Abstract

History has reversed the fascination people have had with mechanized performance for centuries (see Roads [10]), but it is only in our own century—more specifically, only in the last seven years—that we have been able to breach the technological gap which separated us from our desire for an intelligent interactive real-time computer performer. A computer performer that can follow a score, adjust its tempo and dynamics to accommodate the improvisations of a live performer, and, in some instances, contribute to the "ensemble" in ways that simulate what a live human instrumentalist does when performing.

To date, at least four individual approaches to building such a computer performer have come forward. They range in scale from a tool for conducting a computer orchestra, to a fully-synthetic "performer" capable of blending into, learning from, and anticipating the live performer with which it is playing (see Boulanger [8] and [9]). But, have we really simulated the behavior of human performers during a live performance, and how close is "synthetic" performance far then when the performance experience we take out of the laboratory and placed in the arena which live performance? More importantly, what are the demands of live composers on such new compositional resources?

This article centers on my own work and work done in collaboration with my colleagues Bridge Bied and Donald Stevens in creating what we call the "Artificially Intelligent Computer Performer" (AICP).[1] Previous publications (see Band, Stevens, and Zarlino [1], [2], [3], [4], and [5]) have centered on the creation and workings of the program. This essay describes the problems encountered when writing music for this new breed of computer instrument.

Introduction

Those who have worked with technological aids have always possessed a view which argues that the use of such devices should make work easier and give individuals more "freedom." The purpose remains the same in the realm of music.

Ensemble music—a music which requires two or more performers to realize compositions which demands a larger set of contingencies than most other art forms. Individuals wishing to play a composition in isolation, or more instruments must consider, and if, others with the necessary instruments to complete the ensemble are available. They must inquire as to the technical skill of the other instrumentalists and be sure that the literature to be performed is not beyond the expertise of those individuals. They must find a mutually agreeable time and place to rehearse and perform the composition and be confident that there are no overriding interpretative or philosophical differences in opinion that might derail the performance.

These are just a few of the considerations encountered when writing together a music ensemble. The problems grow geometrically as the size of the ensemble increases.

It should come as no surprise that there is a real need for a non-human performer that is able, as a minimum requirement, to satisfy the criteria listed above. In addition, such a performer should be capable of executing any number of instrumental parts, regardless of their difficulty it should be portable, and above all, responsive to the interpretative demands of the live infrumentalist. Earlier alternatives such as pre-recorded accompaniments were static, they did not accommodate the human performer's given interpretation. "Live computer music performance," whereas a human performer plays along with a sequenced accompaniment is not an alternative. Sequences are nothing more than pre-recorded digital files that substitute for magnetic tapes or phonograph recordings. The first portion of any such performance is distinguished from tradition only in the terms of the type of instrument employed, and once the sequence is triggered, tempo, dynamics, etc. are immutable. The interactive real-time intelligent computer performer is the first real alternative to those earlier attempts to satisfy the ensemble needs of individual performers.

The AICP

The AICP is a software program, written in the C language which holds in memory a complete music score. That score is translated into the memory of the computer with an "off the shelf" music notation program.[2] The AICP listens, via midi, to the live performer, performs in real-time, and tracks that performer trying to match the performer's input to the score in its memory. Once it has a "best guess" as to where the live performer is, it responds with its own part which is designated by the score.

The AICP is a Macintosh application which is presently configured to run on a Mac II or SE/30 with four megabytes of RAM. While the notion of such a program is not original in concept, (see Vercoe [12], Vercoe and Pudzota [13], Bloch and Demberg [3], Demberg [8], and Demberg and Muzio [9]) the AICP offers a number of interesting alternatives.

The AICP was developed to assemble a single human performer in a composition of that individual's choosing. It has been designed to accommodate compositions from both the traditional literature as well as contemporary compositions wherein the composer has sought to write for a human performer in combination with the sounds now made possible by technological advances in digital sound synthesis. It performs according to the instructions in the composer's score without stylistic bias. In essence, it is one solution to providing an ensemble to play traditional literature with a single live performer. It is a "live performer" capable of triggering multiple events, via midi, for the contemporary composer who is committed to combine a traditional human instrument(s) with the new capabilities.

