Australasian Computer Music Conference 2000

interFACES

Contents

* Denotes peer review of full paper
† Denotes peer review by abstract

Editors: Andrew Brown and Richard Wilding

Reviewers: Andrew Brown, David Hirst, Gordon Monro, Alastair Riddell, Greg Schiemer, Richard Wilding.

Published by
Australasian Computer Music Conference 2000
C/- Andrew R. Brown
Queensland University of Technology
Victoria Park Rd.
Kelvin Grove, QLD. 4059.
Australia.
Email: a.brown@qut.edu.au

ISBN: 0-646-39885-7

Teaching Electro-Acoustic (EA) Composition to the Uninitiated

Roger Alsop
Fairfield Victoria
email r.alsop@vac.unimelb.edu.au , rogeralsop@hotmail.com

Abstract

This paper looks at the processes used in teaching EA composition to students in the Production Centre at the Victorian College of the Arts. This situation provides a unique and interesting set of problems and opportunities, not least that many of the students have little to no prior interest in EA composition per se. This lack of interest gives the most exciting opportunities to a teacher, the two most pressing being to develop strategies to create an interest in EA and then developing a sense in each student that it is possible for them to be an effective composer in this field.

The VCA provides a rich environment for this to happen as people actively involved in the creative process and who are continuously producing works surround each student. This fever of creativity rubs off on students, who may initially come to learn the more prosaic aspects of theatre production and feel less able or inclined to approach their studies from a creative and personally expressive position. Fortunately this position changes during the student's tenure to that of being a vital cog in a creative engine or, more often, as an inspiration for and developer of fine artistic creations.

The topics discussed here include: the different processes used in introducing composition tools; developing an aural acuity and listening skills; introducing EA concepts and confidence with software; the collaborative approach to composition required by the situation; and how the general audience receives the EA offerings of the students. The software and hardware available, and the reasons for those choices, in the Production Centre is also discussed.

The creative milieu

As the Production Centre supports the Drama, Dance and Music schools, and to a lesser degree the Art and Film and Television schools, the opportunity and the expectation for students to create sound works to be used in each of these media is plentiful. Within this milieu each student is confronted with the varied sonic needs of each art form. Filling these needs can range from providing banal "sound effects" for a theatre piece to a complete EA composition for a dance piece, or vice-versa. This process is done in collaboration with directors, choreographers, and any other commissioning artist.

A major benefit for the EA composer when observing practitioners from other, non-musical, art forms is that the compositional and structural concerns and approaches unique to these practitioners are shown to the EA composer.

The students initial interest

When a student first enters the Production Centre at the VCA he or she may have minimal interest in the EA world, which is what prompted the title of this paper. The students main areas of work while at the VCA are in theatre and dance for which they gain expertise in all areas production; set design, lighting design and operation, costume making, stage and production management and sound design and operation, which is what this paper addresses.

In the structure of the course students spend first year developing skills in all of these areas, then select two that are of special interest to study in second year and then spend much of third year seconded to companies and acting in their chosen field. Their initial understanding of sound deliberately produced to create an effect may be that it is just that, a "sound effect" directed by the script or director. Music for a theatrical event may be what is heard as the patrons enter and leave the auditorium or music from a CD that is considered to enhance the action of the play at strategic points in the narrative. Often the main interest of the students when entering the course is to develop skills sufficient to ensure that they will be capable of working in theatre or similar areas when they have completed their course, and the Production Centre at the VCA has an excellent record in its graduates working in their chosen fields.

For this reason production students often begin from a physical point to their work. Questions such as "Do I need three phase power for my sound application?", "Can I rearrange the tracks on a Minidisk player?" and "What is the best reverb setting for this sound?" have been asked. The first two questions are to do with the physical needs of setting up a sound system for a performance production, they are questions do not require any personal or artistic input from the student. This approach to sound making can result in a prosaic and less adventurous, creative, or expressive approach to developing a sound design for a theatre or dance production. The third question presumes that sound can be affected and presented in a way as to change that sounds nature and the listener's perception of it; it is the beginning of EA composing. From this point the possibilities in manipulating the aural environment to create an affect in the listener become apparent and available to the student. Now the student begins their personal exploration and development as a composer.

The heuristic nature of each students development as an EA composer

In this situation students develop as composers mainly though experimentation. They have a few tools available to them which allow sound manipulation and fit two categories: real time manipulation, for use in the field; and computer based collation and manipulation, which allows the composer to offer a more finished product to the director, choreographer or whoever is commissioning the work. Students are encouraged to create their own sounds through recordings they do themselves and through manipulating the work of others, either commercially available sound effects or the music composed by others.

By making their own recordings skills in studio and field recording techniques are developed. The range of recording equipment available is of very high quality and, though limited in numbers, very broad in scope. Through the use of different microphones and their placement skills in critical listening and sound manipulation are developed. An example of this is the degree of manipulation that is by simply changing the pick-up pattern of a microphone from hyper-cardioid to omni-directional in the same space. Another example of microphone technique which, as a by product develops a more acute sense of listening, is through using shotgun microphones being trained on sources that are not usually heard by themselves, such as the sound of the device on a tram that connects it to the overhead electric wire. This introduction to the more unusual sounds around us creates a wider sonic pallet for the budding composer to use and inherently gives a sense that sound in itself is an easily manipulated source which can be used in creating an affect in the listener.

When students start manipulating the work of others they begin to develop compositional strategies. This is particularly interesting when working with pre-recorded compositions. In this case there is an already existing musical framework that has to be recognised, assessed and then subverted to the needs of the students work. Much of this recognition and assessment is done subconsciously and intuitively without the possible benefit, or obfuscation, of traditional musical thinking, the students simply try to achieve an appropriate aural environment for the commission. From here the students model personal compositional strategies from the pieces they are using, often with the sense that they have discovered the particular strategies themselves.

Composition from an aural position

Most students begin their development as a composer from a position that may best be described as "I may not know music (sound) but I know what I like. This can be a very powerful position for an EA composer to come from. As the broad palette of sound itself does not readily fit the concepts of counterpoint, harmonic rhythm, tension and resolution, musical form, set and specific pitch or any other traditional musical concepts or compositional devices as traditional instrumental music, these concepts or devices do not encumber the nascent EA composer. This does not mean that each of these devices are not used when composing the sound track for a dance or aural environment for a play, just that they are used with the aural affect on the dance or play being the prime consideration. The interesting phenomena is that students readily use these compositional devices from the point of having invented them in order to fulfil the necessities imposed by the commission they are working on, and they then assess the usefulness of each device within the context of that commission.

When teaching these devices are not ignored but introduced as examples present themselves and when problems in a specific commission arise that may best be solved with these devices. By immediately applying musical devices when they are required students generally gain a good grasp of the essential ideas and purposes behind each device and their possible applications. From this point they then develop expertise in areas of traditional musical composition and hone this expertise through listening to the results. Conversely, when the prime relationship to the composition is to serve and enhance another art form issues of compositional technique, style and genre carry less weight and may even not be best not considered.

There are also special problems and questions that may arise for the more experienced EA composer; problems as varied as the quality of recordings to be manipulated or having to learn the intricacies of a computer language, and questions such as: are sounds of similar envelop characteristics more closely related than sounds of similar timbre characteristics; is sound grouping, and its inherent problems, relevant to the current, unique commission and resulting composition; how important is the metaphoric qualities of a particular sound in the composition; and do the metaphoric qualities of a sound mask other qualities of that sound. These questions are tackled by the student within the context and confines of a well-defined commission. It is these confines that allow the developing EA composer to answer these questions first through analysing the work and its relevance to the commission themselves and then through the filter of collaboration with the director, choreographer or musician they are working with.

Non-musical media as the structural element

The sound design student and their collaborator develop EA compositions that usually have as their source a non-musical inspiration. This inspiration may be a choreographer's assignment, a sculptor's presentation, a score or Foley for a film, or a text and its interpretation into a play. Through working with artists in these various fields Production Centre students gain an insight into the many unique, and universal, problems and the possible solutions, encountered in the compositional processes of each of these art forms.

The composer David Ward-Steinman explored the structural aspects of various art forms and, at the time he was at La Trobe University in 1989, had found that there is a structural similarity between art forms, they generally follow a simple ABA structure. In its simplest form the observer is taken from a point of rest to a point of activity which then resolves in a return to a point of rest. This finding is becoming increasingly arguable in the wonderfully diverse time we live in, but can hold true if given considerable latitude in interpretation. The salient point for the EA composer is that each art form will have its own techniques of navigating this ABA, or any other, structure and the composer is then able to learn and apply these techniques to their own creative pursuits.

To use a film as an example of this process: The first step is to examine the intentions of a script and how its structural aspects reflect that intention. This could be the style of language used, the points at which characters are introduced and removed, the changes in the nature of the dialogue, or where the pivotal points in the script reside. The next step is to develop a sonic palette which reflects the intention and the final step is to compose this palette into a sound-scape which enhances both the scripts intention and its structure.

Most films will usually have a point close to the beginning where characters are introduced and the scene is set, followed by a period of exposition where each character is developed and the audience given a more detailed view of their motivations. This period of acclimatisation and rest is then followed by a disruption generated either by the characters themselves or by an unexpected outside agent. Finally the characters negotiate a method of returning to a state of rest at which point the play ends; Ward-Steinmans ABA structure.

This is an extremely simplified analysis, and of course the devil and delights of any artistic expression lie in the detail. The emerging EA composer can give detailed study to the script and, with its creators and interpreters, develop a coherent composition based on the details of the films structure. This relieves the composer of responsibility of developing an idea from scratch, attending to the minutiae of its development alone, its presentation and so on. With these activities accounted for the student can devote their attention to developing an understanding the needs of a piece and to explore their own approaches to fulfilling these needs.

Once a composition has been made the students are able to see the effect of their work, first on their co-participants and then on an audience proper. While the presentation of the composition is usually part of a multi-media event the importance of its contribution to the event is always evident and acknowledged. The VCA is an excellent forum for the emerging artist, it provides a very encouraging environment in which every attempt is made to enhance the talents of each student while at the same time not abiding charlatanism or half-heartedness.

Broadening the audience

The audience base for electronically manipulated sound and EA composition is developing through the regular use of this media in the performance arts. As creators of performance art become more aware of the possibilities available through electronic sound manipulation they become more prepared to both commission works from EA composers and are more able to communicate economically with composers. This is seen in the preparedness of choreographers at the VCA to leave the tradition of using off the shelf pre-recorded music to developing dance specific EA compositions for their work. Performance artists will also try their hand at composing themselves but when doing this most realise the benefits of working collaboratively with a composer who can introduce musical strategies and who can often also double as a technician.

This collaborative approach to EA composition gives the student a unique opportunity to educate, or at least inform, directors, choreographers and traditional musicians and composers, in the possibilities available within the world of EA. Students and their collaborators initial explorations often revolve around the simplest alterations to natural sounds, such as pitch shifting or imposing spatial motion on the sounds. This simple exploration creates a sudden insight into the impact of sound and the possible effect it can have.

Through this introduction to EA practitioners of other performance arts become more versed in the EA arts, often searching out EA music to use in their performances and composers to commission. The result is an increasingly wider audience than the usual enthusiasts and, what is more important, a non-musical audience.

Audio technology in the Production Centre

One of the most pressing requirements of the sound students at the VCA Production Centre is that works must be completed quickly. There is little time to develop software or technology, this is not in the brief of the Production Centre, nor is there time to use software or technology that requires either a long time to learn or produce sounds. For this reason two standard audio editing systems, Cool Edit on a PC, and ProTools on a Macintosh are used. These two platforms were chosen because introduce the students to industry standards and because they offer fairly generic interfaces and procedures for assembling sounds.

The sound manipulation software currently used in the Production Centre includes the standard set of plug-ins bundled with Cool Edit Pro. When used effectively and imaginatively, these algorithms can produce many varied and interesting sounds by altering any kind of pre-recorded sound wave file. On the Macintosh SoundHack, which have different attributes to the Windows plug-ins available in Cool Edit, is used as the sound manipulation tool. This programme requires more time to explore, operating slower than real time, and a greater understanding of the EA lexicon. When this understanding is developed it proves to be a more powerful sound manipulation programme and offers scope for more conceptual and well developed composition processes.

Other programmes include AudioMulch, which is very powerful programme that allows quite sophisticated signal modification, it can also be used to quickly generate dance club style music. This attribute is quite exciting to the early EA composer as it can provide a composition of their own that often fits with their personal aesthetic. This instant gratification helps develop confidence to explore the programme further and to develop skill in using the programme.

A visual composing programme called Coagula is also extensively used by students. This programme allows a student to draw an image on the computer screen that it then translates into an audio signal through sine wave synthesis. Coagula has wonderful advantages in that it too offers instant gratification as well as drawing a direct relationship between a physical and a sonic gesture. This is particularly powerful when working with choreographers and other visual artists.

The hardware available includes a standard set of signal processors, digital and analogue mixing desks and sampling hardware. These items are very important to the developing EA composer as they are simple, often logical in their operation and allow the composer to improvise with sound. This process is specially useful when working on student drama productions as these often require the Production Centre student to improvise a sound track using CDs. The student is encouraged to manipulate the music from the CDs by cross fades, multiple sound tracks and signal modification. The degree to which a student practises and executes these manipulations is dependant mostly on their imagination. It has been noticed that students may first assume that the audio manipulation power available from a mixing desk and two CD players is minimal but after practice and exploration discover that they have a powerful audio manipulation system available to them.

This set of EA devices may seem quite limited at first glimpse but it does fulfil the requirements of a mostly real time studio for the EA composer. The very limitations of the equipment available require a reliance on imagination and artistry to generate high quality product and not a reliance on expensive hard or soft ware to produce interesting sounds. Other considerations when acquiring the equipment available has been chosen to give students as broad a range of the technology available in the performance art industry and that each student be able to equip themselves with a similar studio without too much expense and, as software becomes available it is included in the studios if appropriate.

Teaching strategies

As can be seen from much of the above it is the environment of the VCA and the unique place the Production Centre has in that environment that provides the greatest role in developing the emerging EA composer. Much of the teaching is done on an individual level, assisting students in the needs of each commission as it arrives. There is also a great deal of interaction and communication between the students regarding the discoveries they have made and the strategies they have developed. These strategies may be to do with compositional problems, developing ideas from and for commissions, exploring new ways to use software and to develop communication skills with other performance artists. This interaction between students is encouraged as a philosophy of the Production Centre and is a formal part of the teaching process.

When teaching sound design per se students are first given scripts as the starting point for an EA composition. They are encouraged to look at the script without consideration for any aspect of the performance other than sound. This gives each student an opportunity to look at the intentions and structures of the script and to let their imagination run free in enhancing those aspects.

The results are often astounding, basing the process on ideas and then discovering how to implement them encourages the student to approach learning software and hardware in a very directed way. This means that little time is wasted in less useful learning and students are vitally interested in the information they receive.

Students are given very few assignments as most of their work is in the production of performances. The few assignments that are handed out are based on developing rudimentary ability with the technology and are assessed on the students musical use of the technology. This approach enhances the way students approach their commissions as they develop a more detailed awareness of the musicality of sound and a sense of musicality in the sound designs they make.

Most assessment in the Production Centre is based on how well the commission each student is given is fulfilled and this is assessed by staff in all areas of theatre production practise. The broadness in aesthetics of each staff member means that students have a variety of opinions from which to develop a greater sense of their work.

Conclusion

This approach to teaching EA composition does not follow the usual steps. Students are not considered to have any musical background or knowledge to be expressed through EA compositions, they are not expected to have any interest or ability in creating software for musical application, and these more typical aspects of EA teaching are not a high priority in the teaching process.

What is considered most important in the teaching is a sense that EA hardware and software are expressive tools to be used for artistic expression. From this base many students are developing an interest in the more traditional areas of EA simply for the purpose of increasing their creative abilities.

Modes of Compositional Engagement

Andrew R. Brown
Queensland University of Technology
Victoria Park Rd.
Kelvin Grove, QLD. 4059
Australia
Email: a.brown@qut.edu.au

Introduction

> The composer, in short, is a man of mystery to most people, and the composer's workshop an unapproachable ivory tower.
> Aaron Copeland, 1957:20

A story is recounted by C.F. Ramuz (Cook 1990) of the bewilderment of two women as they sat on park benches outside the house where Stravinsky was working, upon hearing the sounds of Stravinsky's compositional process at the piano and with assorted percussion instruments.

The story demonstrates not only the historically central role of audible feedback during composition but the unusual way in which the composer uses audible feedback during that process. The story illustrates for Cook that;

> when a composer composes at an instrument, he does not listen to what he is playing in quite the manner in which a performing musician does (or in which the composer himself would when performing) (1990: 187).

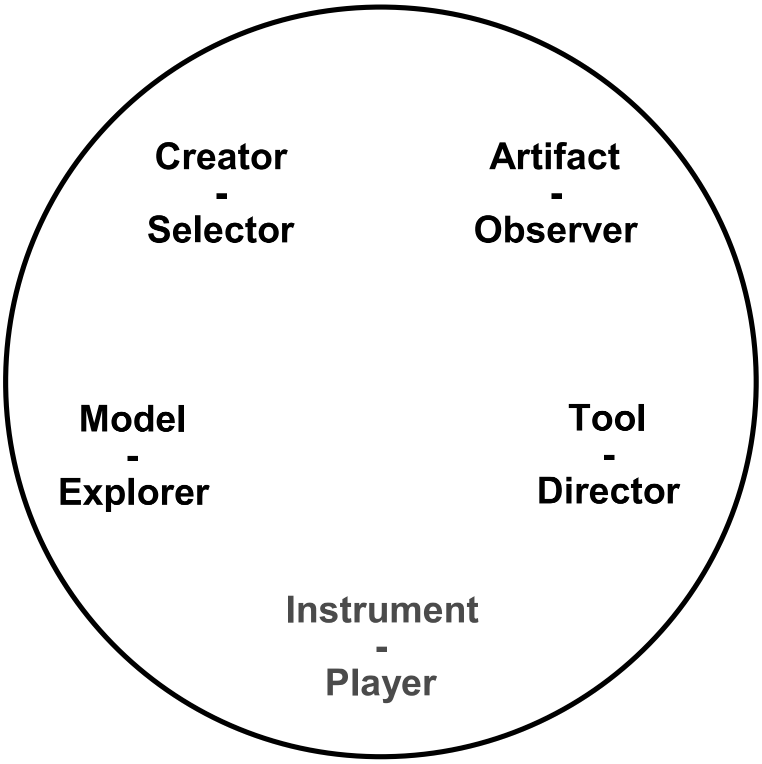
The image of the solitary composer at work creates an almost mystical impression of the compositional process. The activity of a composer interacting with musical materials, as evident with Stravinsky banging away at the piano, is what I characterise as compositional engagement.

In this paper I will present some perspectives on compositional engagement which have arisen from my PhD study into composition processes with computers. I believe that while the specific tasks and outcomes of making music with computers might be different from another medium, such as paper or piano, the composer's engagement is fundamentally similar.

The main point I wish to make is that compositional engagement is multifaceted. There are a range of ways in which composers interact with musical materials and mediums. I refer to these as modes of compositional engagement. A composer works in different modes depending upon the stage the composition is at, the type of material they are working with, the compositional methods they are using, and the medium in which they are working. A composer might work predominantly in one particular mode of compositional engagement, change between modes quickly, or display features of different modes concurrently.

I do not make any claim that the modes of composition are complete nor mutually exclusive, but in describing compositional activity in this way it is hoped that the compositional process and in particular the role of the computer as a compositional medium might become less mystical and more available to debate and understanding.

An overview of the modes of compositional engagement



I have identified five modes of compositional engagement. They depict the composer as:

- Observer,
- Director,
- Player,
- Explorer, or
- Selector.

There are two perspectives from which to view each mode, firstly from the composer's level of involvement, and secondly from the way the medium functions. Engagement combines these perspectives and implies a partnership between the composer and the materials in a medium. The medium can be a computer, it can be another instrument (commonly a piano), it can be a paper score, or it can be the imagination. The composer's perspective is not independent of the medium, but rather exists in relation to it. Therefore, associated with the composer's perspective is the function of the medium.

- Observer - Artefact
- Director - Tool
- Player - Instrument
- Explorer - Model
- Selector- Creator

While in this presentation I will discuss the modes from the perspective of both the composer and medium, I will focus mostly on the composer's side of the relationship, as I feel it would be of most interest to this audience and avoids awkward technical details of computer music systems.

Alignment with previous theories of composition practice

A common understanding of the compositional process is a two stage process of what Roger Sessions (1941) calls *inspiration* and *execution*, where a musical idea is generated and undergoes a deliberate process of development. This characterisation of process I would described as focusing on the *player* and *director* modes of compositional engagement.

Newell, Shaw, and Simon (1962) propose a general theory of problem solving which involves two types of processes, *solution-generating* processes and *verifying* processes. These processes interact where heuristics are used to generate possible solutions which are judged successful or otherwise in the verification process. The problem solving metaphor of composition aligns most closely with my modes of *director* and *selector*, while the processes described by Newell et al. align predominantly with the *director* mode.

A scheme or vision is often seen as an important aspect of the compositional process (Gardner 1980, Solboda 1985). Composers report having a plan about what the composition will be like and so the process of composing involves the filling out and modification of that plan. This process of following a general scheme aligns most closely with what I describe as the *explorer* mode of engagement and includes aspects of the *director* mode. Starting with a scheme assumes that a musical space exists but is not clearly defined, and the activity of the composer is to find a pathway through it.

About the study

The modes of compositional engagement arose during my study of five composers, Steve Reich, Paul Lansky, David Cope, Bridgit Robindoré, and David Hirschfelder. Each of these composers used computers in a significant way as part of their compositional process, and each is an experienced composer and highly regarded in their area. The study examines how computers are used by composers. It became apparent that the diversity of compositional approaches and roles of the computer between these composers sat uneasily with previous, overly simplistic, understandings of the compositional process. The computer was used as a utilitarian tool by some and played like a delicate instrument by others, yet the computer medium was the same platform in each case. This diversity could be partly explained by the malleability of the computing medium itself. With reprogramming, the computer could change, chameleon-like, from one type of tool to another. On closer inspection it became evident that each composer demonstrated a wide range of approaches to the computer, even in the composition of one piece. It's function changed as circumstances demanded – its not unlike a piano shifting function from a tool to test harmonic progressions to a performance instrument. The diversity of function was only partly due to the flexible nature of the computer and more significantly due to the shifting relationship between composer and medium due to changing demands of the compositional process. I suspected that the differences in computer usage were revealing underlying shifts of emphasis in the dialogue between composer and musical material through the medium. The relational dialogue is the process I call engagement. The modes of engagement characterise different aspects of the relationship between a composer and the computer; their musical medium. Many modes of engagement are observable in each composer, but to varying degrees. An individual composer has a preferred mode or modes in which they are most comfortable, and each computer music systems lends itself to certain modes.

Modes of compositional engagement in action

I will now explain each mode of engagement in more detail drawing on examples from the case study composers as illustrations.

Observer - Artefact

As an *observer* the composer stands apart from the compositional process, detached, as an audience. It is almost of state of dis-engagement, where the composing medium, computer or score, is objectified. Often this mode is used by a composer to evaluate the completeness of a work, to attempt to dispassionately assess the development of the work. At times the dis-engagement may be complete that they feel no involvement and make no judgement. The musical material is viewed in a historical sense, not unlike a museum object.

David Hirschfelder provided a indicated that this mode of engagement was important to him, and that it did not come easily but was an perspective he fostered. Hirschfelder comments on the shift between creative involvement and the observer mode.

> It's subjectivity and then objectivity. It's being subject so you are the product and you are that gut reaction, then it's being able to step right outside of that as a listener and being totally objective and saying "How's that working?" Because if you stay subjective and too married to an idea it sometimes makes it difficult to move on. It's necessary to move on because I believe there are two forms of music, finished and unfinished, and that's all. You can always say music is unfinished, but too many people sign off on things too early because they remain subjectively attached. I speak *completely* from experience. I'm guilty of that sin, and probably will be again.
>
> That's why I like to have the time to listen to things after I've let it go. I think I can give myself credit for being able to detach myself a lot earlier than I could when I was younger. When I was younger it would take me weeks or months to get unattached to an idea and see it as others see it. . . I think I can literally detach in about half an hour now. (Hirschfelder 1997:235-259)

When a composer acts as an *observer* they relate to the music as though it were an *artefact* separate from them, static, and complete. They hear it as others would, without imagining elaborations or seeing the potential, they stand critically apart from it and observe its actuality.

