantage of (or has to be compatible with) an underlying intention or an inferred local underlying structure to benefit from a *compositional foresight* to be able to preserve forward motions without predefined temporal structure. This high level approach incites to define generic and extensible mechanisms of temporal guidance to be able to cope with prior knowledge of a structure, the inference of short-term intentions, and / or a reactivity at the event level.

3.1.1 Combining Scenario and Reactive Listening

A preliminary work [32] sketched an architecture guiding the machine improvisation along two different musical dimensions by associating an anticipation module based on the models and architectures presented in subsection 2.3, and a reactive listening module inspired by subsection 2.2. In this framework, for example, a scenario can be defined as a chord progression, and reactive listening can be performed on energy descriptors. Then, during the performance, different processes associated to the different modules (anticipation regarding the scenario, reactive listening, self-listening) continuously modify the activity of the musical memory to retrieve and play the most relevant events on the fly.

A Python system based on the scheduling architecture proposed in [33] is currently being developed to combine dynamic controls and anticipations relative to a predefined or dynamic plan (Figure 3). An *Improvisation Handler* agent [6] embeds music generation models and reacts to dynamic controls by rewriting previously generated anticipations ahead of performance time while maintaining

[5] See "additional resources" at `http://repmus.ircam.fr/dyci2/icmc2017`.

[6] See "additional resources" at `http://repmus.ircam.fr/dyci2/icmc2017`.

consistency when overlaps occur. It introduces a formalism to translate the events coming from reactive listening as well as long-term temporal specifications into music generation queries with different temporal scopes sent to the generative models. This agent handles the pool of time-triggered and event-triggered queries, handles their concurrency and their access to shared data (memory, generation parameters, and execution trace) and manages the reception of new queries while previous ones are still being processed by the generation modules. The queries are dynamically killed, merged, and / or reordered, and launched in due time to produce short-term anticipations satisfying the hybrid temporal guidance.

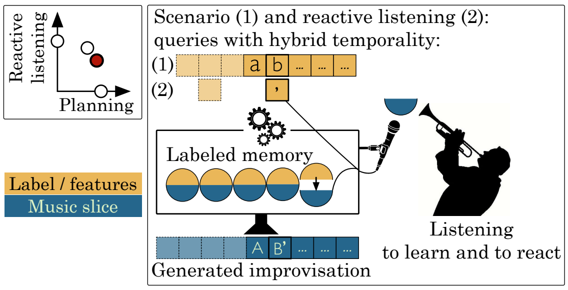

Figure 3. Combining long-term planning and reactive listening.

3.1.2 *Inference of Short-term Scenarios for the Future*

Another direction is that of the inference of short-term scenarios (Figure 4). The models proposed in subsection 2.3 are queried by successive "partial scenarios", i.e. subsequences of a scenario *defined before* the performance. The idea is to enrich the reactive listening introduced in subsection 2.2 so that it discovers underlying structures in what the musician plays to infer possible continuations of this underlying structure. This model inferring short-term scenarios for the future (*e.g.* sequence of chroma, chord progression) to feed a scenario-based music generation engine in real-time will enable the agents to generate anticipative improvisations from an inferred (and not only predefined) underlying structure.

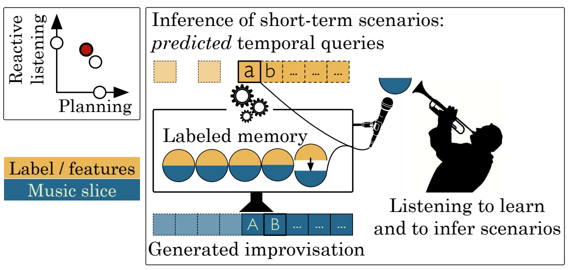

Figure 4. Listening to infer short-term scenarios for the future.

Such a process will take advantage of the anticipatory behavior proposed by the models presented in subsection 2.3 without informing the agent of any prior knowledge regarding the temporal structure of the improvisation. This way, the chain "*specification* $\rightarrow$ *anticipation*" implemented in the agents will extend into the general chain of musical anticipation "*expectation* $\rightarrow$ *prediction* $\rightarrow$ *anticipation*" [34]. These aspects are being investigated through the prism of deep learning by trying to build bridges between the models described in the previous section and a generative graphical probabilistic model currently applied to automated orchestral generation [35]. Our approach is to automatically extract knowledge and temporal structures from series of high-level annotations [36] using recent advances in semi-supervised learning [37].

