From Authored to Produced Time in Computer-Musician Interactions

Arshia Cont, Jean-Louis Giavitto, Florent Jacquemard

From Authored to Produced Time in Computer-Musician Interactions

Arshia Cont, Jean-Louis Giavitto, Florent Jacquemard
MuTant team-project,
Ircam, CNRS, INRIA
1 pl. Stravinsky, 75004 Paris, France
first.last@ircam.fr

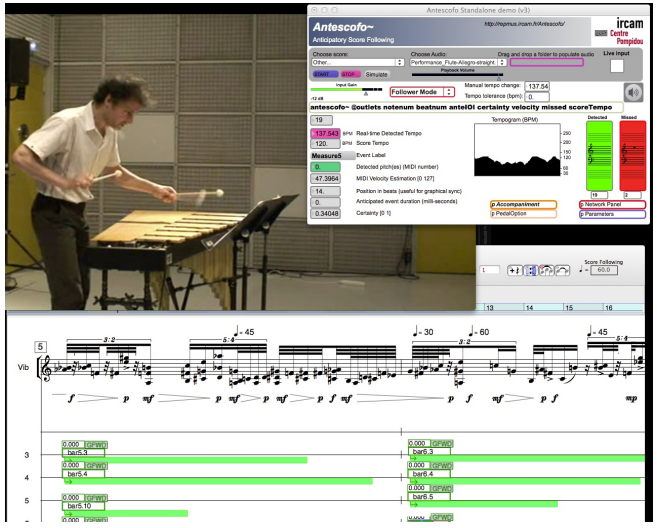

Figure 1: Antescofo: a performance synchronous language for real-time computer-human musical interaction

CHI'13, April 27 – May 2, 2013, Paris, France.
ACM 978-1-XXXX-XXXX-X/XX/XX.

Abstract
Human musicians have since long developed methods and formalisms for ensemble authoring and real-time coordination and synchronization of their actions. Bringing such capabilities to computers and providing them with the ability to take part in musical interactions with human musicians, poses interesting challenges for authoring of time and interaction and real-time coordination that we address in this paper in the context of Mixed Music and the *Antescofo* project.

ACM Classification Keywords
H.5.5 [Sound and Music Computing]: Methodologies and techniques, Systems

Introduction
The act of composing and authoring music is manifest in most civilizations who knew writing, and can be seen as a choreography of events and expectations in time. Western cultures has since long developed strong formalisms to allow sophisticated continuous interactions among several (small ensembles) or hundreds (large orchestra) of musicians. Ensemble music is the result of intrinsic combination of strong language formalisms (for authoring music) and performance mechanisms that allow synchrony, coordination of actions, and robustness of expected results in large ensemble music performances.

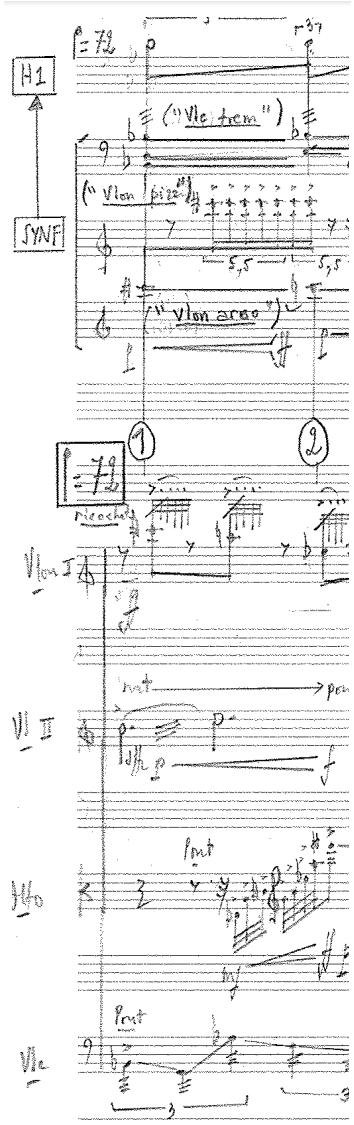

Figure 2: Excerpt of "Tensio" by composer Philippe Manoury for string quartet and live electronics

This paper exposes a research project that aims to bring the power of music authorship and real-time performance to computers and to enable interactive performances between human and computer musicians, in analogy to human ensembles. We show that the bottom-line of such systems is a tight coupling of real-time artificial listening systems with synchronous reactive languages both with heterogeneous and explicit notions of time.

Context and Background

The capacity for real-time synchronization and coordination is a common ability among trained musicians performing a music score that presents an interesting challenge for machine intelligence and CHI when computers are entered into the equation. To understand the challenges, we look at the notions of *Authoring Time* and *Performing Time* among musicians and specifically in the context of *mixed music*, i.e. when humans and computers are performing together.

