Creative Symbolic Interaction

Keynote talk

Gérard Assayag
STMS Lab, IRCAM CNRS UPMC
gerard.assayag@ircam.fr

ABSTRACT

Creative Symbolic Interaction brings together the ad-vantages from the worlds of interactive real-time com-puting and intelligent, content-level analysis and proc-essing, in order to enhance and humanize man-machine communication. Performers improvising along with Symbolic Interaction systems experiment a unique ar-tistic situation where they interact with musical (and possibly multi-modal) agents which develop themselves in their own ways while keeping in style. Symbolic in-teraction 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 is a general idea that makes sense in many artistic and non-artistic domains.

1. INTRODUCTION

Until recently, in the field of musical interaction with machines, engineers and researchers have been con-cerned by fast computer computation and reaction — a logical concern considering available machine speeds and complexity of tasks. However, instantaneous re-sponse is not always the way a musician reacts in a real performance situation. Although decisions are being carried out at a precise time, the decision process relies on evaluation of past history, analysis of incoming events and anticipation strategies. Therefore, not only can it take some time to come to a decision, but part of this decision can also be to postpone action to a later time. This process involves time and memory at differ-ent scales, just as music composition does, and cannot be fully apprehended just by conventional signal and event processing. A symbolic level has to be involved as well.

In order to foster realistic and artistically interesting behaviors of digital interactive systems, and communi-cate with them in a humanized way, we wish to com-bine several means : machine listening [21] — extract-ing 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 — formalized repre-sentations connecting to organized musical thinking [5], analysis and composition. These tools cooperate in ef-fect to define a multi-level memory model underlying a discovery and learning process that contributes to the emergence of a creative musical agent.

In the Music Representation Team at Ircam, after OpenMusic, a standard for computer assisted composi-tion, we have designed OMax [1, 4, 13], an interactive machine improvisation environment which explores this new interaction schemes. It creates a cooperation be-tween heterogeneous components specialized in real-time audio signal processing, high level music represen-tations and formal knowledge structures. This environ-ment learns and plays on the fly in live setups and is used in many artistic and musical performances. OMax exemplifies several trends of current research on interac-tive creative agents capable of adequacy and relevance by connecting instant contextual listening to corpus based knowledge, along with longer term investigation and decision processes allowing to refer to larger-scale structures and scenarios. We call this scheme *Symbolic Interaction*.

Creative Symbolic Interaction brings together the ad-vantages one can get 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 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 strategy makes its way by weaving decisions step after step, either by relating to an overall structural determin-ism, or by jumping in an "improvised" way and generat-ing a surprise. 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, thus it may serve as a productive model for arti-ficial music creativity.

2. GENERAL CONTEXT

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

This type of interaction involves the perception / ac-tion loop and engages the learning process in a renewed design where an agent learns in particular from the reac-tions 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, sig-nal and physical models and representations. Creativity in general is supported at European level by one of its Objective called “Technology and scientific foundations in the field of creativity” in its actions “Intelligent com-putational environments and stimulating human creativ-ity and Enhancing Progress towards formal understand-ing of creativity”.

3. IMPROVISED INTERACTION

To confront the problem of improvised interaction un-derstood as a powerful driver for creativity and situated at the heart of all human activities constitutes one of today’s central challenges in 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 where we wish to integrate artificial listening, learning of mu-sical behaviors, temporal modeling of musical structures and dynamical creative interaction in an architecture for effective experimentation in real time.

Possible applications will be available in varied fla-vors 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 pro-duction 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 installa-tions, in electro-acoustic composition, in the creation of new “variable” formats for music production and distri-bution.

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 appli-cations, 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 no-tion 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 par-ticipation.

4. DYNAMICS OF INTERACTION

4.1 Creative Digital Agents

Our dynamics of creative improvised interaction focuses on the creation, adaptation and implementation of effec-tive 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 autonomous, creative, able to display artistically credible manners in various artistic and edu-cational human setups such as live performance and teaching. These agents mayalso help constitute the per-ceptual and communicative skills of embedded artificial intelligence systems.

The aim is to evolve self-creative agents by the proc-ess of interactive learning from direct exposure to hu-man improvisers, thus creating a retro-action loop (sty-listic 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.

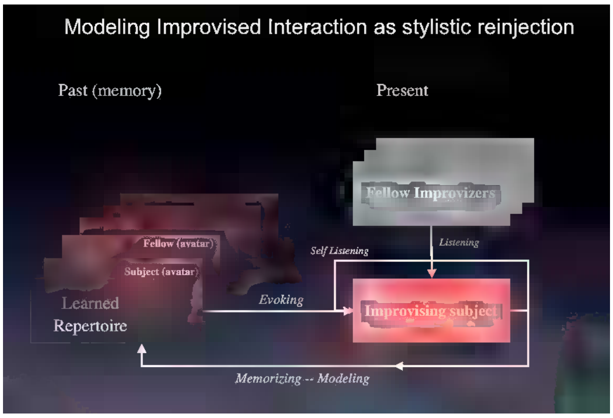

Figure 1. Stylistic Reinjection Paradigm

4.2 Offline Learning

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

In addition to applying these digital skills to live mu-sic 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 resur-rected within a creative or educational framework, thus extending the improvisation paradigm to new narrative and immersive forms in a step unseen yet. We are cur-rently building a project in this direction with EPFL University in Lausanne around the Montreux Festival Archives recently listed by Unesco as part of the World Cultural Heritage.

