Improvising in Creative Symbolic Interaction

Gérard Assayag

Improvising in Creative Symbolic Interaction

Gérard Assayag
Ircam STMS Lab (CNRS, UPMC)

This chapter is a revised version of a key-note conference given at the Onassis Cultural Centre in Sept. 2014 in Athens, Greece[1].

ABSTRACT

Creative Symbolic Interaction brings together the advantages from the worlds of interactive real-time computing and intelligent, content-level analysis and processing, in order to enhance and humanize man-machine communication. Performers improvising along with Symbolic Interaction systems experiment a unique artistic situation where they interact with musical (and possibly multi-modal) agents which develop in their own ways while "keeping in style". Symbolic interaction aims at defining a new artificial creativity paradigm in computer music, and extends to other fields as well : The idea to bring together composition and improvisation by combining mathematical symbolic structures and cognitively inspired dynamical processes belongs to a general scheme that is likely to grow in many artistic and non-artistic domains.

1. INTRODUCTION

Until recently, in the field of musical interaction with machines, engineers and researchers have been concerned by fast computer computation and reaction — a logical concern if one considered the complexity of the tasks at stake and the available hardware performance. However, instantaneous response is not always the way a musician reacts in a living performance situation. Although decisions are taken at a precise moment in time, the decision process relies on evaluation of past history, analysis of incoming events and activation of anticipation strategies. Therefore, not only can it take some time to come to a decision, but the outcome of this decision process can also involve postponing action to a later time. This scenario implies time and memory at different scales, just as music composition does, and it cannot be fully apprehended just by conventional digital signal and event based algorithmics. A symbolic representation level, better formalised with algebraic or geometrical tools, has to be involved as well.

In order to foster realistic and artistically interesting behaviors of digital interactive systems, and communicate with them in a humanized way, we wish to combine several means : machine listening [21] — extracting high level features from the signal and turning them into significant symbolic units ; machine learning [12] — discovering and assimilating on the fly intelligent schemes by listening to actual performers ; stylistic simulation [2] — elaborating a consistent model of style ; symbolic music representation — mathematically formalized representations connected to organized musical thinking in analysis and composition [5]. These tools have to cooperate effectively in order to define a multi-

[1] Originally in ICMC-SMC 2014 proc., “Music technology meets philosophy: from digital ethos to virtual ethos”, A. Georgaki and G. Kouroupetroglou editors, September 2014, Athens, Greece.

level memory model, underlying a discovery and learning process that contributes to the emergence of a creative musical behavior.
In the Music Representation Team at Ircam, after OpenMusic, a standard for computer assisted composition [26], we have designed OMax [1, 4, 13], an interactive machine improvisation environment which explores this new interaction schemes. It creates a cooperation between several heterogeneous components specialized in real-time audio signal processing, high level music representations and formal knowledge structures. This environment learns and plays on the fly in live settings. It has been used in many musical and multi-media performances throughout the world. OMax instantiates several trends of current research on interactive computational creativity, with digital agents showing adequacy and relevance by connecting instant contextual listening to corpus based knowledge, along with longer term investigation and decision mechanisms allowing to refer to larger-scale structures and scenarios [24]. We call this general scheme *Symbolic Interaction*.
Creative Symbolic Interaction brings together the advantages from the two worlds of interactive real-time computing and intelligent, content-level analysis and processing, in order to enhance and humanize man-machine communication. Performers improvising along with Symbolic Interaction systems experiment a unique artistic situation where they interact with a musical (and possibly visual) agent which develops itself in its own ways while keeping in style with the user. It aims at defining a new artificial creativity paradigm in computer music, and extends to other fields as well : The idea to bring together composition and improvisation through modeling cognitive structures and processes [15] is a general idea that makes sense in many artistic and non-artistic domains. It is a decision-making paradigm where a tactical process builds up by weaving decisions step after step, either by relating to an overall structural determinism, or by "improvising" a surprising jump to a musical region that breaks the known continuity but has to be, somehow, explainable with regard to the style and the context. This kind of "improvisation" strategy is observed in the living world, and might be one aspect of intelligence as a way to cope effectively with the unknown and produce novelty, thus it may serve as a productive model for artificial music creativity.

