Activity Report 2012

Project-Team MUTANT

Synchronous Realtime Processing and Programming of Music Signals

RESEARCH CENTER
Paris - Rocquencourt

THEME
Embedded and Real Time Systems

Table of contents

Project-Team MUTANT

Keywords: Audio, Synchronous Languages, Machine Learning, Real-time

MuTant team is located at Ircam-Centre Pompidou and part of the Mixed Research Unit on Science and Technologies of Sound and Music.

Creation of the Project-Team: January 01, 2012 .

1. Members

Research Scientists

Arshia Cont [Team leader]
Jean-Louis Giavitto [Sernior Researcher, HdR]
Florent Jacquemard [Researcher Inria, HdR]

Engineer

Thomas Coffy [ADT Antescofo, Beginning October 2012]

PhD Students

José Echeveste [UPMC]
Arnaud Dessein [UPMC, PhD defended on December 2012]

2. Overall Objectives

2.1. Overall Objectives

The research conducted in *MuTant* is devoted both to leveraging capabilities of musical interactions between humans and computers, and to the development of tools to foster the authoring of interaction and time in computer music. Our research program departs from *Interactive music systems* for computer music composition and performance introduced in mid-1980s at Ircam. Within this paradigm, the computer is brought into the cycle of musical creation as an *intelligent performer* [44] and equipped with a *listening machine* [40] capable of analyzing, coordinating and anticipating its own and other musicians' actions within a musically coherent and synchronous context. Figure 1 illustrates this paradigm. The use of Interactive Music Systems have become universal ever since and their practice has not ceased to nourish multidisciplinary research. From a research perspective, an interactive music systems deals with two problems: realtime machine listening [39], [40] or music information retrieval from musicians on stage, and music programming paradigms [32], [27] reactive to the realtime recognition and extraction. Whereas each field has generated subsequent literature, few attempts have been made to address the global problem by putting the two domains in direct interaction.

In modern practices, the computer's role goes beyond rendering pre-recorded accompaniments and is replaced by concurrent, synchronous and realtime programs defined during the compositional phase by artists and programmers. This context is commonly referred to as *Machine Musicianship* where the computer does not blindly follow the human but instead has a high degree of musical autonomy and competence. In this project, we aim at developing computer systems and language to support real-time intelligent behavior for such interactions.

MuTant's research proposal lies at the intersection and union of two themes, often considered as disjoint but inseparable within a musical context:

1. Realtime music information retrieval and processing
2. Synchronous and realtime programming for computer music

Figure 1. General scheme of Interactive Music Systems

2.2. Highlights of the Year

The Antescofo software and programming language was featured in more than 15 world-premier creations and 30 events worldwide, including its premiers with *New York Philharmonics*, *Orchestre de Paris*, and prestigious venues in USA, Japan, Turkey, Poland, England and more. See website for more details.

3. Scientific Foundations

3.1. Real-time Machine Listening

When human listeners are confronted with musical sounds, they rapidly and automatically find their way in the music. Even musically untrained listeners have an exceptional ability to make rapid judgments about music from short examples, such as determining music style, performer, beating, and specific events such as instruments or pitches. Making computer systems capable of similar capabilities requires advances in both music cognition, and analysis and retrieval systems employing signal processing and machine learning.

In a panel session at the 13th National Conference on Artificial Intelligence in 1996, Rodney Brooks (noted figure in robotics) remarked that while automatic speech recognition was a highly researched domain, there had been few works trying to build machines able to understand “non-speech sound”. He went further to name this as one of the biggest challenges faced by Artificial Intelligence [41]. More than 15 years have passed. Systems now exist that are able to analyze the contents of music and audio signals and communities such as International Symposium on Music Information Retrieval (MIR) and Sound and Music Computing (SMC) have formed. But we still lack reliable Real-Time machine listening systems.

The first thorough study of machine listening appeared in Eric Scheirer’s PhD thesis at MIT Media Lab in 2001 [40] with a focus on low-level listening such as pitch and musical tempo, paving the way for a decade of research. Since the work of Scheirer, the literature has focused on task-dependent methods for machine listening such as pitch estimation, beat detection, structure discovery and more. Unfortunately, the majority of existing approaches are designed for information retrieval on large databases or off-line methods. Whereas the very act of listening is real-time, very little literature exists for supporting real-time machine listening. This argument becomes more clear while looking at the yearly Music Information Retrieval Evaluation eXchange (MIREX), with different retrieval tasks and submitted systems from international institutions, where almost no emphasis exists on real-time machine listening. Most MIR contributions focus on off-line approaches to information retrieval (where the system has access to future data) with less focus on on-line and realtime approaches to information decoding.

On another front, most MIR algorithms suffer from modeling of temporal structures and temporal dynamics specific to music (where most algorithms have roots in speech or biological sequence without correct adoption to temporal streams such as music). Despite tremendous progress using modern signal processing and statistical learning, there is much to be done to achieve the same level of abstract understanding for example in text and image analysis on music data. On another hand, it is important to notice that even untrained listeners are easily able to capture many aspects of formal and symbolic structures from an audio stream in realtime. Realtime machine listening is thus still a major challenge for artificial sciences that should be addressed both on application and theoretical fronts.

In the MUTANT project, we focus on realtime and online methods of music information retrieval out of audio signals. One of the primary goals of such systems is to fill in the gap between *signal representation* and *symbolic information* (such as pitch, tempo, expressivity, etc.) contained in music signals. MUTANT's current activities focus on two main applications: *score following* or realtime audio-to-score alignment [2], and realtime transcription of music signals [20] with impacts both on signal processing using machine learning techniques and their application in real-world scenarios.

The team-project will focus on two aspects of realtime machine listening:

1. **Application-Driven Approach:** First, to enhance and foster existing application-driven approaches within the team such as realtime alignment algorithms and polyphonic pitch transcription. Our contributions on this line correspond to extensions of existing algorithmic approaches to realtime audio alignment and transcription to create new interactive application paradigms with new algorithmic approaches. Arshia Cont's ongoing realtime alignment in *Antescofo* as well as realtime transcription using non-negative factorization methods [20] are examples of this.
2. **Music Information Geometry:** In parallel to concrete applications, we hope to theoretically contribute to the problem of signal representations of audio streams for effortless retrieval of high-level information structures. We have recently shown in [4] that the gap between the symbolic/semantic and signal aspects of music information mostly lies on constructing a well-behaved representational space before any algorithmic considerations, by employing the emerging methods of *information geometry*. Arnaud Dessein's ongoing PhD thesis is focused on this aspect of the project.

