Creating a Computer-Assistant for Performance

John ffitch and Julian Padget

Abstract

Centuries of experience with acoustic instruments means there is a body of knowledge about and understanding of what can be done with a physical instrument, and this knowledge is refined and propagated through the interactions of teachers and pupils. That experience has nothing to tell us about synthetic instruments and the one-off sounds for which such instruments are often created, nor about the instruments' sonic character, which itself is rarely explored, and certainly the intuitive phrasing and shaping applied by a human player is not even attempted. We advocate a new research direction, building on current activities in agent research and in performance knowledge, which could lead us to a new kind of electro-acoustic music, both in public performance and private listening.

1 Introduction

Man has been creating music from before written history; we do not know anything about the performance of the caveman's music, but it is clear that in many cultures the professional musician has been a feature for thousands of years. One feature that this engenders is the emergence of virtuosi performers on a wide range of instruments, where skills are passed from teacher to pupil, and are gradually adapted and improved to meet changing circumstances, such as instrument variation and mechanical inventions that modify the character of the instrument.

From the standpoint of a composer within the Western art tradition this means that one can assume that in writing for say a flute, that the performer will be able to reproduce the intent, except in the most extreme circumstances, without the composer having to explain everything. We can write in the knowledge that the performer knows the traditional notation, and will play in tune with the appropriate interpretation of loudness, crescendos and all the common effects of the instrument. In fact even more strongly, we can assume that the instrumentalist can and will *interpret* the composer's intent in the light of a long tradition. In writing for an unfamiliar instrument the performer can tell the composer what is possible, identify effects that may be what is sought, and act not as a dumb servant but as a collaborator.

There is nothing new or exciting in this, as we would expect nothing less.

In this paper we want to contrast this with the current state of electro-acoustic music, and in so doing identify a technology that can be used to overcome the problems that we have noted. In consequence, we propose a line of research that can make the technical detail of computer-composers work easier, and, also, identify a potential way in which the whole future of 'canned' music in the home, on CD or whatever replaces it, can become in a significant way, alive.

2 The Trials of the Electro-Acoustic Composer

In the current state of the art, the computer-composer has to design the instrument in addition to writing the score (Boulanger, 2000).

2.1 Instrument Problems

While each composer may have a collection of favourite sound designs, on most occasions the instruments (that is the specification of how the sound is created) are made anew for the composition, and the full characteristics of this artificial instrument may not be understood. Furthermore, in the composition only a small range of the potential is typically used.

In the early days of the use of computers to create musical compositions, the long period between writing the program and listening to the resulting sound could serve as an excuse for, or an explanation of the narrowness of the exploration. As our computers become more powerful, and at least some real-time synthesis is possible, it becomes more possible for the composer to begin that exploration. Boulanger has reported that recently the sounds he created for *Trapped in Convert* can be heard in real-time, and he discovered a much wider range of sounds than he had supposed, and so 14 years later he has written *At Last, Free* which uses the self-same instruments but in new sonic regions.

But even so the time scale is insignificant compared with the years of experience of the main classical instruments, or even the more modern additions to the orchestra. In practice it is not possible for a professional composer to discover much about his created instrument in the time available.

Often the result is imaginative music played by what might be conceived as an amateur or student band.

As a particular example of how hard this process can be be refer to the Chaotic Oscillator(Dobson and Fitch, 1995). The authors invented a particular non-linear process which they showed was capable of a range of interesting sounds, as well as wolf-whistles and cracked notes. With significant effort they managed to produce a few examples, but the exploration was a significant effort, and despite the oscillator being available in code for Csound, it has never been released as it is too experimental for musicians ???or composers??? to use. Their related non-linear filter(Dobson and ffitch, 1996) is less prone to instabilities and so is available, but is still mainly used in its linear case.

Until some assistance is found for the exploration and understanding of instruments the electro-acoustic composer will remain hamstrung by the complexities and the detail involved in working with synthesized sounds, which interfere in the larger scale processes of working with the instruments which generate them and beyond that the composition into which they fit. We propose below (section 4.1) a program of work to try to alleviate these problems.

2.2 Score Problems

In addition to the problems of learning to play the instrument, the composer has to consider elements which are either obvious or easy in the physical scenario. In particular, what we are thinking about are the problems of phrasing and emphasis. Returning to our classical flute player we could mark the score with a traditional phrase mark, and the performer would almost without conscious effort play the notes as a recognizable phrase. Without the presence of the performer, our electro-acoustic composer has to describe the low-level elements of the phrase for each note.

It has been known for some time that humans create phrases by a combination of speed changes, legato notes and amplitude envelopes (Clarke and Windsor, 1997). Research teams like that at KTH Stockholm (Friberg and Sundberg, 1995) have investigated this phenomenon in great detail, and have even proposed a performance grammar (in the formal sense). What we would like for our case is a grammar and an implementation of the semantics as transformations to amplitude and timing of notes so as to perform and express on this new instrument.