ICMC 336
of computer instruments and to which it is essential to retain the freedom of interpretation usually attributed to live human performance. In each case, the requirements remain the same for the computer performer: only the constraints (requirements on synthesis hardware and software interfaces) change.

The integrality of human interpretative performance, eliminating the "music-minus-one syndrome," and the means by which human performers achieve that interpretation remain the single most important factor in our work. Modifying the knowledge base upon which a live performer relies has been a massive problem. Human modes of cognition are so numerous that conceptualization of these issues requires specification of the actions and a consequent reduction in the data to be analyzed. In the arts (in particular music performance) this scenario becomes even more complicated. The number of factors performed by each player, physically, as well as mentally, and the myriad of possible consequent actions and strategies, makes the problem all the more enigmatic to conceptization as specific performance circumstances change. Practical considerations have led us to breakdown specific areas of action into modules simulating what we believe are possible accounts of these mental and physical processes. At present these include a "pre-performance score consultation, a pattern-matching algorithm," a "score-matching" algorithm, a library of "possible types of score matches," an "adjust tempo" algorithm, an amplitude-averaging algorithm, and a "lost live performer" algorithm. This method of "forging out" what we believe are some of the most important and most frequent problems encountered during live performance is not a sufficient modeling of the human cognitive capacity during performance, but it seems to be an effective strategy for our present purpose (see Bald, Bevins, Zehler [1] and [2]).

These modules are called up prior to performance or during performance. The "pre-performance score consultation" is called when a score is loaded into the memory of the ACP. It enables the ACP to read a musical score and respond to the notation in that score in a manner not unlike its human counterpart. All other modules are called either at the start of a piece of "possible performance behaviors" which references four possible performance postures in reaction to a "best guess" matrix, "unvarnished," "played through," and "head match" (see Bald, Bevins, Zehler [2]). the "score matching algorithm" is responsible for recognizing where the live performer is in the score. The "adjusting tempo algorithm" is responsible for adjusting the rate of speed at which the accompaniment is performed. The "amplitude adjusting algorithm" scales amplitude values according to the dynamics played by the human performer, and the "lost live performer routine" calls a sequence of actions to be performed by the ACP when it has received sufficient evidence that the live performer has become unresponsive. Our task, then, has been to analyze the processes by which performers make choices when they are actively engaged in a live performance of chamber music and generalize those processes in a meaningful and useful adaptation for the computer.

Composing for the ACP and The-Equatedness of Musicality

When creating any new music oriented performing tool we must consider how and why it should be used. The need for an interactive computer performer has been demonstrated above, but alone we have such a performer it is incumbent on us to justify its usefulness through the creation of a literature which once interests and challenges the composer, human performer, and the listener.

Following in the steps of many others we choose the flute to score

our own composition in combination with the ACP. The combination seemed a logical one. The ACP will only accept monophonic input from the live performer, the flute (disregarding the use of multiphonics) seemed a natural instrument to link the capabilities of the program. The relative simplicity of the pitched signal emitted by the flute makes for easy tracking of pitch through a "pitch-to-midi-convertor," and the availability of the number of skilled and respected flutists supported the decision to write for this combination.

Reasonable theoretical assumptions and the imaginative capacities of composers sometimes make "strange bedfellows." The first two measures of the composition ultimately realized (example No. 1) pointed the way to a composition which, while it would certainly challenge the capabilities of the program, possibly asked more questions about its use as a compositional resource than it answered. The interpretation of this score segment, especially in MIDI data, brings forth a number of interesting problems regarding the program's capabilities and how composers think about the performances they employ.