4.3 Collectivities of Agents

The idea is to create both an artificial expertise of mu-sical practice ("machine musicianship") by this kind of interaction and a rich experience of instantaneous hu-man – digital communication likely to provide an aes-thetic 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 be in effect extended to complex configuration involving potentially a great number of agents — human and artificial — learning and evolving 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 contribu-tory 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 generation of symbolic se-quences, the real-time audio management, and the visu-alization 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]. Sev-eral extensions of the OMax paradigm are currently un-der research or starting in the author’s team (Music Rep-resentation) at Ircam as well as in partner labs at UCSD, EHESS, Inria, UBO, and CNMAT Berkeley. These re-searches include musical information dynamics [11, 16], formal automata [10], probabilistic approaches [8], cog-nitive modeling of memory, accompaniment systems, scenario based improvisation, multi-agent interaction dynamics, temporal adaptation of interaction, digital intelligence and artificial creativity.

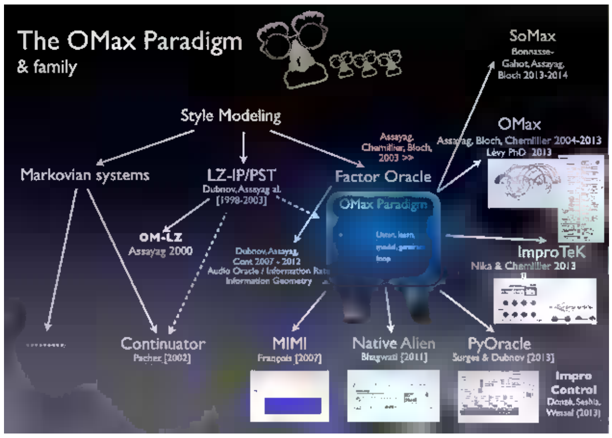

Figure 2. OMax Galaxy

The OMax galaxy connects with active or starting pro-jects 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 interna-tional collaboration have been initiated, in order to stimulate theoretical advances and development of ex-perimental 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 [18] and exploit-ing available prior knowledge possibly through an “in-formed musical decomposition” paradigm [19], taking advantage of scores, annotations, or inferring partial information from learning beforehand stylistically com-patible 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 emerg-ing in a context of musical performance and improvisa-tion. Many formal models may compete and even coop-erate for that purpose. The Factor Oracle model [3, 9] coming from the automata and formal languages re-search field has been massively used in OMax and its derivative with great achievements. Hidden Markov Models, Temporal Grammar Rules Induction, Deep Belief Networks, Concurrent Constraint Calculus, Spatial and Epistemic CCP, Multi-objective Time Series model-ing among others can be explored as well for that pur-pose. Researches at Inria have shown the power of Bay-esian modeling [20] in the case of data sparsity and a new collaboration project will allow the evaluation of integrating sequential formal modeling with probabilis-tic 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, impro-vised interaction logic allows a rich and creative ex-change between human and artificial agents and asks questions on the temporal and spatial collective adapta-tion of interaction at multiple scales. This adaptation takes advantage of the artificial perceptual and cogni-tive environment in order to articulate a proactive con-trol of collective improvised interaction, addressing such issues as internal structure of the agents, memory mod-els, 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 collec-tive 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 situa-tions 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 re-constitutes musical sequences to create new improvisa-tions, compatible in their temporal organization and their different dimensions (harmonic, rhythmic) with the plan. As an example, the plan can be a harmonic grid, and the memory a formal model learned from realiza-tions of this grid and/or others one as well. The Im-provisation will follow the grid while deploying and combining musical material coming from a number of sources. The ImproTek [14] flavor of OMax developed in collaboration between EHESS and the Music Repre-sentation Team at Ircam can already, to a certain extent, simulate such a planned improvisation situation.

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 that case a "floating synchronization" between the data ac-quired in real-time by listening to the input stream and a structured and annotated memory from which it dy-namically reconstructs musical sequences coherent and locally compatible with the input stream for certain mu-sical dimensions. A simple example of such an interac-tion is the automatic accompaniment of a melodic im-provisation, another one is the generation of solos on an improvised series of chord ; in the general case, an arbi-trary number of agents trained on different corpus 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 hu-man experts do. The SoMax flavor of OMax developed in the Music Representation Team at Ircam can already, to a certain extent, simulate such an adaptive behavior linked to an input flow.

Scenario models and flow models can 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 in-teraction, taking into account the formal organization of music as it is learned and modeled into a structured an-notated memory.

5.6 Temporal Adaptation

An adaptive and proactive system of improvised interac-tion takes into account different temporal scales and collective dynamics. This means recognizing and adapt-ing short-term phenomena (reaction, synchronization) as well as long-term phenomena (the emergence of vo-cabulary 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 him to succeed in a com-plex environment which incorporates its own produc-tions, thus involving reflexivity on its own behavior and state. One has to identify the representations and proc-esses 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 struc-tures of memory and control currently implemented as a Factor Oracle, should be extended to sophisticated topo-logical 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 other agents), motivation (wanting to learn or not) and initiative (decision to play or not) in relation to computational models of self and intentional-ity (self-model theory).

6. CONCLUSIONS

Creative Symbolic Interaction has emerged as an ex-tremely productive combination of machine listening, machine Learning and Music structure modeling, in the framework of adaptive interaction dynamics. It is rich of promises for *Improvised Machine Musicianship* (to quote Robert Rowe's famous book Machine Musician-ship) and in general for digital intelligence and creativ-ity.

Acknowledgments

Thanks to the other 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.Ański.Raczy , 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