Director - Tool

Crafting a composition has much to do with articulating a musical statement through controlling and moulding the musical materials. When a composer is consciously manipulating the musical materials to shape them into a desirable form, they are engaged in the compositional mode that I call *director*. This mode is what is most commonly understood as 'compositional' activity, because it involves the externalised manipulation of musical representations. This manipulation can be of a symbol system such as common music notation or a computer programming language, or more direct manipulation of sound material such as digitally sampled waveforms, for example.

When the composer acts as a *director* they usually relate to the compositional medium as a *tool* with which to write or manipulate their musical ideas. The medium is conceived as a canvas upon which they can arrange the music, as a set of processes through which the ideas can pass, or as a lens with which they can highlight and distort the ideas. While the medium has properties that enable or limit the crafting of the music, the dominant conception is that the composer is controlling the creative process with the medium responding to their instructions.

While all composers in this study worked in this directorial mode, Steve Reich provides a clear example of this compositional approach. Reich clearly differentiated the inspiration and perspiration aspects of composing using different tools for each stage. The piano or tape recorder were often used for the generation and selection of ideas and material, these were transferred to the computer for elaboration and development. Most of the time is spent in the working-out stage at the computer. Reich estimated that the balance between work away from or at the computer was divided about “60/40 with the computer getting sixty” percent (Reich, 1997:336-337).

Reich reinforced the directorial nature of his compositional approach with statements about his deliberate control over the process and his views of the computer as a tool used for reasons of efficiency. He describes how when composing *Hindenburg* the computer’s ability to manipulate the pitch of samples enabled him to bend the recorded material to his musical ends, rather than having tempo and pitch determined by the samples, as was the case in *Different Trains* and *City Life*.

> [For *Hindenburg*] I wanted to be able to structure the piece harmonically the way I would structure a piece which was not using samples, where the harmony would be worked out the way I wanted to work out the harmony independent of the sampled material. I wanted to set up a tempo and get a head of steam going rhythmically the way I did with my other pieces. So I decided that in this piece I will change, drastically if need be, any of the sampled material to fit the music. So I start off in three flats and if Herb Morris [who’s voice is sampled] is not in three flats, and he’s not, then I’m going to make him in three flats, and when I want to stretch his voice out then I just stretch it out. So the whole aesthetic is different and the whole technical means of working is better - well, harkens back to the way I was writing if I weren’t working with samples. It’s like having your cake and eating it. (Reich, 1997:537-551).

The mode of composing as a *director* involves the application of skills and techniques in a deliberate way toward a goal, however unclear. The relationship with the compositional medium and representation systems is one of control over tools, where skills and experience are exercised to express the music in that medium.

Player - Instrument

At times the compositional activity might be characterised as improvisatory or intuitive, rather than deliberate or conscious. The *player* mode of compositional engagement can be, much like children’s play, apparently directionless but usually absorbing and enjoyable. Metaphorically it is similar to musical improvisation, and may even appear similar when real-time instruments are involved. It is this *player* mode of engagement in which I believe Stravinsky was operating when his activities caused such bewilderment in the two women on the park bench outside his house. Composers such as Hirschfelder and Reich made regular use of improvisation at the piano as part of the compositional process and the mode of engagement as *player* includes this activity but is not restricted to it. Engagement as *player* need not involve real-time feedback, it is more distinctively characterised by the depth of involvement and the reliance on intuitive knowledge.

Composers engaged in the *player* mode relate to their medium in a personal and intimate way not unlike a performers and their instrument. The composer as *player* considers the computing medium like an *instrument*. The medium becomes a vehicle through which they express their music, and as such, it becomes more than an efficient tool, rather it is considered a partner in the compositional process. The composer as *player* acknowledges the medium as contributor to the compositional process and works with it like partners in a dance where a developed understanding between the partners makes the collaboration appear effortless.

Through building and modifying his own computer music software over decades, Paul Lansky has developed an intimate relationship with the computer as his musical instrument. His absorption in the music and compositional process sustained him over the two years it took to complete *Things She Carried*. Toward the end of writing of that work he described his relationship with the low-level computer programming at the heart of his composition activity.

> Well, I've learned to love it. [Laugh] Maybe out of necessity. I think there was a certain point about 20 years ago when I was about to throw in the towel, and I said 'if I'm going to do this I have to like all aspects of it'. I developed a real love relationship with working at a very low level with materials like that (Lansky, 1997:135-139).

Lansky's compositional processes are far from real time. Yet this does not limit the opportunity to engage with the process as a *player*. In fact, Lansky likes the drawn out process of constructing a compositional process and waiting for the computer to respond with a result. He comments, "I gives me a chance to think about what it is I've done. I'm not performing here, I'm sort of improvising in real-slow time" (Lansky, 1997:74-75).

The intimate nature of the *player* mode of engagement seems to develop with the composer's familiarity with the medium, and with the discipline of composing. For Lansky, familiarity and experience with the medium make the engagement more playful and absorbing, rather than predictable and routine.

> I find that as I get more familiar with the tools so the possibilities expand. You get really comfortable with the tools, and all of a sudden you can do a thousand times more things than you ever did before. Before you could only do a few things, so you'd really polish those and get them right and you'd know what to expect. Now I know all kinds of things and I don't know what to expect at all (Lansky, 1997:92-97).

The composer engaged in the mode of *player* treats musical materials intuitively and they are deeply involved in the making process. They treat their compositional medium as a *instrument* through which they express themselves and consider it a partner in the creative process.

Explorer - Model

A technological medium such as a computer music system and, more fundamentally, a theoretical music structure such as tonal harmony, serialism, or stochastic synthesis provide a conceptual musical space that a composer can explore. In the *explorer* mode of engagement the composer is involved in a predominantly conscious experimentation with musical materials in a semi-guided processes of seeking elements for inclusion in a composition. The goal of this exploration is often not a single point which can be clearly seen, but rather a loosely formed conception with relatively-clear boundaries. This process of compositional exploration within specific constraints is an enjoyable one characteristic of creative activities as Csikszentmihalyi concluded from his studies of creativity. "Paradoxically, it is the abstract rules we invent to limit and focus our attention that give us the experience of untrammelled freedom" (1996: 250). The outcomes may not be preconceived, but the experienced *explorer* can identify a useful discovery when it arises.

The medium for the composer engaged as *explorer* is used as an externalisation of ideas. The composer uses the medium to build models of the exploratory space which can be traversed, the medium can simulate hypotheses, and can bring to life ideas. In this way the medium functions like a physical model does for an industrial designer or a CAD system does for an architect. The medium externalises musical ideas which can then be viewed from many perspectives, reflected upon, and further explored. The medium acts as a cognitive amplifier for the composer in its ability to leverage ideas by building upon the stored knowledge in its symbol systems and existing procedures, and by capturing concepts and freeing up the mind for new thoughts.

The Franco-American composer Bridgitte Robindoré worked in the *Explorer* mode as she collected sound materials for her piece for harpsichord and computer. Her general direction was to collect a variety of noisy timbres for the computer part. “I’m working on the noise spectrum, between pitch and noise” (Robindoré, 1998:515-516). In her pursuit of useful sounds she used both familiar and unfamiliar techniques. The familiar techniques included sampling harpsichord sounds (many of which are inherently noisy) and processing them using established processes such as granular synthesis, frequency modulation synthesis, and fragment looping with the UPIC system.

> I have harpsichord sounds in loop form, some convolutions of them, some grain files and some UPIC work where a lot of it UPIC synthesis. But most is UPIC synthesis based on waveforms which I sampled originally. (Robindoré, 1998:447-449)

While these processes were efficient because of her experience with the UPIC system, her experiences were no less exploratory when working in the unfamiliar techniques of stochastic synthesis via the GENDY software. This technique, developed by Iannis Xenakis, was broadly suitable because it tends to produce noisy timbres. Robindoré’s time with GENDY was both exhilarating and frustrating, not unlike riding a frisky pony. She achieved interesting results but was not always sure she was able to control the direction she was going, let alone retrace the paths she had followed.

The nature of exploring, even familiar territories, means that the composer is often confronted with the unfamiliar. The lack of previous exploration of noisy timbres in western music left Robindoré wanting when asked to articulate what she was seeking.

> It would be nice to have some words to analyse the sounds so you could say, ‘yes, this is what I’m hearing, it’s a frequency band between 2K and 2.5K, its stronger, its noise content is da, da, da.’ But what kind of noises contents are there? (Robindoré, 1998a:183-185).

The inability to adequately express what they have found, either verbally or in the medium in which they are composing means that while the intention to explore may be conscious the process of exploration is, to more or less degree, intuitive.

> There are times though where I’m not quite grasping what the idea is that I need to be expressing, or that I feel. But I have certain sounds which lead me to them, it’s not always from the ideas. For example, the tape piece I have here for you with harpsichord, there are sounds which I know I have to find but I’m not sure yet what the metaphysical idea is behind them is. I can feel that it’s coming. (Robindoré, 1998:366-370).

The *explorer* mode of compositional engagement acknowledges that at times there is uncertainty or fuzziness in the mind of the composer about the direction the composition should go, and that a significant mode of activity is the searching for appropriate material and treatments of material in a loosely structured fashion.

Selector - Creator

The fifth mode of compositional engagement is one where the composer acts more as judge than as maker. Importantly, as *selector* a composer is not judging their own work but work created by the medium for inclusion into their own piece. This mode is perhaps the most reliant of all the modes on the computer as medium, because a computer can be automated to create new material for selection by the composer. However, processes such as aleatoric note selection or serialist processes would also be possible examples of the medium as *creator*.

One composer who actively seeks to have the role as *selector* as an available option during the compositional process is David Cope who has developed Composers Underscoring Environment (CUE) to achieve this end. CUE is music writing software based on a score metaphor using common practice notation, and is the only program of its kind and sophistication in the world at the time of writing. Cope can create music by manually entering notes to the on-screen score or ask CUE to suggest (compose) extensions to the score of a specifiable length.

> CUE has a resident notational program in it that I've written in Common Lisp, but its quite a simple one. So if I'm writing a fairly simple passage I usually use the resident program, which immediately turns all of my notes into what I call events, or things that CUE reads. If I'm doing something very complicated that requires a really intricate notational scheme then I simply bring up Finale and work in it and save the results as a MIDI file, and then load it as a MIDI file [into CUE] and go from there (Cope 1997 Interview 1:224-230).

> [CUE] does two things, it allows me to notate music on a score and see it immediately and play it back through QuickTime musical instruments. The second thing that [CUE] does is that the program attempts to keep track of my musical style and when I want it to it produces, at present extensions, of what I want . . . the important thing about this one is that when it shows you something that's added onto what you've just done, it's not just something built on fractals or mathematical principles of any sort, it is based on the style of music you've been writing, and the music in the database which it's been using. It keeps track of form . . . so that it actually produces contrast and so forth (Cope 1998:671-683).

During the compositional process Cope uses CUE as a score device much of the time, but only uses the generative facilities of CUE when he gets stuck for ideas or thinks the stimulus of a suggestion would be useful. The suggestion options are in a menu at the top of the screen, and Cope comments on his reticence to make use of that menu.

> My best day is never going up here [to the CUE menu]. My second best day is only going up [to the CUE menu] to see what's happening, and my worst day or worst nightmare is when I simply have to go up there because I have no other ideas about what should happen down here [in the score] (Cope 1998:707-710).

In the mode of *selector* the composer gives over responsibility for the generation of material to the compositional medium itself, and then engages as arbiter of taste through selection or rejection of the material generated. The partnership between composer and medium becomes most apparent in this mode as the composer shares creative responsibility with the medium.

Phenomenological perspective

The modes of compositional engagement are states with blurred boundaries, a particular type of compositional activity might show characteristics of two or more of the modes. Compositional activity usually oscillates between modes, at times frequently. However, as can be seen by the examples above, composers and mediums can be characterised as being more clearly in one mode rather than others. For this reason I believe it is reasonable to draw the distinction between them.

The modes are not sequential in any developmental sense, but they are arranged in the circular diagram according to a phenomenological progression. The modes at the top of the circle, *observer* and *selector*, are the modes where the composer takes a position of most distance from the work. They intentionally objectify the musical material in order to assess it. The processes of composition here are quite formal and abstract.

At the bottom of the circle is the mode of *player* in which the composer is most deeply involved in the musical material at an intuitive and emotional level. While engaged as a *player* the composer makes subjective judgements about the compositional direction, and feels most as-one with the medium and representational systems which are employed. The significance of the composers relation to their medium was also noticed more cautiously by Slodoba in his studies in composition,

> I suspect that an important component of compositional skill is a degree of 'trust' in one's medium—a certainty that the habitual processes of generation will yield material which is richer than one first sees, and which, even if initially unsatisfactory, usually contains within it discoverable properties which can be used to profit (Sloboda 1985:138).

The modes of *director* and *explorer,* half way down the circle diagram, combine aspects of the analytical and intuitive. They are modes where compositional processes are both conscious and unconscious, and the relationship with the medium is at arms-length, its influence is acknowledged but controlled.

Conclusion

Descriptions of compositional process have been largely developmental, beginning with a musical idea or theme followed by a process of extending and elaborating. The common conception seems to have been that the initial idea or theme provided direction for the work which was then articulated. In this presentation I have elaborated on that understanding. Firstly, to suggest that based on the composers in my study the initial idea, while it may involve a concrete theme, consists more often of a vague schema or plan often expressed in emotional or metaphysical terms. The process of composition following from that idea is one of clarification rather than realisation.

Secondly, and more central to this presentation, is the notion that the process of composition is a complex mixture of different relations between the composer and the musical materials, which I have called modes of compositional engagement. Five modes have been proposed which at times overlap and in which a composer may spend just seconds or many hours. Each of the five modes can be viewed from the perspective of the composer or of the compositional medium, which in my study was primarily the digital computer. Therefore each mode has a two part title reflecting the importance of the partnership between the composers and their mediums. The five modes are:

- Observer - Artefact
- Director - Tool
- Player - Instrument
- Explorer - Model
- Selector- Creator

The viewing of compositional practice from these multiple perspectives I suggest will assist in better understanding the compositional process and in particular the role of the compositional medium in contributing to that process. In addition it is not unreasonable, I believe, to hope that compositional curricula and pedagogy might benefit by clearer articulation of different models of compositional activity and practice. Also that the development of compositional mediums, including theoretical systems, music representations, and software environments, might be enhanced by better understanding the ways a composer might engage with them.

References

Copeland, A. 1957. *What to listen for in music*. New York: McGraw Hill

Cook, N. 1990. *Music, Imagination, and Culture*. Oxford: Clarendon Press

Csikszentmihalyi, M. 1996. *Creativity: Flow and the psychology of discovery and invention*. New York: Harper Collins.

Gardner, H. 1980. Composing symphonies and dinner parties. *Psychology Today*. 14, 4:18-27.

Hirschfelder, D. 1997. Telephone Interview. 12 November 1997. Melbourne, Australia.

Lansky, P. 1997. Telephone Interview. 28 October 1997. Princeton, USA.

Newell, A., Shaw, J. C., and Simon, H. A. 1962. The process of creative thinking. In *Contemporary approaches to creative thinking*. (eds.) Gruber, H., Terrell, G., and Wertheimer, M. New York: Prentice Hall

Reich, S. 1997. Telephone Interview. 11 November 1997. Vermont, USA.

Robindoré, B. 1998. Personal Interview. 2 Februrary 1998. Paris, France.

Robindoré, B. 1998a. Personal Interview. 17 April 1998. Paris, France.

Sloboda, J. A. 1985. *The Musical Mind: The cognitive psychology of music*. Oxford: Clarendon Press.

BRISBANE NOCTURNE - An algorithmic composition using SoftStep.

Warren Burt
Freelance composer.
PO Box 2154
St. Kilda West, Vic. 3182
waburt@melbourne.dialix.com.au

Abstract

This paper describes the structures used in *Brisbane Nocturne*, a real-time interactive composition made using John Dunn's "Softstep" algorithmic Windows software. I have been involved in the design of Softstep, first as a beta tester, then more recently as a consultant and module designer. The piece, made initially as a demonstration of some of the algorithmic functions of the software, quickly assumed a life of its own, and became music as well as a demonstration of various structures. The paper describes the functioning of the following Softstep modules: Ball, Rhythm Generator, Bio-Sequencer, Probability Generator, Fractal Generator, Table Write, Page Sequencer, Corruption, Delay, Chaos Generator, Image, and Number Generator, all of which are used to generate aspects of the piece. Aspects of some of these modules (Probability and parts of Chaos Generator) have been designed by me, although the bulk of the work on the program (99%+) is by John Dunn. The functioning of each of these modules will be demonstrated with examples from the piece. The result of combining all these structures is a composition with, hopefully, great subtlety, and this kind of subtlety is proposed as one of many possible ways forward for algorithmic composition.

Brisbane Nocturne Overview

In mid 1999, I became a beta tester for John Dunn, while he was developing SoftStep, an algorithmic sequencer based MIDI controller for the Windows environment. I had been using his earlier DOS based Kinetic Music Machine for many years, and was delighted that he was finally porting that over to Windows. My involvement as a beta tester quickly escalated to that of a consultant as I began plying him with suggestions, which he responded favourably to. Finally, I began designing and suggesting a few modules for the program, and eventually, many of my compositional dreams of the past 20 years became reality, thanks to John's programming skill, hard work, and musical intuition. I am now very excited by all the possibilities made available in SoftStep. It has a unique set of possibilities not found in any other program for any other platform.

Brisbane Nocturne is a piece made specially for this conference and this paper. It uses John Dunn's Softstep to control the synthesis engine in Martin Fay's Vaz Modular sound synthesis software. The piece is performed on a Pentium II Windows 98 laptop running both programs simultaneously, which are linked by Hubi's Midi loopback driver. Originally, the piece started off just as a series of demonstrations of various algorithmic function modules in Softstep. However, somewhere along the line, it also became a piece of music which pleases me greatly. Perhaps this is because of the way simple things combine to make a complex whole. I remember back in the 1970s, Dary John Mizelle, surely one of the most important and yet mysteriously underrated complexist composers around, said to me that he liked to pile structure on top of structure until chaos resulted. I think that is what might be going on in this piece, and that, in addition to the pleasant FM timbres, may be why I find the piece so attractive.

The piece, in 13-tone equal temperament, consists of three strands of music. I like the way the bluesy, "compressed" intervals of 13-tone tuning combine with the FM timbres used in this piece. Two of the musical strands are timbrally similar, while one is both registrally and timbrally different. The first strand is a two voice canon, in which the leading voice consists of a low melody and a high melody, each produced by a different algorithmic process. These two melodies are switched between randomly, making a composite melody which behaves differently in the high and low registers. This composite melody is then played canonically by two voices, the second voice being delayed by 61 clock ticks (each clock tick is equivalent to a demi-semi-quaver here) and is transposed up by seven 13-tone semitones.

The second strand is structurally similar to the first, but uses different algorithms. That is, it consists of a low melody and a high melody which are switched between to make a registrally varying composite melody. Here, however, different algorithms than were used in the first strand produce the low and high source-melodies, and the two source-melodies are switched between in such a way that a gradual transformation from one to the other takes place. This composite melody is then also played canonically by two voices, but here the second voice is delayed by 53 clock ticks, and is transposed up by five 13-tone semitones.

Additionally, the timbres of the two FM strands are slightly different. The first strand is made by having the modulating oscillator (a triangle wave) two 12-tone semitones higher than the carrier (also a triangle wave), while the second strand has the modulating oscillator two 12-tone semitones lower than the carrier. In the second strand, both oscillators are triangle waves as well. Finally these two voices are processed through resonant filters, panners, and amplifiers, all controlled by low frequency sine-wave functions, so that the timbre, spatial position and amplitude of each voice slowly and continuously change.

The third strand uses an additive synthesis tone made with Nicholas Fournel's Virtual Waves software. This is controlled by an algorithm called the Thue-Morse sequence, which in this case, produces arpeggiated melodies with self-similar properties. This voice is played in a high register, and is only played occasionally in the piece. Its function is to act as a contrast to the more continuous sounding of the first two strands.

Strand 1

SoftStep has a number of different rhythm generators, and this piece uses three of them. The first strand is controlled by a cute little generator called a ball. In this, a "ball" bounces around a frictionless four sided space, its speed controlled by x and y directional inputs. Each time the ball touches one side of the space (the space may be sized to any rectangular dimensions desired), a pulse is given off. Further, contact with each side of the space generates a unique pulse, which can be extracted (in any combinations desired) with a Mask Logic module, so that one ball can control a number of different functions in different, but related ways. (I guess if I wanted to use techno-jargon, I'd describe the ball as a physically modelled virtual rhythm generator, but I prefer to think of it as just fun.) Here, the speed of the ball is controlled by a mouse pad controller (an x-y axis position controller), and each aspect of the structure of the first strand is controlled by a different combination of side-contacts of the ball.

As mentioned earlier, each of the first two strands is made up of a melody consisting of high and low register components. The first strand's high melody is made with the Bio-Sequencer module. This is a module which reads DNA or protein sequences and outputs the result as numerical information. It is a very complex module, but here I use it very simply. There are twenty amino acids which make up proteins. So if a protein sequence is being read, values from 1 to 20 are output, as each new amino acid is read on each new clock pulse. Here the protein sequence of the human blue cone pigment gene is simply read out as a series of 20 values to select between 20 possible high pitches. (Free sources of genetic sequences are the Swiss Protein Data Bank www.ebi.ac.uk/textonly/swissonly.html and the NIH GenBank www.ncbi.nlm.nih.gov/. There will undoubtedly be a race to have the first composition using the complete human genome, once mapping of it is completed!) Rhythmically, the Bio-Sequencer is controlled by the right wall of the ball module. The Midi Voice Out is triggered off by a more complex combination of the left, right and bottom walls of the ball, but the module is set in sustain mode so that only new pitch information will trigger off a new pitch. This prevents repeating pitches, something I found useful for these particular melodies. Those wishing to know more about the use of biological data in music are referred to "Life Music, The Sonification of Proteins" by John Dunn and Mary Anne Clark
http://geneticmusic.com.

(Demonstrate Example 1, the Ball and the Bio-Sequencer.)

The low melody of strand one is made with a combination of the Probability module and the Page module. The Probability module, one of my suggestions, allows a user to set up any probability table they want. That is, any of 128 elements can have its own probability of occurrence. One simply draws the probability curve one wants, and the output is weighted in that way. The module can then select, in real time, between 128 of these curves that are preset by the user, or, using the Table-Write module, any of these 128 curves may be reset in real time. The Page module is a sequencer module, where a preset sequence (drawn by the user in the "Fill" utility module), is accessed either sequentially or randomly. Again, it is possible to switch in real-time between 128 sequences. To make this melody, the Page module is switched between three tables, which give the closest possible 13-tone equivalents to an ancient Greek Dorian diatonic mode, and the major and minor scales. Every 47 notes the scale that is being used changes. (Keeping the lower component of this composite melody within some notion of "diatonicism" is my concession to the notion that bass lines ought to be simpler, harmonically, than melodic lines!) The choice of notes from these diatonic scales is made by the Probability module, which is using a probability curve that is continually being rewritten. The two modules doing the rewriting are a Fractal generator generating a Mira fractal to determine which step of the probability table is to be rewritten, and a Chaos generator set to the "Burt Shift" algorithm to determine what value to set the chosen step to. (The "Burt Shift" algorithm is a shift-register-feedback circuit which is similar to that used in my "Aardvarks IV" hardware machine, and also similar to that used in Greg Schiemer's "Monophonic Variations".) So the resulting low component of the melody will always have some notes of the scale played more than some other notes, but which notes these are will be always changing. Suffice it to say that the melody will NOT simply sound like the result of equally weighted random numbers, but will hopefully have a much richer moment to moment structure than that produced by a simple random generator.

Those wanting the details of the implementation of fractals in SoftStep are referred to the help file, and to Robert Greenhouse's "The Well-Tempered Fractal."

(Demonstrate Example 2, the Page module and low melody, strand 1.)

These two melodies, the low and the high, are now switched between, using the Corruption module. This is a module which allows either switching between or adding of two incoming streams of information. The switching or adding is controlled by a probability setting. That is, one can set probability to, say, 25%, and then there would be a 25% chance that for each new event, an element from the second input (here called "offset") would be either substituted for or added to the corresponding element from the first input. This module was suggested by an article by Laurie Spiegel, called "An Information Theory Based Compositional Model".

In this piece, the two melodies are put into the two inputs of the Corruption module. Then, any top-side contact on the ball module generates a trigger, which selects a different random number between 1 and 100. (This is a simple equally weighted random number generator. (Okay, for the pedants, it's a pseudo-random number generator, but it's close enough for jazz.)) This determines the probability of choice between the melodies until the next top-contact of the ball, at which time this probability changes. So the high melody is triggered off by right-side contacts of the ball, the low melody is triggered off by left- and bottom-side contacts of the ball, and the probability of switching changes with top-side contacts of the ball. This gives us the possibility of a wide range of selections from, and combinations of, the various rhythm pulses generated by the ball.