2. GENERAL CONTEXT

Improvised interaction between humans and digital agents is a recent field of studies in the context of artificial creativity, which convenes several active research issues: interactive learning practices, whose models are built at the very time of interaction, and whose generated outcomes inflect the very conditions of this interaction ; artificial perception, based on and helped by this interaction ; modeling of social and expressive interaction between human and / or digital agents, in the anthropological, social, linguistic, and IT dimensions.

This type of interaction involves the perception / action loop and engages the learning process in a renewed recursive design where an agent learns, among other sources, from the very reactions of other agents to its own creative productions.
Research on human and artificial creativity in sound and music raises a lot of interest worldwide and has been developing seriously these last years with many technical progresses in artificial listening, epistemic modeling, artificial intelligence, machine learning, signal / physical models and formal algebraic and geometrical representations of music structures. Creativity in general is supported at the european level by one of the recent objectives of the E.U. Commission called "Technology and scientific foundations in the field of creativity" in its actions "Intelligent computational environments and stimulating human creativity" and "Enhancing Progress towards formal understanding of creativity".

3. IMPROVISED INTERACTION

To confront the problem of improvised interaction, seen as a powerful driving force for creativity at the heart of all human activities constitutes one of today's central challenges for digital intelligence in sound and music computing. We envision it in the realm of interactions between physical, digital and human worlds, in a music information dynamics setup [11]. We wish indeed to integrate artificial listening, learning of musical behaviors, temporal modeling of musical structures and dynamical creative interaction in an architecture for effective experimentation in real time, following the succesful Omax research program. Premices of the answers to this challenge are already available in child projects of Omax, such as SoMax and ImproTek [14, 25], the latter a collaboration between Ircam and EHESS[2].

Possible applications will be available in varied flavors on-line and off-line. They are likely to change the situation in the artistic relationship between human and artificial agents. Off-line listening to and learning from large music databases will feed generative systems for composition and performance and ease the addition of creative functionalities to software for multi-media production and post-production, digital games and cinema, access to audiovisual heritage. Integrating listening and learning in the very process of artistic interaction makes it possible to program software agents with the skills to react in real-time to human performance in man-machine improvisation setups, in multimedia installations, in electro-acoustic composition, in the creation of new "variable" formats for music production and distribution.

Architectures integrating concurrent agents and logical constraints [17] capturing the idea of scenario provide opportunities for powerful generativity in situations ranging from popular music to digital games, web applications, generative cinema, by providing a control layer for the interactive generation of new temporal sequences through the progression of the game, the incarnation and the evolution of a character, the behavior of the user.

In this sense, symbolic interaction will disrupt the notion of public access to recorded music, with potentially major impact on cultural heritage and industries, shifting more and more the public from passive to genuine participation.

4. DYNAMICS OF INTERACTION

4.1 Creative Digital Agents

The dynamics of creative improvised interaction that we defend in this article focuses on the creation, adaptation and implementation of effective and efficient models of artificial listening, learning, interaction and automatic creation of musical content in order to allow the formation of digital music agents that succeed in being ayutonomous, creative, able to display artistically credible manners in various artistic and educational human setups such as live performance and teaching. These agents may also help constitute the perceptual and communicative skills of embedded artificial intelligence systems.

The aim is to evolve self-creative agents by the process of interactive learning from direct exposure to human improvisers, thus creating a retro-action loop (we call it *stylistic reinjection*) through the simultaneous exposure of humans to the "improvised" productions of the digital agents themselves. This involves a complex dynamics of time and space evolving human / digital communication.

