A Behavioral or Actor-Based Paradigm of Sound Production and Studio Composition

Kevin Elliott
MediArts/ New Media Research
The Banff Centre for the Arts
Banff, Alberta Canada
kellien@acs.ucalgary.ca

Abstract
Dynamic and non-linear forces of sound design and composition will assume increasing importance in the near future. Emerging technologies will enable high-bandwidth media systems capable of rich instruction, and it seems likely that an increasingly tech-literate society will embrace and demand content that offers personal control and involvement. Examples given frequently include virtual reality, interactive video and games, and navigable movies/music/theater books. Composers and musicians seeking to take advantage of the possibilities of this digital culture will be able to develop new relationships between sound materials and other mediums, as well as with their audiences. This paper discusses an approach to tools and working methods for this new environment, based on practical production and research experience at The Banff Centre for the Arts.

1 Introduction: linear vs. non-linear

The established, familiar models of music composition and sound distribution are linear: the music score, the tape recording or compact disc, the concert hall performance. Some movement has been made toward exploring non-linear models. In past centuries, and the early decades of this century, composers fiddled with dice-game modular scores. The effort has gathered momentum in recent decades, as digital technologies such as MIDI, hypermedia, and random access storage have allowed composers to explore interactive music performance systems, CD-ROM or laser disc, and audio installations that respond to audience input or the environment.

There are a number of artistic motives for adopting non-linear forms and methods. Underlying most of them is a desire to break down the long-standing social, professional, and logistical barriers of formal music-making and its process. The linear world has imposed a structure that creates hierarchies of creativity: creators vs. consumers, performers vs. audiences vs. composers, intuition vs. intellect vs. skill. Composers and artists working in non-linear mediums are deconstructing these hierarchies, and creating an aesthetic framework that supports collaboration, personal involvement, and individual creativity.

The trend towards the non-linear is notably well-developed in music and sound, specifically computer music and electroacoustics, as compared to other art disciplines and communication mediums. This is in part because the hierarchical relationships of music making and sound production are so heavily structured that they demand (or can accommodate) change, and in part because the technical means of working with sound are relatively mature. In other words, composers and musicians have long wanted or needed to involve interactive processes in their work, and have had relatively easy access to the resources needed to do so.

1.1 Digital culture

Nonetheless, current examples of non-linear technology and content, even in the sound domain, are of primitive quality and limited distribution. Composers working with interactive intentions are constrained by tools that are either clumsy (MIDI instrument-sets, modems-and-phone-lines), rare (high-end DSP-engines, fast data networks), or both clumsy and rare (computer-controlled laser discs, non-keyboard input devices).

6A.2 174
ICMC Proceedings 1993
This is about to change. It is widely recognized that a communications and computing revolution is taking place, as the functions of computer, home entertainment system, television, and telecom services merge and mutate. This is happening because people crave responsive, interactive access to information and technology, and want to have personal control of their environment and creative/creative spaces. Industry is scrambling to support the vision of a media-connected society in which information, cultural material, entertainment, and social communication is delivered interactively at the individual’s home, in the workplace, and on the move.

At The Banff Centre for the Arts, the networked society described above is referred to as the digital culture. We view it as an opportunity for artists of vision to play a seminal role in developing an enlightened and humanized technological landscape. The Banff Centre’s New Media Research initiative proposes that artists can define, inform, and design the required new technologies in critical ways that science cannot. Composers and artists working with sound will be at the forefront of this process.

2 Current digital audio practice

Typical approaches to the application of digital and computer technology in audio productions involve emulation in the digital domain of familiar analog studio practices. DAT replaces 2-track tape, digital play-list editing replaces tape splicing, computer-based console automation extends digital precision and manipulation to the “analog” gestures of the engineer, and MIDI simplifies or coordinates the parametric control of audio processing devices. There are obvious advantages and disadvantages to this kind of computerization in the sound studio, much as there are known advantages and disadvantages to the use of computer word processors in place of typewriters or pencil and paper.

In a typical audio production or composition situation, the artist attempts to translate an abstract, visual image or musical intuition into reality through a carefully-calculated technical process. Each component aspect of the visual image is constructed parametrically, and in isolation. A simple visual gesture such as causing a sound to move realistically from point to point in apparent space, or changing the size or reflectivity of an imagined acoustic environment may require dozens of detailed technical manipulations. These manipulations (fader moves, EQ pot settings, adjustments of delay or reflection parameters) will usually have little or no intuitive relationship to the intended musical result (changing the emotional import of the sound, moving it around). The process of translating creative intention into technical implementation is often filtered through imprecise language (artist to engineer), which results in a further estrangement of the intention from the result.