3.2 Multi-dimension and Multi-level Approach

Most current automatic music improvisation systems focus on generating melodies by modeling a style from one-dimensional sequences (usually sequences of pitches). However, when improvising, musicians use information from several dimensions (pitch, harmony, rhythm...) to guide and structure their improvisation. The development of their improvisations is based on intuition over the local context enriched with background knowledge gathered from all their musical experience [38]. Moreover, in a situation of collective improvisation, each musician is also influenced by the style of the other musicians and their expectations of what they are going to do. Therefore, a generative system has to take the multidimensional aspect of music into account and also to emulate a collective improvisation situation. On top of that, when improvising in an idiomatic context (such as a jazz standard), a musician also uses the hierarchical aspect of the chord progression to properly construct their improvisation upon multiple level of organization. We want to be able to exploit this multi-level structure of a chord progression in order to enrich the possibilities of music generation.

First, we designed a system able to train over multidimensional sequences to generate improvisation [39]. This system combines:

- interpolated probabilistic models [40] representing the multidimensional aspect of music which are trained on a large corpus. These models can be n-grams on one dimension, or conditional models such as "which melody can be played when this chord is played?". They represent the multidimensional knowledge of the system and can profit from advanced smoothing and optimization techniques [41];
- a model melodic memory with links to places sharing a similar musical context trained on a smaller corpus or with a live musician's playing.

During the generation process, the multidimensional probabilistic models guide the navigation through the memory; what is generated follows the contextual logic of the improvisation while enriching its musical discourse with multidimensional knowledge. This way, we can also create hybrid improvisations mixing different corpora for the multidimensional knowledge and for the memory.

We then designed a system creating multidimensional improvisations based on interactivity between dimensions via message passing through a cluster graph [42], each dimension being represented by the model described above. The communication infers some anticipatory behavior on each dimension now influenced by the others, creating a consistent multidimensional improvisation. These systems were evaluated by professional improvisers during listening sessions and received good feedback, first regarding how a multidimensional knowledge can help performing better navigation in the memory and second how communication through message passing can emulate the interactivity between musicians.

Finally, to consider the multi-level organization of chord progressions in the particular case of idiomatic improvisation, we used a hierarchical grammar [43] to represent an analysis of chord progressions revealing chord functions and theme sections. The multi-level structure of this grammar is then exploited to guide the music generation in a way inspired by the strategy described in subsection 2.3, using a multi-level scenario using the nature of each chord as well as its function and its role in higher level hierarchical structure. This method ensures the consistency of the improvisation regarding the global form of the tune, and shows how the knowledge of a corpus of chord progressions sharing the same hierarchical organization can extend the possibilities of music generation [44]. [7] Future work will consist in using variable length sequence modeling to automatically infer this multi-level structure from a corpus.

4. LET THE MUSIC(IANS) PL(/S)AY

The DYCI2 project is led in close and continuous interaction with expert musicians. The associated artistic residencies, studio sessions, and public performances enable to focus on targeted scientific issues by exploring different idioms and repertoires: Bernard Lubat and "la Compagnie Lubat" ("jazzed-up songs", jazz, scat, and free improvisation), Jovino Santos Neto (Brazilian music and jazz), Kilema (marovany zither), Charles Kely (marovany zither and jazz), Louis Mazetier (stride piano), Michelle Agnès Magalhaes (composed contemporary improvisation), Rémi Fox (jazz, funk, and generative improvisation), Hervé Sellin and Georges Bloch (jazz), Benoît Delbecq, Doctor Bone a.k.a. Ashley Slater, Jozef Dumoulin (jazz, electronic music, generative improvisation), Marta Gentilucci (contemporary music), Pascal Mabit (jazz, generative improvisation), and Rodolphe Burger (songs, indie rock, blues). More than 15 demonstrations, public performances, and concerts using the successive versions of the prototypes were performed since 2015, among them: *Secret Heroes* with Benoît Delbecq, Jozef Dumoulin, Doctor Bone a.k.a Ashley Slater, and Gilbert Nouno, Ircam Manifeste Festival, Grande Salle, Centre Pompidou, Paris, June 22, 2016; Bernard Lubat, concert at Collège de France, Paris, May 27, 2016; *Edith Piaf, Elizabeth Schwarzkopf, Billie Holiday, tre donne del 1915*, multimedia performance from and by Hervé Sellin and Georges Bloch, Festival interazionale "Pietre che cantano", L'Aquila, August 21, 2015; Rémi Fox, performance at Montreux Jazz festival, July 17, 2015.

Beyond validation, these collaborations are intended to move the topic from "simulation" to "stimulation", exploiting the successive versions of the models, seeking then to divert them and perpetuate these diversions in order to participate in the creative processes of the musicians. These interactions are an integral part of the iterative development of the models and of the software prototypes: the assessments from these experts of the concerned idioms are gathered to validate and / or refine the incremental scientific and technical steps, and to motivate the future developments. The collaborations we carried out so far led to hours of filmed and documented music sessions, listening sessions, and interviews. [8]

The numerous ongoing musical collaborations experiment guided improvisation in various musical contexts, from improvisation on jazz standards and composed improvisation to generative and free improvisation. Among them, the results of the project aim at being used in a creative project in association with EPFL MetaMedia Center and Montreux Jazz Festival around Montreux Digital Heritage archive: a huge collection of 5000 hours of concert in multitrack audio and video, listed as UNESCO World Heritage.