Authoring Time

Authoring of musical time during the act of music composition is not merely an analysis of rhythmic content. Even in the most classical forms of music writing, composers deal with *heterogeneous Models of Time* that are well developed and understood among musicians [4]. In the context of *Mixed Music*, the heterogeneous nature of time becomes even more relevant due to the diversity of computational processes at stake. Figure 2 shows an excerpt of a recent piece by composer Philippe Manoury for a string quartet and live electronics. Similar to classical music scores, each staff (or line) represent different voices and the horizontal dimension represents (musical) time. The score for (human) instruments (lower four staves) is superposed to that of computer actions (top staves) despite their different computational nature.

Each computer staff is labeled with its appropriate program module specifying the nature and destination of computerized action, as well as the temporal progression of their parameters along the instrumental events. Note that they comprise metric/symbolic events as well as continuous parameters. The instrumental events (destined for human musicians) also go under various time-scales. To enable computer-human interactions in such contexts a dedicated time-aware language is needed to address such processes in real-time.

Time in Performance and Interaction

A music score such as the one in Figure 2 asks for predictable behavior among components while leaving several timing parameters to the stake of musicians for interpretation. During human performances of ensemble music, the perception of musical time undergoes an active listening strategy in which listener's expectations about future events play a role as important as the events themselves. Contrary to basic speech recognition, the temporal structure of musical expectation is a dynamic structure. Musicians are capable of synchronizing with each other and coordinating their actions despite all variances inherent in musical interpretation. This is achieved not only by visual/aural cues but also by on-line evaluation and anticipation of temporal structures.

In mixed music, a computer musician should be capable of coordinating its actions in interaction with the outside environment (the human musicians) via an active listening. In this context, timeliness is not a matter of performance optimization but a *semantic property*. Achieving *predictable behavior* and *scalibility* in such systems is far from being achieved in traditional interactive music systems.

Computer-Human Musical Interactions in Antescofo

A system capable of interacting with live human musicians should at least have two components: (1) A *Real-time Listening Machine*, capable of decoding required parameters (position, tempo, dynamics, etc) for synchronization and coordination of events, and (2) A *Synchronous Reactive Machine* for authoring and scheduling of computerized action and coordination thereof. Whereas the two topics have been extensively studied in the literature, their joint dynamics has rarely been addressed.

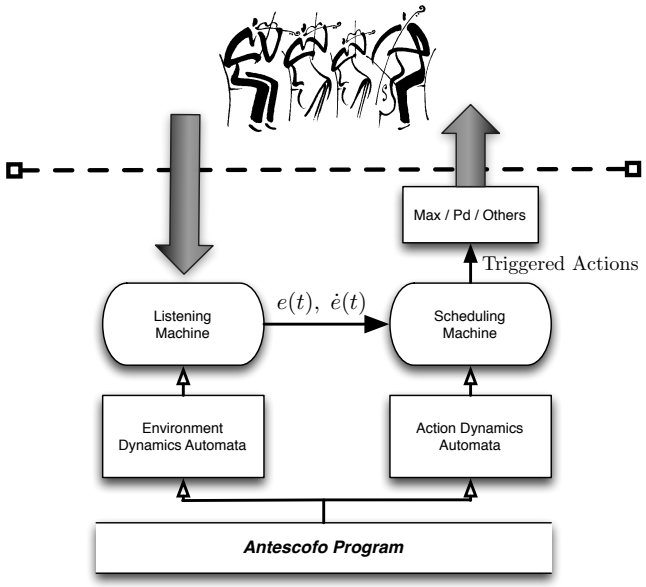

Figure 3: Antescofo's Runtime Components

Antescofo [6] is the codename for a system developed in our team that attempts to enable computer-human musical interactions in the context of mixed music. *Antescofo* provides an abstract programmer's model for an artificial musician in an ensemble with musical real-time constraints. Its runtime system is composed of two virtual machines as discussed above and shown in Figure 3. The *listening machine* handles artificial perception of the environment whereas the *Scheduling Machine* handles the interactions between computer programs and the outside environment dynamics. The two are jointly described using *Antescofo*'s dedicated language which attempts to describe a musical score. It's of high importance to note that it is the joint dynamics of the two virtual machines that enable *Antescofo* to undertake its role interactively. Below we will discuss important design aspects of *Antescofo* with regards to time.

Artificial Listening Machine

The listening machine in *Antescofo* is a state-of-the-art real-time audio-to-score alignment system using machine learning techniques which is fully described in [7]. For this abstract, it is important to note that the models employed in modeling time in the listening machine make use of *hybrid* graphical models for implicit and explicit handling of time. The listening machine itself is comprised of a coupling two agents working collaboratively and competitively to decode incoming events from the environment $e(t)$ and their tempo/speed $\dot{e}(t)$.