3.2. Synchronous and realtime programming for computer music

The second aspect of an interactive music system is to *react* to extracted high-level and low-level music information based on pre-defined actions. The simplest scenario is *automatic accompaniment*, delegating the interpretation of one or several musical voices to a computer, in interaction with a live solo (or ensemble) musician(s). The most popular form of such systems is the automatic accompaniment of an orchestral recording with that of a soloist in the classical music repertoire (concertos for example). In the larger context of interactive music systems, the "notes" or musical elements in the accompaniment are replaced by "programs" that are written during the phase of composition and are evaluated in realtime in reaction and relative to musicians' performance. The programs in question here can range from sound playback, to realtime sound synthesis by simulating physical models, and realtime transformation of musician's audio and gesture.

Such musical practice is commonly referred to as the *realtime school* in computer music, developed naturally with the invention of the first score following systems, and led to the invention of the first prototype of realtime digital signal processors [28] and subsequents [31], and the realtime graphical programming environment *Max* for their control [37] at Ircam. With the advent and availability of DSPs in personal computers, integrated realtime event and signal processing graphical language *MaxMSP* was developed [38] at Ircam, which today is the worldwide standard platform for realtime interactive arts programming. This approach to music making was first formalized by composers such as Philippe Manoury and Pierre Boulez, in collaboration with researchers at Ircam, and soon became a standard in musical composition with computers.

Besides realtime performance and implementation issues, little work has underlined the formal aspects of such practices in realtime music programming, in accordance to the long and quite rich tradition of musical notations. Recent progress has convinced both the researcher and artistic bodies that this programming

paradigm is close to *synchronous reactive programming languages*, with concrete analogies between both: parallel synchrony and concurrency is equivalent to musical polyphony, periodic sampling to rhythmic patterns, hierarchical structures to micro-polyphonies, and demands for novel hybrid models of time among others. *Antescofo* is therefore an early response to such demands that needs further explorations and studies.

Within the MUTANT project, we propose to tackle this aspect of the research within three consecutive lines:

- **Development of a Synchronous DSL for Real Time Musician-Computer Interaction:** Ongoing and continuous extensions of the *Antescofo* language following user requests and by inscribing them within a coherent framework for the handling of temporal musical relationships. José Echeveste's ongoing PhD thesis focuses on the research and development of these aspects. Recent formalizations of the *Antescofo* language has been published in [6].
- **Formal Methods:** Failure during an artistic performance should be avoided. This naturally leads to the use of formal methods, like static analysis or model checking, to ensure formally that the execution of an *Antescofo* program will satisfy some expected property. The checked properties may also provide some assistance to the composer especially in the context of "non deterministic score" in an interactive framework.

3.3. Off-the-shelf Operating Systems for Real-time Audio

While operating systems shield the computer hardware from all other software, it provides a comfortable environment for program execution and evades offensive use of hardware by providing various services related to essential tasks. However, integrating discrete and continuous multimedia data demands additional services, especially for real-time processing of continuous-media such as audio and video. To this end interactive systems are sometimes referred to as off-the-shelf operating systems for real-time audio. The difficulty in providing correct real-time services has much to do with human perception. Correctness for real-time audio is more stringent than video because human ear is more sensitive to audio gaps and glitches than human eye is to video jitter [43]. Here we expose the foundations of existing sound and music operating systems and focus on their major drawbacks with regards to today practices.

An important aspect of any real-time operating system is fault-tolerance with regards to short-time failure of continuous-media computation, delivery delay or missing deadlines. Existing multimedia operating systems are soft real-time where missing a deadline does not necessarily lead to system failure and have their roots in pioneering work in [42]. Soft real-time is acceptable in simple applications such as video-on-demand delivery, where initial delay in delivery will not directly lead to critical consequences and can be compensated (general scheme used for audio-video synchronization), but with considerable consequences for Interactive Systems: Timing failure in interactive systems will heavily affect inter-operability of models of computation, where incorrect ordering can lead to unpredictable and unreliable results. Moreover, interaction between computing and listening machines (both dynamic with respect of internal computation and physical environment) requires tighter and explicit temporal semantics since interaction between physical environment and the system can be continuous and not demand-driven.

Fulfilling timing requirements of continuous media demands explicit use of scheduling techniques. As shown earlier, existing Interactive Music Systems rely on combined event/signal processing. In real-time, scheduling techniques aim at gluing the two engines together with the aim of timely delivery of computations between agents and components, from the physical environment, as well as to hardware components. The first remark in studying existing system is that they all employ static scheduling, whereas interactive computing demands more and more time-aware and context-aware dynamic methods. The scheduling mechanisms are neither aware of time, nor the nature and semantics of computations at stake. Computational elements are considered in a functional manner and reaction and execution requirements are simply ignored. For example, *Max* scheduling mechanisms can delay message delivery when many time-critical tasks are requested within one cycle [38]. *SuperCollider* uses Earliest-Deadline-First (EDF) algorithms and cycles can be simply missed [35]. This situation leads to non-deterministic behavior with deterministic components and poses great difficulties for preservation of underlying techniques, art pieces, and algorithms. The situation has become worse with the demand for nomad physical computing where individual programs and modules are available but no action

coordination or orchestration is proposed to design integrated systems. System designers are penalized for expressivity, predictability and reliability of their design despite potentially reliable components.

Existing systems have been successful in programing and executing small system comprised of few programs. However, severe problems arise when scaling from program to system-level for moderate or complex programs leading to unpredictable behavior. Computational elements are considered as functions and reaction and execution requirements are simply ignored. System designers have uniformly chosen to hide timing properties from higher abstractions, and despite its utmost importance in multimedia computing, timing becomes an accident of implementation. This confusing situation for both artists and system designers, is quite similar to the one described in Dr. Edward Lee's seminal paper "Computing needs time" stating: "general-purpose computers are increasingly asked to interact with physical processes through integrated media such as audio. [...] and they don't always do it well. The technological basis that engineers have chosen for general-purpose computing [...] does not support these applications well. Changes that ensure this support could improve them and enable many others" [30].