[2] Ecole des hautes études en sciences sociales

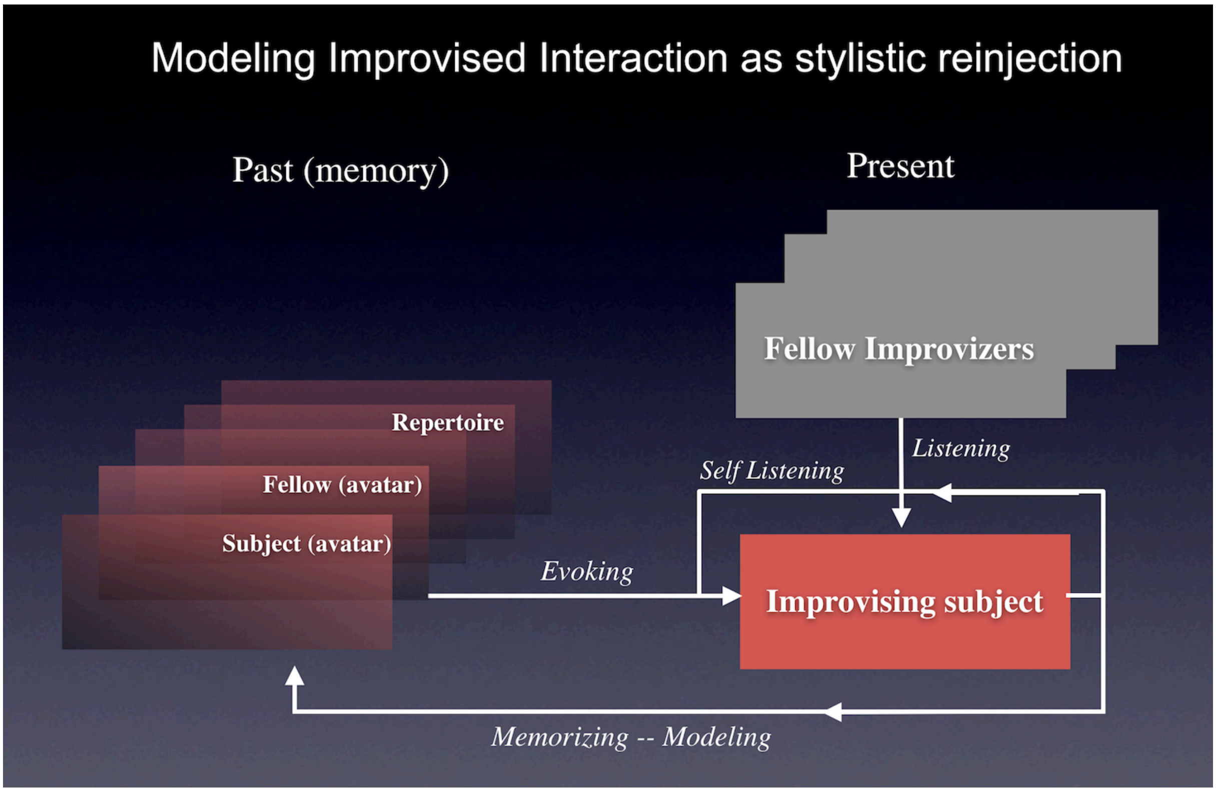

Figure 1. Stylistic Reinjection Paradigm

4.2 Off-line Learning

Even in the case of real-time performance, an offline learning process based on large musical corpora can be anticipated so as to stylistically "color" digital individualities or position the experience within a particular genre (classical, jazz, traditional, electro etc.)

In addition to applying these digital skills to live music situations where digital and human agents interplay, the game can be extended to innovative applications such as interaction of users of all sorts with audio-visual heritage archives which would be dynamically resurrected within a creative or educational framework, thus extending the improvisation paradigm to new narrative and immersive forms in a direction unseen yet. We are currently building a project in this direction with EPFL[3] University in Lausanne around the Montreux Jazz Festival Archives recently listed by Unesco as part of the World Cultural Heritage[4].

4.3 Collectivities of Agents

The idea is to create both an artificial expertise of musical practice ("machine musicianship") by this kind of interaction and a rich experience of instantaneous human – digital communication likely to provide an aesthetic satisfaction to the user, to broaden its sound and music production means, to "talk" with him by imitation or contradiction, and, in general, to stimulate and boost the musical experience individually and collectively. This human-agent interaction will in effect be extended to complex configuration potentially involving a great number of agents — human and artificial — evolving and learning from each other. As the experience will take form, autonomous digital music personalities, able to intervene credibly in complex situations of interaction with humans and other agents, will emerge.

An artificial entity in a creative audio-musical context subsumes itself as a collection of elementary contributory and competitive components, capable of interactive learning, implementing the artificial listening tasks, the discovery of short and long-term temporal structures, the modeling of style, the

[3] Ecole Polytechnique Fédérale de Lausanne
[4] See e.g. the MJF Digital Project at claudenobsfoundation.com

generation of symbolic sequences, the real-time audio management, and the visualization and human-machine interface functions as well.