(Demonstrate Example 3, strand 1 switching.)

Finally, to thicken the plot, and add an element of traditional structure to this, the pitch and rhythm information for the composite melody is also delayed by 61 clock ticks, and is applied to a Midi Key out module (a slave to a Midi Voice out - it allows polyphonic pitches on the same Midi channel, but doesn't send any pan, program, or controller information) which is transposed up by seven 13-tone semitones. This gives a canon at a "somewhat-less-than-a-fifth" interval (646 cents, to be precise).

(Demonstrate Example 4, canonic imitation, strand 1.)

Strand 2

The second strand is controlled by the Rhythm generator. This is a clock which enables 7 synchronised rhythm patterns of up to 32 pulses to be generated and combined in many different ways. Each of the 7 rhythm patterns can have any combination of off and on pulses and can be of any length up to 32 pulses. Using the Mask Logic modules, these can be combined in any way desired. And the clock of the rhythm generator itself can be controlled from any external signal, allowing for tempo modulations of any desired degree of flexibility. In this piece, the seven patterns are set to lengths of 11, 13, 17, 19, 23, 29 and 31 pulses: prime numbers that ensure maximum length sequences resulting from the combining of their pulses. Each element of control of the second strand uses a different combination of these 7 patterns.

The high melody component of the second strand is controlled by a Chaos generator. This one uses the Henon attractor to generate its values. It is triggered off by a 13 against 17 pulse pattern from the Rhythm generator, and chooses 14 pitches for one very high octave of pitches in 13-tone equal temperament.

(Demonstrate Example 5, Rhythm Generator and strand 2 high melody.)

The low melody is controlled by an Image Generator module. This is a module which can read any 128 x 128 pixel graphic (bitmap) image and read colour values either singly, or in combination for each pixel. It can generate its own Mandelbrot and Julia Dragon images internally, or accept images from the Fractal generators, or use an externally generated image. I was tempted to use an image of the cartoon character Bill the Cat for this function, but Bill just didn't have the pitch variation I was looking for (ack!), so I settled on an image I generated using the "Toy Universes" cellular automata in James Gleick's classic "Chaos" program. The pixel to be read from the Image Generator is chosen with inputs to its x and y parameters. These can be counters for linear readings, or random generators for samplings of various areas of the image. Here, I'm using two counters, one of which is reading an 11:19:31 pulse from the Rhythm generator, the other of which is reading a 17:23:29 pulse from the Rhythm generator. This produces a quasi-random walk across the image, moving generally in a downward diagonal direction, but deviating from a straight diagonal line according to the combination of the pulses. The values the image produces from this quasi-random walk across it are then scaled so they control a range of 16 low adjacent pitches in 13-tone equal temperament.

(Demonstrate Example 6, low melody produced with Image Generator, strand 2.)

As in strand 1, these two components are put into a Corruption module, and a composite melody is then produced by switching between them. A counter (controlled by a 17:23:29 pulse from the Rhythm Generator) simply counts from 0 to 127. This count is put into the probability control on the Corruption generator. Low numbers from the count mean there will be a greater probability of the high melody notes being chosen. The higher the count gets, the greater will be the probability of low melody notes being chosen. So for each time through the count, there will be a different transition from mostly high melody notes to all low melody notes. Note that the probability generator reads all values of 100 and above as a 100 percent probability that the second input will be read. Therefore, this progression, using a count from 0 to 127, is biased in favour of the low melody notes.

(Demonstrate Example 7, switching to produce strand 2 composite melody.)

Finally, this second strand of composite melody is also delayed canonically, this time by 53 clock pulses, using the Delay modules, and the result is transposed up five 13-tone semitones, for transposition upward of "less-than-a-perfect-fourth" (462 cents, to be exact).

(Demonstrate Example 8, strand 2 canon.)

Strand 3

In performance the tempo of these two voices is constantly changed. The ball is used to control the tempo of strand 1, while a bar control at the bottom of the screen is used to control the tempo of strand 2. But on top of these two, a third strand is sometimes added. This is a high melody of contrasting timbre (and one which doesn't pan) which is controlled by the Number Generator module. This is an expanded implementation of the Thue-Morse sequence, a number sequence which exhibits great self-similarity.

A good, simple explanation of the Thue-Morse sequence can be found in Gustavo Diaz-Jerez' article "Fractals and Music" in the October 1999 issue of "Electronic Musician." In this implementation, the sequence can count in any base from 2 to 127, any step size can be used, and one can begin from many different points in the sequence. And each of these parameters can be re-set in real time. For example, I randomly change the step size between 1 and 2 while generating my melody. This produces a melody with two sizes of basic generating intervals, 13-tone "seconds" and "thirds". Further, the output of this sequence is added to an offset and then fed back into the input of the Clock (the third, and simplest type of rhythm generator used in this piece) which is driving the Number Generator. This makes low notes have a shorter duration than high notes. In performance, the offset is changed with a second bar control at the bottom of the screen, changing the tempo of this Thue-Morse melody. This is a melody of upward arpeggios, and the pitches that begin each new upward arpeggio themselves form an upward arpeggio, etc. I included this melody as a timbral foil to the first two strands, and also to have one very clear example in the piece of one melody that was unambiguously produced by one clearly heard algorithm.

(Demonstrate Example 9 - Number Generator melody.)

Conclusion

In *Brisbane Nocturne,* I use a wide variety of algorithms to control different parameters of the piece. Normally, I wouldn't do this - my propensity is to thoroughly explore one kind of algorithm in a piece. However, the desire to show a number of different functions of SoftStep at once made me choose to structure the piece in this way. Conceptually, the structural result should probably be, quite simply, a mess. However, my ears tell me it's not. I hear both variety and coherence in this piece - of a sort that intrigues and satisfies me. I think what is happening here is that the end result of combining all these individually simple structures is a musical object which has a lot of subtlety to it, and this subtlety works on a variety of levels. So, for my ears, the piece has a satisfying structure.

However, there's a larger issue here, which is what I see as our increasingly complex understanding of the worlds of chaos, randomness and algorithmic structure. In the early aleatoric works of Cage and Xenakis, for example, one algorithm is often used. This gives the works great structural purity, and despite their sometimes chaotic surfaces, often a great sense of underlying simplicity. Now, with so many kinds of algorithms so easily available, we can quite freely combine them into structures which will possess an underlying sense of structural intricacy. (Whether the resulting music will be any better or not is up for grabs, I'll admit, and is probably irrelevant to the point I want to consider here.) In this progression I see an analogy (a fairly shoddy analogy, its true, but....) to what happened to musical structure in the mid-18th century. In the first works of the Mannheim school, the contrapuntal complexities of the Baroque were cast aside in favour of simpler homophonic structures. By the middle-late decades of the century, in the works of such composers as
C. P. E. Bach, contrapuntal thinking began to work its way back into the music, and structures of greater complexity were made, resulting finally in the structures of the late-Classical music of the 1780s and '90s. In my view, both the early works of Cage and Xenakis, and the earliest of minimalist works also were made in this same "paring-away" manner. These works, for this analogy, are the equivalent of the early Mannheim works. Now, with our greater understanding of how probability works,

and of the nature of a plethora of different deterministic and non-deterministic structures, we are perhaps on the verge of creating works which will be the structural analogies to the middle period works of C. P. E. Bach. Producing works with these tools that could be considered analogs to the complex late-Classical works of the 1780s and beyond is a task that could provide some fun for the next few years.

Thanks

A huge thank you to John Dunn for his inspiration and effort in developing SoftStep. Also immense thanks to the Music Fund of the Australia Council for their 1998-2000 Composers' Fellowship, during which time I worked on this software, this piece, and this paper.

References

Burt, Warren. 1975. "Aardvarks IV - A Real-Time Electronic Music Performance Machine." Unpublished Masters Thesis, University of California, San Diego.

Diaz-Jerez, Gustavo. 1999. "Fractals and Music." *Electronic Musician* October:108-113.

Dunn, John. 1999-2000. "SoftStep" Help Files. http://geneticmusic.com

Dunn, John and Clark, Mary Anne. 1997. "Life Music: The Sonification of Proteins" now included in "SoftStep" help files, ibid. http://mitpress.mit.edu/e-journals/Leonardo/isast/articles/lifemusic.html

Gleick, James and Rucker, Rudy. 1991. "Chaos - The Software" Autodesk, Inc.

Greenhouse, Robert. 1995. "The Well Tempered Fractal" User Manual.

Schiemer, Greg. 1989. "Monophonic Variations." NMA 6. NMA Publications, Melbourne.

Spiegel, Laurie. 1998. "An Information Theory Based Compositional Model." Leonardo Music Journal 7(January).

Making Computer Music Meaningful In Schools

Steven Dillon
Griffith University
Queensland, Australia.

Abstract

This paper examines the variety of modes of interaction with computers that occur within a school context and relates these to three locations of meaning: the personal, social and cultural. The computer is examined as a tool for teaching and developing teaching materials, as an instrument and as a cognitive amplifier. These ideas are examined through a philosophical argument and supported by vignettes drawn from a participant observation case study that focused upon the meaning of music to students in a school and the processes which give access to that meaning. The central argument revolves around the idea that music making in schools can be 'cutting edge' creative works rather than merely 'school music' and push the boundaries of aesthetic product. When computers are used in these kinds of ways in school music previously abstract concepts and the opportunity to produce high quality music become accessible to children. Finally, the paper argues that through experiencing making music in these highly 'technologised' ways we come to gain an understanding and be critical of how art is made in our every day lives. This serves as a demystification of the processes of music making within the media that pervades our lives.

Making computer music meaningful in schools.

Music is no stranger to technology. The piano key, the trumpet valve and indeed the system of western music notation are all brilliant technologies that had a profound effect upon how humanity expressed itself through music. Each of those technologies was seized upon by composers and performers alike to increase the range and expressiveness of the music they made. In this paper, I propose that the so-called new technologies of the synthesizer and the computer are new surfaces in which we inscribe meaning. Through examining a philosophical shift in the way, we view the modes of engaging with technology in music education, I hope to argue that music making in schools can become more relevant meaningful and 'cutting-edge'. To illustrate the practical application of the principles discussed here I will also draw narrative vignettes taken from a recent long-term participant-observation cases study (Dillon, 2000).

The philosophical argument.

Technology is often described as a tool. Brown (1999) suggests that this is a limiting metaphor. In education, this is often the case. Teachers who are often unfamiliar users of technology use the word in this manner. Those in technology related studies also perceive the human machine interaction as a user and tool metaphor. Historically, artists have a different kind of relationship with technology. This kind of creative rather than reactionary relationship is of interest here. We do not applaud the piano-key mechanism for its technical beauty. Although in engineering terms it may have been a past century equivalent to a major increase in cheap computer memory, or speed. Rather this innovation increased attention to the dynamic and expressive qualities of music of the Romantic period. Likewise, the trumpet valve increased the range and flexibility of brass instruments. The response from composers was to compose music that stretched the aesthetic boundaries of those instruments and of music itself. The important idea here is the focus upon stretching aesthetic boundaries rather than the technical brilliance of the tool. Brown (1997) suggests that we perceive technology in multiple ways- not merely as a tool. Amongst the modes of engagement he suggests are computer/ synthesizer as instrument, as cognitive amplifier (Papert, 1980), and as a tool.

What these ideas imply are that the technology takes on different roles and different modes of engagement depending upon our needs and uses for them. In musical terms the implications are quite far reaching. Firstly, to consider the computer and the synthesizer as an instrument requires the traditional notion of a developmental pedagogy and a repertoire that stretches and expands the expressive qualities of that instrument. The notion of cognitive amplifier or assistant suggests that the technology has input to the process of making music that reaches beyond that of a passive tool. It functions as a way of thinking that is external to the mind and body, a kind of assistant in the process. The traditional aspect of tool is still apparent – it does a job that it is designed to do functionally and that is proportional to the effort we afford it. A hammer and nail analogy, the technology has strength where we physically may not. It enables quite specific functions. When we consider the notion of technology having these kinds of characteristics and our modes of engagement with it being varied then we can see that the tool metaphor alone is inadequate for artists and indeed educators. Furthermore, we can see limitations of our preparation for engagement with technology- skills and in the *repertoire* and methods of learning about this-*Education.*

Cage- limitations of notation as opposed to magnetic tape and digital.

John Cages suggests that he '*finds it quite curious that, given such wonderful technological means for creativity, our schools continue to fix their minds on historical means of creating music.* (Reimer & Wright (1992:48). He is referring to the technology of written notation and traditional western art music forms of composition that focus upon this technology for creative music making. In reference to Western harmony he suggests that '*the ideas of harmony and its reference to particular steps in the field of sound,* [as].. *inadequate for the composer, who will be faced with the entire field of sound*. (Cage as cited in Reimer & Wright, 1992:46).He calls these '*limiting sound producing mechanisms, that do not allow for the vast possibilities of such materials as magnetic tape and other contemporary devices for sound production*. (Cage as cited in Reimer & Wright (1992:44). In the study that these statements appear, and are discussed, they are considered radical and perhaps the most opposite to the existing viewpoints of music educators, philosophers and composers. The synthesizer and the computer have had an indisputably large influence on music made in the last 30 years. If we consider the notion of technology as tool being a limiting metaphor then these views become even more clearly relevant to contemporary music making and education. As the body of materials- *repertoire*, produced by these means increases in volume and significance our need to understand, access and develop music in these ways also becomes more necessary.

Where is technology now?

The eminent Australian saxophonist Dr Peter Clinch once compared the synthesizer's position to that of the saxophone at the beginning of the century. Highly used as an instrument in popular music but little organised repertoire and pedagogical development. An instrument played by clarinet players and one that was often excluded from large-scale art ensembles. Even now, this seems to be an accurate description of the computer and synthesizer position. The limitations of traditional sound production and notation systems are also apparent as the cognitive amplifiers of the last centuries. Even as tools for making music including traditional notation, the digital medium is faster, clearer and has greater capacity for assisting and giving more immediate feedback to the composer. The digital medium has made the once abstract concepts of harmony and sound production more ready to hand, immediate and available. Many concepts of music making that were reserved for more experienced users of the technology of notation are now available to children in accessible form.

Meaning and music making

In my own research into the meaning of music to students in a school context (Dillon, 2000), I have suggested that music has three primary meanings. Firstly, a personal meaning that refers to the intrinsic meaning that is experienced when we make music through personal engagement. The flow of personal experience (Csikszentmihalyi, 1994). Secondly that of social-meaning that we get making music with others in ensembles. Something described by musicians in the study as a wordless meaning, a shared communication (Dillon, 2000). Finally, cultural meaning- when the music that is made and the music maker is affirmed and valued as an artist and for their art by the community. This is a reciprocal interaction between culture and artist. Both the artist and product are contribute to the distinctive character of its 'culture' and are projected by the community as such. I suggestion of this study that meaning in music experience is gained from engagement with all of these interactions. For music to be meaningful students need access to music making experiences that afford the opportunity to experience the full range of meaning from personal through social to cultural.

Implications for practice

What are the implications for practice within a school context?

Firstly, the computer and synthesizer need to be considered instruments in their own right. The AMEB has recently included a pedagogy for 'keyboard' referring to electronic keyboard in its examination structure and in Victoria the Synthesizer is included as an instrument for Solo performance in the Year twelve examination syllabus. This is some evidence of the inclusion of the synthesizer as an instrument; the acceptance of the computer as instrument is a more difficult instance although QUT in its New Media degree includes it as a principal instrument. There is also a need to assemble and commission repertoire or its 'sound project' equivalent so that students of the instrument can learn progressively through repeatedly deeper engagement with the expressiveness of the instrument.

Secondly, The role of computer/ synthesizer as cognitive amplifier for composition arrangement and realisation of musical ideas needs to become part of the process of music making. There needs to be recognition initially of the 'concretizing' aspects of the media that allows it to make previously abstract concepts of large-scale orchestration and arrangement accessible to children. Beyond this, the notion of using these media to create new music rather than simply recreate what Cage call historical music, become increasingly possible. Such ideas as aleatoric music, Algorithmic composition, Granular synthesis, Minimalism and a myriad of even newer forms are available to young composers through digital medium.

Thirdly, the use of the computer as a tool for doing rapid or complex computations frees the composer to experiment with ideas and receive immediate aural and visual feedback. The composer can focus upon expression rather than technical manipulation.

In relation to meaningful experience, these ideas suggest that students need access to all of these modes of engagement and within the framework of personal, social and cultural experiences. It means that technology must be available to be engaged with in all of these ways. There needs to be the opportunity to make and perform music using this technology and to develop its expressive capabilities to the same kinds of levels as any instrument. It means students need access to e ensemble experiences – synthesizer ensembles, electro acoustic ensembles, sound creation projects and installations and exhibitions of sonic realisations . In a cultural sense, the school community needs to value the person and the work that is made through these kinds of interactions. It needs to do this in pragmatic ways- using music of this kind to project something about the school community. Furthermore, it must also explore purely aesthetic ways, where the instrument and works are given the opportunity to push the boundaries of aesthetic development in the domain.

Case study example

In 1995, I had the opportunity to begin observing how these philosophical ideas worked in practice. I began a participant-observation case study at an independent pre –preparatory to year-twelve school in the eastern suburbs of Melbourne. The following is a description of how I observed those philosophical aspects were addressed in curriculum, co curriculum and as part of the school community.

As an Instrument

The school hired keyboard teachers that were proficient synthesizer players and sound programmers. The instrumental program offered private lessons on the instrument and created a student-centred pedagogy for development. Each synthesizer was required to play in a synthesizer ensemble and a number of electro acoustic ensembles/bands, projects. The instrument was offered with the year-eight classroom-instrumental program and used in classroom music ensembles from grades four to twelve.

As a cognitive amplifier

Within the instrumental program, composing using computer was offered as a private study in the same way as an instrument. Within the classroom program year ten featured synthesizer and computer use within the 20th century composition elective and the computers/ synthesizers were used throughout the classroom program (P-12) as a mode of composing and organising musical materials. There was a conscious effort made for the technology to be ‘transparent’, that is, not used merely for its ‘new’ attractiveness but for its functionality and ready access to experience.

As a tool

Digital software and hardware in the music department was used to make resource materials by teachers and students, publish print music and to give repetitive- drill and response training in aural skills and music notation theory. Not to mention the use of the digital technology to demonstrate and make concrete abstract concepts of harmony and musical structure.

Access to meaning

In relation to access to meaning, students had access to personal engagement though classroom use and private tuition. Students had access to social interaction, through classroom ensembles and school ensembles. Cultural interaction through electronic music concerts and multi media church services and ceremonial events. These events featured music composed and arranged by teachers and students. Those compositions were commissioned with a brief to serve the pragmatic function of the ceremony and to push the boundaries of the art form. Because of the lack of repertoire for the instruments, they were forced to arrange and compose music of their own. This made the process even more meaningful as the students were affirmed by the school community for their unique contribution to school ceremony and by the domain for their contribution to the development of the idiom. This idea is amplified by the school minister who describes the role of a synthesizer ensemble at a school ceremony:

> What it said to the students was that there was something almost tribal about this. That they could create their own rhythm, their own message, their own particular way of being musical , it wasn't imposed from anywhere else it actually happened through the creativity of the staff at the time and the students who worked with those people. I thought it was an extraordinary time. We all commented later how passionate it was and how people if they didn't remember anything else about that night, remembered that music. (Minister)

This is a powerful affirmation of the meaningfulness of this kind of music making within a school community. Remember that we are not describing students re creating a Beethoven quartet, but students work. The comment is not about a fine performance but a cutting edge, relevant and expressive one.

Discussion of implications

What I am arguing here is that we need to move beyond the limitations of ‘technology as tool’. We need to view the modes of engagement with technology as multiple. Then we can both develop the boundaries of creative music making and enhance our access to meaningful and expressive music making experiences. This means that within a school context, music making can be truly creative not merely re-creative. The computer and the synthesizer have the capacity to give access to experiences in music making that previously were only available to an elite few who could navigate the complex representation system/technology. Cage argues that this is a kind of democratisation of the process of making music, where conductor, notation and even performers are no longer necessary (Reimer & Wright, 1992:49). This is reflected strongly in the musical community in popular music, music from a variety of non-western cultures and in contemporary art music. This new access and democratisation makes it possible for cutting-edge music to be made by young composers in a school. For music as it is found in the wider community to be simulated in a school, demystifying the process of music making in media (Schusterman, 1992) and indeed music making per se. It gives access to understanding and experiencing complex musical ideas first hand. To achieve this the first step is to consider a shift to a philosophy that expands the view of technology and human interaction in music making. Next, there is a need to respond to this shift in thinking in the simple ways described in the case study example.

Concluding remarks

So how does this approach make computer music meaningful? Members of the electronic music ensemble at the school suggested that they felt that as a group they were on the ‘cutting edge of electronic chamber music’. They attributed this feeling to the sense of uniqueness, control over repertoire development and function in the community. This had a profound effect on the confidence and self-esteem of the participants. The school minister described it as a sense of belonging, ‘*which has had the effect of raising their own awareness of who they are.’* The students reported experiencing this sense of belonging on two fronts, firstly the deep relationships developed through playing together in an ensemble and secondly, the relationship with the community gained through interaction and purpose. The dual roles of performing music for purely aesthetic and pragmatic purpose had an effect upon their personal feelings about community and the feelings of the community about them. What I am arguing here is that if we open the school curriculum to view our modes of engagement with technology in music education as multiple as suggested by this paper, then we are making way for more meaningful relationship with expressive music making and surely this is the goal of music education.

References

Brown, Andrew R. 1997. "Changing Technologies, Changing Minds: Taking account of music technologies in the curriculum." Gifford, Brown and Thomas (Ed.s). *Conference Proceedings of the ASME XI National Conference*. Brisbane: Australian Society for Music Education.

Brown, Andrew R. 1999. "Music, Media and Making: Humanising digital media in music education." *International Journal of Music Education* 33: 3-9.

Csikszentmihalyi, M. 1994. *Flow: The Psychology of Happiness.* Australia: Random Century Group.

Csikszentmihalyi, M. 1996. *Creativity*, New York: Harper-Collins.

Dewey, John. 1980 (first edition 1934). *Art as Experience.* USA: Pedigree Books.

Dillon, Steven C. 1995. "The student as maker: An examination of making in music education and the implications for contemporary curriculum development." MEd Thesis. Victoria, Australia: La Trobe University.

Dillon, Steven C. 1996. "The Student as Maker: A Narrative for a Pragmatist Aesthetic." *Proceedings ASME XI National Conference*. Brisbane, Australia.

Dillon ,Steven C. 2000. "The student as maker: the meaning of music to young people in a school." Unpublished PHD dissertation. Melbourne, Australia: Latrobe University. (Please note data has been drawn from this thesis that is due for submission late 2000).

Papert, Seymour. 1981. *Mindstorms, Children, Computers and Powerful ideas.* USA: Basic Books.

Reimer, Bennet & Wright, Jeffrey, E., Eds. 1992. *On the Nature of Musical Experience.* Colorado: University Press of Colorado.

Schusterman, Richard. 1992. *Pragmatist Aesthetics : Living Beauty Rethinking Art.* Massachusetts: Blackwell Publishers.

About the Author: Steve Dillon is a lecturer in Music education at Griffith University in Brisbane. He is in the final stages of research for his doctoral thesis at La trobe University. He has many years of experience as a teacher and head of performing arts. He continues to work professionally as a contemporary composer and a vocal performer, most recently as *Realisatuer* and composer for the international music-theatre production *Atlantis*. He has published a number of commercial music resource books and is a partner in a Synthesizer and music education publishing company (http://www.explodingart.com.au) which publishes an electronic journal *Xart* and encourages and promotes works in these areas.

On the Computer Recognition of Solo Piano Music

Simon Dixon
Austrian Research Institute for Artificial Intelligence
Schottengasse3, A-1010 Vienna, Austria
simon@ai.univie.ac.at
Phone: +43-1-533611222 Fax :+43-1-533611277

Abstract

We present work towards a computer system for the automatic transcription of piano performances. The system takes audio files containing polyphonic piano music as input, and produces MIDI output, representing the pitch, timing and volume of the musical notes. The aim of this work is not to reduce the performance data to common music notation, but to extract the performance parameters for a quantitative study of musical expression in piano performance. Standard signal processing techniques based on the short time Fourier transform are used to create a time-frequency representation of the signal, and adaptive peak-picking and pattern matching algorithms are employed to find the musical notes. In order to perform large scale testing, the test process is automated by synthesizing audio data from MIDI files using high quality software synthesis, and comparing results with the original MIDI data. The test data used is Mozart piano sonatas performed by a concert pianist.