Despite all shortcomings, one of the main advantages of environments such as *Max* and *PureData* to other available systems, and probably the key to their success, is their ability to handle both synchronous processes (such as audio or video delivery and processing) within an asynchronous environment (user and environmental interactions). Besides this fact, multimedia service scheduling at large has a tendency to go more and more towards computing besides mere on-time delivery. This brings in the important question of hybrid scheduling of heterogeneous time and computing models in such environments, a subject that has had very few studies in multimedia processing but studied in areas such simulation applications. We hope to address this issue scientifically by first an explicit study of current challenges in the domain, and second by proposing appropriate methods for such systems. This research is inscribed in the three year ANR project INEDIT coordinated by the team leader (started in September 2012).

4. Application Domains

4.1. Application Domains

- **Authoring and Performing Interactive Music.** The combination of both realtime machine listening systems and reactive programming paradigms has enabled the *authoring* of interactive music systems as well as their realtime performance within a coherent synchronous framework called Antescofo. The module, developed since 2008 by the team members, has gained increasing attention within the user community worldwide with more than 30 prestigious public performances yearly. The outcomes of the proposed research will enhance the interactive and reactive aspects of this emerging paradigm as well as creating novel authoring tool for such purposes. The outcome of the ANR Project INEDIT (with LABRI and GRAME and coordinated by team leader), will further extend the use-cases of *Antescofo* for interactive multimedia pieces with more complex temporal structures and computational paradigms.
- **Music Post-Production.** Outcomes of our recognition and alignment paradigms can improve and ease existing workflows employed by audio engineers for mixing and editing using commercial Digital Audio Workstations (DAW) in post-production. We have recently initiated collaborations with audio engineers at Ircam and Paris Superior Music Conservatory (CNSMDP) to define the framework [8] and we will continue to develop and integrate our tools into their daily workflow.
- **Realtime Music Information Retrieval** We will apply our information geometric approach to well-known and complex MIR problems. A glance of such problems is presented in [5]. Such applications can be used as front-end of many high-level MIR applications such as audio summarisation, audio finger printing, and automatic annotation tools. Besides such low-level enhancements, our information geometric approach can address the well-known (and still to be solved) problem of audio queries over a database.

- **Automatic Accompaniment/Creative Tools for Entertainment Industry** Technologies developed by MUTANT can find their way with general public (besides professional musicians) and within the entertainment industry. Recent trends in music industry show signs of tendencies towards more intelligent and interactive interfaces for music applications. Among them is reactive and adaptive automatic accompaniment and performance assessment as commercialized by companies such as *MakeMusic* and *Tonara*. Technologies developed around *Antescofo* can enhance interaction between user and the computer for such large public applications. We hope to pursue this by licensing our technologies to third-party companies.

5. Software

5.1. Antescofo

Participants: Arshia Cont, Jean-Louis Giavitto, Florent Jacquemard, José Echeveste.

Antescofo is a modular polyphonic Score Following system as well as a Synchronous Programming language for musical composition. The module allows for automatic recognition of music score position and tempo from a realtime audio Stream coming from performer(s), making it possible to synchronize an instrumental performance with computer realized elements. The synchronous language within Antescofo allows flexible writing of time and interaction in computer music.

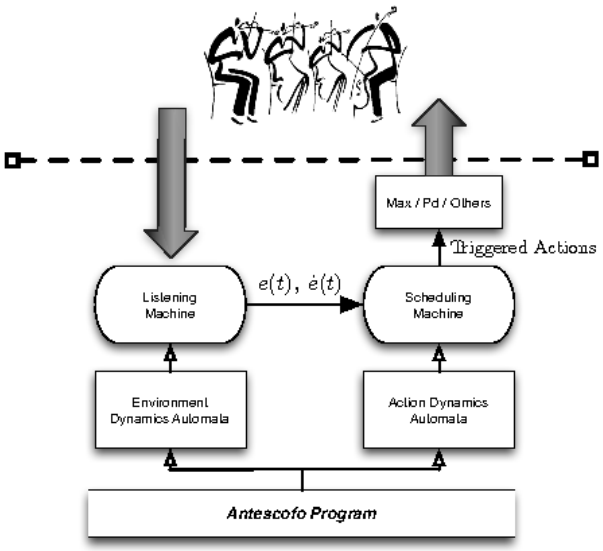

Figure 2. General scheme of Antescofo virtual machine

Antescofo is developed as modules for *Max* and *PureData* real-time programming environments.

5.2. Antescofo Visual Editor

Participants: Thomas Coffy [ADT], Arshia Cont, José Echeveste.

The Antescofo programming language can be extended to visual programing to better integrate existing scores and to allow users to construct complex and embedded temporal structures that are not easily integrated into text. This project has started in October 2012 thanks to Inria ADT Support. The foundations of a visual editor is set and the goal is to release a standalone editor for Antescofo programs in 2013.

6. New Results

6.1. Information-Geometric Approach to Real-time Audio Change Detection

Participants: Arnaud Dessein, Arshia Cont.

We developed a generic framework for real-time change detection of audio signals using methods of information geometry. The present method is limited to generative models of audio signals based on generic exponential distribution families. The proposed system detects changes by controlling the information rate of the signal as they arrive in time. The method also addresses shortcomings of traditional approaches based on cumulative sums which assume known parameters before change. This is achieved by calculating exact generalized likelihood ratio test statistics with complete estimation of unknown parameters in respective hypothesis [9]. The interpretation of this framework within a dually flat geometry of exponential families provide tractable algorithms for online use. Results are presented for speech segmentation into different speakers and polyphonic music segmentation.

6.2. Real-time Polyphonic Music Recognition

We investigated real-time recognition of overlapping music events in two context of dictionary-based detection and real-time alignment:

6.2.1. Real-time detection of overlapping sound events using non-negative matrix factorization

Participants: Arnaud Dessein, Arshia Cont.