4.4 The OMax Galaxy

A great part of these capacities are already available in the OMax environment, which has already been used in a number of public performances throughout the world and is established as a well known reference in the realm of *improvised machine musicianship* [5, 6]. Several extensions of the OMax paradigm are currently under research or starting in the author's team (Music Representation) at Ircam as well as in partner labs at UCSD, EHESS, and CNMAT Berkeley. These researches include musical information dynamics [11, 16], mathematics of formal automata [10], probabilistic approaches [8], cognitive modeling of memory [25], accompaniment systems, scenario based improvisation, multi-agent interaction dynamics, temporal adaptation of interaction [22], digital intelligence and artificial creativity.

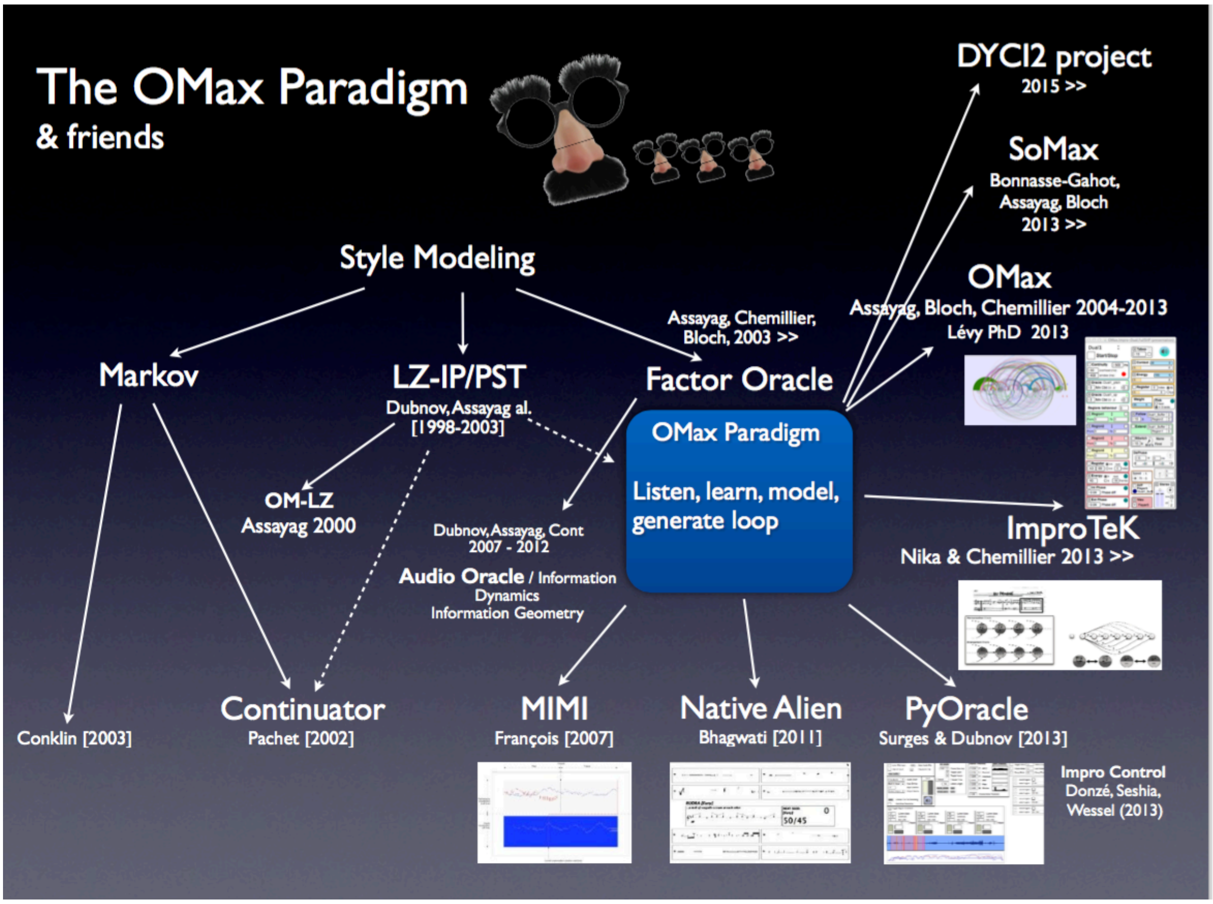

Figure 2. OMax Galaxy

The OMax galaxy of researches connects with active or starting projects in several of these labs with studies in knowledge models and decision-making strategies for synchronous and asynchronous agents caught in collective creative action, elicitation of augmented listening by the use of prior knowledge, co-adaptation of perception and action in an interactive learning environment, integration of multi-dimensional and multi-scale aspects of musical structures.