[7] See "learning of multi-dimensional structures" at `http://repmus.ircam.fr/dyci2/icmc2017`.

[8] See "performances with expert musicians" and "qualitative interviews" at `http://repmus.ircam.fr/dyci2/icmc2017`.

Acknowledgments

This work is made with the support of the French National Research Agency, in the framework of the project "DYCI2: Creative Dynamics of Improvised Interaction" (ANR-14-CE24-0002-01). Access to the Montreux Jazz Festival Database provided by courtesy of EPFL Metamedia Center Project Montreux Digital Heritage Archive.

5. REFERENCES

[1] A. A. Nugraha, A. Liutkus, and E. Vincent, "Multichannel audio source separation with deep neural networks," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 10, pp. 1652–1664, Jun. 2016.

[2] D. Di Carlo, K. Déguernel, and A. Liutkus, "Gaussian framework for interference reduction in live recordings," in *Proc. of the Audio Engineering Society Conference on Semantic Audio*, 2017.

[3] G. Surges and S. Dubnov, "Feature Selection and Composition Using PyOracle," in *Proc. of the Artificial Intelligence and Interactive Digital Entertainment Conference*, 2013, pp. 114–121.

[4] A. R. François, I. Schankler, and E. Chew, "Mimi4x: an interactive audio–visual installation for high–level structural improvisation," *International Journal of Arts and Technology*, vol. 6, no. 2, pp. 138–151, 2013.

[5] S. Dubnov, G. Assayag, and R. El-Yaniv, "Universal classification applied to musical sequences," in *Proc. of the International Computer Music Conference*, 1998, pp. 332–340.

[6] G. Assayag and S. Dubnov, "Using factor oracles for machine improvisation," *Soft Computing*, vol. 8, no. 9, pp. 604–610, 2004.

[7] G. Assayag, G. Bloch, M. Chemillier, A. Cont, and S. Dubnov, "OMax brothers: a dynamic topology of agents for improvization learning," in *Proc. of the ACM workshop on audio and music computing for multimedia*, 2006, pp. 125–132.

[8] B. Lévy, G. Bloch, and G. Assayag, "OMaxist Dialectics," in *Proc. of the International Conference on New Interfaces for Musical Expression*, 2012, pp. 137–140.

[9] J. Moreira, P. Roy, and F. Pachet, "Virtualband: interacting with stylistically consistent agents," in *Proc. of the International Society for Music Information Retrieval Conference*, 2013, pp. 341–346.

[10] F. Pachet, P. Roy, J. Moreira, and M. d'Inverno, "Reflexive loopers for solo musical improvisation," in *SIGCHI Conference on Human Factors in Computing Systems*, ACM, Paris, 2013, pp. 2205–2208.

[11] A. Einbond, D. Schwarz, R. Borghesi, and N. Schnell, "Introducing CatOracle: Corpus-based concatenative improvisation with the Audio Oracle algorithm," in *Proc. of the International Computer Music Conference*, 2016, pp. 141–147.

[12] A. Papadopoulos, P. Roy, and F. Pachet, "Avoiding plagiarism in Markov sequence generation," in *Proc. of the AAAI Conference on Artificial Intelligence*, 2014, pp. 2731–2737.

[13] D. Herremans, S. Weisser, K. Sörensen, and D. Conklin, "Generating structured music for bagana using quality metrics based on Markov models," *Expert Systems with Applications*, vol. 42, no. 21, pp. 7424–7435, 2015.

[14] A. Eigenfeldt and P. Pasquier, "Considering vertical and horizontal context in corpus-based generative electronic dance music," in *Proc. of the International Conference on Computational Creativity*, 2013, pp. 72–78.

[15] A. Donzé, R. Valle, I. Akkaya, S. Libkind, S. A. Seshia, and D. Wessel, "Machine improvisation with formal specifications," *Proc. of the International Computer Music Conference*, pp. 1277–1284, 2014.

[16] F. Ghedini, F. Pachet, and P. Roy, "Creating music and texts with flow machines," in *Multidisciplinary Contributions to the Science of Creative Thinking*. Springer, 2016, pp. 325–343.

[17] A. Cont, S. Dubnov, and G. Assayag, "Guidage: A fast audio query guided assemblage," in *Proceedings of International Computer Music Conference*, 2007, pp. 252–259.

[18] C. Wang and S. Dubnov, "Guided music synthesis with variable markov oracle," *Proc. of the Artificial Intelligence and Interactive Digital Entertainment Conference*, pp. 55–62, 2014.