Non-negative matrix factorization (NMF) methods have naturally found their way since their inception to sound and music processing. This work is an extension to our previous work in [1] on Real-time Music Transcription using sparse NMF methods. We investigate the problem of real-time detection of overlapping sound events by employing NMF techniques. We consider a setup where audio streams arrive in real-time to the system and are decomposed onto a dictionary of event templates learned off-line prior to the decomposition. An important drawback of existing approaches in this context is the lack of controls on the decomposition. We propose and compare two provably convergent algorithms that address this issue, by controlling respectively the sparsity of the decomposition and the trade-off of the decomposition between the different frequency components. Sparsity regularization is considered in the framework of convex quadratic programming, while frequency compromise is introduced by employing the beta-divergence as a cost function. The two algorithms are evaluated on the multi-source detection tasks of polyphonic music transcription, drum transcription and environmental sound recognition. The obtained results in [20] show how the proposed approaches can improve detection in such applications, while maintaining low computational costs that are suitable for real-time.

A specialized version of NMF for Real-time Music Transcription is exposed in Arnaud Dessein's PhD thesis [9].

These methods will be subject to software development in 2013.

6.2.2. Robust Real-time Polyphonic Audio-to-Score Alignment

Participant: Arshia Cont.

The *Antescofo* system is polyphonic since 2009 but its use in highly polyphonic and noisy concert environments have been challenging. To overcome this, we have studied more robust inference mechanisms. As a results, the previous inference mechanism based on maximum a posteriori of Viterbi Forward variables in mixed semi-Markov and Markov chains in [2] were abandoned in favor of a more robust method based on *importance resampling* on state-space models and smoothing of variable-order hybrid chains. This has led to robust real-time alignment and the employment of the system in various Piano performances in 2012. Further extensions are currently under study.

6.3. Real-time Multi-object Detection for Music Signals

Participants: Philippe Cuvillier [Master 2 ATIAM], Arshia Cont.

Multiple-object detection and tracking has been widely used in applications such as missile tracking and radar and has given birth to several formalisms such as Random Finite Sets [33]. Such formalisms can be seen as extensions to existing probabilistic inference mechanisms with explicit birth and death stochastic mechanisms for multiple source tracking.

In this work we aim at studying such formalisms in the case of real-time music signal processing. The idea is to track multiple sources (instruments, audio flows) from one source of observation. This approach can be beneficial to two main applications in real-time music listening:

- Extension of existing audio-to-score [2] or audio-to-audio alignment [7] mechanisms (currently based on one source) to multiple objects can address the following short-comings of existing approaches: explicit consideration for asynchrony of parallel sources; robustness to uncertainties on one or more voices.
- Studying the classical *Partial Tracking* applications in audio processing within the RFS context can lead to better results in low-level sinusoidal partial tracking of sounds.

Early studies of such formalisms are exposed in [25]. Concrete applications will be exposed in 2013.

6.4. Antescofo Language Extensions and Performance Fault-Tolerance

We have improved the *Antescofo* framework widely used for mixed instrumental and live electronic computer music. The new framework paves the way for future language extensions and paves the way for future research regarding performance fault-tolerance, synchronization mechanisms and formal verifications.

6.4.1. Antescofo Language Extensions

Participants: José Echeveste, Jean-Louis Giavitto, Florent Jacquemard, Arshia Cont.

To further extend the *Antescofo* language, the system has been formally modeled as a network of parametric timed automata in [29]. The model obtained provides operational semantics for the input scores, in particular the interaction between the instrumental and electronic parts and the timing and error handling strategies mentioned below. This approach would enable better authoring of time and interaction during programing/composing, permits to use state of the art software verification tools for the static analysis of Antescofo scores and also provides means to address critical aspects of musical performances in real-time.

In parallel, a new grammar for the score language and a new architecture have been designed for *Antescofo*, taking into account new demands from the community such as addition of timed variables in the language, dynamic time processes, time-conditional constructs, and more.

6.4.2. Performance Fault-Tolerance and Synchronization Mechanisms

Participants: José Echeveste, Jean-Louis Giavitto, Arshia Cont.

We formalized the timing strategies for musical events taking into account the variability of environment signals (musicians) and their effect on computer events programmed in *Antescofo*. The result of this work is presented in [15], where new block attributes in the language determine expected behavior in case of environment changes in real-time (errors, timing discrepancies, etc.). These additions have been implemented in the current version of the system and are widely used by the user community.

6.5. Temporal Analysis and Verification of Interactive Music Scores

Participants: Léa Fanchon [Master 2 École Centrale], Florent Jacquemard.

Léa Fanchon's Masters thesis, under the supervision of Florent Jacquemard, [26] presents an analysis module that complements the real-time score authoring and performance in *Antescofo*, with the aim of exploring possible behavior of authored programs with respect to possible deviations in human musician performance. This work employs formal methods for temporal automata networks using linear constraint inference techniques commonly in use for task scheduling and circuit verifications.

Obtained results pave the way for future works in formal verification of interactive multimedia applications, being one of the first of its kind in computer music literature, and provides the following input to programmers and artists using *Antescofo*:

- Evaluation of robustness of the program with respect to the environment's (musician's performance) temporal variations,
- Feedback to programmers/artists on critical synchronization points for better programming.

An article describing this work is currently in preparation for a submission to a computer music conference.

6.6. Formal study of Antescofo as a Reactive System

Participants: Guillaume Baudart [Master 2 ATIAM], Florent Jacquemard, Marc Pouzet [ENS], Jean-Louis Giavitto, Arshia Cont.

An *Antescofo* score/program can be considered as a specification of a reactive system through its coupling of a machine listening with a real-time synchronous language. In his master thesis under the supervision of Florent Jacquemard and Marc Pouzet (team Parkas), Guillaume Baudart has has studied the links between the reactive system of *Antescofo* and existing synchronous languages such as *Lucid Synchrone* [36] and *Reactive ML* [34]. The reactive engine of a preliminary version of *Antescofo* was developed in both languages and their structures were compared.