1 Introduction

This paper addresses the problem of extracting musical content from audio data. More specifically, we consider the task of ascertaining performance parameters from polyphonic piano music, that is, calculating which notes were played *(pitch*), when they were played *(timing*), and how loud they were played *(velocity*[1]).

This is closely related to the task of automatic transcription, whereby musical notation is created from audio recordings of music. However, common music notation is not suitable for accurately representing a musical performance, as it does not describe the velocities of individual notes, nor the precise timing of onsets and off-sets, using instead a more abstract discrete set of notated durations. In order to study musical expression, one of the goals of the current project, such extra details are required. In this sense, the accuracy required in this project is much greater than for a transcription system, which filters out the expressive details as if they were noise. On the other hand, several tasks necessary for transcription into common music notation are not addressed in this paper which would be required for a complete transcription system. Examples of these tasks are: rhythm understanding, quantization, key finding, note naming and the physical layout of musical symbols on a page[2] .

The input to the system is digital audio, taken from CD's or synthesised by a high quality software synthesizer. Processing begins with time-frequency analysis based on the windowed short time Fourier transform, from which peaks in the time-frequency terrain are extracted, creating frequency tracks which are then grouped as partials of musical notes. The output is a symbolic representation of the music, in MIDI format, representing the musical events detected from the performance. Thus

[1] MIDI terminology is used throughout this paper.

[2] see (Dixon and Cambouropoulos, 2000; Cambouropoulos, 2000) for recent work on some of these tasks.

the system could function as a front end for an automatic transcription system, for content-based indexing and retrieval, or for a performance analysis system. In section 2, we briefly review the previous work in this field, and then in section 3 describe the system itself. The following section outlines the testing methodology, and in section 5 we present the results obtained to date and an evaluation of the system's performance. The concluding section contains a discussion of the results, and lists possible directions of further work.

2 Related Work

Most related work consists of various attempts at building an automatic transcription system, each of which tackles a different subset of the transcription problem (Moorer, 1975; Piszczal-ski and Galler, 1977; Chafe et al., 1985; Mont-Reynaud, 1985; Schloss, 1985; Watson, 1985; Kashino et al., 1995; Martin, 1996; Klapuri, 1998). The current state of the art is that pitch and timing for a known instrument playing one note at a time can be detected quite reliably, and is commercially available in hardware and software realisations. However, rhythm understanding even in this limited case is still quite poor in these commercial systems, most of which re-quire a strict metrical performance synchronized with a pre-specified beat. Polyphonic transcription has only been performed successfully with restrictive conditions placed on the input data. Space does not permit reference to more than a small number of the approaches used[3].

The pioneering work of Moorer (Moorer, 1975) used comb filters and autocorrelation to perform transcription of very restricted duets. The input data was allowed to contain no more than two notes sounding simultaneously, and note combinations which shared common frequency components (eg. octaves) were not allowed, in order that the frequency components could be interpreted unambiguously. The range of notes was restricted to two octaves. Schloss (Schloss, 1985) developed useful time domain techniques for accurate estimation of onset times in his work on transcription of untuned percussion, but did not address pitch extraction. Martin (Martin, 1996) allowed up to 4 voices in the input data, but it was restricted to the chorale style of J.S. Bach, in which the notes have relatively long duration and change simultaneously, which made it possible for him to segment the signal in the time domain into "musically constant" sections. Furthermore, octave intervals were not allowed, and the note range was restricted to under 2 octaves (f_0 = 123 - 440Hz). Klapuri (Klapuri, 1998) allowed a 5 octave fundamental frequency range (65 - 2093Hz), but required example notes covering the complete range of each instrument in order to train the system. Good results were achieved for the stated test examples; it is not clear how the system would perform on more complex musical examples.

The only work explicitly concerned with extraction of performance parameters is that of Scheirer (Scheirer, 1995; Scheirer, 1997), who also dealt with solo piano music, and built a sys-tem which required that the musical score be provided to guide the system. Knowing in advance which notes were played, means that the system is only required to search for onsets, off-sets and velocities of notes at known frequencies and within quite small time windows. Furthermore, the interactions between notes (eg. shared frequency components) can be predicted and avoided (as is done in Scheirer's system). Thus, although the techniques developed in this system are relevant to the "blind" transcription problem, they are not immediately reusable, and some are not at all.

[3] see (Klapuri, 1998) for a more complete review.

3 System Description

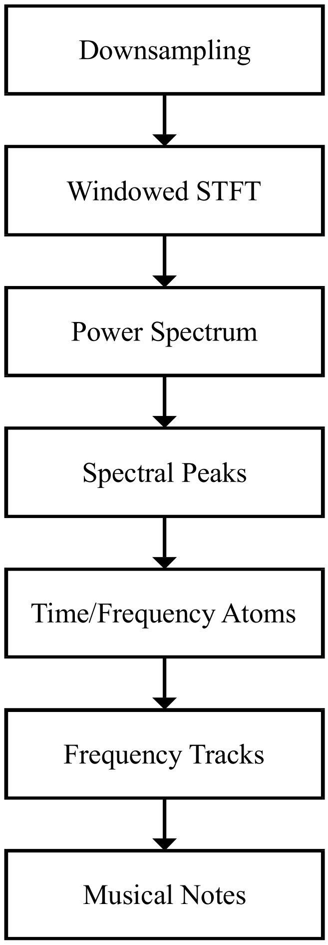


Figure 1: Data-processing stages

We now briefly describe the stages of processing performed by our system, shown diagrammatically in Figure 1. The first stage of processing consists of down-sampling the signal to a 12kHz sampling rate, in order to decrease the processing time of subsequent stages. The use of a lower bandwidth signal does not appear to affect results adversely.
A time-frequency representation is then created using a windowed short-time Fourier transform. By default, a window of 4096 samples (341 ms) is used, containing 230ms of signal shaped by a Hamming window and zero-padded to fill the window. Command line parameters can be used to select windows with different sizes and shapes.
The complex frequency domain data is then converted into a magnitude squared (power) spectrum and an adaptive peak-picking algorithm finds spectral peaks, which give an initial estimate of the significant frequency components in each window of the signal. These peaks are represented on a logarithmic frequency scale, to conform with the musical representation of pitch in semitones.

Figure 2 shows the resulting power spectrum and a typical problem in time-frequency analysis of music: in order to get a sufficiently good frequency resolution to determine pitch accurately, the time resolution is poor, and a large amount of overlap is seen between notes in scale passages. This is not an insurmountable problem for the transcription of piano music, since the piano has a very sharp attack, and by using overlapping windows a reasonably accurate estimate of onset times can be obtained.

A current extension of the system, aimed at improving the time-frequency resolution trade-off, uses the rate of change of phase in the FFT filterbank channels, rather than the centre frequency of the channels, in order to obtain a more accurate estimate of frequency, and in turn allow a smaller window size to be used. This idea was first used in the phase vocoder of (Flanagan and Golden, 1966), and has since been used in many computer music applications (Dolson, 1986).

The peaks in the power spectrum are then isolated by finding at each time point the local maxima in the frequency dimension which are above a minimum threshold and which contain at least 1% of the total power of the signal at that time, giving a set of atoms of energy localised in time and frequency. The time-frequency atoms are then searched in the time dimension, in order to find the peaks in time corresponding to note onsets, and these onsets are followed until the power drops below the minimum threshold, which determines the offset time of the note.

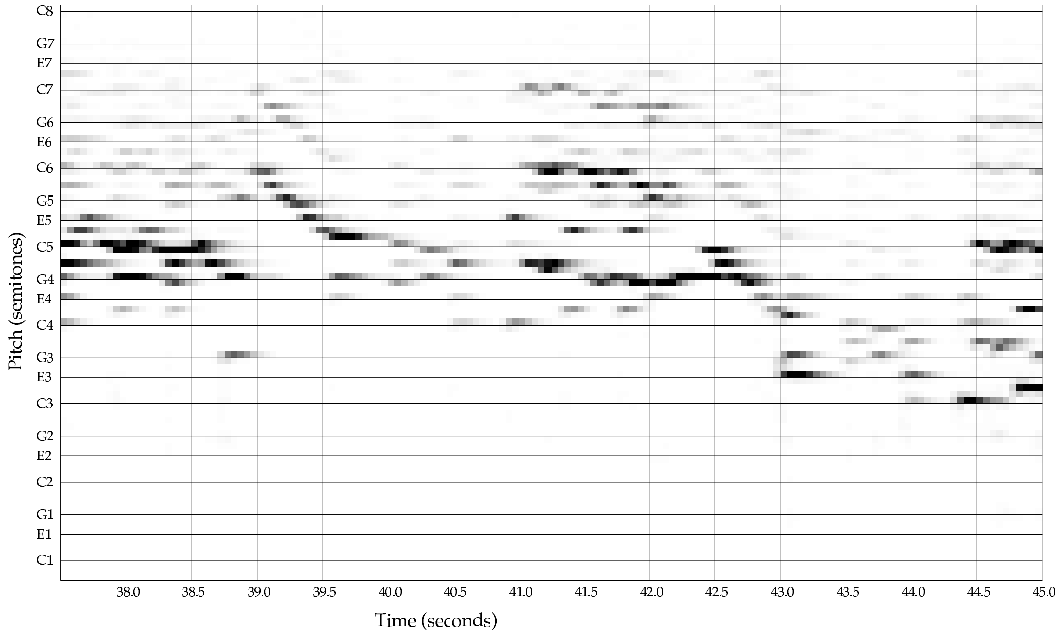


Figure 2: Power spectrogram from STFT with Hamming window, showing 7.5 seconds of Mozart's Piano Sonata in C major, KV279, 3rd movement.

Finally, the velocity is determined from the peak power, which occurs at the onset time of the note. The frequency tracks calculated in this way represent partials (harmonics) of the musical notes. A few simple rules are employed to eliminate rogue frequency tracks, such as those with very short durations, which may be, for example, caused by transients at note onsets.

The final step is to interpret the frequency tracks as musical notes, which is done by finding a set of fundamental frequencies which provides the best explanation for the observed frequency data, relative to an implicit generic model of musical instrument tones. That is, we combine partials which occur simultaneously, are harmonically related, and do not fall outside the expected spectral envelope for piano tones. Due to considerable variability in the individual notes of a piano, this final condition cannot be particularly stringent, but is necessary in order to be able to detect intervals such as octaves. The resulting output, in MIDI format, is then evaluated as described in the following sections.

4 Test Methodology

One of the greatest difficulties in building audio analysis systems is the lack of high quality tagged test data. In the field of speech recognition there are large corpora of tagged audio, making possible the use of statistical and machine learning methods such as hidden Markov models in analysis of audio with speech content. Music has no similar large corpora of data, so these methods cannot be used. Similarly, thorough quantitative testing is also made difficult by the lack of suitable test data. Although it is becoming easier to generate test data using, for example, a Bösendorfer SE290 computer-monitored piano or a Yamaha Disklavier, a further hindrance to gathering large data sets is the legal issues associated with copyright of professionally performed musical data.

In order to test our system, we require a symbolic representation of the audio content of the test data, which is more than just the musical score; we also need the expressive details of timing and dynamics. These are rarely, if ever, available in conjunction with audio recordings. Manual transcriptions can be performed, but are extremely time-consuming to produce, and cannot guarantee a high degree of accuracy or precision. In the literature to date, testing is usually performed manually, using simple musical excerpts, normally no more than 30 seconds long. However, we have developed and tested our system using music from the standard classical repertoire, in this case Mozart piano sonatas, performed by a professional pianist. These provide a realistic source of noisy data, and a far more difficult data set than has been attempted elsewhere.

In order to ensure that the system was tested with a wide range of musical situations, a large data set was obtained, consisting of 13 complete Mozart piano sonatas (KV 279-284, 330-333, 457, 475 and 533). These were performed on a Bösendorfer SE290 computer-monitored piano, and were converted to MIDI format using conversion software. As the original audio data from the performances was not available, the audio data was then generated from the MIDI files using a software synthesis program, Timidity. A number of different synthesizer voices were chosen to test the sensitivity of the algorithm to the instrument timbre. The accuracy of the note recognition system was tested by comparison of pairs of MIDI files – the input files from which the audio data was generated, and the output files of detected notes.

5 Results

A matching algorithm was developed to pair events in the input file with corresponding events in the output file, under the constraints that the notes must have the same pitch and onset times differing by no more than a small error margin (70ms). The results are evaluated in terms of false positives (FP = the number of notes reported by the system that were not played) and false negatives (FN = the number of notes played that were not reported by the system). An incorrectly identified note (eg. wrong pitch) is counted as both a false positive (the wrongly re-ported note) and a false negative (the note that should have been reported), which is perhaps an unnecessarily harsh evaluation metric. The error figures are combined with the following formula into a single percentage score (where is the number of correctly identified notes):

$$Score = \frac{N}{FP + FN + N}$$

In Figure 3 we present results computed for the complete test data set for three different synthesizer voices.

Voice	N	FP	FN	Score
acpiano	95443	32053	11016	68.9%
britepno	87331	18185	19128	70.1%
honky	93777	8227	12682	81.8%
acpiano*	95914	21433	10545	75.0%

Figure 3: Results for 3 different piano sounds

The voices are labelled with their names from the General MIDI Specification, representing a standard acoustic piano (acpiano), a brighter sounding piano (britepno), and a honky-tonk piano sound (honky), respectively. These results were generated with a single set of parameters, and show a recognition accuracy of around 70 - 80%. The differences between these three results demonstrate a weakness of the current system, that it is sensitive to the amplitude and timbre of the instrumental sound which is used. To demonstrate this more clearly, we tuned the parameters to obtain better values for the acoustic piano sound, as shown in the last row of the table (acpiano*).

6 Discussion

Although the system is far from complete, the preliminary results are positive. The stated aim of this paper is to recognise piano notes, but as yet, the piano-specific assumptions have not been built into the system, and the software framework developed is suitable for a range of instruments. It is clear that results can be improved by modelling the sound source accurately, as is done by (Klapuri, 1998). One planned extension of this work is to build a recognition module that is specific to acoustic grand pianos.

But before tuning the system to particular instrument models, we intend to address the problem of sensitivity of results to the particular instrument by using dynamic modelling (Dixon, 1996) to automatically determine suitable parameter values from the audio data, rather than requiring the user to determine these values by trial and error. This is particularly important for the more general problem of transcription of music from unknown instruments, and is a more elegant approach than hard-wiring instrument parameters into the recognition algorithms. One way to implement dynamic modelling is the use of artificial intelligence iterative improvement algorithms, so that the system can learn automatically to improve its performance, using feedback from the evaluation part of the system.

At this point in time, the use of synthesised data was a pragmatic necessity, in order to make large-scale testing possible. Other authors have used the same method even for small-scale tests (Martin, 1996). The performance of the system on non-synthesized input data (ie. recordings from acoustic piano) will be tested as the data becomes available.

Another area requiring further work is the evaluation function. Currently, the matching algorithm uses a fixed time tolerance, giving a binary result, which should be replaced by a more gracefully degrading accuracy value. Similarly, the note identification is an all-or-nothing value, which could be improved by classifying incorrectly identified notes into common error types (eg. octave errors). Finally, the evaluation function should also take the perceptual strength of errors into account, as the input data contains a number of notes which are neither perceptually nor physically detectable from the audio signal, which are currently counted as errors when not detected by the system. For example, a key adjacent to a played key is sometimes brushed sufficiently hard to be detected by the monitoring system, even though no audible sound is produced.

Further development of the system is required to assess and improve the accuracy of the dynamics and offset times reported by the system. These are much more difficult problems than onset detection. For example, the physical release of a key may occur long after the note is no longer audible, making the offset time irrelevant. Similarly, the amplitude of a note, which is represented by the velocity of the hammer as it hits the string(s), is dependent on other factors not represented, such as interactions (coupling) between the vibrations of different strings and also of the piano frame and soundboard. Finally, we plan to extend this work to investigate expressive timing. One approach would be to use time domain analysis to identify the onsets of notes more accurately. We also intend to link this work with recent work on beat tracking (Dixon and Cambouropoulos, 2000), in order to separate the different levels of expressive timing in the music, that is to separate local changes, such as the displacements of events from their nominal temporal position, from more general changes, such as an increase in the average tempo.

Acknowledgements

This research is part of the project Y99-INF, sponsored by the Austrian Federal Ministry of Education, Science and Culture in the form of a START Research Prize and support to the Austrian Research Institute for Artificial Intelligence. We also thank the L. Bösendorfer Company, Vienna, for the performance data used in these experiments.

References

Cambouropoulos, E. 2000. From MIDI to traditional musical notation. In *Proceedings of the AAAI Workshop on Artificial Intelligence and Music: Towards Formal Models for Composition, Performance and Analysis*. AAAI Press. To appear.

Chafe, C., Jaffe, D., Kashima, K., Mont-Reynaud, B., and Smith, J. 1985. Techniques for note identification in polyphonic music. In *Proceedings of the International Computer Music Conferenc*e. Computer Music Association, San Francisco CA.

Dixon, S. 1996. A dynamic modelling approach to music recognition. In *Proceedings of the International Computer Music Conferenc*e, pages 83–86. Computer Music Association, San Francisco CA.

Dixon, S. and Cambouropoulos, E. 2000. Beat tracking with musical knowledge. In *ECAI 2000: Proceedings of the 14th European Conference on Artificial Intelligenc*e. IOS Press. To appear.

Dolson, M. 1986. The phase vocoder: A tutorial. *Computer Music Journal* 10(4):14– 27.

Flanagan, J. and Golden, R. 1966. Phase vocoder. *Bell System Technical Journa*l, 45:1493–1509.

Kashino, K., Nakadai, K., Kinoshita, T., and Tanaka, H. 1995. Organization of hierarchical perceptual sounds: Music scene analysis with autonomous processing modules and a quantitative information integration mechanism. In *Proceedings of the International Joint Conference on Artificial Intelligenc*e.

Klapuri, A. 1998. Automatic transcription of music. Master's thesis, Tampere University of Technology, Department of Information Technology.

Martin, K. 1996. A blackboard system for automatic transcription of simple polyphonic music. Technical Report 385, Massachussets Institute of Technology Media Laboratory, Perceptual Computing Section.

Mont-Reynaud, B. 1985. Problem-solving strategies in a music transcription system. In *Proceedings of the International Joint Conference on Artificial Intelligenc*e. Morgan Kaufmann.

Moorer, J. 1975. *On the Segmentation and Analysis of Continuous Musical Sound by Digital Compute*r. PhD thesis, Stanford University, CCRMA.

Piszczalski, M. and Galler, B. 1977. Automatic music transcription. *Computer Music Journal* 1(4):24–31.

Scheirer, E. 1995. Extracting expressive performance information from recorded music. Master's thesis, Massachusetts Institute of Technology, Media Laboratory.

Scheirer, E. 1997. Using musical knowledge to extract expressive performance information from audio recordings. In Okuno, H. and Rosenthal, D., editors, *Readings in Computational Auditory Scene Analysi*s. Lawrence Erlbaum.

Schloss, W. 1985. *On the Automatic Transcription of Percussive Music: From Acoustic Signal to High Level Analysi*s. PhD thesis, Stanford University, CCRMA.

Watson, C. 1985. *The Computer Analysis of Polyphonic Musi*c. PhD thesis, University of Sydney, Basser Department of Computer Science.

Boolean Networks for the Generation of Rhythmic Structure

Alan Dorin
School of Computer Science & Software Engineering,
Monash University, Clayton, Australia 3800
aland@cs.monash.edu.au http://www.cs.monash.edu.au/~aland

Abstract

This paper describes an interactive multi-track MIDI sequencer for generating polyrhythmic patterns. Patterns are determined by the state of a set of autonomous Boolean networks, each node of which corresponds to a single musical event. This event will be triggered if the Boolean node is active and the sequencer step corresponds to its location in the pattern. Users may alter the networks which generate the patterns and change the note information they trigger whilst a sequence plays. Thus the sequencer may be used in live performance or in a studio. Additionally the sequencer may be left to generate its own patterns based on the behaviour of the Boolean networks.

Introduction

The generation of complex patterns from simple rules is a constant source of fascination for practitioners in the fields of computer generated imagery and sound [1,2,3,4,5,6,7,8]. This interest takes root in a deeper drive to comprehend the world around us. Where we may, we reduce complex outcomes to the combination of simple principles. We look for causal connections between otherwise mysterious and unpredictable events. We seek order in chaos and, in our own work, seek *complexity for free*.

This paper describes a musical event generator which has application as a multi-track interactive sequencer and generator of polyrhythms. The complex patterns it generates emerge from the neighbourhood interactions of binary switching elements. The tool is intended to work as an instrument for composition or as a stand-alone rhythm generator which, once established, runs without interference from a human user.

The combination of multiple independent rhythmic patterns gives an emergent and intricate shimmer with a "life" of its own. This phenomenon is quite irreducible to its component parts. It is this aspect of rhythm which first attracted the author to the idea of generating interwoven processes in a manner reminiscent of the functioning of a living thing.

Steve Reich hits the nail on the head when he states "Sometimes everything just comes together and suddenly you've created this wonderful organism" (Schwarz 1997). The resulting piece of music is alive, an organism, in the sense that a multitude of interacting processes combine to produce a transient dynamic entity with an energy not apparent on the dissecting table. Michael Nyman (1999) also uses the term musical "organism" in a discussion of the work of Phillip Glass.

Xenakis' percussion work *Pleiades* (1990) and Reich's pieces (for example *Drumming*, *Music for 18 Musicians* and *Sextet* [1987, 1997, 1986]) are good examples of the dynamism which may emerge from the careful orchestration of simple patterns. Of course Reich is best known for his use of *phasing*- a term coined to describe the procedure of playing repetitive patterns of different lengths against one another so that they gradually shift out of (and then back into) phase. Music from African, Chinese and Cuban sources is also frequently based around complex polyrhythmic structures (Copland 1999).

Reich was, in his earlier days at least, specifically interested in the *process* by which his music arose. “Once the process is set up and loaded it runs by itself” (Nyman 1999) The simplicity of a process such as playing a repetitious drum pattern, or the operation of mechanical tape players, contrasts strongly with the complexity of the result. Here then the composer has demonstrated a kind of power over sound which is not a reflection of their ability to conceive of intricate rhythms, so much as it illustrates an understanding of the origins of complexity in simplicity. A fascination with these ideas has played a part in the production of the software described shortly in this paper.

The behaviour of Wolfram’s one-dimensional cellular automata (CA) (1984) is also of relevance to the present discussion. Wolfram established a closed loop of automata which is usually depicted as a row of cells. The two end cells are assumed to be neighbours in the same way that adjacent cells within the line neighbour one another.

Wolfram’s cellular automata are binary state machines–they may be in the state *on* or *off*. The state a given automaton will be in at a discrete time step immediately following the present one is determined by its present state and the present state of its two neighbours. All automata in a row are updated synchronously.

It is possible to exhaustively document the patterns Wolfram’s connected machine produces by running through all the possible rules governing the state changes of its automata. Wolfram found four general types of behaviour in his machines. He found machines which:

(i) move into a homogeneous state (limit-point);
(ii) move into simple, separated, periodic structures (limit-cycle);
(iii) produce chaotic aperiodic patterns (strange attractors);
(iv) produce complex patterns of localized structures.

The relationship between these behaviours and the production of music or imagery using procedural means has been discussed in my paper, “A Classification of Physical Processes for Virtual-Kinetic Art” (Dorin 1999). Various authors and composers have produced systems which generate musical structures from cellular automata (McAlpine, et. al. 1999; Dorin 1999 “Liquiprisms”).

Wolfram’s CA’s are, like any dynamical system, capable of producing patterns for use as rhythms in music. However for various reasons cellular automata were not suitable candidates for the work presented in this paper. Instead, Boolean networks were used as the source of rhythmic complexity. The reasons for this are explained in the next sections. Following this, the present system for generating musical patterns is explained in some detail. The tool’s visual interface is also described since its production required numerous decisions to be made regarding the effective presentation of a large amount of information. Finally, general thoughts are given on the utility of the system, some future work is proposed and conclusions drawn.

What are Boolean networks?

Boolean networks are a connected set of binary state machines (nodes) similar to the automata in Wolfram’s CA’s. Nodes in a Boolean network may be in one of two states: *on* or *off*. In this paper only *synchronous* Boolean networks will be considered. These are networks in which the

node states are updated simultaneously, just as they were with Wolfram's CA's.
The future state of a node depends on the states of nodes in the network designated as that node's *inputs*. A node may feedback its own state as a self-input. Additionally, inputs may be received from outside the network. A Boolean network which receives no input from the "outside world" is an *autonomous* network. An autonomous, synchronous Boolean network is clearly a special case of the CA discussed in the previous section.