5. SYMBOLIC INTERACTION

From there on, we wish to move in the direction of a more powerful and versatile instantiation of symbolic interaction. A series of projects involving several PhD and post-doc works and several national and international collaboration have been initiated, in order to stimulate theoretical advances and development of experimental environments in a number of directions.

5.1 Perceptual Skills

It is thus necessary to be able to empower digital agents with the ability to analyze complex auditory scenes in real time and to extract musical structure from them by discovering time and space regularities through new geometrical paradigms[18] and exploiting available prior knowledge through an "informed musical decomposition" mathematical signal processing scheme [19], taking advantage of existing scores, annotations, or inferring partial information from learning beforehand stylistically similar corpora.

5.2 Learning Skills

Interactive learning of musical structures stems from data provided by the sequential process of listening, learning symbolic models that capture high-level multi-dimensional and multi-scale musical structures emerging in a context of musical performance and improvisation. Many formal models may compete and even cooperate for that purpose. The Factor Oracle model [3, 9] coming from automata and formal languages theory has been intensively used in OMax and its derivative with great musical achievements[5]. Hidden Markov Models, Temporal Grammar Rules Induction, Deep Belief Networks, Concurrent Constraint Calculus, Spatial and Epistemic Calculus, Multi-objective Time Series modeling among others can be explored as well for that purpose. Researches at Inria have shown the power of Bayesian modeling [20] in the case of musical data sparsity and a new collaboration project[6] will allow the evaluation of integrating sequential formal modeling with probabilistic methods in the case of complex musical discourses when one lacks training data, trying to combine this way computational efficiency and robustness of learning.

5.3 Interaction Logic

In this perceptual and knowledge framework, improvised interaction logic allows a rich and creative exchange between human and artificial agents and asks questions on the temporal and spatial collective adaptation of interaction at multiple scales. This adaptation takes advantage of the artificial perceptual and cognitive environment in order to articulate a proactive control of collective improvised interaction, addressing such issues as internal structure of the agents, memory models, knowledge and control capabilities. There is a need to go beyond the conventional static and predetermined approaches and be able to adapt in real-time the models, representations and learning methods of interaction, taking into account different temporal scales and collective dynamics [7], engaging attention, comprehension and decision skills. It will be possible in this way to construct intelligent multi-agent systems well equipped for dealing credibly with more complex musical situations involving a variety of styles.

5.4 Scenario Models

An artificial improvising musical agent may have in certain cases a planned strategy or a scenario. It has to improvise in an explicit harmonic and rhythmic context by exploiting a priori information that is structuring on one hand the training corpus and on the other the current context of improvisation providing they are labeled by a common symbolic vocabulary, e.g. harmonic, textural, or timbral descriptors [14]. The improvisation process guided by a control sequence can be described as the relationship between an external plan and a structured and annotated memory from which it dynamically reconstitutes musical sequences to create new improvisations, compatible in their temporal organization and their different dimensions (harmonic, rhythmic) with the plan. As an example, the plan could be a harmonic grid, and the memory a formal model learned from actual realizations of this grid and/or

[5] Uzeste Jazz Festival, Montreux Jaz Festival Workshop, Ircam Manifeste Festival etc.

[6] DYCI2 : Creative Dynamics of Improvised Interaction, dyci2.ircam.fr

others grids as well. The improvisation will follow the grid while deploying and combining musical material coming from a number of sources in the memory. The ImproTek [14] flavor of OMax developed by Marc Chemillier and Jérôme Nika in collaboration between EHESS and the Music Representation Team at Ircam can already, to a certain extent, simulate such a planned improvisation situation. It is worth noting that in this model there is some access to the future, as the scenario is known in advance, and the algorithmical machinery can obviously take advantage of this characteristics even (and especially) in a real-time setting.

5.5 Flow Models

However, an improvising agent may act freely with no information on the future as in the previous case. When there is no known scenario, the generative process can nevertheless be oriented causally by an input stream, typically produced by a human musician or by an other digital intelligence, themselves improvising freely or with a defined strategy. The process maintains, in this case, a "floating synchronization" between the data acquired in real-time by listening to the input stream and a structured and annotated memory from which it dynamically reconstructs musical sequences coherent and locally compatible with the input stream for selected musical dimensions. A simple example of such an interaction is the automatic accompaniment of a melodic improvisation, another one is the generation of solos on an improvised series of chord ; in the general case, an arbitrary number of agents trained on different corpora will be able to co-improvise by listening to each other and adapting flexibly to each other with regards to pitch, rhythm, harmony, or texture and timbre, as would human experts do. The SoMax flavor of OMax developed by Laurent Bonnasse-Gahot in the Music Representation Team at Ircam can already, to a certain extent, simulate such an adaptive behavior linked to an input flow [25].