This study reveals the particularities of musical applications of reactive systems specific to *Antescofo* (see [24]). *Reactive ML* allows dynamic constructions but real-time performance can not be guaranteed especially when the machine listening is combined with the reactive system. On the contrary, *Lucid Synchrone* does not easily allow dynamic process creation. Each language specificity leads to strong considerations in the program/score structure for the artists. This work will be continued in 2013 to further strengthen ties between the reactive aspects of *Antescofo* and that of synchronous languages.

6.7. Tree Structured Presentation of Symbolic Temporal Data

Participants: Florent Jacquemard, Michael Rusinowitch [Project-team Cassis], Luc Segoufin [Project-team Dahu].

In traditional music notation, in particular in the languages used for the notation of mixed music such as Antescofo DSL, the durations are not expressed by numerical quantities but by symbols representing successive subdivisions of a reference time value (the beat). For this reason, trees data structures are commonly used for the symbolic representation of rhythms in computer aided composition softwares such as OpenMusic (developed at Ircam). It is therefore worth studying the applications in rhythm notation of existing formalisms for recognizing, querying, transforming and learning sets of tree structured data.

In 2012 we have studied several classes of tree recognizers which could be of interest in this context. First, with Michael Rusinowitch we have proposed in [16] a novel class of automata computing on unranked trees, which are context free in two dimensions: in the the sequence of successors of a node and also along paths. Second, we studied with Luc Segoufin [21] automata and logics computing on data trees and their relationship. Data trees are unranked ordered trees where each node carries a label from a finite alphabet and a datum from some infinite domain.

7. Partnerships and Cooperations

7.1. National Initiatives

7.1.1. ANR

7.1.1.1. INEDIT

Title: Interactivity in the Authoring of Time and Interactions

Project acronym: INEDIT

Type: ANR Contenu et Interaction 2012 (CONTINT)

Instrument: ANR Grant

Duration: September 2012 - September 2015

Coordinator: IRCAM (France)

Other partners: Grame (Lyon, France), LaBRI (Bordeaux, France).

Abstract: The INEDIT project aims to provide a scientific view of the interoperability between common tools for music and audio productions, in order to open new creative dimensions coupling *authoring of time* and *authoring of interaction*. This coupling allows the development of novel dimensions in interacting with new media. Our approach lies within a formal language paradigm: An interactive piece can be seen as a virtual interpreter articulating locally synchronous temporal flows (audio signals) within globally asynchronous event sequence (discrete timed actions in interactive composition). Process evaluation is then to respond reactively to signals and events from an environment with heterogeneous actions coordinated in time and space by the interpreter. This coordination is specified by the composer who should be able to express and visualize time constraints and complex interactive scenarios between mediums. To achieve this, the project focuses on the development of novel technologies: dedicated multimedia schedulers, runtime compilation, innovative visualization and tangible interfaces based on augmented paper, allowing the specification and realtime control of authored processes. Among posed scientific challenges within the INEDIT project is the formalization of temporal relations within a musical context, and in particular the development of a GALS (Globally Asynchronous, Locally Synchronous) approach to computing that would bridge in the gap between synchronous and asynchronous constraints with multiple scales of time, a common challenge to existing multimedia frameworks.

7.1.2. Other National Initiatives

The team participated to the CLASYCO network on DSL for simulation, supported by the RNSC (réseau national des systèmes complexes).

Jean-Louis Giavitto participates to the SynBioTIC ANR Blanc project (with IBISC, University of Evry, LAC University of Paris-Est, ISC - Ecole Polytechnique).

7.2. International Research Visitors

7.2.1. Visits of International Scientists

Miller S. Puckette is a professor of computer music in University of California San Diego (UCSD) and author of *Max* and *PureData* real-time programming environments for interactive arts. He participated in May 2012 in the MuTant Real-time Multimedia Computing Seminars (available on the web) and contributed to the team's knowledge of multimedia real-time scheduling challenges and paradigms.

James McCartney is a senior researcher in Apple Core Audio project and author of the audio synthesis and algorithmic composition programming environment *SuperCollider*. He visited *MUTANT* in November 2012 and participated in the MuTant Real-time Multimedia Computing Seminars (available on the web). He is interested in robust scheduling of heterogeneous computing for real-time multimedia applications.

David Rizo is lecturer at the University of Alicante, Spain. He is interested in music information retrieval and classification of musical genres by combining audio and symbolic descriptors. He visited *MuTant* in March 2012 and participated in a session of the MaMux seminar dedicated to trees and hierarchical structures in computer music.

Masahiko Sakai is a professor at the University of Nagoya and director of the Sakabe/Sakai computer science laboratory of the department of computer science and mathematical informatics of Nagoya University. He visited *MuTant* in April 2012.

Yoshiharu Kojima is an research fellow of the Japan society for the promotion of science. He has made a two months post-doctoral visit in *MuTant* in October and November 2012 on the application of term rewriting techniques to the formalization of musical processes, under the institutional program for young researchers overseas visits of the graduate school of information science at Nagoya University.

8. Dissemination

8.1. Scientific Animation

8.1.1. Community Initiatives

The Brillouin Seminar series on *Information Geometry* is coordinated by MUTANT in partnership with LIX and THALES. It gathers 80 international researchers on the topic from various disciplines. In 2012, we organized 6 talks. Videos are available on the seminar website.

Jean-Louis Giavitto is the management team of the GDR GPL (Génie de la programmation et du logiciel), responsible with Etienne Moreau of the "Languages and Vérification" pole of the GDR. He is also and expert for the ANR DEFI projects and a reviewer for FET projets for the UC.

8.1.2. Editorial Boards

Jean-Louis Giavitto is the redactor-in-chief of TSI (Technique et Science Informatique), a french scientific journal published by Lavoisier and Hermes Science.

8.1.3. Participation in Program Committees of Conferences

Florent Jacquemard has been member of the program committees of the 6th International Joint Conference on Automated Reasoning (IJCAR 2012), June 26th to July 1st, 2012, Manchester, UK, and of the International Workshop on Trends in Tree Automata and Tree Transducers (TTATT 2012), a satellite workshop of the 23rd International Conference on Rewriting Techniques and Applications (RTA 2012), June 2, 2012, Nagoya, Japan.