In writing ensemble pieces another feature of the physical world is apparent which currently does not have a counterpart in the virtual world. Writing for a small ensemble very little is needed by the composer to indicate in which part the tune lies, or which parts are adding textures rather than foreground events. We can call this the problem of balance between the components. Our overworked computer-composer has to adjust each note-event for balance between the parts, in addition to the existing problems of overall amplitude, as typified by ensuring that the audio material remains within the 16 bits of precision of CD/DAT technology.

Constructing performance details such as we have identified, phrasing and balance, can be a tedious exercise and could act to deter the use of the technology. While an established composer may merely find this tiresome, for the would-be composer it is likely to dampen any enthusiasm, and precipitate a change to other activities. But if balance and phrasing are not undertaken it is all too possible that some inadvertent masking of musical material will take place, or at worst, a form of homogenization through equal weighting that makes the piece less interesting.

3 Some Computer Science

A branch of artificial intelligence which has been particularly active recently is called Agents (Nwana, 1996). The idea is to have a collection of autonomous and probably distributed computer programs, cooperating to solve a particular problem. Each such agent may have particular areas of expertise, which they offer to the collection of agents whenever they see a subproblem for which they are suited. The decisions on whether to accept of reject such help rests with the other agents who are also attempting to solve the problem.

A particularly revealing application of agent technology can be seen in artificial societies, and the collaboration between them and humanity(Padget, 1999). It is from this area that ideas about the creation of language as a phenomenon between agents, who wish to cooperate but are initially without shared experience on which to impose a shared language. It has been shown by Steels (1996) that language can emerge from the self-organization of such artificial life.

That seems a distance from the problem posed in section 2, but we maintain that this research into artificial life is exactly what is needed to fulfill our artistic needs. By creating artificial 'living' programs it might be possible to accelerate the learning of how to play the instruments, and with a little more mechanism attack the other problems.

4 The Synthetic Ensemble

The solution we are advocating to both the scoring and instrumental design questions is to analyze the situation from the standpoint of computational *Agents*(Wooldridge and Jennings, 1994). Important characteristics of agents are quoted as autonomy, responsiveness, proactiveness and social ability: all of these have important parts to play in music making.

We observe that music performance is characterized by a stylized form of social interaction, where music replaces conversation (to use anthropomorphic terms, but one could argue they are indistinguishable from an ontological perspective). Furthermore, the agents that take part in music performance make decisions about exactly how to play a note or a phrase, within the framework of a score (perhaps) and sometimes within the temporal framework defined by a conductor. Thus, in music-making we can observe instantiations of the abstract characteristics of agents listed above, and which sustain our belief in the value of exploring the problem of musical synthesis within an agent framework(Boulanger and ffitch, 2001).

The autonomous agents in a musical performance are the individual instrumentalists, and possibly a conductor who is interpreting the intentions of the composer, or adding an individual stylistic variation.

4.1 The Synthetic Performer

So we are advocating that the synthetic instruments should be played by synthetic players in the guise of agents. These performance agents should explore their virtual instruments, and offer examples and feedback to the composer. This conversation needs to be at a moderately high level of abstraction, labelling sonic regions as interesting, nasty, acceptable and so forth, and the numerical parameters which capture this should not be immediately communicated to the composer. Similarly the performer agent can learn the relationship between the amplitude parameter and the output amplitude achieved. This is in preparation for being able to interpret instructions to play at uniform loudness for a scale, or to adjust for later balancing. In effect the agent would start the process of becoming a proficient player, or perhaps even a virtuoso.

4.2 The Conductor

The conductor agent exists to mediate in the performance between the composer's intentions and the vocabulary of the performers. It is the conductor who should correct balance issues, instruct performers as to the placing of phrases, and act as a communication conduit to the composer. We view the interaction to be at a musical level; instructions like *play this section staccato* are greatly to be preferred to *for this section reduce the note duration by 5%, unless you are the melody in which case that should be 4%*, or even worse a list of detailed event-by-event numerical changes.

5 The Current State

This suggestion for a future vision is not just wishful thinking. We have described elsewhere (ffitch and Padget, 2001) the details of the progress we have made towards this aim; here we just give a general view of this work.

We have started by concentrated on phrasing. As we know that human players create phrases by speeding up and slowing down over the duration of a phrase, with an amplitude envelope, usually of the crescendo-diminuendo variety, this seems a simple starting place, but one which can also demonstrate the validity (or otherwise) of the approach.

We have an experimental version of Csound (Boulanger, 2000) which, out of real-time, reacts to the identification of phrases. It then rewrites the score to apply some simple time modulation and amplitude variation. The command format allows the composer to select from a small number of particular phrasing methods. So far we have only used this system to experiment with the kinds of action which are musically interesting.