The state of a node in a Boolean network at a future time is governed by a logical rule or Boolean function which operates on the node's inputs. Examples include the NOT, AND, OR and XOR functions which have become so familiar in this digital age. The set of states a system passes through as its nodes are updated is known as its *trajectory*. Kauffman (1993) discusses the trajectories of autonomous Boolean networks with exactly two inputs to each node. The first important concept to grasp is that any actual network has a finite number of binary nodes. Therefore the machine as a whole has a finite number of possible states (specifically 2^N where N is the number of nodes). Since the network is deterministic, if it returns to a state from which it previously emerged (which it must do eventually since there are a finite number of possible states), the network must be in a *limit cycle*. The number of steps in a limit cycle may range from one (a limit point) to 2^N.

The set of machine states which lead the network to fall into a particular limit cycle is the *basin of attraction* for that limit cycle. The limit cycle is the *attractor* for the basin. A network must have at least one limit cycle but some networks have many, each with its own basin of attraction. Some basins are very broad (a large number of machine states lead to their attractor), others are quite narrow. Whilst the machine is within the basin of attraction but not yet within its limit cycle the machine is said to be in a *transient*.

Besides basins of attraction, the other major property of interest is the susceptibility of a particular attractor to disturbance. Some limit cycles are more stable than others. That is, even if some nodes in the system have their state flipped, the basin of attraction around the attractor ensures the trajectory of the system returns from its transient to the previous limit cycle. Sometimes, even major structural alterations to a network such as changes to node connections or node transition rules do not upset limit cycles.

Why use Boolean networks?

Of course multi-layered rhythms may be constructed and edited by hand in a conventional sequencer. There is no disputing the value of this technique. However the current task is the generation of such patterns using the computer as an able assistant, even a master. A system is sought which may be influenced or guided by a user but which does not require complete user-specification.

Some CA rule sets are known to be effective producers of complex patterns and cyclic activity. For example Conway's *Game of Life* falls into this category (Gardner 1970, 1971). Some researchers have used CA's like this for music event generation (McAlpine, et. al. 1999) or generated their own CA's with desirable properties (Dorin 1999 "Liquiprism"), however the mapping from CA to music is almost arbitrary. Why should the CA rules produce music? The researchers may as well revert to Cage's dice or coins (Ford 1997) or any of a myriad of other physical systems, and then attempt an arbitrary mapping from arbitrarily chosen system to musical events– an idea this author feels was exhausted in the sixties.

Effective transition rules for cellular automata are notoriously difficult to come up with. Wolfram's studies and the studies of Sims (1992) and Langton (1986), as well as this author's own investigations, indicate that of the enormous range of possible rules, the majority reliably lead a system to a fixed point or short cycle, rather than to one in which lengthy cyclic or complex patterns emerge. Boolean networks reliably produce repeating patterns which may be easily altered without causing the system to fall into uninteresting limit points–although these are not completely eliminated, and nor should they be!

It is therefore primarily the properties discussed in the previous section which make autonomous Boolean networks appropriate for the present task. As mentioned, Boolean networks reliably fall into limit cycles. These may be used as simple repeating patterns to generate musical events. Transient changes may be introduced which alter a limit cycle temporarily, thereby adding variation to a pattern without removing it altogether. Alternatively, changes to the network may be introduced which produce more complex effects on the cycle. These changes may be readily carried out in real-time by a human composer as the network is running, resulting in the immediate feedback so helpful in musical composition. They may also be carried out automatically by the computer.

Changes to the transition rules of a Boolean network do not dramatically change the *kind* of behaviour of the system. This is not true for cellular automata transition rules. For example, even a minor change to the rule set of Conway's *Game of Life* may result in the total extinction of all patterned activity on the grid. Such behaviour is not (usually) desirable from a musical standpoint.

The Boolean sequencer

The Boolean sequencer of this paper presents to the user a variable number of *tracks*. Each track contains a variable number of *nodes*, each of which corresponds to a measure in a musical bar. The system works like a conventional MIDI step-sequencer in loop mode. Each track is assigned a MIDI channel and each node a pitch, velocity and duration. Tracks (and nodes) may be muted or left to sound. The sequencer regularly steps through each node in the bar playing its assigned note at the appropriate time before looping back to repeat the pattern.

Unlike the notes in a conventional sequencer, the state of each node in the Boolean sequencer is determined by the state of its neighbours in the previous bar. Each node is connected to the two nodes adjacent to it in the pattern (first and last notes are also connected to one another) and assigned a transition rule: OR; XOR; NOT or AND. When the sequencer step corresponds to a particular node, that node will only sound its note if it is in the *on* state (and the track is not muted). *Off* notes remain silent. Node states are updated according to their individual transition rules at the end of each bar.

In user-mode the sequencer's node types and note information may be altered interactively. This allows the composer to experiment with different Boolean networks and the rhythms they produce. The user interface of the Boolean sequencer is described in the following section. The system may be left to run its own course in an automatic mode which relies on random perturbation of the Boolean networks to move the system through complex wandering trajectories.

The Boolean sequencer's U.I.

This section focuses on the visual design of the sequencer interface. Users may interact with the interface using the computer keyboard and mouse. In future versions MIDI keyboard operation will be incorporated also.

It was decided that a traditional piano-roll interface such as that found on most off-the-shelf-software would not be suitable for the Boolean sequencer. Reasons for this included the need to:

677 represent node types (XOR, AND etc.);
678 have the sequencer loop rapidly across a single bar whilst it is playing (This accentuates the need to clearly indicate which note is sounding at any given time, allowing the user to distinguish a sound from many others and identify its visual representation);
679 indicate whether a given node was sounded/will sound during the current bar.

In addition, some conventional interface requirements were relevant to the Boolean sequencer including the need to:

- represent pitch, velocity and duration;
- represent node temporal ordering;
- allow simple editing of node parameters during sequence playback;
- represent the current location in the bar;
- represent multiple tracks simultaneously.

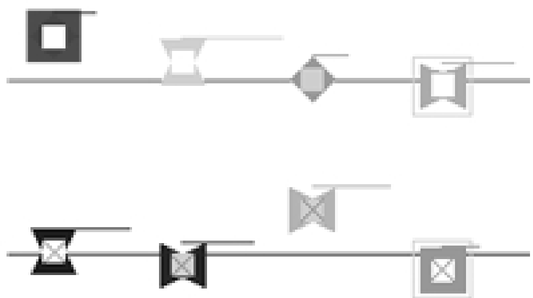
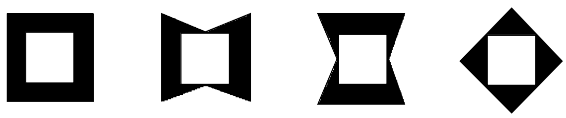


Fig. 1: Snapshot of the Boolean sequencer interface

Figure 1 shows two sequencer tracks of four nodes each. The (top) track indicated by the bright red center-line is currently selected for editing. The (top left) node indicated by the solid red square is the currently selected node. Hollow yellow squares indicate the nodes (right most) in the sequence which are currently being triggered. The rule determining the behaviour of a node (its type) is indicated by the icon used to represent it. There are presently four node types (fig. 2).



OR XOR NOT AND

Fig. 2: Sequencer iconography

--[An unscientific aside: The "logic" for developing these icons came from considering a node's behaviour in the context of the sequencer. Whilst the reasoning requires a little creativity, once learned the node representations are (hopefully) easy to remember…

A NOT node switches on if neither of its neighbours is on at the previous time step. Hence its sides are hollow to indicate that the adjacent nodes were not on. An XOR node requires either of its neighbours to be on at the previous time step (its sides are flat) but not both neighbours (its top and bottom are hollow). An OR node requires either of its neighbours to be on at the previous time step (its sides are flat) but it is not fussed if both are on (its top and bottom are flat). The AND node requires *both* of its neighbours to be on (its sides, top and bottom actively poke out from the icon's centre).

Hopefully the reader can find some sense in this description. If not, he/she will just have to learn the node icon shapes by heart to use the sequencer effectively. Thus ends the unscientific aside.]--

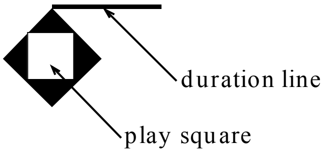


Fig. 3: Node components

The horizontal lines from the node tops (fig. 1&3) indicate a note's duration. The grey-level of a node indicates its MIDI velocity ranging from black (velocity=0) to nearly white (velocity=127). A track which has been muted will display with a red X drawn in the *play square* of each node. The play square is drawn in yellow when a node is in the *on* state, otherwise it is left vacant.

The horizontal line through each track indicates middle C. The positions of nodes above and below this indicate their relative pitch. Exact pitches are not visualized–this sequencer was intended primarily for users who *listen*.

As mentioned earlier, the yellow square in each track surrounds the currently triggered step in the sequence. The square shifts up and down with the current node pitch as the sequence progresses. This method of indication was used instead of the vertical bar found in many sequencers because a vertical bar (which travels horizontally as the sequence progresses) gives no visual indication of the musical event which it triggers. By having the indicator move up and down with pitch and horizontally with time, a visual indication of the pattern's flow is given.

Comments on using the sequencer, its present and future interface

The number of tracks, number of nodes per track and the playback speed of the sequencer can all be determined by the user. In the current system they cannot be varied after initiation but it is anticipated this feature will be incorporated eventually. Although it is possible to operate the interface with a mouse, the computer keyboard shortcuts provided to alter the sequencer's parameters are simpler to activate during playback. All aspects of nodes and tracks may be altered via the appropriate keyboard commands in real-time.

Presently the pitch, duration and velocity of a node may be altered incrementally up or down using p/P, d/D and v/V shortcuts. This is clearly an ineffective means of specifying note parameters. Instead, these values might be read directly from a MIDI keyboard as a user plays a note. Nevertheless, to date MIDI keyboard specification of note data has not been required and the feature is presently unimplemented.

At present all sequencer tracks are updated synchronously. The system therefore runs like a conventional sequencer or drum machine in this regard. It is hoped in future that each track will have its own speed parameter so that patterns may shift in and out of phase with one another in the fashion of the rhythms used by Reich.

A further anticipated addition to the system is the incorporation of an arpeggiator mode. Nodes in a sequence might be taken as commands to raise or lower a note (or notes) in a complex pattern determined by the Boolean network.

Extensive experimentation with the present system has revealed it to be a versatile tool for rhythm construction. Whilst this author finds careful tweaking to be needed, especially of note pitches, other composers might find pleasing patterns form almost at random.

With percussive sounds the sequencer excels at complex rhythm production and is easily tweaked to maintain interest through variation. Complex developing sounds and samples may also be triggered by the sequencer. As these are truncated, triggered and re-triggered in unusual patterns complex timbral/rhythmic effects emerge.

This author has found the use of the NOT and AND node types alone to be sufficient to generate a diversity of multi-bar patterns. An individual track/network regularly falls into limit cycles of two to four bars using only these node types. The addition of node types which distinguish between left and right neighbours is also being considered to produce patterns which trigger events that move left and right through the bar.

Bar lengths of less than five nodes/beats per network seem to fall into very short and uninteresting limit cycles. More interesting results may be achieved using networks with more nodes than this although beyond about ten nodes per network there is no noticeable increase in the complexity of pattern produced. In the future it would be interesting to document the behaviour of networks with respect to the number of nodes they contain.

Implementation

The Boolean sequencer was implemented on a Macintosh G3 running MacOS 8.5. It was written in C++ using Metrowerks Code Warrior, the OpenGL library for Macintosh v1.1, and Opcode's OMS 2.3.8. OMS and OpenGL need to be installed to run the sequencer.

Conclusion

Like any art-making tool, one view is that it is inevitably up to the artist to make the most of its capabilities. Like any algorithmic means of art production, its output may be taken "as is" to represent the process which underlies it. In this case the result may be worthy in and of itself, or simply irrelevant. Take your pick! Whichever view you hold, Boolean networks have been shown to be a viable means of producing complex patterns which are nevertheless easily manipulated by a user. These patterns may be usefully employed in the production of complex, changing rhythmic structure.

Bibliography

Copland, A. 1999. "What To Listen For In Music", first published, McGraw-Hill, 1939, new edition, Mentor Books.

Dahlstedt, P., and Nordahl, M. 1999. "Living Melodies: Co-evolution of Sonic Communication", in Proceedings, *First Iteration*, Dorin & McCormack (eds). CEMA, Monash University, Australia, pp. 56-67.

Dorin, A. 1999. "Classification of Physical Processes for Virtual-Kinetic Art", in Proceedings, *First Iteration*, Dorin & McCormack (eds). CEMA, Monash University, Australia, pp. 68-79.

Dorin, A. 1999. "Liquiprism", in *Process Philosophies*, curated by Dorin & McCormack. First Iteration Conference, Monash University, Australia.

Ford, A. 1997. *Illegal Harmonies, Music in the 20th Century*. Hale & Iremonger.

Gardner, M. 1970. "Mathematical Games: The Fantastic Combinations of John Conway's New Solitaire Game 'Life'." *Scientific American* 223(4): 120-123.

Gardner, M. 1971. "Mathematical Games: On Cellular Automata, Self-Reproduction, the Garden of Eden and the Game 'Life'." *Scientific American* 224(2): 112-117.

Hiller, L. 1981. "Composing with Computers: A Progress Report", *Computer Music Journal* 5(4).

Kauffman, S.A. 1993. *The Origins of Order - Self Organization and Selection in Evolution*. Oxford: Oxford University Press.

Langton, C.G. 1986. "Studying Artificial Life with Cellular Automata." *Physica 22D*. North-Holland, pp. 120-149.

Laske, O. 1981. "Composition Theory in Koenig's Project One and Project Two." *Computer Music Journal* 5(4).

McAlpine, K., Miranda, E., and Hoggar, S. 1999. “Making Music With Algorithms: A Case-Study System.” *Computer Music Journal* 23(2):9-30.
Nyman, M. 1999. *Experimental Music, Cage and Beyond*. 2nd ed. Cambridge: Cambridge University Press.
Prusinkiewicz, P., and Lindenmeyer, A. 1990. *The Algorithmic Beauty Of Plants*. Springer Verlag.
Reich, S. 1986. *Sextet/Six Marimbas*. Elektra / Asylum / Nonesuch Records.
Reich, S. 1987. *Drumming*. Elektra / Nonesuch Records.
Reich, S. 1997. *Music For 18 Musicians*. Nonesuch Records/Warner Music.
Schwarz, K.R. 1997. “Music For 18 Musicians, Revisited.” Liner notes in CD, *Music For 18 Musicians.* Reich, S. Nonesuch Records/Warner Music.
Sims, K. 1992. “Interactive Evolution of Dynamical Systems.” *Toward a Practice of Autonomous Systems: Proceedings of the First European Conference of Artificial Life.* Varela & Bourgine (eds). Massachusetts: MIT Press, pp. 171-178.
Turing, A.M. 1952. “The Chemical Basis of Morphogenesis.” *Philosophical Transactions of the Royal Society* 237:37-72
Turk, G. 1991. “Generating Textures on Arbitrary Surfaces Using Reaction – Diffusion.” SIGGRAPH 91. *Computer Graphics* 25(4):289-298.
Wolfram, S. 1984. “Universality and Complexity in Cellular Automata.” *Physica 10D.* North-Holland, pp1-35
Xenakis, I. 1990. *Pleiades*. Grammofon AB BIS.

Multi-feature Musical Instrument Sound Classifier

I. Kaminskyj
Electrical & Computer Systems Engineering
Monash University
Clayton Campus, Victoria, Australia.
Email: kaminskyj@eng.monash.edu.au

Abstract

A classification system has been developed which utilises three features to classify isolated monophonic musical instrument sounds: RMS amplitude envelope, constant Q transform (CQT) frequency spectrum and multidimensional scaling analysis (MSA) trajectories. Nineteen musical instruments of definite pitch were used to test the system, covering the note range C3-C6 and representing the major musical instrument families and subfamilies. Musical instrument sound classification was performed using the leave-one-out classification scheme with different implementations of Nearest Neighbour classification (NNC) (varying k values). The classification results were evaluated in terms of classification accuracy and reliability. The individual feature classifier results showed that the MSA trajectory feature provided the best classification results, followed by the CQT frequency spectrum, and then the RMS amplitude envelope. Combining the individual classifiers together to form one classification system provided further improvement, yielding an overall accuracy of 82% and reliability of 83%.

Introduction

Until only very recently, most research in the area of automatic music transcription has focused on the automatic identification of the notes being played by musical instruments, without regard to which instrument played which note (Brown 1992 and Foster et al. 1982). Its really only been in the last 5 or so years where interest has begun to occur in the problem of automatic identification of isolated and/or continuous musical instrument sounds; a problem closely interrelated with that of automatic music transcription.

Kostek and Wieczorkowska (1996) evaluated the use of artificial neural networks (ANN) and rough sets to reliably distinguish isolated musical instruments tones.

For the rough set algorithm, experiments were carried out with sounds from nine orchestral wind instruments. The algorithm input comprised vectors of fourteen features: six steady-state as well as the eight from the attack. Experiments including modified leave-one-out tests yielded an accuracy rate of 65%. For the neural network algorithm, four wind instruments were used for training and testing. Recognition results at 97% were achieved.

Additional experiments (Kostek & Krolikowski 1997) were carried out where training of the ANN was performed using either the left or right channel recordings of four slightly different wind instruments; the McGill University Master Samples (MUMs) CD recordings all being stereo. Classification results for fourteen features ranged from 94.4–100%.

Brown (1998) performed a series of experiments relating to the computer identification of continuous flute, clarinet, oboe and saxophone sounds using autocorrelation and cepstral coefficients as features. She found the best pairwise classification results (90.6 – 100%) were obtained using cepstral coefficients based on a constant Q transform calculated for short segments of each instrument sound using a k-means algorithm and a Bayes decision rule.

Martin et al. (1998) applied a statistical pattern-recognition technique to the classification of isolated musical instrument tones within a taxonomic hierarchy. Thirty-one perceptually salient acoustic features were extracted from the output of a human auditory model based on the log-lag correlogram. 1023 isolated tones from fifteen orchestral instruments were used to test the system, recorded from the MUMs CDs.

Using 70%/30% splits between training and test data, the classifiers distinguished pizzicato from sustained tones with an accuracy of 99%. Instrument families were identified with an accuracy of 90%, while individual instruments were identified with an accuracy of 70%. Martin et al. claimed that these results compared favourably with human performance on the same task.

Ichiro (2000) designed a computer-based classifier to recognise simple continuous tones from the MUMs CDs of musical instrument sounds. 39 different timbres from 23 orchestral instruments played at different pitches were used as samples for an exemplar-based learning system that incorporated a k-NNC with genetic algorithm.

The weighted features used during classification included moments calculated from both the steady-state and dynamically changing spectral envelopes as well as spectral irregularity and tristimulus features describing the spectral envelopes. While the recognition accuracy varied greatly between the instruments, the average overall accuracy was 68%.

Although the results obtained previously are encouraging, they are generally limited to a restricted number and type of musical instruments. In some cases, only wind instruments are used, while in others, only orchestral instruments. The classification system described in this paper uses a more general selection of instruments together with a novel approach to combine individual feature classifiers, applicable particularly where limited data sets are available. The combination process makes use of the application of weighted ranked individual classifier results together with the confusion matrix as a mechanism for determining the relative confidence that can be applied to each classifier classification result. The classification results obtained for nineteen musical instrument show the merit of the approach taken.

Data collection

Sounds from nineteen musical instruments of definite pitch were chosen for system testing. 517 recordings of these instruments were taken directly from the MUMS CDs. The instruments included: guitar (plucked string), violin, cello and double bass (bowed string), piano (struck string), flute (air reed wind), accordion, clarinet, saxophone (single mechanical reed wind), oboe and bassoon (double mechanical reed wind), organ (air/mechanical reed wind), trumpet, trombone, French horn, and tuba (lip reed wind) and xylophone, glockenspiel and marimba (percussive definite pitch) (Olson 1967).

The note range used was C3-C6 of the equally tempered musical scale. A three-octave note range was chosen to limit the system complexity, to provide practical limits on the number of recordings required and to simplify instrument selection. Not all instruments covered the complete note range, but each had at least some notes falling within it. Recordings were made from the MUMs CDs using the SoundBlaster WaveStudio program that allowed recording, display, editing and playback of sounds. All recordings comprised 2 seconds of edited sound.

Data preprocessing (feature extraction)

Once the raw musical instrument recordings were edited, they needed to be preprocessed before classification. The aims of this stage of the system were (a) to reduce the amount of data that needed to be handled by the classifier(s), as well as (b) to perform feature extraction, so as to make the classification task as easy as possible. The three features discussed below were evaluated for their effectiveness in instrument classification. They were chosen either because they had been used successfully in previous research in solving similar problems (Kaminskyj 1995) or had proven important in allowing humans to discern between musical instrument sounds (Handel 1995) or simply looked intuitively promising (Hourdin et al.1997).

RMS amplitude envelope

Many schemes exist for measuring the amplitude envelope of a waveform. The approach used in this research involves the calculation of the short-term RMS value over 2-3 periods of the fundamental frequency (Kaminskyj 1996).

CQT frequency spectrum

Many spectral representations are possible for musical waveforms. The representation used in this research is the constant Q transform developed by Brown (1991), because of its logarithmically spaced spectral bins which align with musical scale note frequencies.

MSA trajectories

MSA was performed on harmonic extracts of CQT spectra of musical instrument sounds (Kaminskyj 1999). Principal component analysis (PCA) (Jollife 1986) was applied to the resultant CQT spectral snapshots to determine the most significant attributes characterising each sound. The resultant sound trajectories yielded a "tone signature" for each instrument sound.

Classification algorithms

For "curse of dimensionality reasons" (Pandya & Macy 1996), the classification approach adopted for this work uses three *k*-NNCs utilising the Euclidean distance measure to classify musical instruments using different features and the leave-one-out classification scheme (Duda & Hart 1973).

With this scheme, all recordings were used to test the classifiers, but only one at a time. As each recording was classified, the remaining recordings formed labelled reference samples to be used to compare against the test recording. In this way, all recordings were available to evaluate the classification system performance. In addition, a large number of labelled reference samples were available to aid test sample comparison and classification.

With the leave-one-out scheme, classification was obtained by determining the *k* closest reference samples to the test sample, with *k* values of 1, 3 and 5 explored. Either weighted or unweighted majorities were then used to determine the most likely instrument to have produced the test sample. Classification of instruments sounds occurred using either the raw feature waveforms (Direct) or alternatively, after PCA was performed on these waveforms.

Classifier results were evaluated in terms of accuracy and reliability (Pandya & Macy 1996). Classification accuracy indicates how many of the test samples were correctly classified. Classification reliability on the other hand, provides an indication as to the confidence that can be placed on the classifier results; in other words, if a classifier gives a clarinet result for example, how often is this result correct. A classifier with good classification accuracy but poor reliability for certain instruments is of limited use for these specific instruments.

Initially, the classification system was devised such that a test sample was compared against all reference samples. This approach however resulted in misclassifications which were obviously incorrect (eg. C3 marimba tone misclassified as a glockenspiel, whose note range doesn't extend below G6). As it has been assumed that the note of the test sample is known *a priori* by the classification system, such an obvious misclassification had to be somehow avoidable.

As an end to avoiding such misclassifications, octave search limits were then applied when performing comparisons between the test and reference samples for each of the three feature classifiers. Octave search limits invariably improved the classification results obtained. For the most promising input feature type (direct or PCA) for each individual classifier, the search limits were then further reduced to ± 5 notes, ± 4 notes, ± 3 notes (half octave), ± 2 notes and ± 1 notes. In this way, the optimum search limits for each individual feature classifier was determined. For example, a ± 2 note limit applied to a G3 test tone meant that comparisons for this test tone only occurred with F3, F3#, G3, G3# and A3 reference tones.

Results

Tables 1-6 show the classification accuracy results for the individual feature classifiers. Table 7 shows the results obtained when combining the individual feature classifier results in a variety of ways, including using the confusion matrix (Pandya & Macy 1996) instrument accuracy (Acc) and reliability (Rel) values as combination weights (See Tables 8-10).