Jean-Louis Giavitto was co-chair of the Spatial Computing Workshop (SCW2012) a satellite workshop of AAMAS, 5 June 2012, Valencia, Spain. He was also member of the Program Committees of: AFADL 2012 (Approches Formelles dans l'Assistance au Développement de Logiciels, 11-13 January 2012, Grenoble), IPCAT 2012 9th International Conference on Information Processing in Cells and Tissues 31st March – 2nd April 2012, Trinity College, Cambridge, GDS track at GECCO 2012, the Generative and Developmental track at the Genetic and Evolutionary Computation Conference, 7–11 July 2012, Philadelphia, FHIES 2012 Foundations of Health Information Engineering and Systems, an international symposium co-located with FM 2012, August 27-28, Paris, MeCBIC 2012, 6th Workshop on Membrane Computing and Biologically Inspired Process Calculi 8th September 2012, Newcastle.

8.1.4. Participation in Conferences

- José Echeveste, Jean-Louis Giavitto and Florent Jacquemard attended the 2012 International Computer Music Conference (Ljubljana, Slovenia, 9-14 September) and gave a talk on fault tolerance and language extensions of *Antescofo* [15].
- Jean-Louis Giavitto attended the:
  - Journées du GDR GPL (comité de direction)
  - SCW 2012, Valencia Spain, 5 June 2012 (chair)
  - UCNC 2012, Orleans and gave a tutorial on MGS

8.1.5. Invited Talks

Arshia Cont gave an invited talk at the Computer-Human Interactive Performance Symposium in University City London (UCL) in June 2012.

Arshia Cont and Jean-Louis Giavitto gave an invited talk in the Produce Time Symposium in June 2012 at IRCAM.

Jean-Louis Giavitto:

- tutorial on MGS at the 11th International Conference Unconventional Computation and Natural Computation UCNC 2012, 37 September, Orléans.
- *Non-Standard Multiset*, for the mini symposium Let's Imagine the Futur for the retirement of Jean-Pierre Banâtre, 7–8 Novembre, Rennes.
- *Programmation spatiale : application à l'analyse des contes de fées et à l'analyse musicale*, IPAC seminar, 8th March, Nancy.
- *Le projet AutoChem*, presentation closing the ANR Blanc project, at the NAR Symposium, January 2012.

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

Licence :

- Arshia Cont, Audio Technology Review, 2h/week, 3rd year Bachelors, Conservatoire Nationale Supérieure de la Musique et de la Danse de Paris (CNSMDP), Métier du Son, Paris, France.

Master :

- Arshia Cont, Machine Learning for Music, 18h, Master ATIAM, Université Pierre et Marie Curie (UPMC), Paris, France.

8.2.2. Supervision

PhD : Arnaud Dessein, *Computational methods of information geometry with real-time applications in audio signal processing*, Université Pierre et Marie Curie (UPMC), December 2012, Arshia Cont (co-directed with Gérard Assayag, IRCAM).

PhD in progress :

- José Echeveste, Synchronous Languages for Computer Music Composition and Performance, Started September 2011, co-directed by Jean-Louis Giavitto and Arshia Cont.

8.2.3. Juries

- Arshia Cont participated in the defense committee for Augustin Lefèvre, "Dictionary learning methods for single-channel audio source separation", ENS Cachan, September 2012.

- Arshia Cont participated in the defense committee for Mauricio Toro, "Structured Interactive Scores", University of Bordeaux 1 (LaBRI), September 2012.
- Florent Jacquemard was a examinator for the PhD of Hèdi Benzina, "Enforcing Virtualized Systems Security", ENS Cachan (LSV), December 2012.
- Jean-Louis Giavitto was a reviewer for the thesis of:
  - Héctor Fernández, "Flexible Coordination through the Chemical Metaphor for Service Infrastructures", University of Rennes 1 (IRISA), June 2012.
  - Jose David Fernández Rodríguez, "The Evolution of Diversity in the Structure and Function of Artificial Organisms", Universidad de Maálaga, Spain, January 2012.

8.3. Popularization

8.3.1. Cultural Events, Highlights

The Antescofo software developed and maintained by the team, has been featured in various prestigious international events with the following highlights:

- 13 World-premiers for live electronics and instruments by international composers such as Kaija Saariaho, Philippe Manoury, Steve Coleman and Marco Stroppa including one opera. See complete list.
- More than 30 world-class public performances around the world including: Antescofo premier with *New York Philharmonics* (June 2012), Lucerne Festivale (August 2012), Salle Pleyel (September 2012), Warsaw Autumn Festival (September 2012), Royal Albert Hall in London (August 2012), Vienna Festwochen (June 2012), Opéra Comique de Paris (June 2012), and more.
  See partial list.
- Two Giga-Hertz Prizes for two pieces strongly featuring *Antescofo* (life-time award for Emmanuel Nunes' *Einspielung*, young category for Lara Morciano).
- A documentary movie on Antescofo produced by Inria. (watch)

8.3.2. International Public Press

BBC Radio program featuring *Antescofo* on the occasion of a live performance in Royal Albert Hall.

8.3.3. National Public Press

Major public press mentions for *Antescofo* in 2012:

- *Antescofo: Au plus près de la partition*, Science et Avenir, July 2012.
- *L'ordinateur qui joue comme un musicien*, La recherche, June 2012.
- Quand l'informatique maîtrise le temps de l'interprétation musicale, France Culture Online , July 2012.
- Antescofo à l'avant-garde de l'informatique musicale, InterStices, November 2012.
- Sous le signe du calcul (with F. Reichenmann) in DocSciences, special issue on " Alan Turing : la pensée informatique" (DocSciences is a journal for college students published by the Rectorat de Versailles), June 2012. Published also on the InterStices web site, September 2012.
- À propos des nouveaux langages de programmation, audio podcast, InterStices, July 2012.

8.3.4. Exhibitions

- The *MuTant* team's *Antescofo* software was the demo feature for Inria's stand at the 13th exhibition for Mathematical Cultures and Games (Salon cultures et jeux mathématiques) in June 2012.
- *MuTant* members hold a stand during the "Fêtes de la science" on *Antescofo* at Inria Rocquencourt in October 2012.
- Jean-Louis Giavitto has participated to a public debate of the Bar des Sciences with G. Longo and G. Dowek on "Turing aurait 100 ans – De Turing au Qbit", 7th November.
- Jean-Louis Giavitto has participated to a public debate following the projection of Codebreaker: Alan Turing with C. Villani, B. Chazelle and J. Lassègue at the movie theater Grand Action, 17th December.