3 Towards non-linear audio

Extending the consideration of digital audio practice to computer composition and sound design, it appears that there are two distinct working methods: real-time and non-real-time. The real-time model includes MIDI-mediated or MIDI-accompanied performance, audio processing (e.g. pitch shifting, delays, convolution) in live performance, and live mixing of multitrack tape or live source materials. The non-real-time model includes audio technique as described in the previous section, as well as traditional computer synthesis and scoring (e.g. CSound, HMSL), and MIDI sequencing. Hybrid methods have evolved, as well. For example, there are compositions and systems in which structured sound material is prepared or composed out of real-time (e.g. samples, sound bites, MIDI sequences) are triggered in real-time performance or interactive situations.

It should be apparent that non-real-time methods lead to linear results. Materials prepared out of real-time will generally be reproduced later in a linear manner, by performers or through technological means. A case might be made that some non-real-time compositions produce non-linear results, as in algorithmic processes that respond to some form of random or chance stimulus.

Real-time methods may create either linear or non-linear results, depending on the nature of real-time stimulus or interaction. The imposing digital culture, with its promise of rich, interactive and immersive experience, will demand real-time performance.

4 Requirements of non-linear audio

In the digital culture, new resources and working methods will be required for the soundcreation, control and manipulation. In turn, these will lead to new paradigms of composition and sound design. Artists working with non-linear sound at The Banff Centre, especially in the context of virtual environment and interactive multimedia systems, are helping to define these new resources and paradigms. The following section outlines a number of needs and intentions that have emerged thus far.

ICMC Proceedings 1993 177 6A.2
Sound objects of arbitrary duration. In a non-linear or navigable sound space, performers or participants need freedom to move at their own pace, perhaps stopping altogether at an interesting point. It is not enough to merely loop, repeat or sustain a sound or musical figure at such a dwelling point. The sound must evolve as it continues, or respond in some way to the act of lingering.

Sound objects linked to images or other artifacts of interactive space. Elements in immersive environments are not frozen points; they are dynamic and may be approached or experienced in numerous ways. The sounds related to such elements (or "free-standing" sound objects without relational links) must sustain similar dimensionality.

Sound objects defined in 3-space. Participants in digital culture expect sounds (and images) to be presented with tangible localization and rich spatial resolution. The also wish to control the spatial display of sounds.

Real-time control of sound processing. The processing of sound elements (e.g. filtering, loudness contour, timbral modification etc.) must respond to interactive stimuli.

Relational control of sound processing. The processing of sound elements should also respond to relationships defined with other components of the interactive space.

Control of sound at "local" and "global" levels. It will be helpful to have the ability to apply similar processes and controls to individual sound objects, definable groups of sound objects, or an entire soundscape.

Intuitive interfaces that extend the body. Different sound objects and dynamics will demand a variety of interfaces, both physical and in software, to match interactive possibilities with bodily expression.

5 Paradigms of sound production

Without wishing to attempt a detailed classification, I will identify some familiar examples of sound production paradigms that may be contrasted with paradigms to come.

The strict non-real-time method may be referred to as the compositional paradigm: the content is defined and structured in advance, and played out in a linear fashion. In the real-time domain, there are at least two common models: the improvisational paradigm, in which content is created and/or controlled spontaneously by performers or participants; and the algorithmic paradigm, in which some defined system of rules acts upon material according to the input of performers or participants. All of these existing paradigms involve procedural methodology.

5.1 The behavioral paradigm

Various approaches to sound creation in digital culture will evolve in response to the requirements of non-linear audio enumerated above. One example of a suitable paradigm is based on the notion of "actors" with sound characteristics defined as behaviors. In contrast to the procedural paradigms of current practice, the behavioral paradigm is based on concepts of object-oriented design.

In this model, sound elements become dynamic objects, actors with lives of their own on an imaginary acoustic stage. The stage is a virtual aural space, modeled through physical description (perhaps using 3D drawing and texturing tools). Many such modeled acoustics may coexist, mingle, overlap, or morph from one to another. Sound objects are introduced into these acoustic environments, vested with behaviors and dynamic audio characteristics that represent a holistic understanding of a musical intention. Three examples of common production scenarios will suggest how this paradigm is relevant in the production of non-linear sound materials.