Table 1: RMSPCA

NNC k Value	Limits	% accuracy Majority (unwghtd)	% accuracy Majority (weighted)
1	none	39%	N/A
3	none	41%	39%
5	none	43%	43%
1	octave	46%	N/A
3	octave	45%	46%
5	octave	48%	48%

Table 2: RMS Direct

NNC k Value	Limits	% accuracy Majority (unwghtd)	% accuracy Majority (Weighted)
1	none	48%	N/A
3	none	50%	48%
5	none	49%	49%
1	octave	56%	N/A
3	octave	58%	56%
5	octave	56%	57%
1	± 5	56%	N/A
3	± 5	60% (best)	56%
5	± 5	55%	59%
1	± 4	56%	N/A
3	± 4	59%	56%
5	± 4	55%	58%
1	± 3	57%	N/A
3	± 3	56%	57%
5	± 3	53%	56%
1	± 2	54%	N/A
3	± 2	54%	54%
5	± 2	48%	52%
1	± 1	46%	N/A
3	± 1	41%	46%
5	± 1	30%	38%

Tables 8-10 show the confusion matrices for the best individual feature classifiers in Tables 1-6, while Table 11 shows the confusion matrices for the best combined feature classifier in Table 7.

Discussion

The results in Tables 1-6 indicate that the best classification results for the RMS amplitude envelope NNC were obtained using the direct waveform feature (accuracy 60% & reliability 60%). On the other hand, for the CQT spectral NNC and MSA trajectory NNC, best results were obtained with the PCA features (CQT accuracy 66% & reliability 68%; MSA accuracy 75% & reliability 76%). It is perhaps not surprising that the latter two feature NNCs both favour the PCA features as both are closely related together; the MSA trajectories being derived from the CQT spectral snapshots over the evolution of each musical instrument sound.

Of the three features, the MSA trajectories provide the most powerful discrimination for the nineteen instruments used for testing. This is perhaps not too surprising as these trajectories provide the most comprehensive amount of information regarding both the temporal and spectral characteristics of each individual instrument sound.

Table 3: CQTPCA

NNC k Value	Limits	% accuracy Majority (unwghtd)	% accuracy Majority (weighted)
1	none	62%	N/A
3	none	58%	62%
5	none	58%	60%
1	octave	65%	N/A
3	octave	62%	65%
5	octave	61%	63%
1	± 5	65%	N/A
3	± 5	63%	65%
5	± 5	62%	63%
1	± 4	66%	N/A
3	± 4	64%	66%
5	± 4	62%	63%
1	± 3	66% (best)	N/A
3	± 3	64%	66%
5	± 3	61%	63%
1	± 2	65%	N/A
3	± 2	62%	65%
5	± 2	57%	62%
1	± 1	61%	N/A
3	± 1	55%	61%
5	± 1	45%	55%

Table 4: CQT Direct

NNC k Value	Limits	% accuracy Majority (unwghtd)	% accuracy Majority (Weighted)
1	none	51%	N/A
3	none	50%	51%
5	none	49%	51%
1	octave	53%	N/A
3	octave	53%	53%
5	octave	52%	53%

Table 5: MSAPCA (normalised trajectory)

NNC k Value	Limits	% accuracy Majority (unwghtd)	% accuracy Majority (weighted)
1	none	67%	N/A
3	none	68%	67%
5	none	66%	68%
1	octave	71%	N/A
3	octave	71%	71%
5	octave	71%	72%
1	± 5	72%	N/A
3	± 5	72%	72%
5	± 5	70%	73%
1	± 4	74%	N/A
3	± 4	74%	74%
5	± 4	71%	72%
1	± 3	74%	N/A
3	± 3	75% (best)	74%
5	± 3	70%	73%
1	± 2	73%	N/A
3	± 2	70%	73%
5	± 2	67%	70%
1	± 1	70%	N/A
3	± 1	62%	70%
5	± 1	49%	59%

Table 6: MSA Direct

NNC k Value	Limits	% accuracy Majority (unwghtd)	% accuracy Majority (weighted)
1	none	53%	N/A
3	none	51%	53%
5	none	51%	51%
1	octave	56%	N/A
3	octave	55%	56%
5	octave	55%	55%

Table 7: Combined Classifier

NNC k Value	Confn Matrix Wghtg Method	% accuracy Majority (unwghtd)	% accuracy Majority (weighted)
1	none	78%	N/A
3	none	80%	80%
5	none	75%	81%
1	Rel	78%	N/A
3	Rel	81%	81%
5	Rel	77%	82% (best)
1	Acc	78%	N/A
3	Acc	80%	78%
5	Acc	75%	81%
1	Rel+Acc	78%	N/A
3	Rel+Acc	80%	79%
5	Rel+Acc	74%	79%

The varying search limit study results of Tables 2, 3 and 5, indicate that for the RMS amplitude envelope feature, best results are obtained for a roughly octave search limit.

For the CQT and MSA features however, best results are obtained for a half octave search limit, before extraneous confusions with reference samples further away in frequency became a problem. These results tend to highlight that the proximity effect (the range of notes over which features remain roughly invariant and instrument confusions minimal) is feature dependent.

The combined feature results highlight that the confusion matrix reliability weighting provides the best classification results [*k=5*, weighted majority] (accuracy 82% and reliability 83%), although this is only marginally better than the results obtained for the non-confusion matrix weighted results. As reliability indicates the confidence that can be placed on any given classifier results, it is not surprising that of the three confusion matrix weightings, it provides the best results. What is surprising is that such little improvement was obtained by using the confusion matrix weightings. This leads to the conclusion that the rankings of the individual classifiers beyond the first choice provide very useful information when combining the individual classifier results together.

For the combined feature classifier, only two instruments of the nineteen tested did not provide reasonable classification results (deemed a minimum of 67% classification accuracy and reliability). These two instruments were cello and saxophone, although even cello came very close (accuracy 65% & reliability 100%). This result indicates that for these two instruments at least, the three features chosen were not adequate enough to allow reliable discernment.

It is interesting to note that few researchers have paid any attention to the issue of reliability of their classification results; most focused purely on the overall system accuracy.

Table 8: CONFUSION MATRIX RMS Direct: ±5 NOTE LIMIT k=3 UNWEIGHTED MAJORITY

	Acc	**Bas**	**Cel**	**Cla**	**Dbb**	**Flu**	**Glo**	**Gui**	**Hor**	**Mar**	**Obo**	**Org**	**Pia**	**Sax**	**Tro**	**Tru**	**Tub**	**Vio**	**Xyl**	%Acc
Acc	22		1	1	2	3									2	1		2		65%
Bas		12		1					2					1	1		2			63%
Cel	1		5	1				2	3			5			1		2			25%
Cla	1	1		30										4						83%
Dbb	1				11							2						3		65%
Flu	6					16					1	1				1				64%
Glo							6													100%
Gui			2					33					1						1	89%
Hor	1			1					21			1		1	2					78%
Mar							2			23			5						7	62%
Obo	1	1									17	1		1	3	1	1	1		63%
Org	2	4	4	1	1	4			1		1	12		1	1	2		3		32%
Pia								3		11			20						3	54%
Sax		1		13										22	1					59%
Tro		2		1	1				2			1			17	3	1			61%
Tru	3														7	13		2		52%
Tub		3		1					1		2				4		9			45%
Vio	2				6	1					1				2	3	2	13		43%
Xyl								1		5			4						10	50%
%Rel	55	50	42	60	52	67	75	85	70	59	77	52	67	73	41	54	53	54	48	60/60

Table 9: CONFUSION MATRIX CQTPCA: ± 3 NOTE LIMIT k=1NNC

	Acc	Bas	Cel	Cla	Dbb	Flu	Glo	Gui	Hor	Mar	Obo	Org	Pia	Sax	Tro	Tru	Tub	Vio	Xyl	%Acc
Acc	22		1		1	1						2	1	1	2	2		1		65%
Bas		12				1			2				2				2			63%
Cel	1		10	1	7					1										50%
Cla	1		2	22							1	1		4	1	3		1		61%
Dbb			1	2	8			2						2	1		1			47%
Flu					1	21			1		1				1					84%
Glo							5											1		83%
Gui	1				1	1		28			2		2	1	1					76%
Hor									21						6					78%
Mar								2		28		1	1				2		3	76%
Obo				5		1		4			10	1		3	1	1		1		37%
Org	5		1			1						26		1		3				70%
Pia								4	2				28	1					2	76%
Sax	2	1		2	1			3	2		5		1	15	2	1		2		41%
Tro									8					1	19					68%
Tru	1			2							1	1				19		1		76%
Tub		2							1								17			85%
Vio	3		3	1							2	2		2		1		16		53%
Xyl								1					2	2					15	75%
%Rel	61	80	56	63	42	81	100	64	57	97	45	76	76	45	56	63	77	70	75	68/66

Although in general, overall reliability tracks overall accuracy quite closely, this is not necessarily the case for an individual instrument. As reliability directly indicates the confidence that can be placed on the results of a classifier for a given instrument, it is of primary importance in quantifying the performance of a classifier; accuracy may simply not be enough.

For example, consider clarinet in Table 8. For this classifier, clarinet was accurately classified 83% of the time, which at first glance is encouraging. Unfortunately, when this classifier provided a clarinet result, it was only right 60% of the time; considerably worse than one would expect from just the accuracy figure.

Table 10: CONFUSION MATRIX MSAPCA: ±3 NOTE LIMIT k=3 UNWEIGHTED MAJORITY

	Acc	Bas	Cel	Cla	Dbb	Flu	Glo	Gui	Hor	Mar	Obo	Org	Pia	Sax	Tro	Tru	Tub	Vio	Xyl	%Acc
Acc	28					1									2	2		1		82%
Bas		14				1			2						1	1				74%
Cel			12	2	3			1								1	1			60%
Cla				25		1					1			8		1				69%
Dbb					15	1												1		88%
Flu	1				1	13		1	3					1	4			1		52%
Glo							6													100%
Gui			2					32					2				1			86%
Hor									24						3					89%
Mar										31			2						4	84%
Obo				3							13	1		3	1	4		2		48%
Org				1								31				5				84%
Pia								5		5			25						2	68%
Sax	1			9		2		2			3			17	2	1				46%
Tro									6						22					79%
Tru											3					22				88%
Tub																	20			100%
Vio	1				2	1					3				2			21		70%
Xyl							1			2			1						16	80%
%Rel	90	100	86	63	71	65	86	78	69	82	57	97	83	59	59	61	87	81	73	76/75

Table 11: CONFUSION MATRIX NNC Combined: k=5 WEIGHTED MAJORITY, Confusion Matrix weighted (reliability)

	Acc	Bas	Cel	Cla	Dbb	Flu	Glo	Gui	Hor	Mar	Obo	Org	Pia	Sax	Tro	Tru	Tub	Vio	Xyl	%Acc
Acc	27					3						2			1	1				79%
Bas		16				1											2			84%
Cel			13	1	2			2								1	1			65%
Cla				31		1					1			3						86%
Dbb					14							1			1			1		82%
Flu					1	22			1						1					88%
Glo							6													100%
Gui								35					1				1			95%
Hor									25						2					93%
Mar							1			33			1						2	89%
Obo	1			3							18	1		1		1		2		67%
Org												33				4				89%
Pia								3		2			31						1	84%
Sax		1		10		1		1	2					21	1					57%
Tro									7						21					75%
Tru											1					24				96%
Tub																	20			100%
Vio	2				3						1	1			1			22		73%
Xyl								1		3			1						15	75%
%Rel	90	94	100	69	70	79	86	83	71	87	86	87	91	84	78	75	83	88	83	83/82

Similar results are obtained for bassoon in Table 8, French horn in Table 9 and trombone in Table 10.

Conclusion

The classification results obtained for the combined classifier system provide encouragement that the approach taken bears merit, particularly where limited data sets are available. This has been achieved using a comprehensive, yet general selection of instruments (rather than just wind or orchestral), which should in turn make the results obtained more generally applicable across a wide range of instruments. It is also surprising that the results were obtained for so many instrument using only three features. This shows in particular, the strength of the MSA trajectory feature for discriminating between different musical instruments.

The results achieved question the validity of modeling a musical instrument classification system exclusively on the capabilities of the human auditory system. Given the limitations humans have demonstrated in identifying musical sounds, such approaches necessarily restrict system performance to at best mimic the results achieved by humans for such tasks; yet never possibly surpass them.

Although the system described is principally intended for off-line analysis of musical instruments sounds, Ichiro (2000) recently demonstrated that using similar classification algorithms with current day computing facilities, real-time performance is achievable.

The work described is but one more step towards the development of a robust musical instrument classifier. Much work still remains. Possible directions for future work include exploring (a) new features, eg. spectral centroid and cepstral coefficients (b) the generalisation capabilities of the combined classifier (isolated and continuous tones) and (c) extending the system to handle polyphonic sounds.

Bibliography

Brown, J. C. “Calculation of a Constant Q Spectral Transform”. J. Acoust. Soc. Am.*,* 89 (1) 1991, pp. 425-34.

Brown, J. C., "Musical Fundamental Frequency Tracking using Pattern

Recognition Method", J. Acoust. Soc. Am., 90 (3) 1992, pp. 1394-1402

Brown, J. C., "Musical instrument identification using autocorrelation coefficients", Proc. Int. Symp. Musical Acoustics, Washington, 1998, pp. 291-5.

Duda, R. O., and Hart, P. E., “Pattern Classification and Scene Analysis”, Wiley-Interscience, 1973

Foster, S., et al "Toward an Intelligent Editor of Digital Audio: Signal Processing Methods", Computer Music Journal, 6 (1), 1982 pp. 42-51

Fujinaga, I., and MacMillan, K., "Realtime recognition of orchestral instruments", accepted for Proc. Int. Comp. Music Conf. 2000.

Handel, S., “Timbre perception and auditory object identification”, in *Hearing*, Moore, B. C. J., New York: Academic Press, 1995.

Hourdin, C., Charbonneau, G., et al. “A Multidimensional Scaling Analysis of Musical Instruments' Time-Varying Spectra”, Computer Music Journal, 21/2, 1997, 40-55.

Jollife, I. T., “Principal Component Analysis”, Springer-Verlag, 1986.

Kaminskyj, I., “Multidimensional Scaling Analysis of Musical Instrument Sounds’ Spectra”, Proc. ACMA Conf. 1999 Wellington, NZ, pp. 36-39.

Kaminskyj, I., and Materka, A., “Automatic Monophonic Musical Instrument Sound Identification System”, Proc. ACMA Conf 1996, Brisbane, Aust., pp. 41-46.

Kaminskyj, I., and Materka, A., "Automatic Source Identification of Monophonic Musical Instrument Sounds", *IEEE Int. Conf. Neural Networks 1995*, Vol. 1, Perth, WA, pp. 189-194.

Kostek, B., and Wieczorkowska, A., "Study of Parameter Relations in Musical Instrument Patterns", Proc. 100th Audio Eng. Soc. Conv., Preprint 4173, Copenhagen, 1996.

Kostek, B., and Krolikowski, R., “Application of Artificial Neural Networks to the Recognition of Musical Sounds”, Archives of Acoustics, 22 (1), pp. 27-50, Poland, 1997.

Martin, K. D., and Kim, Y. E., "Musical instrument identification: A pattern-recognition approach", 136th Meeting J. Acoust. Soc. Am., Oct. 1998.

Olson, H. F., “Music, Physics and Engineering”, 2nd Ed, Dover Publications, 1967.

Pandya, A. S., and Macy, R. B., "Pattern Recognition with Neural Networks in C++", CRC Press, 1996.

The Byzantine Dome: The Sound Dome and Wassily Kandinsky

Theodore Lotis
London, U.K.
E-mail: T.Lotis@city.ac.uk
http://www.digital-music-archives.com/lotis.htm

Abstract

This paper examines the use of sound domes for the diffusion of electronic and mixed music. These sound domes should have as their prototype, the Byzantine dome, an architectural achievement for the better diffusion of sound and light. The construction of sound domes can be proved exceptionally useful for the resolution of most of the inconveniences we encounter during the diffusion of electronic or mixed music. As an archetypal form, the dome is found in different cultures. Its use can be helpful in our endeavour to create a global sonic space culture, which can merge into it the various local characteristics and qualities.

We will attempt to focus on the application of sound domes for the transference of the composed spatial articulation into the listening environment, simultaneously examining the equivalent resultants in the form of painting as this was established by the painter Wassily Kandinsky. The theory of Kandinsky concerning form, can be useful as a guide for spatialization. The "point" – primordial element of space -, the "line" and the "plane" can be diffused through the sound dome without any previous reflections.

Introduction

A few years ago a group of musicians tried to organize a concert in the Byzantine monument of Rotonda which is situated in the city of Thessaloniki. This effort confronted the strong opposition of the Greek Orthodox Church and a part of its citizens. Finally, the concert was never performed. This example shows the intention of several musicians to put into practice the acquired knowledge of architecture and acoustics in order to perform contemporary musical works.

The construction of sound domes containing loud speakers for the diffusion of music, can be proved exceptionally useful for the resolution of most of the inconveniences we encounter during a concert of electroacoustic or mixed music. The construction of the dome as a means of tracing the utmost form of art that is, the search for the divine, commences in the distant past and ends up in the Byzantine dome; an architectural achievement for the diffusion of sound and light.

In this paper we will attempt to focus on the application of sound domes for the transference of the composed spatial articulation into the listening environment, simultaneously examining the equivalent resultants in the form of painting as this one was established by the painter Wassily Kandinsky.

Sound Domes

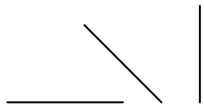
According to Kant, space and time are two "a priori" forms. This was a starting point that led Einstein to the conception of a 4-dimensional reality. It wouldn't be excessive for us to consider, that elements of this reality existed in the Byzantine architecture and art with its peak, the Byzantine dome. The conception of space was not synonymous with the one of size. Space was not formed by the means of perspective rather than, the diffusion of light. Equally, time was not the linear duration that comes between a starting point and an end, but rather had an almost static entity, which embraces both.

Since the beginning of the 70's, Leo Kupper has been trying to complete a well-tempered sound instrument for the diffusion of electronic music through a sound dome. His installations in Grenoble in 1973, Bonn and elsewhere, led the way to a meaningful approach of spatialization. His aim has been the completion of a space modulation system with which the interpreter will be able to articulate the musical space structure. This system will function with strict and pre-determined laws on equal terms to the ones governing the pitch modulation system.

In his article "The Well Tempered Space Sound Instrument. A New Musical Instrument," published in 1991, L. Kupper considers the creation of such a system feasible. Although he mentions in detail the importance of psychological time for the perception of musical space, he is not indicating us how the psychological time will become measurable, in order to be used as one of the parameters in the space modulation system.

The Sound Dome And Wassily Kandinsky

In W. Kandinsky's treatise "Punkt und Linie zu Flache" (Point-Line-Plane) (1926), we come across many references regarding the characteristics and parameters of sound. We also encounter some elements, which even though are originally related to the form of painting, they could have been used for the spatialization of musical space. As for the painter, the primordial component of the form is the "point". W. Kandinsky considers the point as the most laconic, ephemeral form, which is characterized by excessive immobility. Then, the motion overwhelmed the point's immobility by leaping from a static to a dynamic status and thus, creating the "line". The "line" may have any direction. We will stick to the three basic archetypal types (Kandinsky 1926):

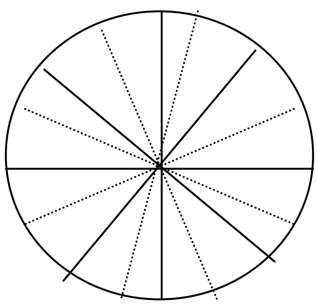


horizontal direction: cold form: potentiality of cold motion
vertical direction: hot form: potentiality of hot motion
diagonal direction: cold/hot form: potentiality of cold/hot motion

This is the point where we come across the simplest contradiction found in the form: the immobility of the point versus the motion of the line. Then, the creation of the "plane" is constructed by the combination of multi-directional lines.

These three basic components of the painting are also used in the musical form. "Point" can be considered any static percussive sound with a minimum resonance. "Line" can be considered the diffusion towards one direction of any sonic morphology, through a loud speaker. Finally, "plane" can be the multi-directional diffusion of soundscapes and images, through a combination of concentric or non-concentric loud speakers.

Another very important element of painting that we also encounter in the diffusion of musical space, is the continuum of condensation-rarefaction, which is related to the changes of temperature in the continuum of cold-hot (Kandinsky 1926).



- Concentric changes of temperature
- Concentric changes of direction

The continuum of condensation-rarefaction is qualified by the seven criteria, which are defined by P. Schaeffer: masse, dynamique, timbre harmonique, profil melodique, profil de masse, grain, allure. Here follows an example for the diffusion of a distant space (Chion 1983):

DISTANT SPACE

→ CONDENSATION

→ HOT

→ masse : nodal group
dynamique : cyclic/resonance
timbre harmonique: complex/continuous
profil melodique : torculus
profil de masse : delta
grain : compact/discontinuous
allure : irregular

The above example is a guide for the spatialization of the composed space through a sound dome rather than, an attempt to create a space scale. The application of the three previously mentioned elements (point, line, plane) during the diffusion of the musical space, is clear and apparent. The architectural structure of the sound dome is based on the hemispherical or spherical form (two opposite-sited sound domes). The possibility of placing (geometrically or not) the loud speakers in any point of the periphery of the dome, allows the interpreter the transition from the simplest spatial component that is the point, to more complex components such as the sonic planes or soundscapes. Thus, the interpreter not only diffuses the composed space but also controls the diffused space (Smalley 1991).

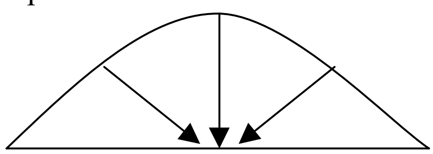
Despite all these, it is exceptionally difficult, if not impossible, for a general space modulation system with specific scales and articulation structures, to be formed. The changes of temperature in the cold-hot continuum but also the combinations of condensations and rarefactions are boundless. The morphological transformations (changes in spectrum, texture) which result to spectral transformations (Smalley 1986), are not perceived by the listeners in the same way. L. Kupper admits that the perception of spectral changes and therefore the perception of space scales, is closely related to the listener's attention (Kupper 1991).

At least, the creation of a general space modulation system, is not possible without the acceptance of some very important and necessary compromises. The well-tempered pitch modulation system, on which all the horizontal and vertical structures of the instrumental music are based, neglects the microtones, the slightest timbral changes and the extremely complex rhythmical sequences which are found in the non-western music. In the same way, a well-tempered space modulation system will not contain the microscopic transitions, which are taking place inside the above continua.

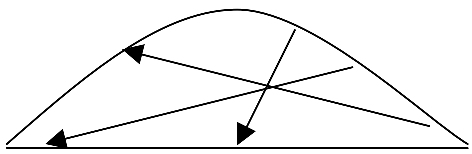
It is more essential for the interpreter to be given the freedom to use these continua as auxiliary elements for the diffusion of the musical space. The architectural structure of the sound dome offers him such a freedom. Sonic lines and directions, starting from one or more loud speakers (concentric or not) situated in different points of the dome and also, countless combinations of sonic planes and images, can all be directed towards the listener without any previous reflections. From the immobility of the "point" to the daedalian combinations of motion, the structure of the sound dome, which hovers over the audience, diffuses the sound directly to the listener. As for the mixed music, the dome with the instruments which are placed in the same level with the audience, form a sphere, a unique instrument for the spatial diffusion. This sphere embraces the listener and thus triggers off the closest connection between the spatialization and the perception of the musical space.

The sphere and therefore its half – the hemisphere – is the perfect geometrical form. It symbolizes at the same moment the infinitude and the "point" that is, the space and its primordial structural element. The sphere itself is a point consisting of other points (loudspeakers) which outline its periphery. The expansion of this periphery is boundless.

The placing of the loud speakers across the periphery of the dome can be:
1) Geometrical or concentric. In this way, the diffusion of the "lines" and the "planes" is also geometrical.
2) Non-geometrical or non-concentric. Thus, the interpreter can distract the uniformity of the concentric "lines" and "planes".



1) concentric placing



2) non-concentric placing

Conclusion

The example of the cancelled concert in the Byzantine monument of Rotonda, should not discourage us from the practise of the acquired knowledge into the contemporary musical expression.

The shape of the dome, for thousands of years now, symbolizes the utmost architectural construction for the quest of the divine. As an archetypal form, the dome is found in different cultures. Its use can be helpful in our endeavour to create a space culture. This space culture can merge into it the various local characteristics and qualities. A very important factor for the creation of this general space culture is the construction of permanent sound domes that will lead to the end of the Italian scenes or at least, their maintenance as museums.

Just like in the painting form as in the musical one, the articulation of the real and the psychological time, needs an expansion of its expressive means. We may have to stop confusing the notion of space with the one of size. We may have to stop measuring the time with rhythmical sequences. We may, finally, have to use an expressive means that will contain at the same moment the "point" and the infinitude, the diminution and the magnification, the starting point and the ending.