9. Bibliography

Major publications by the team in recent years

[1] A. CONT. *Realtime Multiple Pitch Observation using Sparse Non-negative Constraints*, in "International Symposium on Music Information Retrieval (ISMIR)", Victoria, Canada, 2006, http://hal.inria.fr/hal-00723223.

[2] A. CONT. *A coupled duration-focused architecture for realtime music to score alignment*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", 2010, vol. 32, n^{o} 6, p. 974-987, http://articles.ircam.fr/textes/Cont09a/.

[3] A. CONT. *On the creative use of score following and its impact on research*, in "Sound and Music Computing", Padova, Italy, July 2011, http://articles.ircam.fr/textes/Cont11a/.

[4] A. CONT, S. DUBNOV, G. ASSAYAG. *On the Information Geometry of Audio Streams with Applications to Similarity Computing*, in "IEEE Transactions on Audio, Speech and Language Processing", May 2011, vol. 19, n^{o} 4.

[5] A. DESSEIN, A. CONT, G. LEMAITRE. *Real-time polyphonic music transcription with non-negative matrix factorization and beta-divergence*, in "Proceedings of the 11th International Society for Music Information Retrieval Conference (ISMIR)", Utrecht, Netherlands, August 2010.

[6] J. ECHEVESTE, A. CONT, J.-L. GIAVITTO, F. JACQUEMARD. *Formalisation des relations temporelles entre une partition et une performance musicale dans un contexte d'accompagnement automatique*, in "Colloque Modélisation des Systèmes Réactifs (MSR)", Lille, France, November 2011, (Accepted).

[7] N. MONTECCHIO, A. CONT. *A Unified Approach to Real Time Audio-to-Score and Audio-to-Audio Alignment Using Sequential Montecarlo Inference Techniques*, in "Proceedings of International Conference on Acoustics, Speech and Signal Processing (ICASSP)", Prague, Czech Republic, May 2011.

[8] N. MONTECCHIO, A. CONT. *Accelerating the Mixing Phase in Studio Recording Productions by Automatic Audio Alignment*, in "12th International Symposium on Music Information Retrieval (ISMIR)", Miami, Florida, October 2011.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[9] A. DESSEIN. *Méthodes Computationnelles en Géométrie de l'Information et Applications Temps Réel au Traitement du Signal Audio*, Université Pierre et Marie Curie - Paris VI, December 2012, http://hal.inria.fr/tel-00768524.

Articles in International Peer-Reviewed Journals

[10] A. BOUHOULA, F. JACQUEMARD. *Sufficient completeness verification for conditional and constrained TRS*, in "Journal of Applied Logic", 2012, vol. 10, n^{o} 1, p. 127-143, http://dx.doi.org/10.1016/j.jal.2011.09.001.

[11] A. CONT. *Synchronisme musical et musiques mixtes: du temps écrit au temps produit*, in "Circuit : musiques contemporaines", May 2012, vol. 22, n^{o} 1, http://hal.inria.fr/hal-00698922.

[12] J.-L. GIAVITTO, H. KLAUDEL, F. POMMEREAU. *Integrated Regulatory Networks (IRNs): Spatially organized biochemical modules*, in "Journal of Theoretical Computer Science (TCS)", May 2012, vol. 431, p. 219-234 [*DOI :* 10.1016/J.TCS.2011.12.054], http://hal.inria.fr/hal-00769275.

Articles in National Peer-Reviewed Journals

[13] A. CONT. *Modélisation anticipative des systèmes musicaux. Reconnaissance, génération, synchronisation et programmation synchrone temps réel en informatique musicale*, in "Technique et Science Informatiques (TSI)", May 2012, vol. 31, n^{o} 3, p. 311-335 [*DOI :* 10.3166/TSI.31.311-335], http://hal.inria.fr/hal-00699290.

Articles in Non Peer-Reviewed Journals

[14] J.-L. GIAVITTO, F. REICHENMANN. *Sous le signe du calcul*, in "DocSciences", June 2012, vol. 14, n^{o} Alan Turing : la pensée informatique, p. 12-15, DocScience est une revue éditée par le CRDP de l'Académie de Versailles à destination des lycéens et de leurs professeurs., http://hal.inria.fr/hal-00769278.

International Conferences with Proceedings

[15] A. CONT, J. ECHEVESTE, J.-L. GIAVITTO, F. JACQUEMARD. *Correct Automatic Accompaniment Despite Machine Listening or Human Errors in Antescofo*, in "ICMC 2012 - International Computer Music Conference", Ljubljana, Slovenia, IRZU - the Institute for Sonic Arts Research, September 2012, http://hal.inria.fr/hal-00718854.

[16] F. JACQUEMARD, M. RUSINOWITCH. *Rewrite Closure and CF Hedge Automata*, in "7th International Conference on Language and Automata Theory and Application", Bilbao, Spain, Lecture Notes in Computer Science, Springer, 2013, http://hal.inria.fr/hal-00767719.

[17] A. SPICHER, O. MICHEL, J.-L. GIAVITTO. *Arbitrary Nesting of Spatial Computations*, in "Spatial Computing Workshop, IFAMAAS (International Foundation for Autonomous Agents and Multiagent Systems)", Valencia, Spain, 2012, p. 25-32, http://hal.inria.fr/hal-00769286.

[18] A. SPICHER, O. MICHEL, J.-L. GIAVITTO. *Spatial Computing in MGS*, in "11th International Conference Unconventional Computation & Natural Computation (UCNC 2012).", Orléans, France, Springer, 2012,

vol. 7445, p. 63-69, Additional material is available from http://www.spatial-computing.org/mgs/tutorial [*DOI :* 10.1007/978-3-642-32894-7_7], http://hal.inria.fr/hal-00769284.