Example No. 1

For a live human performer, deciding this passage presents little challenge. In fact, the only note of clarification needed would be an explanation of the "X" row heads as percussive attacks using key steps (an idiosynsty of the composer). But what of the ACP? What knowledge does it come to a rehearsing or performance with to sort out even as tame a departure from convention as this? As mentioned above we have tried to equip the ACP with the necessary knowledge to gain as much from a printed score as its human counterpart. Upon loading a score into the program a "pre-performance score consultation" is carried out by the ACP. This consultation allows the ACP to learn the performance specificities of the composition: tempo, dynamics, what the common practice use of "key" or "note" is, where tonic cadences are found (if it exist), to acknowledge the value of the rests (and the least note as dictated by the composition in conjunction with the tempo), decide on the time of prep tempo fluctuation to be allowed, locate possible "jump points," etc. The score, as the composer's blueprint for performance, is the source for any possible performance of, and the guide to any deviation from, the composition under consideration. In order to monitor the feedback loop which is obligatory for any ensemble situation the artificially intelligent performer must have the capability to keep track of and make decisions about how the performance is progressing. Consequently, we include modules such as the library of "possible performance behaviors," which, through a number of different scenarios (during performance), justifies the actions of the human performer when that player deviates from the notated score in ways that do not significantly undermine the progress of the composition. The "amplitude-adjusting algorithm," called during the slack time between ticks of the computer's clock, balances the amplitude of the signal controlled by the ACP when the human performer deviates from the dynamics indicated in the score.
Unfortunately, these variations still allow a great many features of performance to go unnoticed. They do not acknowledge the importance of or relationship of other subparts which contribute to interpretive performance. A communication gap exists between the composer's intentions for performance, the score, and the actual physical requirements in bringing forth sound from acoustic instruments.

Looking again, at example 1, in addition to the usual parameters for pitch, duration, time, and amplitude, the ACF must ask what does the "notated silly" line in the bass part with the superscript breve to be interpreted; how are the smaller notes ("quaver notes") to be included in the performance; what is the baseline of the notes denote; what does the dot above the note mean; and how is the fermata without the time signature different in performance from the earlier unencumbered fermata. To be sure, clarifying questions arise as well, but we must recognize the fact that our accepted music notation involves reference to a large number of performance techniques that are implicit in the definitions refined in the creation of the composer's score. These techniques are representable from the composer's instructions and are understood as such by virtue of the call for a particular kind of sound. We must also understand what consequences these implicit techniques have in performance and how we may bridge the gap, between the physicality of performance and the composers score.

The composers notations remain the score in such a way as to define his expectations for performance in MIDI, many of these questions are already answered for the ACF. Functions most easily utilized in the time and pitch space where they incorporate the duration of notes, prolongation of individual note values, and change of pitch and amplitude with time supply some of the information necessary for an acceptable performance. When mode of attack indications outside the time and pitch domain or develop complex variations of these values things become fuzzier.

We can easily program many of these note values as such harmonic rhythms; but the relationship is not reciprocal. The operation of our ACF is predicated on a fixed signal represented by a MIDI stream through parameters of pitch and time. Any distortion of that signal lead to significant problems regarding the capture of event-based data and the correlation of live performance to score. In addition, while encoding MIDI values for many of these techniques as possible a sort, it is no so simple, it only avoids our asking ourselves more difficult questions.

Along with the fixed position we find other problems posed by demands composers make on traditional instruments. Our notation reflects modifications of the fundi temporal components through the introduction of notes, "breathings" and harmonics, an alteration of the natural harmonic series of the instrument (example 2). These examples carry with them the performer's expectations which lie beyond simple parametric domain coordinates.

An analysis of these timbral methods reveals the inclusion of number comprising was a system in which the demands of flow of a continuous stream of well-defined event-based data for which MIDI protocol was essentially unsuited. The world of live performance composer takes for granted by his physical consequences these limitations make on performers and their instruments. They are an inconvenience, a mark, a noisy line and path too slow to communicate effectively through the MIDI bandpass, rending MIDI an inadequate source of communication for such activities. Instrumental performers have traditionally compensated for the difficulties encountered in producing such timbral transformations, but these intuitive practices are learned ad hoc, through expressions specific to the instrument under consideration. Any serious inquiry as to how in general such a phenomenon begins a broad question. That question asks if the encoding of the pitch, time, timbre (in some elementary way), and velocity components of a work result in sufficient information for the interactive performance of a composition faithful to the perception of the work by the composer? It is especially evident that real-time MIDI variations in acoustic instruments are not being read with any errors in ways that give a specification such as MIDI to communicate these variations effectively any useful way. As we attempt to produce more and more impressive machine performances, the knowledge base on which we draw for them begins to look inadequate.