Using the sound dome for these purposes does not mean that we will be forced to create a general space modulation system. In any case, the pressing urge of the western civilization to decode every single evolution, should not lead us to a far-fetched syntax of such a system. By the end of this millennium there are a lot who consider typography to be the most important cultural invention. Unfortunately, this invention also encloses its degeneration since anything that is recorded ceases to evolve.

Acknowledgements

This paper is part of a general research program concerning the composed and diffused space. This research is generously supported by the "ALEXANDER S. ONASSIS PUBLIC BENEFIT FOUNDATION".

References

Chion M. 1983. *Guide des Objets Sonores*. Ed. Buchet/Chastel.

Kandinsky, Wassily. 1926. *Punkt und Linie zu Flache*. Ed. A. Langen.

Kupper, Leo. 1991. "The Well Tempered Space Sound Instrument: A New Musical Instrument". *L'Espace du Son II*. Ed. Musiques&Recherches. pp.92-98.

Smalley, Denis. 1986. "Spectromorphology and Structuring Processes". *The Language of Electroacoustic Music.* Ed. S. Emmerson. Macmillan.

Smalley, Denis. 1991. "Spatial Experience in Electroacoustic Music". *L'Espace du Son II*. Ed. Musiques&Recherche.

Boolean logic as a harmonic filter

Dr. Greg Schiemer
Sydney Conservatorium of Music
Sydney, Australia.
gregs@conmusic.usyd.edu.au

Abstract

The author describes the motivation and method for generating harmonically related sets using boolean logic. The paper details an implementation in which boolean filters are used to generate scales and documents the process with musical examples.

1. Preamble

Boolean logic is a useful tool for culling surplus information. Its musical legacy dates back to some of the earliest composers involved in computer music including Xenakis, who developed the musical concept of set theory[1], and Martirano who built some of the first live performance music systems that used digital logic[2]. Boolean concepts have also been applied by composer-instrument-builders such as Burt[3], Vine[4] and the author[5] who built interactive hardware based on random logic integrated circuits.

The desire to use boolean instructions as a harmonic filter for MIDI note numbers was motivated by my involvement in the design of the MIDI Tool Box, purpose-built hardware for real-time musical applications using MIDI. Here minimal memory resources necessitated investigation of any means that might allow a composer to impose musical constraints on a system using a minimum of information. In the context of MIDI, where it takes 320 to transmit or receive a single byte, it is practical to apply such a harmonic filter in real time.

My own compositional use of boolean logic was based on the notion of getting more using less, and represents a paradigm that has now perhaps been overtaken since the arrival of the internet. Boolean logic is nowadays used in search engines as a way to extract what is sufficient from a superabundance of information.

2. The boolean concept of a harmonic filter

The set of MIDI note numbers [0,1,2,..,127] can also be expressed as sets of modulo-2 numbers; musically this representation allows the set to be thought of as two interleaved sets of whole-tone scales, a set of even numbers $[0,2,4,....,126]_{even}$ interleaved with a set of odd numbers $[1,3,5,....,127]_{odd}$. Figure 1 shows a two-octave section of a music keyboard with both interleaved whole-tone sets. MIDI note numbers in this figure are shown in decimal.

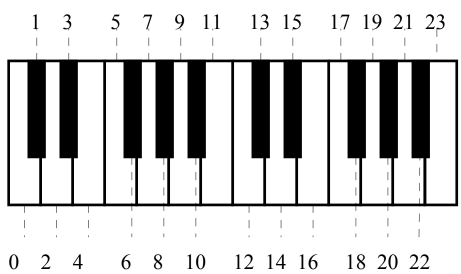


Fig. 1

Boolean operations are used to define members of either whole-tone scale set by

1 Xenakis, I [1992]: Formalized Music, Pendragon Press

2 Pellegrino, R. [1983]: Sal-Mar Construction (Martirano), pp. 95-98

3 Aardvarks Machine in Burt, W [1991]: Experimental Music in Australia using Live Electronics CMR, Vol. 6 Part 1, eds. P. Nelson & S. Montague, pp. 168-171

4 Vine, C. [1978]: Patent Little Marvel, New Music Magazine, Latrobe U., ed. W. Burt, pp. 6-7

5 Tupperware Gamelan in Schiemer, G. [2000]: Improvising Machines: Spectral Dance and Token Objects, Leonardo Music Journal, ed. N.Collins, Vol.9, pp.107-114

masking out or hiding those bits that do not belong to the set.
Because they allow modulo-2 calculations to be performed very efficiently, boolean operations are included as part of the instruction set of most microprocessors and digital signal processors.

In the harmonic filter, two types of boolean operation are used: logical AND and logical OR. Appendix A.1. illustrates both these as 1-bit binary operations; Appendix A.2 illustrates boolean AND as a 7-bit binary operation.

Logical AND is used to define the whole-tone set of even numbers using an even numbered mask 111 1110 [126]$_{even}$ while logical OR is used to define the whole-tone set of odd numbers using an odd numbered mask 000 0001 [1]$_{odd}$.

Other masks can be used to hide non-members of the various subsets of both whole-tone sets. This allows more sophisticated harmonic filters to be developed. For example, the even mask 111 1100 [124]$_{even}$ and the odd mask 000 0011 [3]$_{odd}$ will both produce steps of a major third, while the even mask 111 1000 [120]$_{even}$ and the odd mask 000 0111 [7]$_{odd}$ will both produce steps of a tritone.

The boolean concept of harmonic filter potentially allows any chord to be expressed as an intersection of the even and odd sets of whole-tone scales or their sub-sets. For example, a major triad can be expressed as the intersection of [0, 4]$_{even}$ and [7]$_{odd}$ for the lower octave and [12, 16]$_{even}$ and [19]$_{odd}$ for the upper octave as shown in Fig. 2, while the minor triad can be expressed as the intersection of [0]$_{even}$ and [3, 7]$_{odd}$ for the lower octave and [12]$_{even}$ and [15, 19]$_{odd}$ for the upper octave as shown in Fig.3.

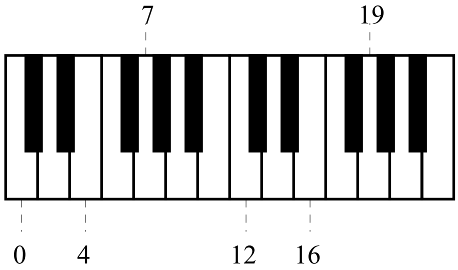


Fig. 2

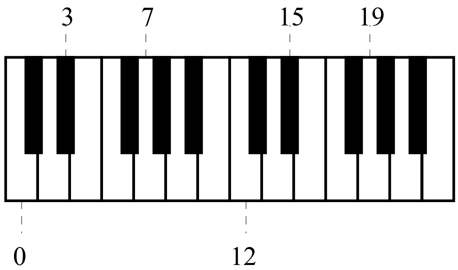


Fig. 3

The technique of defining harmonic filters using boolean operations, though computationally efficient, requires a simple modification to work with the set of modulo-12 numbers, called pitch classes, which has been the foundation of western harmony for the past few centuries. Integer division allows a set of pitch classes [0,1,2,...,11] to be extracted from a set of MIDI note numbers; with a divisor of 12, the quotient is the pitch class, the dividend is the octave. Once the harmonic filter has been applied, the filtered pitch class is then recombined with (ie. multiplied by) the octave to produce a filtered version of the original MIDI note number. Only one pitch class extraction is required for each MIDI note number received or transmitted. This allows harmonic filtering to be applied in conventional 12-note-per-octave harmony, and makes it practical for real-time musical applications.

However before further examining the feasibility of conventional chords within the 12-note- per-octave system, it seemed necessary to explore the musical possibilities of modulo-2 numbers first. Pitch class extraction would limit such exploration to chords formed from octave-limited scales ie. scales confined to a range of one octave. There is plenty we already know about harmony produced from octave-limited scales. There is much more to learn about harmony generated using boolean operations.

3. Boolean harmonic filter algorithm

The following algorithm was developed. For each note :

- MIDI note number is duplicated once for odd mask & once for even mask
- odd mask is ORed with MIDI note number to form odd partial result
- even mask is ANDed with MIDI note number to form even partial result
- both partial results are interrogated to detect membership of either set

These steps are repeated for every MIDI note number until two octaves have been played.
Appendix B shows the relevant extract of a documented listing of the algorithm as it was implemented in HC11 assembler code for the MIDI Tool Box (MTB).

3.1. Method

The above implementation of the boolean filter algorithm was cross-assembled and down-loaded into the MTB. The program was run several times. Each time the program was run, new values for the even mask and the odd mask were manually entered in the program prior to restarting.

The musical behaviour of boolean operations was observed by listening to the effect of various logical AND and logical OR masks on a series of MIDI note numbers played in ascending order. Without using pitch-class and octave separation, the boolean filter algorithm described above was found to generate interesting scales. Several of these are octave-limited, most are non-octave-limited.

Data samples shown in Appendix C were recorded as a screen dump on the MTB each time the program was run. These form a representative selection of samples of the musical scales that can be synthesised using this technique. They are shown in hexadecimal (ie. modulo-16) notation. Some of these scales can be produced using more than one set of masks. For instance, example C.1, C.2, C.3 & C.4 are the same scale produced using different masks; likewise, for example C.5, C.6 & C.8

Scale steps are expressed chromatically, where a single step represents a minor 2nd (m2), two steps, a major 2nd (M2), three steps, a minor 3rd (m3) and so on. The first few examples are chromatic ie. consisting of all minor 2nds. All examples shown have a span of two octaves, spanning note numbers from two octaves below middle C (36 or, in hexadecimal, \$24) to middle C (60 or \$3C). This allows sufficient steps to determine whether or not the scale is octave-limited.

Because the MTB screen dump is in hexadecimal a few of the more interesting examples are also presented here using staff notation. One is octave-limited the others non-octave-limited.

Figure 4 (corresponding to example C.5) is like the octotonic scale except that it contains two semitone steps for each whole-tone step. Like the octotonic scale, it is octave-limited. Figure 5 (example C.10) is a non-octave-limited scale with three whole tone steps for each pair of semitone steps. Figure 6 (example C.19) is another non-octave-limited scale in which a whole tone step is separated from a minor 3rd by one semi-tone step in one direction and two semi-tone steps in the other direction.

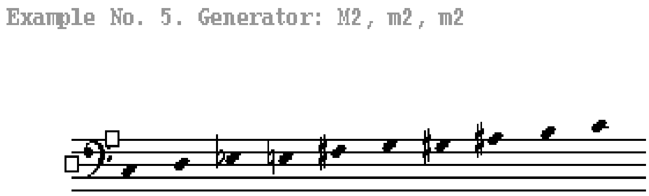


Fig. 4

Fig.5

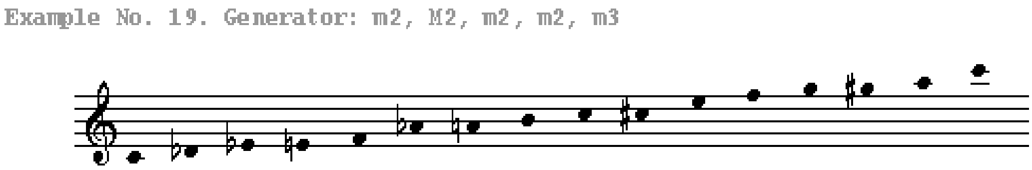


Fig.6

4. Future development

4.1. Applications

One of the proposed changes to the harmonic filter algorithm included separation of MIDI note number into pitch class and octave as outlined in 2. This would provide a way to define groups of notes using just two bytes making it possible to identify chord patterns quickly in live performance. By comparing a note played with a filtered version of the same note, one can establish whether or not the filtered and unfiltered versions are identical; chordal identity of a note can be established in less time than it takes to transmit or receive the next byte of MIDI.

With or without the inclusion of pitch class, the algorithm would also be useful in algorithmic compositions to filter random notes in a harmonically ordered way. Subtle changes to the harmonic field can be made by applying ramp, and triangular functions etc. to manipulate one or other, or both masks in realtime. Examples of this can already be demonstrated.

4.2. Platforms

There is no reason for this algorithm to be limited to the MIDI Tool Box even though this has been an ideal platform for my own development and compositional work with MIDI. Others hopefully will find ways of implementing this algorithm by creating Max objects or writing grammars using the Bol Processor. In such computing environments where MIDI and audio signal processing meet, harmonic filters have applications that go beyond the availability of 12 notes in every octave.

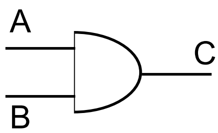
5. Conclusion

Regardless of computer platform or software application, boolean operations provide a new way to interpret or manipulate harmonic information in real time. For me, any further musical experimentation with these operations as harmonic filters will not be confined to the 12-tone equal tempered system.

Appendix A.1. 1-bit

Boolean functions use binary logic.
Figure 7a shows a logic circuit called a logical AND-gate, where both inputs must be true (ie. = 1) for the output to be true. Figure 7b shows the truth table for an AND-gate. Figure 8a shows a logic circuit called a logical OR-gate, where, if either input is true, then the output must be true. Figure 8b shows the truth table for an OR-gate.

Logical AND gate



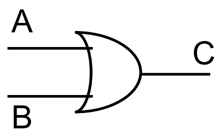
[a] AND circuit

inputs		output
A	B	C
0	0	0
0	1	0
1	0	0
1	1	1

[b] AND truth table

Fig. 7

Logical OR gate



[a] OR circuit

inputs		output
A	B	C
0	0	0
0	1	1
1	0	1
1	1	1

[b] OR truth table

Fig. 8

Appendix A.2. 7-bit

Boolean operations are simply a collection of binary gates that operate in parallel. Figure 9 shows how an odd MIDI note number 7 (binary 000 0111 in input register A) is transformed using boolean AND operation with a 7-bit binary mask of 111 1110 (in input register B) to produce an even MIDI note number 6 (binary 000 0110 in output register C). As long as the value in register B remains 111 1110, then no matter what 7-bit binary value is contained in register A, the output register will always contain an even number, ie. the least significant bit of register C will always equal zero.

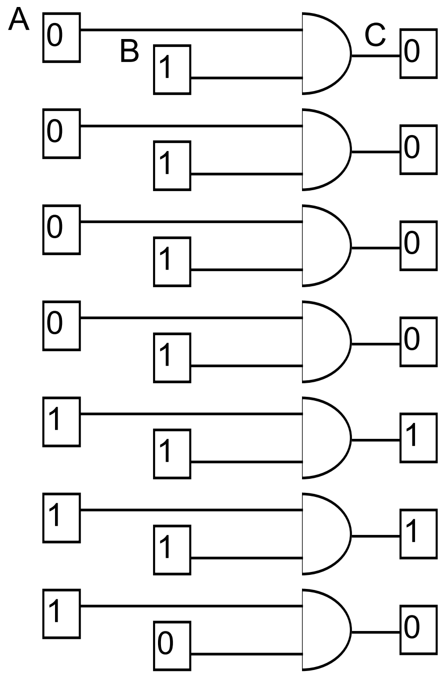


Fig. 9

The same principle of 7 binary gates that operate in parallel applies for odd numbers using the boolean OR operation.

Appendix B HC11 Assembler Code

```
MIDI_COMMAND    EQU     $90     (144 : MIDI Status Byte)
STARTING_NOTE   EQU     $24     (36 :  MIDI note number)
AUDIBLE_LEVEL   EQU     $7F     (127 : MIDI key velocity)
NOTE_LENGTH     EQU     $18     (note length approx. 0.25 sec)
NOTE_FILTER1    EQU     $7F     AND_MASK
NOTE_FILTER2    EQU     $0      OR_MASK

        ORG     $3000
STATUS_BYTE     RMB     $1      (1-byte variable)
NOTE_NUMBER     RMB     $1      (1-byte variable)
KEY_VELOCITY    RMB     $1      (1-byte variable)
DURATION        RMB     $1      (1-byte variable)
EVEN_MASK       RMB     $1      (1-byte variable)
ODD_MASK        RMB     $1      (1-byte variable)
*_____________________________________________________
        ORG     $2000
        JSR     POLLM1          Select Polling [ROM Tool]
        JSR     INIT_REGS       Setup program controls
*                               Start of Main Program
MAIN    JSR     SET_LENGTH      (Re-)set Note length
        JSR     KON             Turn MIDI Note ON
        JSR     HOLD            Hold it on
        JSR     KOFF            Turn MIDI Note OFF
        JSR     NEXT_NOTE       Generate next note
        BRA     MAIN            Repeat the main loop

*______ User Subroutine 6
*       Generate next Note
NEXT_NOTE
        LDAA    NOTE_NUMBER     Now try next note
        INCA                    and keep this one
        ANDA    #$7F            within the limits
        STAA    NOTE_NUMBER     of Note Numbers
        JSR     PITCH_FILTER    and filter it
        RTS

*______ User Subroutine 7
*       Note-numbers pass through filter only when these are members
*       of either set of pitches defined by the two mask patterns

PITCH_FILTER
        LDAA    NOTE_NUMBER     Get Note-number
        TAB                     in AccA & AccB
        ANDA    EVEN_MASK       Check for set of
        CBA                     EVEN Whole-tones
        BEQ     NOTE_PASS
        TBA                     Get Note-number
        ORAA    ODD_MASK        Check for set of
        CBA                     ODD Whole-tones
        BEQ     NOTE_PASS
        INC     NOTE_NUMBER     Try again with
        BRA     PITCH_FILTER    the next one

        NOTE_PASS
*                               Next one worked!
        ANDA    #$7F            Keep within limits
        STAA    NOTE_NUMBER     of Note Numbers
        RTS
```

Appendix C Musical examples

C.1 AND = $7F; OR = $00

```
            Chromatic
       0  1  2  3   4  5  6  7   8  9  A  B   C  D  E  F        ascii
7000  24 25 26 27  28 29 2A 2B  2C 2D 2E 2F  30 31 32 33  $%&'()*+,-./0123
7010  34 35 36 37  38 39 3A 3B  3C
```

C.2 AND = $7F; OR = $01

```
            Chromatic
       0  1  2  3   4  5  6  7   8  9  A  B   C  D  E  F        ascii
7000  24 25 26 27  28 29 2A 2B  2C 2D 2E 2F  30 31 32 33  $%&'()*+,-./0123
7010  34 35 36 37  38 39 3A 3B  3C
```

C.3 AND = $7E; OR = $00

```
            Chromatic
       0  1  2  3   4  5  6  7   8  9  A  B   C  D  E  F        ascii
7000  24 25 26 27  28 29 2A 2B  2C 2D 2E 2F  30 31 32 33
7010  34 35 36 37  38 39 3A 3B  3C
```

C.4 AND = $7E; OR = $01

```
            Chromatic
       0  1  2  3   4  5  6  7   8  9  A  B   C  D  E  F        ascii
7000  24 25 26 27  28 29 2A 2B  2C 2D 2E 2F  30 31 32 33  $%&'()*+,-./0123
7010  34 35 36 37  38 39 3A 3B  3C
```

C.5 AND = $7E; OR = $02

```
            M2 m2 m2 M2  m2 m2 M2 m2  m2 M2 m2 m2  M2 m2 m2 M2
            m2 m2
0  1  2  3   4  5  6  7   8  9  A  B   C  D  E  F        ascii
7000  24 26 27 28  2A 2B 2C 2E  2F 30 32 33  34 36 37 38  $&'(*+,./0234678
7010  3A 3B 3C
```

C.6 AND = $7E; OR = $03

```
            M2 m2 m2 M2  m2 m2 M2 m2  m2 M2 m2 m2  M2 m2 m2 M2
            m2 m2
       0  1  2  3   4  5  6  7   8  9  A  B   C  D  E  F        ascii
7000  24 26 27 28  2A 2B 2C 2E  2F 30 32 33  34 36 37 38  $&'(*+,./0234678
7010  3A 3B 3C
```

C.7 AND = $7E; OR = $04

```
            m2 m2 m2 m2  M2 M2 m2 m2  m2 m2 M2 M2  m2 m2 m2 m2
            M2 M2
       0  1  2  3   4  5  6  7   8  9  A  B   C  D  E  F        ascii
7000  24 25 26 27  28 2A 2C 2D  2E 2F 30 32  34 35 36 37  $%&'(*,-./024567
7010  38 3A 3C
```

C.8 AND = $7E; OR = $05

```
            M2 m2 m2 M2  m2 m2 M2 m2  m2 M2 m2 m2  M2 m2 m2 M2
            m2 m2
       0  1  2  3   4  5  6  7   8  9  A  B   C  D  E  F        ascii
7000  24 26 27 28  2A 2B 2C 2E  2F 30 32 33  34 36 37 38  $&'(*+,./0234678
7010  3A 3B 3C
```

C.9 AND = $7E; OR = $06

M2 m2 m2 M2 M2 M2 m2 m2 M2 M2 M2 m2 m2 M2 M2

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	ascii
7000	24	26	27	28	2A	2C	2E	2F	30	32	34	36	37	38	3A	3C	$&'(*,./024678:<
7010																	

C.10 AND = $7E; OR = $07

M2 m2 m2 M2 M2 M2 m2 m2 M2 M2 M2 m2 m2 M2 M2

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	ascii
7000	24	26	27	28	2A	2C	2E	2F	30	32	34	36	37	38	3A	3C	$&'(*,./024678:<
7010																	

C.11 AND = $7E; OR = $08

M2 M2 m2 m2 m2 m2 m2 m2 m2 m2 M2 M2 M2 M2 m2 m2
m2 m2

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	ascii
7000	24	26	28	29	2A	2B	2C	2D	2E	2F	30	32	34	36	38	39	$&()*+,-./024689
7010	3A	3B	3C														

C.12 AND = $7E; OR = $09

M2 M2 m2 m2 m2 m2 m2 m2 m2 m2 M2 M2 M2 M2 m2 m2
m2 m2

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	ascii
7000	24	26	28	29	2A	2B	2C	2D	2E	2F	30	32	34	36	38	39	$&()*+,-./024689
7010	3A	3B	3C														

C.13 AND = $7E; OR = $0A

M2 M2 M2 m2 m2 M2 m2 m2 M2 M2 M2 M2 M2 m2 m2

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	ascii
7000	24	26	28	2A	2B	2C	2E	2F	30	32	34	36	38	3A	3B	3C	$&(*+,./02468:;<
7010																	

C.14 AND = $7E; OR = $0B

M2 M2 M2 m2 m2 M2 m2 m2 M2 M2 M2 M2 M2 m2 m2

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	ascii
7000	24	26	28	2A	2B	2C	2E	2F	30	32	34	36	38	3A	3B	3C	$&(*+,./02468:;<
7010																	

C.15 AND = $7D; OR = $01

m2 M2 m2 m2 M2 m2 m2 M2 m2 m2 M2 m2 m2 M2 m2 m2
M2 m2

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	ascii
7000	24	25	27	28	29	2B	2C	2D	2F	30	31	33	34	35	37	38	$%'()+,-/0134578
7010	39	3B	3C														

C.16 AND = $7D; OR = $02

Chromatic

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	ascii
7000	24	25	26	27	28	29	2A	2B	2C	2D	2E	2F	30	31	32	33	$%&'()*+,-./0123
7010	34	35	36	37	38	39	3A	3B	3C								

C.17 AND = $7D; OR = $03

m2 M2 m2 m2 M2 m2 m2 M2 m2 m2 M2 m2 m2 M2 m2 m2
M2 m2

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	ascii
7000	24	25	27	28	29	2B	2C	2D	2F	30	31	33	34	35	37	38	$%'()+,-/0134578
7010	39	3B	3C														

C.18 AND = $7D; OR = $04

m2 m2 m2 m2 m2 m3 m2 m2 m2 m2 m2 m3 m2 m2 m2 m2
m2 m3

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	ascii
7000	24	25	26	27	28	29	2C	2D	2E	2F	30	31	34	35	36	37	$%&'(),-./014567
7010	38	39	3C														

C.19 AND = $7D; OR = $05

m2 M2 m2 m2 m3 m2 M2 m2 m2 m3 m2 M2 m2 m2 m3

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	ascii
7000	24	25	27	28	29	2C	2D	2F	30	31	34	35	37	38	39	3C	$%'(),-/0145789<
7010																	

C.20 AND = $7D; OR = $06

m2 m2 m2 m2 m2 m3 m2 m2 m2 M2 m2 m3 m2 m2 m2 m2
m2 m2

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	ascii
7000	24	25	26	27	28	29	2C	2D	2E	2F	30	31	34	35	36	37	$%&'(),-./014567
7010	38	39	3C														

C.21 AND = $7D; OR = $07

m2 M2 m2 m2 m3 m2 M2 m2 m2 m3 m2 M2 m2 m2 m3

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	ascii
7000	24	25	27	28	29	2C	2D	2F	30	31	34	35	37	38	39	3C	$%'(),-/0145789<
7010																	

C.22 AND = $7D; OR = $08

m2 m3 m2 m2 m2 m2 m2 m2 m2 m2 m2 m3 m2 m3 m2 m2
m2 m2

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	ascii
7000	24	25	28	29	2A	2B	2C	2D	2E	2F	30	31	34	35	38	39	$%()*+,-./014589
7010	3A	3B	3C														

C.23 AND = $7D; OR = $09

m2 m3 m2 M2 m2 m2 M2 m2 m2 m3 m2 m3 m2 M2 m2

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	ascii
7000	24	25	28	29	2B	2C	2D	2F	30	31	34	35	38	39	3B	3C	$%()+,-/014589;<
7010																	

C.24 AND = $7D; OR = $0A

m2 m3 m2 m2 m2 m2 m2 m2 m2 m2 m2 m3 m2 m3 m2 m2
m2 m2

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	ascii
7000	24	25	28	29	2A	2B	2C	2D	2E	2F	30	31	34	35	38	39	$%()*+,-./014589
7010	3A	3B	3C														

C.25 AND = $7D; OR = $0B

m2 m3 m2 M2 m2 m2 M2 m2 m2 m3 m2 m3 m2 M2 m2
m2 m2

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	ascii
7000	24	25	28	29	2B	2C	2D	2F	30	31	34	35	38	39	3B	3C	$%()+,-/014589;<
7010																	

Introducing jMusic

Andrew Sorensen
GET Systems
Chatswood, Sydney, Australia

Andrew R. Brown
Queensland University of Technology
Victoria Park Rd.
Kelvin Grove, 4059. Australia

Abstract

This paper introduces the jMusic compositional language. jMusic is a package of Java classes which provides an environment for non real time music composition. jMusic consists of a music data structure for event organisation along the lines of Common Music and HSML, however it relies more heavily than these on the common practice notation score metaphor in an attempt to provide easy access to composers familiar with that environment. The music data structure is practical for analysis as well as composition, and jMusic reads and writes standard MIDI files to facilitate interaction with existing computer-based music systems. jMusic includes audio classes for sound synthesis, and signal processing, along the lines of Csound and Cmix, however its object oriented nature and the integration between compositional and synthesis functions provide a platform for efficiently combining music-event and audio processes. jMusic is not a scripting language but a direct extension of the Java programming language. This allows access to the full functionality of Java, including the ability to write applets for Internet usage, and for applications to be run unchanged on a wide range of platforms (even with graphical interfaces). Learning jMusic is learning Java, so computer musicians can leverage to mutual advantage jMusic and Java experience. This paper outlines the jMusic data structure and built-in functionality, includes code examples, and shows how jMusic can be extended by users. jMusic is available for free public download.

