Music Interaction With Style

François Pachet
SONY CSL – Paris
e-mail : pachet@csl.sony.fr

Interacting versus Learning

The Continuator system may be seen as a realization of Ray Kurzweil's prophecy [Kurzweil, 1999], which predicts that "Human musicians routinely jam with cybernetic musicians".

Indeed, musical systems have traditionally been of one of two categories: interactive systems, such as the Karma musical workstation [Kay, 2000], [Risset & Van Duyne, 1996] take into account user input such as keystrokes or chords, but are not capable of learning, and use pre-programmed musical styles. On, the other hand, many musical learning systems have been designed to reproduce music "in the style of X" (e.g. [Cope, 1996], but these systems are intrinsically non interactive.

The goal of the Continuator project is to fusion both world, i.e. design interactive musical instruments that are able to learn. To this aim, we have addressed and solved the following issues:

- The ability to learn musical styles in real time, without a priori musical knowledge,
- The ability to take into account user input in real time during the generation process,
- The definition of interacting modes that give users degrees total control on the music generated, while enhancing their musical expressiveness.

To address these issues, we have developed a number of innovations with regards to prior art in musical interaction [Pachet, 2002]:

- A *robust and efficient representation* of musical phrases, taking into account polyphony, noise and arbitrary rhythm structures,
- An *Extended Multi-layer Markov model* to learn efficiently arbitrary corpuses of musical phrases in arbitrary styles. This model allows to generalize from musical phrases which are not exactly similar, and thus speeds up the learning phase drastically. As a result, the system is able to respond immediately to musical phrases in unknown styles.
- A Biasing mechanism that forces the Markov generation to specific harmonic regions. This mechanism allows users to control in real time the generation of the system, and therefore avoid the mechanical sounding effect of traditional music generation systems. This is achieved by introducing a probabilistic scheme in the Markov generation process based on a compromise between Markovian probability (the most expected continuation) and a fitness function (the most fitting continuation with regards to external input). Thanks to this probabilistic scheme, the Markov generation becomes *elastic*, and can be twisted by the user in a flexible way.

New Musical interaction modes

The Continuator's core system can be used in two primary modes:

- *Continuation or question/answer*. In this mode, the user plays musical phrases. Continuator detects phrases endings using a dynamical threshold, and produces a continuation to the input phrase in the style learnt so far,
- *Collaboration*. In this mode, the Continuator plays an "infinite" stream of music in some learnt style (for instance jazzy chord sequences). The user plays a melody and Continuator tries to adapt in real time its generation to the users input.

A Whole class of Interaction Systems

More generally, the Continuator is one instantiation of a larger class of interactive systems that are able to learn. The model developed in the Continuator (basically a real time Markov generator that is "twistable") is a key element for many systems that have to satisfy two classes of constraints:

- constraints of consistency with respect to a given language;
- constraint of fitness with a given context.

This class of systems includes for instance music playlist generation systems, which have to produce stylistically consistent music programs, which are consistent with user taste. Stylistic consistency may be modelled with a Markovian process, while user taste is typically a non-Markovian fitness function.

Another typical and non musical situation where the Continuator is experimented is dialog generation systems. From a given agent's viewpoint, and "interesting" dialog takes place when:

- The other agent expresses a consistent individuality, but also
- Remains focused on a shared subject of attention.

Here also, the consistency of personality can be successfully modelled as a Markovian process, but the concentration on a given subject is a typical non-Markovian fitness function. The Continuator proposes a unified paradigm to model and implement these systems. Furthermore, the compromise between the Marvovian and non-Markovian forces is explicated as a parameter, and can be controlled by the user. This parameter (called "attachment") allows users to change the personality of the interactive system (more reactive, or more stylistically consistent).

References

KURZWEIL, R. 1999. The Age of Spiritual Machines, Phoenix.

PACHET, F. 2002. Playing with Virtual Musicians: the Continuator in practice, IEEE Multimedia 9:3, pp. 77-82.

COPE, D. 1996. Experiments in Musical Intelligence Madison, WI: A-R Editions.

KAY, S. 2000. The Korg Karma music work station, Korg Inc.

RISSET, J.-C. AND VAN DUYNE, S. 1996. Real-time Performance Interaction with a Computer-Controlled Acoustic Piano, CMJ 20(1), pp. 62-75.