Similar problems have arisen in the past. Our colleagues engaged in signal processing and sound synthesis struggled for years to find appropriate algorithms to imitate the complexity of tones found in ordinary musical instruments like the violin. The motion of the bow, v/ . and the rubbing of the strings in a constantly changing velocity and pressure while simultaneously changing the left hand frequency of vibration, makes it impossible to find a "synthetic" solution to the problem of generating such a sound through the synthesis methods available at the time. Chosen real-time "sampling" was a breakthrough that changed the way we thought about the problem and broaden our view with which to model real-world systems. The approach of sampling accommodates our psychophysical knowledge of certain perceptual properties of acoustic instruments and mode of attack transitions, allowed us to produce extremely realistic copies of acoustic instruments while "cheating" on the digital outputs necessary for such sounds (see Straun [1]).

The idea of "crossing" on data necessary for unfailing a particular source type of event may be necessary step in our quest to recreate the basic elements of human behavior and human interaction. The problem is not this. It should be clear that the result of artificially supplying MIDI to ACP with information which will allow it to interpret all the beavers of the ambient sound transmission of our human performer would give us an acceptable performance, a performance which might even be used as a successful example of a Turing test, but it is the same sort of "typing" that has taken place in other situations and does not enter the essential question about the processes of the mind during human performance. We accept programs like the ACP as members of the instrument ensemble only with the limitations that other, more traditional, instruments have, "confusing the means with the end." In effect, we don't want to expend human knowledge base for this we are now with quite another story.

**Example 2:**

![Breathy.png](attachment:Breathy.png)

ICMC 338
The interaction between human performer and "synthetic performer, to date, is an elementary approximation of the actual teaching possibilities used by human performers. The following excerpt from our composition (example 3) shows the fluxus part accompanied by the AICP (for convenience the AICP part has been garthered into a "grand staff" without regard for the various timbres it is using). When faced with the coordination of two parts such as these, we must ask ourselves what information each of the performer has to one and the other, and how that information can be effectively transmitted to the AICP. These inquiries, once again, beg a broader question, a question which asks what a performative performance and how may we reproduce it under the constantly changing dynamic of live performance?

To understand performative performance we need to begin with a more "event-based" data structure and its effectiveness for performance in the area of computer music. The time and pitch transients which define "gestural data structure" must be incorporated into the "performance model" which can produce effective life-like performances. Research must continue to concentrate on parameters beyond the detection of what results as individual musical lines or parts and focus on the coordination of those parts as they mesh with the composer's score.

When we begin to speak of the "musicability of a performance" conversation usually turns away from the technical aspects of a performance (the mechanics of playing the instrument). We talk about "putting beyond the notes," but what does this really mean? If we are speaking of "interpretive performance" as simply the grouping of phrases, periods, sections, etc. in coordination with changes in tempo, then a simple segmentation of a composition based on tempo varying will be sufficient to model this activity. Our belief is that we are talking about something much more complicated and that as we multiply the number of people involved in a composition the problem becomes exacerbated. The consequence of these actions on the performance actually given and the means by which this information can be related to the AICP is a terribly difficult and tedious task. But these issues which, for the most part, are overlooked by researches, including ourselves, because of the effectiveness with which most tracking algorithms can be made to work.

Returning to example 3, we note that in order to do this, each player must have their own channel, but in the future, trying to break through to the next climactic passage. These types of performance patterns are abundant throughout music literature and depend on each note of the passage being interpreted with an individual nuance of its own. While we expect the flutes as part of its more creative role to take on this responsibility, the mmusical input to the AICP can only be encoded with one set of parameters for playing this passage. These parameters will seem to be interpreted differently each time as the live flute's reading of the passage changes, but the part of the AICP will never really initiate as original interpretation of it's own. it will never set a new tempo, rephrase a passage, initiate a ritard, etc. Conclusion and Speculations

In the final analysis, tools for interactive performance such as the AICP are a major step forward in the collaboration between people and machines, but they lack many of the essential ingredients that go into a truly "musical performance." In order to get to the bottom of this problem we need to model the actual mental processes in question. When we can truly ask the correct questions about this process of interactive interactive performance we get to the heart of what "musical knowledge" really is. This use of musical knowledge is what allows composers to use the human/machine collaboration in a truly imaginative and creative way. For now we will content to use the tools we have available to us and open up to be as resourceful as possible in programming them to make performances which are "natural.

We continue to probe the questions within our minds for alternate ways of viewing these problems and in so doing are humbled by how much yet there is to know.

Bibliography


ICMC 339