[19] J. Nika, "Guiding human-computer music improvisation: introducing authoring and control with temporal scenarios," Ph.D. dissertation, Université Pierre et Marie Curie Paris 6 - Ircam, 2016.

[20] L. Bonnasse-Gahot, "An update on the SOMax project," Ircam - STMS, Tech. Rep., 2014, Post-doctoral report. [Online]. Available: http://repmus.ircam.fr/_media/dyci2/somax_project_lbg_2014.pdf

[21] D. Bailey, *Improvisation: its nature and practice in music*. Da Capo Press, 1993.

[22] J. Nika, M. Chemillier, and G. Assayag, "Improtek: introducing scenarios into human-computer music improvisation," *ACM Computers in Entertainment*, vol. 14, no. 2, pp. 1–27, 2017.

[23] D. Bouche, "Processus compositionnels interactifs : Une architecture pour la programmation et l'exécution des structures musicales," Ph.D. dissertation, Université Pierre et Marie Curie, Paris, 2016.

[24] 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.

[25] W. Prinz, "Perception and action planning," *European Journal of Cognitive Psychology*, vol. 9, no. 2, pp. 129–154, 1997.

[26] P. E. Keller, "Joint action in music performance," *Emerging Communication*, vol. 10, p. 205, 2008.

[27] G. Hickok, "Computational neuroanatomy of speech production," *Nature Reviews Neuroscience*, vol. 13, no. 2, pp. 135–145, 2012.

[28] J. Bonnerave, "Improviser ensemble. De l'interaction à l'écologie sonore," *Tracés. Revue de Sciences humaines*, no. 18, pp. 87–103, 2010.

[29] P. Bourdieu, *Esquisse d'une théorie de la pratique, précédée de trois études d'ethnologie kabyle*. Le Seuil, 1972.

[30] B. Lortat-Jacob, "Formes et conditions de l'improvisation dans les musiques de tradition orale," in *Musiques, une encyclopédie pour le XXIe siècle. Volume 5: L'unité de la musique*, J.-J. Nattiez, Ed. Actes Sud / Cité de la musique, 2007.

[31] Y. Citton, "Politics as Hypergestural Improvisation in the Age of Mediocracy," in *The Oxford Handbook of Critical Improvisation Studies*, G. E. Lewis and B. Piekut, Eds. Oxford: Oxford University Press, 2013.

[32] A. Chemla-Romeu-Santos, "Guidages de l'improvisation," Master's thesis, Master ATIAM - Ircam, UPMC, 2015.

[33] J. Nika, D. Bouche, J. Bresson, M. Chemillier, and G. Assayag, "Guided improvisation as dynamic calls to an offline model," in *Proc. of the Sound and Music Computing conference*, 2015, pp. 427–434.

[34] D. B. Huron, *Sweet anticipation: Music and the psychology of expectation*. MIT Press, 2006.

[35] L. Crestel and P. Esling, "*Live Orchestral Piano*, a system for real-time orchestral music generation," in *Proc. of the Sound and Music Computing Conference*, 2017, pp. 434–442.

[36] P. Esling and C. Agon, "Time-series data mining," *ACM Computing Surveys*, vol. 45, no. 12, pp. 1–34, 2012.

[37] A. Rasmus, M. Berglund, M. Honkala, H. Valpola, and T. Raiko, "Semi-supervised learning with ladder networks," in *Advances in Neural Information Processing Systems*, 2015, pp. 3546–3554.

[38] M. Crispell, "Elements of improvisation," in *Arcana : Musicians on Music*, J. Zorn, Ed., 2000, pp. 190–192.

[39] K. Déguernel, E. Vincent, and G. Assayag, "Using multidimensional sequences for improvisation in the OMax paradigm," in *Proc. of the Sound and Music Computing Conference*, 2016, pp. 117–122.

[40] S. A. Raczyński, S. Fukayama, and E. Vincent, "Melody harmonisation with interpolated probabilistic models," *Journal of New Music Research*, vol. 42, no. 3, pp. 223–235, 2013.

[41] S. F. Chen and J. Goodman, "An empirical study of smoothing techniques for language modeling," Harvard University, Tech. Rep. TR-10-98, 1998.

[42] D. Koller and N. Friedman, *Probabilistic graphical models : principles and techniques*. MIT Press, 2009.

[43] N. Chomsky, *Studies on semantics in generative grammar*. Walter de Gruyter, 1996, vol. 107.

[44] K. Déguernel, J. Nika, E. Vincent, and G. Assayag, "Generating equivalent chord progressions to enrich guided improvisation : application to Rhythm Changes," in *Proc. of the Sound and Music Computing Conference*, 2017, pp. 399–406.