Scenario models and flow models could (and should) be combined to simultaneously take advantage of the ability of the first to manage improvisation plans and of the second to flexibly adapt to the contingencies of listening and interaction. Such a system would flow naturally in interaction with a moving contexte while taking into account the formal organization of music as it is learned and modeled into a structured annotated memory, off-line or in-line in the very time of interaction.

5.6 Temporal Adaptation

An adaptive and proactive system of improvised interaction takes into account different temporal scales and collective dynamics. This means recognizing and adapting short-term phenomena (reaction, synchronization) as well as long-term phenomena (the emergence of vocabulary and higher level forms) and involves updating multi-scale hierarchical representations of interaction (engaging cognitive representations, or mental states). The formalization of a consistent and versatile scheme of interaction to address these areas is challenging and requires the development of a global approach in order to reach “expressiveness” and “style”.

5.7 Memory, Knowledge, Control

A creative artificial agent capable of listening, learning and performing improvised musical interaction with humans and other artificial agents must have a minimal "cognitive" structure allowing her to succeed in a complex environment which incorporates its own productions, thus involving reflexivity on its own behavior and state. One has to identify the representations and processes best suited to model this internal structure and its activation by external stimuli.

It may consist in a memory and a knowledge network operated at different time scales (echoic memory, long term memory) and at different activation and control levels, whether they be implicit (procedural or reflex) or explicit (episodic or semantic memory). Coupled structures of memory and control currently implemented as a Factor Oracle, should be extended to sophisticated topological devices, possibly using SOM (self organizing maps). These would contribute to the formation of a semantic memory through the automatic organization of a topology of musical objects. It would contribute as well to the constitution of reflexive processes simulating awareness (curiosity triggered by a stimulus), attention (listening or not to other agents), motivation (wanting to learn or not) and initiative (decision to play or not) in relation to computational models of self and intentionality (self-model theory).

6. CONCLUSIONS

Creative Symbolic Interaction has emerged as an extremely productive combination of machine listening, machine learning and music structure formalisation, in the framework of adaptive interaction dynamics. It is rich of promises for *Improvised Machine Musicianship* (to quote Robert Rowe's famous book *Machine Musicianship*) and in general for digital intelligence and creativity.

Acknowledgments

Thanks to the others *OMax Brothers*: M. Chemillier, S. Dubnov, G. Bloch, B. Lévy, L. Bonnasse-Gahot, J. Nika.

Improvised Machine Musicianship projects described in this document partly funded by Agence nationale de la recherche (ANR) projects ImproTech, SOR2 (finished) and DYCI2 (starting).

7. REFERENCES

[1] G. Assayag, G. Bloch, M. Chemillier, A. Cont, S. Dubnov "Omax Brothers : a Dynamic Topology of Agents for Improvization Learning", *Workshop on Audio and Music Computing for Multimedia*, ACM Multimedia 2006, Santa Barbara, 2006

[2] G. Assayag, S. Dubnov, and O. Delerue. "Guessing the composer's mind: applying universal prediction to musical style". *Proceedings of the ICMC,* Bejing, China, 1999

[3] G. Assayag, S. Dubnov "Using Factor Oracles for machine Improvisation", *Soft Computing, vol. 8, n° 9,* Septembre, 2004

[4] Assayag, G. , Bloch, G. « Navigating the Oracle: a Heuristic Approach », *Proceedings of the ICMC*, The In. Comp. Music Association, Copenhagen 2007.

[5] Assayag, G., Bloch, G., Dubnov, S., Cont, A., "Interaction with Machine Improvisation", in *The Structure of Style,* Springer Verlag, K. Burns, S. Argamon, S. Dubnov (Eds), pp. 219-246, 2010

[6] T. Blackwell, O. Bown & M. Young. "Live Algorithms: Towards Autonomous Computer Improvisers". In J. McCormack & M. d'Inverno, eds, *Computers and Creativity*, Springer Berlin Heidelberg, 2012, pp. 147–174.