Scientific Books (or Scientific Book chapters)

[19] J. BEAL, O. DULMAN, J.-L. GIAVITTO, A. SPICHER. *Proceedings of the Spatial Computing Workshop (SCW 2012) colocated with AAMAS (W21)*, IFAMAAS (International Foundation for Autonomous Agents and Multiagent Systems), June 2012, 60, http://hal.archives-ouvertes.fr/hal-00769288.

[20] A. DESSEIN, A. CONT, G. LEMAITRE. *Real-time detection of overlapping sound events with non-negative matrix factorization*, in "Matrix Information Geometry", F. NIELSEN, R. BHATIA (editors), Springer, 2012, p. 341-371 [*DOI :* 10.1007/978-3-642-30232-9_14], http://hal.inria.fr/hal-00708805.

Research Reports

[21] J. DIMINO, F. JACQUEMARD, L. SEGOUFIN. *FO2(<,+1,) on data trees, data tree automata and an extension of BVASS*, Laboratoire Spécification et Vérification [Cachan] - LSV , MUSYNC - Inria Paris-Rocquencourt , Sciences et Technologies de la Musique et du Son - STMS , DAHU - Inria Saclay - Ile de France, October 2012, 32, http://hal.inria.fr/docs/00/76/92/49/PDF/notes.pdf.

Scientific Popularization

[22] A. CONT. *L'ordinateur qui joue comme un musicien*, in "La Recherche", June 2012, n^{o} 465, p. 68-72, http://hal.inria.fr/hal-00701916.

[23] A. CONT, F. JACQUEMARD, P.-O. GAUMIN. *Antescofo à l'avant-garde de l'informatique musicale*, in "Interstices", November 2012, http://hal.inria.fr/hal-00753014.

Other Publications

[24] G. BAUDART. *Antescofo : Vers une programmation synchrone*, Master ATIAM, Université Pierre et Marie Curie (UPMC) et IRCAM, Paris, September 2012, 46, http://hal.inria.fr/hal-00730443.

[25] P. CUVILLIER. *Suivi de partition : étude du cadre multi-objet pour l'inférence de position*, ATIAM - UPMC, December 2012, 49, http://hal.inria.fr/hal-00762015.

[26] F. LÉA. *Temporal Analysis of Mixed Intrumental/Electronic Music Scores*, Ecole Centrale Paris, Châtenay-Malabry, November 2012, 51, http://hal.inria.fr/hal-00762004.

References in notes

[27] G. ASSAYAG, A. GERZSO (editors). *New Computational Paradigms for Computer Music*, Sciences de la musique, Editions Delatour, 6 2009.

[28] G. DI GIUGNO, J. KOTT. *Système 4X : processeur numérique de signal en temps réel*, Ircam, 1981, http://articles.ircam.fr/textes/DiGiugno81a/.

[29] J. ECHEVESTE, A. CONT, F. JACQUEMARD, J.-L. GIAVITTO. *Antescofo: A Domain Specific Language for Realtime Musician-Computer Interaction*, in "Discrete Event Dynamic Systems", 2012, in submission.

[30] E. A. LEE. *Computing Needs Time*, in "Communications of the ACM", May 2009, vol. 52, n^{o} 5, p. 70-79, http://chess.eecs.berkeley.edu/pubs/615.html, http://dl.acm.org/citation.cfm?id=1506409.1506426.

[31] E. LINDEMANN, M. STARKIER, F. DÉCHELLE. *The Ircam Musical Workstation : hardware Overview and Signal Processing Features*, in "ICMC: International Computer Music Conference", Glasgow, 1990, http://articles.ircam.fr/textes/Lindemann90a.

[32] G. LOY, C. ABBOTT. *Programming languages for computer music synthesis, performance, and composition*, in "ACM Computing Surveys (CSUR)", 1985, vol. 17, n^{o} 2, p. 235–265.

[33] R. P. S. MAHLER. *Statistical multisource-multitarget information fusion*, Artech House, 2007.

[34] L. MANDEL, F. PLATEAU. *Interactive Programming of Reactive Systems*, in "Proceedings of Model-driven High-level Programming of Embedded Systems (SLA++P'08)", Budapest, Hungary, Electronic Notes in Computer Science, Elsevier Science Publishers, April 2008, p. 44–59, http://www.lri.fr/~mandel/papiers/MandelPlateau-SLAP-2008.pdf.

[35] J. MCCARTNEY. *SuperCollider: a new real time synthesis language*, in "Proceedings of the International Computer Music Conference", 1996, http://www.audiosynth.com/icmc96paper.html.

[36] M. POUZET. *Lucid Synchrone, version 3*, in "Tutorial and reference manual. Université Paris-Sud, LRI", 2006.

[37] M. PUCKETTE. *The Patcher*, in "Proceedings of International Computer Music Conference (ICMC)", 1988, p. 420-429.

[38] M. PUCKETTE. *Combining Event and Signal Processing in the MAX Graphical Programming Environment*, in "Computer Music Journal", 1991, vol. 15, p. 68–77.

[39] R. ROWE. *Interactive music systems: machine listening and composing*, MIT Press, Cambridge, MA, USA, 1992.

[40] E. D. SCHEIRER. *Music listening systems*, MIT Media Lab, 2000, http://web.media.mit.edu/~tristan/Classes/MAS.945/Papers/Technical/Scheirer_Thesis.pdf.

[41] B. SELMAN, R. A. BROOKS, T. DEAN, E. HORVITZ, T. M. MITCHELL, N. J. NILSSON. *Challenge Problems for Artificial Intelligence (Panel Statements)*, in "AAAI/IAAI, Vol. 2", 1996, p. 1340-1345.

[42] R. STEINMETZ. *Analyzing the multimedia operating system*, in "MultiMedia, IEEE", 1995, vol. 2, n^{o} 1, p. 68 -84, http://dx.doi.org/10.1109/93.368605.

[43] R. STEINMETZ. *Human perception of jitter and media synchronization*, in "Selected Areas in Communications, IEEE Journal on", 1996, vol. 14, n^{o} 1, p. 61–72.

[44] B. VERCOE. *The Synthetic Performer in the Context of Live Performance*, in "Proceedings of the ICMC", 1984, p. 199–200.