Similarly, live performers emphasize the main melodic line by playing a little louder, and playing just before the beat. Our system can imitate this, in a rigid way.

We are also starting a more computational experiment to construct a computer agent architecture that is capable of learning the instruments written in Csound, and so react in a high-level manner to instructions from a composer or conductor, commands like “louder”, “more staccato”. At the time of writing this work is only just starting and it is premature to say much about its success. We are using a standard agent software system, and initially are not contemplating real-time operation.

6 The Vision

The process we have described so far has been aimed at the composer, a person for whom we have a special respect. But the underlying agent technology and the autonomous agents can be used for an even more exciting possible future.

Imagine a scene some years from now:

> Our music-lover is considering listening to a musical work from his collection, either from a DVD or downloaded from a broadband network in an extended MPEG-4 format. He picks up the control console and directs that the music begins. Somehow the performance seems a little too relaxed for his mood, so he uses the console to ask the conductor agent to change the music to be more strident; within moments the music has changed as the conductor agent acts on the instructions from the communication network. The individual players who are dealing with the individual lines each changes their performance on command from the conductor. The listener considers a fine adjustment, but on reflection feels that this conductor has captured his intent, which is not so surprising as the conductor has been learning this individual's means of expression, and have been adjusting the unseen model accordingly. As the music moves to the slow movement a small request for a more prominent bass line is made.

Clearly this is currently well beyond our capabilities, but if the music is encoded not as simple, stereo (or more) channels, but as individual instruction streams for each instrument, rather like the premix in an orchestral recording, together with general control information, the performance agents can adjust within a range set by the composer or music producer for personal preference, room acoustics, or temporal circumstances.

There are a range of interesting questions in this scenario. The Human Computer Interaction elements will need careful design; the composers will need to give ranges of performance detail (possibly assisted by composer assistant agents), and there are human problems concerning whether the listener should be able to override the artist's intentions. But what this is suggesting is that what we currently have as 'canned' music could be a live experience, perhaps never repeating the same performance, but doing so in a manner much more determined than the arbitrary application of random numbers (see ffitch (2000) for the current state of that). The system could also make concerts involving tape music more compelling, as one would hear the sound projectionist's interpretation, without the additional problems of diffusion.

Most of the technology already exists, and is waiting to be used to this end. Already real-time modification of synthesis has been demonstrated (*et al*, 2000); we are suggesting that such systems can be controlled by agents that have learnt something of the musical world.

7 Conclusion

We have advocated a programme of research which applies the significant advances made in Software Agents to music, and especially electro-acoustic composition. A number of the problems which limit the development of composers can be alleviated by such an approach.

We have made a small start in this work, concentrating on phrasing, but with will (and funding!) this could revolutionize not only electro-acoustic composition, but also the consumption of music within the domestic environment.

This is our vision for an artistic future.

References

Richard Boulanger, editor. *The Csound Book: Tutorials in Software Synthesis and Sound Design.* MIT Press, 2000.

Richard Boulanger and John ffitch. Teaching software synthesis: from hardware to software synthesis. In *KlangArt 1999 - New Music Technology*, 2001. to apopear.

E. Clarke and W. L. Windsor. Expressive timing and dynamics in real and artificial music performances: using an algorithm as an analyitical tool. *Music Perception*, 15:127–152, 1997.

R. W. Dobson and J. P. Fitch. Experiments with chaotic oscillators. In *ICMC'95: Digital Playgrounds*, Banff, Canada, pages 45–48. ICMA and Banff Centre for the Arts, September 1995.

Richard Dobson and John ffitch. Experiments with non-linear filters; discovering excitable regions. In *On the Edge*, pages 405–408. ICMA, ICMA and HKUST, August 1996.

F. Pachet *et al.* ??? In Ioannis Zannos, editor, *ICMC2000*, pages ??–?? ICMA, August 2000.

John ffitch. *The Csound Book: Tutorials in Software Synthesis and Sound Design*, chapter 16: A Look at Random Numbers, Noise and Chaos with Csound. (ed. Boulanger) MIT Press, February 2000.

John ffitch and Julian Padget. Towards computer-assistant performance in electro-acoustic music. In *ICMC2001*. ICMC, 2001. Submitted.

A. Friberg and J. Sundberg. *Grammars for Music Performance.* KTH, 1995.

H. S. Nwana. Software agents: An overview. *Knowledge Engineering Review*, 11(3):205–244, September 1996.

Julian Padget, editor. *Collaboration between Human and Artificial Societies.* Number 1624 in LNAI. Springer-Verlag, 1999.

Luc Steels. Synthesising the origins of language and meaning using co-evolution, self-organisation and level formation. Technical report, Artificial Intelligence Laboratory, Vrije Universiteit Brussel, 1996.

M. Wooldridge and N. Jennings. *Agent Theories, Architectures, and Languages: A Survey*, volume 890 of *LNAI.* Springer-Verlag, 1994.