

# Musical Accompaniment Systems

Christopher Raphael  
School of Informatics  
Indiana University  
Bloomington, IN 47408

February 6, 2008

In February of 2007 we gave a live talk and demonstration of our musical accompaniment system, *Music Plus One*, in the Indiana University Jacobs School of Music — one of the top music schools in the United States. In this demonstration two violinists from the JSOM performed complete movements from the Beethoven and Sibelius concerti with flexible and responsive orchestral accompaniment provided by our MPO system. We would like to present video and audio from this demonstration, which documents some notable successes in our work with accompaniment systems. However, we would also like to describe some failures, highlighting challenging open problems in this area. Our presentation will include a live demonstration of the system in action.

Our system is composed of three parallel processes. First, the system *hears* and interprets the live incoming audio through a hidden Markov model whose hidden states correspond to positions within the musical score. Next, this running commentary produced by the hearing module serves as input to a prediction and scheduling engine. This prediction engine continually estimates and reestimates the time of the pending orchestra event, each time new information becomes available. These predictions serve as a trail of breadcrumbs that guide the third component of our system, the time-stretched resynthesis of a previously-recorded accompaniment.

Each of these three components performs reliably in a wide range of contexts, leading to a system that is robust enough for many realistic accompaniment scenarios. However, each of these components eventually falters, or performs naively, in cases we would like to add to our repertoire. For instance, while the hearing module reliably accommodates fast, rubato, and inaccurate playing with a monophonic solo instrument, the performance degrades as we treat vocal music, with its high degree of variability, as well as the complex polyphony of piano music. In both of these cases the musical score does not provide a clear enough description of the audio we encounter in practice.

Another open problem lies with our prediction mechanism. In this context future note onset times are estimated from a Kalman-filter-like model for musical timing, trained to each new piece and player. The model does well at copying the interpretation of the soloist, and, in doing so, exhibits some degree of musicality through imitation. However, some of the necessary musicality for the accompanist cannot be learned from the solo player. In such cases our system performs naively. Especially challenging is the ever-fluid opera context which we will demonstrate. This context requires the orchestra to perform as a musically equal partner, rather than with slavish obedience.

Finally, we seek a wide range of beautifully played and expertly recorded orchestral accompaniments to extend the range of our system. While a small number of such audio recordings are commercially available, no such “accompaniment-only” recordings exist for the vast majority of music of interest. For this reason, source separation is an especially attractive possibility for us. Furthermore, our work with score-matching automatically gives a correspondence between the symbolic score (the notes) and the audio data. Methods that use this correspondence to “desolo” the audio is a third ongoing and unsolved area of interest.

**topic: musical accompaniment systems**

**preference: oral**