Navigable music. This term implies a musical construct that may be explored at the whim of an audience or performer. This poses problems of structure. How long does a section or "cue" last? When and where does the piece end? How will the rhythm or texture of the work be controlled? Using the behavioral approach, a composer could define individual sound objects (instead of melodic phrases, for instance) with characteristics such as collections of note data, a timbral envelope of variable duration and scale, a location in virtual space, and linkage to an interface gesture or trigger. Collections of individual sound objects could be organized in multi-dimensional space and relative time to create a larger segment or complete work (hierarchies of musical structure). The audience or performer has freedom to explore the musical architecture within the constraints of an interface designed by the composer (or possibly by the user).
Interactive video. A characteristic problem here is to establish appropriate relationships between images and sound, when the transitions from image to image, scene to scene, cannot be predicted. The problem is compounded if the imagery is multi-layered (independent layers) or if individual elements of imagery are subject to transformation. Behavioral modeling of sound allows the designer to attach sound material and dynamics to each discrete visual element. This can include relational protocols for transitions between images. Acoustic environments are defined as a separate class in the object hierarchy, so that the placement of sound elements (dialogue, sound effects) on different stages (a forest, a living room, a gymnasium) is not only possible but trivial.

Virtual environment. A participant is exploring a mythical landscape based on real-world samples (visual and aural) of a babbling brook, a sub-alpine meadow and a waterfall. The scene is populated with a number of creatures (a snake, a bird, a bear, a cat), some of them speaking dialogue or making musical sounds. The participant may choose to inhabit the body of any of these creatures, and observe the environment through the senses and the intellectual disposition of the creature. The sound design problems are complex: the composer must not only create a convincing soundscape with numerous variables of perspective, dynamics, mix and processing, but must also transform this infinitely variable soundscape through the arbitrary choice of creature sensory filters and point of view. The permutations are practically uncalculable, and the task cannot be accomplished by procedural methods. With behavioral modeling, each sound element may be defined acoustically in relatively simple terms, relationships specified amongst them, and the whole system is free to evolve according to a manageable process that reflects real-world models.

6 MixNet**: an enabling environment

One of the principal design objectives of the MixNet sound processing and production environment currently under development at The Banff Centre for the Arts and State University of New York at Buffalo (SUNY) is to support new paradigms of composition and sound design, such as the behavioral model proposed here. The MixNet system is presented elsewhere in these proceedings (see P2.01 "MixNet**: — A Comprehensive Real-time Automated Production System: Progress Report"), but a brief summary of relevant features is provided here.

In essence, MixNet is a GUI that makes calls to signal processing resources via a distributed object environment that supports real-time operation. MixNet includes a tool kit of basic interface objects and templates required to accomplish standard audio studio tasks such as multi-track recording, automated mixing and processing, and multi-channel wave form editing. It will also encompass novel features such as arbitrary user-configurability, distributed tasking among multiple media service "engines," connectivity with other production mediums and environments, and varying scales of data representation. MixNet and its underlying operating system are object-oriented environments, with the implied attributes of encapsulation, abstraction, polymorphism, class hierarchies and inheritance.

Key features of MixNet that support the proposed behavioral paradigm are the binding of processes and control characteristics (methods) to individual sound objects, and the unification of handling of mixed data types (e.g., control signals, sound files, code modules, MIDI data, live inputs). The binding of processes to sound objects (expressed differently: a unified approach to sound treatment at all levels, from minute detail to grand design) allows the definition of audio characteristics and musical detail without explicit reference to linear placement or finite duration. The uniform handling of mixed data types allows the creation of relationships and linkages with other sound and media elements, interface components, spatial locations, etc.

7 Summary

The imminent digital culture of high-bandwidth connectivity, distributed media computing, and interactive/disseminated communication will create new possibilities and demands for sound creation, composition and music production. The behavioral paradigm proposed here, which presents sound objects as actors upon a virtual acoustic stage, may be helpful to composers or sound designers working in the new medium. The object-oriented MixNet sound production environment is being developed to implement the features required to support this paradigm.

8 Acknowledgments

Thanks are due to the community of composers, computer musicians, artists, technologists and sound designers who have contributed personal experience and ideas to the development of expertise in interactive systems and immersive environments through their work at The Banff Centre for the Arts. In particular, I wish to acknowledge the work of Dorota Blazczak, principal sound designer and audio programmer for virtual environments, and Rick Bigley, project manager at Banff for the MixNet environment.