

Demonstration of Music Plus One — A Real-Time System for Automatic Orchestral Accompaniment

Christopher Raphael

School of Informatics
Indiana Univ.
Bloomington, IN, 47408

Abstract

We demonstrate a system that creates a real-time accompaniment for a live musician performing a non-improvisatory piece of music. The system listens to the live player by performing a hidden Markov model analysis of the player's acoustic signal. A belief network uses this information, a musical score, and past rehearsals, to create a sequence of evolving predictions for future note-onsets in the soloist and accompaniment. These predictions are used to guide the time-stretched resynthesis of prerecorded orchestral audio using a phase vocoder.

Past work on musical accompaniment systems (Dannenberg 1984), (Dannenberg & Mukaino 1988), (Grubb & Dannenberg 1997), (Vercoe & Puckette 1985), (Baird, Blevins, & Zahler 1993), (Raphael 2001), (Raphael 2002), has focused mostly on generating MIDI or other sparsely-parameterized accompaniments for live musicians. MIDI is particularly well-suited to piano and percussion music since, in these cases, the MIDI representation captures much of the interpretative quality of the performance. Most other acoustic instruments are considerably less deterministic and their MIDI equivalents are currently much less convincing. We present here a system that generates a complete *orchestral* accompaniment that *follows* a live player and *learns* and assimilates the player's interpretation through a series of rehearsals. Unlike our previous efforts, this system creates the accompaniment by synthesizing, in real time, an orchestral accompaniment using an actual audio recording.

Our system contains three separate modules called "Listen," "Predict" and "Synthesize" which perform tasks analogous to the human's hearing of the soloist, anticipating the future trajectory, and actual playing of the accompaniment.

Listen is based on a hidden Markov model that tracks the soloist's progress through the musical score. Each note in the score is modeled as a series of states, while the note models are chained together in left-to-right fashion to form the hidden process. Listen can be used "off-line," to estimate the onset times of each note in the score given the entire acoustic data file, or "on-line," by delivering real-time onset time estimates. In the on-line version these estimates are delivered

with as little latency as the local ambiguity of the acoustic data will allow.

Predict models our prior and learned notions of musical interpretation through a collection of hundreds of Gaussian random variables whose dependency structure is expressed by a belief network. The observable variables correspond to the note onset times estimated by Listen as well as the known accompaniment onset times. These variables are connected through a layer of hidden variables corresponding to unobservable quantities such as local tempo and rhythmic stress. During a rehearsal phase we learn the musical interpretations of soloist and accompaniment by estimating parameters for the distributions of the hidden variables using the EM algorithm. During live performance this module anticipates the future musical evolution by computing the most likely time of the next unplayed accompaniment note given all currently observed variables. Thus our future predictions are constantly being reassessed as new information is observed. These predictions form the basis on which the accompaniment is synthesized.

The Synthesize module begins with a digital recording of the orchestral accompaniment playing without the soloist. In some of our experiments we have used "Music Minus One" recordings which are created for the purpose of providing canned accompaniment for a live player. We create an "index" into this audio using an offline score matching algorithm (Raphael) which is hand-corrected for precision. Our playback of this recording will be warped in a time-variable way consistent with the real-time analysis of the live player's acoustic signal, as estimated through Listen, and the learned musical interpretation, as represented in Predict. This is accomplished through a phase vocoder. The phase vocoder synthesizes a time-varying trajectory through the orchestral performance by simulating the local frequency content, computed through a short-time Fourier transform, while preserving the phase continuity from frame to frame for each frequency.

This system has been used in a variety of traditional classical music with many musicians including students at the Indiana School of Music. In addition, the system creates interesting possibilities for composers writing for acoustic instruments and electronica. In this latter context the system provides virtually unlimited technical facility both in terms of playing fast notes and coordinate complex rhythms. Our

web site,

http://xavier.informatics.indiana.edu/~craphael/music_plus_one
provides examples of both of these scenarios.

We will provide a live demonstration of this system.

References

Baird, B.; Blevins, D.; and Zahler, N. 1993. Artificial intelligence and music: Implementing an interactive computer performer. *Computer Music Journal* 17(2):73–79.

Dannenberg, R., and Mukaino, H. 1988. New techniques for enhanced quality of computer accompaniment. In *Proceedings of the International Computer Music Conference, 1988*, 243–249. Int. Computer Music Assoc.

Dannenberg, R. 1984. An on-line algorithm for real-time accompaniment. In *Proceedings of the International Computer Music Conference, 1984*, 193–198. Int. Computer Music Assoc.

Grubb, L., and Dannenberg, R. 1997. A stochastic method of tracking a vocal performer. In *Proc. Int. Comp. Music Conf.*, 301–308. Int. Computer Music Assoc.

Raphael, C. Aligning music audio with symbolic scores using a hybrid graphical mode. *Machine Learning* to appear.

Raphael, C. 2001. A probabilistic expert system for automatic musical accompaniment. *Jour. of Comp. and Graph. Stats.* 10(3):487–512.

Raphael, C. 2002. A bayesian network for real-time musical accompaniment. In Dietterich, T.; Becker, S.; and Ghahramani, Z., eds., *Advances in Neural Information Processing Systems, NIPS 14*. MIT Press.

Vercoe, B., and Puckette, M. 1985. Synthetic rehearsal: Training the synthetic performer. In *Proceedings of the International Computer Music Conference, 1985*, 275–278. Int. Computer Music Assoc.