[7] Canonne, C., Garnier, N., "A Model for Collective Free Improvisation", *Proc. MCM'03*, Springer, Paris, France, 2011.

[8] Conklin, D. "Music Generation from Statistical Models", *Proc. of the AISB 2003 Symposium on Artificial Intelligence and Creativity in the Arts and Sciences,* Aberystwyth, Wales, 2003, pp. 30-35

[9] M. Crochemore, L. Ilie, E. Seid-Hilmi "The Structure of Factor Oracles"*, Int. J. Found. Comput. Sci.* 18(4), 2007, pp.781–797

[10] A. Donze, S. Libkind, S.A. Seshia, D. Wessel, "Control improvisation with application to music". *Technical report No. UCB/EECS-2013-183. EECS* Department, University of California, Berkeley, 2013.

[11] Dubnov, S., Assayag, G., Cont, A., "Audio Oracle Analysis of Musical Information Rate", *Proc. IEEE Semantic Computing Conference*, ICSC2011, Palo Alto, CA, 2011, pp. 567-571

[12] S. Dubnov, G. Assayag, O. Lartillot, G. Bejerano "Using Machine-Learning Methods for Musical Style Modeling", IEEE Computer, vol. 10, n° 38, Octobre, 2003

[13] Lévy, B., Bloch, G., Assayag, G., "OMaxist Dialectics : Capturing, Visualizing and Expanding Improvisations", *Proc. NIME 2012*, Ann Arbor, 2012, pp. 137-140

[14] J. Nika, M. Chemillier, "ImproteK, integrating harmonic controls into improvisation in the filiation of OMax," in Proceedings of the International Computer Music Conference, pp. 180–187, 2012.

[15] J. Pressing, "Cognitive processes in improvisation". *Advances in Psychology*, Vol. 19, 1984, pp. 345–363

[16] Surges, G. and Dubnov, S. "Feature Selection and Composition using PyOracle." *Workshop on Musical Metacreation, Ninth AAAI Conference*, Boston, MA. October 14-15, 2013.

[17] Assayag, G., Truchet, C. (Eds) *Constraint Programming in Music*, ISTE Ltd and John Wiley & Sons Inc, 256 p., 2011

[18] Cont, A., Dubnov, S., Assayag, G., "On the Information Geometry of Audio Streams with Applications to Similarity Computing", *IEEE Transactions on Audio, Speech, and Language Processing,* Aug. 2011, vol. 19, n° 1, pp. 837-846, 2011

[19] Marchand, S., Badeau, R., Baras, C., Daudet, L., Fourer, D., Girin, L., Gorlow, S., Liutkus, A., Pinel, J., Richard, G., Sturmel, N., and Zhang, S. "DReaM: a novel system for joint source separation and multi-track coding". *133rd Audio Engineering Society (AES) Convention*, San Francisco, California, USA, October 2012.

[20] S.A. Raczyński, E. Vincent and S. Sagayama. "Dynamic Bayesian networks for symbolic polyphonic pitch modeling" *IEEE Transactions on Audio, Speech and Language Processing*, 2013.

[21] W. Wang, Ed. "Machine Audition: Principles, Algorithms and Systems", Information science reference, Hershey*,* New York, 2010

[22] K. Sanlaville, G. Assayag, F. Bevilacqua, C. Pelachaud. « Emergence of synchrony in an Adaptive Interaction Model ». *Intelligent Virtual Agents 2015 Doctoral Consortium*, Aug 2015, Delft, Netherlands. IVA Doctoral Consortium, 2015.

[23] J. Nika, J. Echeveste, M. Chemillier, J.L. Giavitto. « Planning Human-Computer Improvisation ». *International Computer Music Conference*, Sep 2014, Athens, Greece. pp.330, 2014.

[24] J. Nika, M. Chemillier. « Improvisation musicale homme-machine guidée par un scénario temporel ». *Technique et Science Informatique, Numéro Spécial RenPar\'13*, Hermes, 2015, 7-8 (33), pp.651-684.

[25] L. Bonnasse-Gahot «An update on the SOMax project, Soma v0.1.3 » Ircam – STMS Internal report, 2014

[26] J. Bresson, C. Agon, G. Assayag. « OpenMusic – Visual Programming Environment for Music Composition, Analysis and Research » *ACM* MM '11: Proceedings of the 19th ACM international conference on Multimedia, 2011