Synchronous Domain-Specific Programming Language

To specify mixed scores, *Antescofo* is equipped with a reactive coordination language which is similar to *synchronous* languages for programming real-time embedded systems such as ESTEREL [3]. Programs in this language, briefly, describe both environmental dynamics (expected human musician events) and computer actions. Figure 4 shows the equivalent of

```
BPM     72
TRILL    ( 8500 ) 0.5                IA...EVT-1
GFWD              arco
{
         syn_voices 1      pgm 40  bnk 7     vol 122
         syn_voices 1      56      120       @b(2)
         syn_voices 1      50      120       @b(2)
         syn_voices 1      46      120       @b(2)
2.0       syn_voices 1     53      120       @b(0.36)
         syn_voices 1      52      120       @b(0.36)
         syn_voices 1      58      120       @b(0.36)
0.36     syn_voices 1      69      120       @b(1.5)
         syn_voices 1      63      120       @b(1.5)
         syn_voices 1      59      120       @b(1.5)
1.5      syn_voices 1      66      120       @b(0.8)
         syn_voices 1      65      120       @b(0.8)
         syn_voices 1      62       120      @b(0.8)
}
CFWD          h1_trans     @grain 30ms
{
         0         0
         2.0       300
         0         0
         0.46      800
         0         0
         1.5       -500
}
GFWD     0.5       Pizzicati
{
         Bsyn_voices       4       vol 127
1/5.5    Bsyn_voices 4     85      127       100
1/5.5    Bsyn_voices 4     85      127       100
1/5.5    Bsyn_voices 4     85      127       100
1/5.5    Bsyn_voices 4     85      127       100
1/5.5    Bsyn_voices 4     85      127       100
1/5.5    Bsyn_voices 4     85      127       100
1/5.5    Bsyn_voices 4     85      127       100
}
NOTE     0         0.5
TRILL    (8100)    0.5
NOTE     0         0.5
TRILL    ( 7500)   0.5                                 IA...EVT-2
```

Figure 4: Program corresponding to the excerpt of "Tensio" (Fig. 2)

Figure 2 in *Antescofo*. Computer-musician interactions are defined logically by launching an action in reaction to environmental events with specified timing constraints. Temporalities of events and actions are organized globally in a program, like in traditional music scores. Some additional constructs, like error handling strategies, can also be specified. See [8, 9] for more details.

Timing Verification

Writing reliable mixed scores is a difficult task for composers. A particularly challenging issue in this context is to be able to predict the behavior of the system running the score, whatever the timing deviations in the musicians' performance. We are working on static analysis procedures for Antescofo scores, in order to assist composers with indications on some time critical aspects of their programs. For this purpose, we follow two verification approaches based on timed automata models [1] of mixed scores [9]. The first approach consists of replacing the delays in the model of the instrumental score by parameters representing possible timing in performance situations. Then a linear constraint on these parameters is inferred using *Imitator* employing inverse methods [2]. The constraint ensures that the ordering of actions defined by the score is preserved. A second approach is to provide a quantitative evaluation of the robustness to performance variations similar to [5] employing timed-game theory.

Perspectives

Antescofo as system and language is nowadays widely used in public concerts throughout the world. However, the problems posed by coupling real-time computing, perception and coordination between components and the environment are from being achieved. Among interesting challenges, one can mention coordination of different levels of computing, control and signal for instance, with heterogeneous timing constraints that has pushed us to look at *Globaly Asynchronous Locally Synchronous* approaches to computing. Furthermore, the joint dynamic approach in *Antescofo* can be leveraged theoretically and practically as Cyber-Physical Systems.

References

[1] Alur, R., and Dill, D. L. A theory of timed automata. *Theor. Comput. Sci. 126* (1994), 183–235.

[2] André, É., Fribourg, L., Kühne, U., and Soulat, R. IMITATOR 2.5: A tool for analyzing robustness in scheduling problems. In *18th Int. Symp. on Formal Methods*, vol. 7436 of *LNCS*, Springer (2012), 33–36.

[3] Berry, G., and Cosserat, L. The esterel synchronous programming language and its mathematical semantics. In *Seminar on Concurrency*, Springer (1985), 389–448.

[4] Boulez, P. *Penser la Musique Aujourd'hui*. 1964.

[5] Bouyer, P., Fahrenberg, U., Larsen, K. G., and Markey, N. Quantitative analysis of real-time systems using priced timed automata. *Com. ACM 54*, 9 (2011), 78–87.

[6] Cont, A. Antescofo: Anticipatory synchronization and control of interactive parameters in computer music. In *Proc. of Int. Comp. Music Conf. (ICMC)* (8 2008).

[7] Cont, A. A coupled duration-focused architecture for realtime music to score alignment. *IEEE Trans. Patt. Anal. and Mach. Intelligence 32*, 6 (2010), 974–987.

[8] Cont, A., Echeveste, J., Giavitto, J.-L., and Jacquemard, F. Correct Automatic Accompaniment Despite Machine Listening or Human Errors in Antescofo. In *ICMC* (2012).

[9] Echeveste, J., Cont, A., Jacquemard, F., and Giavitto, J.-L. Antescofo: A domain specific language for realtime musician-computer interaction. *Discrete Event Dynamic Systems* (2012). (submitted).