Introduction

jMusic is a programming library written for musicians in the Java programming language. While still relatively new, we hope jMusic will be a library that is simple enough for new programmers (as many musicians are) but sophisticated enough to enable composers to accomplish comprehensive work, whatever form that may take. jMusic is designed to be used as a compositional medium, therefore it is primarily designed for musicians—not computer programmers.

jMusic is a music research project based at Queensland University of Technology (QUT) music program in Brisbane, Australia. The project was begun in 1987 as part of Andrew Sorensen's Master of Arts degree program, at which time it focused on algorithmic composition using Markov chains. Since that time it has been rewritten and expanded to be a more generic system capable of supporting, or being extended to support, most compositional processes.

jMusic is designed to assist the compositional process by providing an open but partially structured environment for musical exploration; it has also been effective to support musical analysis and computer music education. The jMusic data structure uses metaphors from established musical practice (scores and instruments) in order to provide familiarity at the early stages of learning, in keeping with Alan

Kay's advice that "simple things should be simple and complex things should be possible." We hope that jMusic is simple to learn, but powerful to use. As well, interfacing jMusic with other music software is facilitated by easy importing and exporting of MIDI files and audio files. This means that a users' current knowledge and tools are not discarded when working with jMusic. Because jMusic has full access to the Java language and support structures, jMusic applications can be as extensive as Java allows (and that is very extensive).

Composing in jMusic is programming in Java, not in a Meta-language or scripting environment as provided by Csound and Cmix/Minc. This means that the full power and cross platform independence of Java is maintained, it also means that the more you know about Java programming the more useful jMusic will be to you. Learning jMusic can be a fun way to gain Java programming skills while focusing on making music. Currently jMusic works in Java versions 1.1 and higher; even though later releases are available we have tried to maintain compatibility with the widest possible user base (including Personal Java on PDAs, Applets in web Browsers, and operating systems without 'cutting edge' Java support). When designing jMusic we felt that any additional complexity imposed by having to work directly in Java was outweighed by the advantage of the widespread availability of general Java support materials and the advantage to musicians of learning programming skills useful beyond the musical domain. However, we have made a concerted effort in the core jMusic classes to hide programming complexity, such as exception handling, where possible.
jMusic is an open source package distributed under the GNU General Public Licence. It is being developed to support computer music making and the development of shared tools. We hope that this will encourage people to become a part of the jMusic community and contribute to, as well as benefit from, the development of jMusic. Any music created with jMusic is not itself covered by any licence and can be freely distributed or sold by the composer. Extensions to the language and application built using the jMusic libraries are subject to the Licence which, in short, requires that source code be available for others to utilise further.

Music, Composition, and the Computer

The marriage of music, composition and the computer may seem like an unlikely partnership, bringing together the arts, mathematics and the sciences in what many would see as an unholy trinity. It is not however as unnatural as it might at first appear. Many great scientists have also been exceptional artists and vice versa, the most famous of whom is unquestionably Leonard Da Vinci. Music and technology have in fact had an intimate relationship throughout time with technological advances aiding music in many diverse ways. The printing press is one example, as are the ability to shape iron over a flame, or the phonograph.

What makes computer music significantly different from older technological developments is that the use of computers in music making affects each of the compositional, instrument building, and performance processes. Computers have the ability to radically change the way in which we compose and the compositional processes that we use due to their ability to be programmed.

The twentieth century has been marked by significant technological innovation, particularly with the advent of the computer, whose influence has spread beyond the confines of the office to include activities such as sport, entertainment, and the arts. In these roles, the digital computer is often used to provide a means for modelling the complexity of real world patterns and associations. Machine abstractions of the world are becoming increasingly common, providing humanity with artificial representations of real world events and phenomena within digital micro-domains. The potential for designing machine abstractions of artistic structures has driven many artists to explore digital micro-domains as a means to create digital tools designed to investigate structural formalisms.

It seems to us that the function of the computer in music composition is to expose aspects of the music. That is, to help structure and give form to the musical ideas of the composer, not unlike the function of a score. Further, the computer also exposes the musical assumptions of the composer through the choice of representation system and associated methods of representational manipulation. In this regard the computer is implicated by requiring the composer to specify their musical intentions. Thirdly, the computer as digital sound medium creates or renders the sonic attributes of the composition highlighting certain characteristics over others. We hope that jMusic can provide a link from the musical past and present to the musical future, by providing familiar structures and functions at the outset and allowing the more experienced jMusic composer to evolve new musical structures and ideas. We seek to provide a pathway for the evolution of musical ideas, rather than a radical break from them

Some jMusic Basics

jMusic is currently a non real time environment. The compositional process follows a cycle, familiar to computer music composers, of write, compile, listen, rewrite. In support of a non real time process we could begin by giving a summation of composition over the centuries and all the non real time work that has been completed before now, but suffice it to say that non real time has had a successful run of it up until now, and there is still an argument that says non real time has benefits that outweigh its disadvantages. It is imagined that jMusic will soon become almost real-time, and work is underway for real time Java Applet support. As a result, jMusic outputs music in standard MIDI file format (SMF) or audio file (at present only .au format).

Packages

The classes that make up jMusic are organised into core classes in the lib directory (those kept under the jm directory are the very core classes) and additional classes are organised in surrounding directories such as util(ities), app(lication)s, and inst(ruments). The distinction is made to keep separate the *development* and *use* of jMusic. In general, users will program new compositional utilities, instruments, and applications without being concerned with the core classes in the lib directory.

```
JMusic
   |
   +---- lib
   |       |
   |       +----jm
   |             |
   |             +---- music (Constant values. e.g., C4, PIANO, REST.)
   |             |       |
   |             |       +---- data (Contains Note, Phrase etc.)
   |             |       |
   |             |       +---- tools (Contains data processes.)
   |             |
   |             +---- midi (Converting between jMusic and MIDI files.)
   |             |
   |             +---- audio (High level audio operations.)
   |             |       |
   |             |       +---- io (File reading and writing classes)
   |             |
```

```
|              |       |
|              |       +---- synth (Synthesis modules)
|              |
|              +---- GUI (Graphical user interface components.)
|
+---- inst (jMusic audio instruments.)
|
+---- jmUtil (Composing helper classes, e.g. Score Visualisation.)
|
+---- apps (Complete programs that use jMusic.)
|
+---- demos (Example code.)
|
+---- doc (JavaDoc html files with the details of jMusic classes.)
```

The compositional process

The basic process behind non real time computer music composition, including jMusic follows these steps:

1) Decide what it is that you want to compose.
2) Work out what compositional techniques you will need to use to achieve this.
3) Write a Java program explaining your compositional techniques to the computer.
4) Compile and run the program.
5) Play back your new composition as a MIDI or audio file.

Step 3, writing the program, involves some more specific jMusic understanding:

The basic sound objects in jMusic are called Notes. The traditional conception of notes, that they are discrete sound objects with serval dimensions (pitch, duration, loudness . . .) holds true for jMusic. Composing in jMusic involves constructing and organising note objects. Notes can be collected as a sequence of events into a Phrase, a phrase is monophonic and continuous, and each phrase has a start time. Phrases which will sound using the same instrument can be collected into a Part. Several parts can be collected into a score. This structure is outlined in more detail below.

Time in jMusic is measured in beats. Notes have independent rhythmValues (crotchet, quaver . . .) and duration, allowing for gaps between notes (staccato) or overlapping notes. Working in clock time in jMusic is a simple process of calculating beats per minute into seconds (working at 60 bpm makes this trivial as one beat = one second).

Pitch of notes is measured in one of two ways, as MIDI pitch numbers (i.e., middle C = 60) or as frequency for audio output (i.e., 440hz).

jMusic and other systems?

JMusic is most similar in conception and design to Common Music, major differences being the use of Java rather than Lisp and jMusic's CPN data structure metaphor rather than a unique data structure. There are other non real time computer assisted composition languages with similar aims including PatchWork (Laurson 1996) and OpenMusic (Assayag & Agon 1996) both of which focus, like jMusic, on event-based structuring ahead of signal processing. These languages rely heavily on a graphical representation as a way of 'assisting' the composer. We feel that such interfaces are generally more limiting to structural possibilities than a text-based environment, however, jMusic does provide visualisation tools but they simply display output and are not interactive. HMSL, Csound, Cmix, Supercollider and others provide score and instrument separation similar to jMusic, and provide much more extensive audio processing feature than jMusic currently has. This reflects a difference in priorities between event-based composition and signal processing and we hope that with MIDI file output musicians may, for example, choose to do event generation in jMusic and convert the score to Csound for audio rendering. Presently, jMusic is extensively used in conjunction with MIDI sequencers and synthesizers in a two step process of write and render.

Music Data Structure

The core of the jMusic data structure are the five classes in the jm.music.data package contains four Java source code file:

Note.java - The Note.java file contains one class called Note.
Phrase.java - The Phrase.java file contains one class called Phrase (Phrases in jMusic are monophonic).
CPhrase.java - The CPhrase.java file contains one class called CPhrase (C stands for chord).
Part.java - The Part.java file contains one class called Part.
Score.java - The Score.java file contains one class called Score.

These classes form the backbone of the jMusic data structure which allows the composer to create, analyse, manipulate music. The next section looks at these classes in more detail.
The musical information in jMusic is stored in a hierarchical fashion based upon a conventional score on paper.

```
Score (Contains any number of Parts)
   |
   +---- Part (Contains any number of Phrases)
           |
           +---- Phrase (Contains any number of Notes.)
                    |
                    +---- Note (Holds information about a single musical event.)
```

Notes

The jm.music.data.Note class is the basic note structure used by jMusic. Note objects contain lots of useful attributes:

- Pitch - the note's pitch
- Dynamic - the loudness of the note
- RhythmValue - the length of the note (eg., Crotchet)
- Pan - the notes position in the stereo (or more) spectrum.
- Duration - the length of the note in milliseconds
- Offset - an deviation from the 'normal' start time of the note

The Note class also has a collection of methods which carry out various manipulations on the data.

Phrases

The Phrase class is a little more complicated than note objects but can be simply explained as being voices. A piano part is a single part but can have multiple voices. Take a Bach Fugue as an obvious example. Phrase objects really only contain a single important attribute, a list of notes. Every Phrase object contains a list of notes that you can add to, subtract from and move all over the place. The list that does all this is called a *vector* and is a Java class found in the util package java.util.Vector. Phrases include:

- noteList – A vector of notes
- startTime – The beat position at which the phrase begins
- title – The name of the phrase

CPhrases

The CPhrase class allows the composer to construct homophonic musical structures easily; that is, have jMusic play chords. A chord is defined as a group of notes that share the same onset time and duration, but differ in other respects; in particular they have different pitches. The CPhrase class can be used just like the Phrase class, but behind the scenes CPhrase structures are converted into Phrases. This becomes important only when you work with relatively complex musical structures in jMusic.

Parts

Part is a class that, surprisingly enough, holds the notes (in phrases) to be played by an instrument. A part contains a vector of phrases. A part also has a title ("Violin 1" for example), a channel, and an instrument (in MIDI, a program change number - in audio an index in the instrument array). Again don't worry about the details too much just take a look. Parts include:

- phraseList – A vector of phrases
- progChg – The instrument used for notes in this part
- title – The name of the phrase

Scores

The Score class represents the top level of our data structure and it contains a vector of parts, and also has a title (name). Scores include:

- partList – A vector of parts
- title – The name of the score

jMusic constants

The JMC class contains a collection of constants (unchangeable variables) which help the computer refer to musical events using musical language. In this paper we will introduced some key words which jMusic uses to refer to musical attributes. Constants allow your code to be more human readable and 'musical'. For example, you can specify a pitch using the constant C4 rather than the number 60. Behind the scenes Java translates the constant C4 to 60.

The constants are set up in the class called JMC and to use them you must import that class. You will notice that most jMusic code has this line toward the start: `import jm.JMC;`

Secondly, when the class is declared it needs to implement the JMC class so the constants can be used. Therefore class declarations in jMusic tend to look like this:

```
public final class AudioScale
implements JMC{
```

There are currently jMusic constants for pitch, rhythmic values, dynamics, and MIDI program changes. In this paper we will examine each in turn.

Pitch

Note constants are written in pitch-class/octave notation, such that middle C is C4, the D above it is D4 and the B below it is B3. Accidentals are indicated with S for sharp and F for flat, eg., C sharp above middle C is CS4 or DF4. The pitch constants are converted into integer numbers for absolute pitch as used in MIDI. C4 (middle C) = 60. The range is from G9 - CN1 (C negative one).

Pitch can also be specified as frequency for jMusic audio output. Middle C, for example, becomes a frequency of 261.63. The constants convert equally tempered pitches to frequency. The syntax is FRQ[n] where n is the MIDI note number to be converted.

Rhythmic Value

Full names of abbreviations for English and American terms are provided for most common rhythmic values. jMusic is based on a beat pulse where one beat is a value of 1.0 and all other rhythms are relative to that. A rhythmValue can have several constants that equal it: `CROTCHET, C, QUARTER_NOTE, and QN.`

Dynamic

jMusic allows integers from 0-127 to indicate the dynamic (loudness) of a note. The constants, shown below, specify some set levels using abbreviations for the common Italian terms for dynamics (pianissimo etc). For example, `SILENT = 0, MF = 70, and FFF = 120;`

Panning

There are jMusic constants for panning notes across the stereo image. jMusic supports more than two (stereo) outputs but the constants only support the normal two channel output. Here are a few of the constants which relate the panning of a note: `PAN_CENTRE = 0.5, PAN_LEFT = 0.0, PAN_RIGHT = 1.0;`

Duration Articulation

In many cases jMusic notes will have a duration which relates to the rhythmicValue. By default this is 85% of the rhythm value. So a crotchet note will have a rhythmicValue of 1.0 and a duration of 0.85. The duration constants are designed to be multiplied by the rhythmicDuration (CROTCHET * STACCATO), that way they can be applied to any rhythmicValue. There are a few common Italian terms for articulation defined as constants for the duration of a note. For example, `STACCATO = 0.2, LEGATO = 0.95.`

Timbre (MIDI Program Change)

The General MIDI set of program changes are included as jMusic constants. Each constant equates to a integer number from 1-127. These will only be relevant when playing back MIDI files generated from jMusic on General MIDI equipment (such as PC sound cards or Quicktime Musical Instruments). Most orchestral and rock band instruments are part of the GM set and you should be lucky enough to guess the name in most cases as several alternate wordings are provided form many instruments. For example, `PIANO = 0, EPIANO = 4 or ELECTRIC_PIANO = 4 or ELPIANO = 4.`

Code syntax

JMusic follows all the convention of Java syntax, so there is little need to cover that in detail here. Below is an example of a class that generates a jMusic score of one octave ascending chromatic scale, and outputs it as both a MIDI file and .au audio file using a simple triangle waveform oscillator.

```
import jm.JMC;
import jm.music.data.*;
import jm.midi.*;
import jm.audio.*;
import jm.audio.synth.*;
import jmInst.*;


public final class Scale implements JMC{
public static void main(String[] args){
 Score s = new Score("JMDemo - Audio Scale");
 Part p = new Part("Flute", 1, 0);
 Phrase phr = new Phrase(0.0);

 TriangleInst ti = new TriangleInst(44100);
 Instrument[] ensemble = {ti};

 for(short i=0;i<12;i++){
  Note n = new Note(i+60, Q, 127, 0.5);
  phr.addNote(n);
 }

 p.addPhrase(phr);
 s.addPart(p);
 s.writeMIDI("Scale.mid");
 s.writeAU("Scale.au", ensemble);
}
}
```

Standard Methods

As well as a data structure for music, jMusic provides a collection of methods (procedures) for operations on the data. Each class—Note, Phrase, CPhrase, Part, and Score—has methods for acting on its data.

Get and Set

There has been a deliberate attempt to make the jMusic classes operate and use syntax similar to the Java libraries. In keeping with this idea each of the class attributes can be retrieved with a method call starting with 'get'. As in getPitch for the Note class. To define the value of object attributes there are methods beginning with 'set'. As in setStartTime for a Phrase.

Manipulations

There are methods for making structural changes to the jMusic data. These aim to be fundamental processes from which the user can construct more complex methods. For example, the user might combine copy and transpose methods to effect a melodic sequence. Some of the methods in the music.data package include:

- copy
- repeat
- size
- getEndTime
- fadeIn
- fadeOut
- append
- empty
- compress

Audio Instruments

There are a whole set of audio classes in jMusic. They are arranged into three subdirectories, each of which should be imported. The jm.audio.data directory contains the audio structure information and supports reading and writing audio files. The jm.audio.synth directory contains various synthesis functions and the oscillator class. The jmInst directory contains the jMusic Instruments. The first order of audio business is to declare an instance of the Instruments used to play the jMusic score. In the code above this involves creating an instance of the TrangleInst which produces a simple triangle waveform. The instruments need to be collected into an array. This array is called 'ensemble' in the example, and only contains one instrument. More complex scores may use several instruments.

Creating an audio instrument involves extending the Instrument class and by chaining together audio components, either built-in processes or those developed by the user. For example, the SimpleSampleInst class is chain of the following audio processes;

- AUIn
- resample
- envelope
- sampleOut

The details of creating audio modules and processes is beyond the scope of this paper, but is not dissimilar to creating instruments in Csound or Supercollider.

Obtaining jMusic

JMusic is available for downloading from the jMusic web site:
http://www.academy.qut.edu.au/newmedia/jmusic/
The web site contains installation instructions and an extensive range of tutorials to help users get started. In addition the web site has a list of suggested references on computer assisted composition, related topics, links to similar computer music projects, and examples of music created with jMusic.

Conclusion

In this paper we have provided an overview of the jMusic language. In doing so we have described the aims of jMusic and how it relates to other computer music languages.

The data structure and other aspects of jMusic have been introduced and some of the main issues in the jMusic composing process have been explored. The jMusic project continues to develop, and we hope that more musicians may make use of jMusic and that their feedback and coding efforts might contribute to its ongoing development. Most of all we hope that jMusic might be a pathway for many musicians to explore algorithmic composition and that plenty of interesting music will result.

References

Assayag, G. & Agon, C. 1996. “Open Music Architecture.” In *Proceeding of the 1996 International Computer Music Conference, Hong Kong*. San Francisco: ICMA. pp.339-340

Laurson, Mikael. 1996. *Patchwork: A visual programming language and some musical applications. Studia Music 6*. Helsinki: Finland.

Diffusion: Realisation, Analysis and Evaluation

Ian Stevenson
Audile™ Design
2/46 Pitt Street
Ringwood Victoria 3134
Audile@bigpond.com

Introduction

This paper attempts to deal with various aspects of the theory and practice of music diffusion. Last year's conference was centred on space and I want to focus on the interfaces by which we achieve spatialisation and the interface between spatialised sonic material and our perception and understanding. Generally speaking diffusion is the process whereby the inherent spatial properties in a piece of audio may be realised through electro-acoustic means during concert presentation. It may also be a process whereby spatial properties are imposed on material where no strong spatiality is present. The term itself implies a particular approach characterised by a diffuse/ de-correlated sound-field. I would like to argue that while this approach has many benefits, and probably results from an understanding of the practical limitations of physical electro-acoustics, it is by no means the only answer. What follows is a wide ranging discussion of the various aspects informing the practice, analysis and evaluation of electro-acoustic diffusion.

Realisation

I will start with the practical aspects, the realisation of concert diffusion. Last years conference dealt with this issue in some detail. I would like to bring together some ideas from my experience. Firstly, I think it's important to have some understanding of the history of electroacoustic spatialisation and more broadly the history of the use of space in musical practice.

The very nature of the acoustic transmission of sound within air as a medium invites composers to manipulate spatial properties in performance. As soon as more than one performer is present there is a spatial element. A quick fire survey of the Western canon would identify antiphon in medieval church music particularly Gregorian chant; followed on with more elaborate applications of antiphonal composition in the Baroque and classical periods; St. Mark's in Venice with its spatially opposed organs was the site of some development in the late Renaissance / early Baroque, of the antiphonal devices especially associated with the Gabrielis, Andrea and his nephew Giovanni. They pioneered a polyphonic style where the various melodic lines were presented by spatially distinct choral and instrumental groupings. Other famous examples of the use of spatial effects include the antiphonal choral effects in J.S.Bach's St. Matthew Passion and Mozart's Serenade in D for 4 Orchestras with its motivic interplay between spatially separated instrumental groups. The inevitable spatial effects of Mahler's Symphony No.8 in E flat 'Symphony of a Thousand' are really more a result of the sheer expanse of performers than an effect by design. Finally, the spatial location of musical sources has often been exploited in the theatre. There are many examples of on or off stage bands in opera. All this has led to an expectation of spatial properties in musical performance.

With the advent of electro-acoustic devices for the recording, transmission and re-play of musical material, it is not surprising that a more extensive presentation of space was expected by the public than that which could be encoded by variations in level, timbre and reverberation in a mono signal. It took sixty years for the music publishing industry to release stereophonic recordings although the theory of stereo spatial encoding had been around for at least twenty years prior. Popular music performers were quick to adopt recordings as the site for their compositional process with strange and clumsy use of stereo exemplified by such works as The Beatles "Sgt Pepper's" album.

Of course, I have jumped over the infiltration of electroacoustic devices in to the music academy. The potentially liberating invention of the loudspeaker held a unique fascination for composers inclined toward modernism. It was not until after the second world war, of course, in the spirit of redevelopment and turning away from the past, that the use of loudspeakers in musical composition became more common. The use of the tape recorder as an adjunct to acoustic instruments or as an alternative, as in the work of Pierre Schaffer, became more common. However, electroacoustic composition with space as an integral musical component is a distinct form. Another quick fire list of the more famous examples includes (excuse my pronunciation): Stockhausen's "Gesang der Jünglinge" of 1956 which includes five loudspeaker locations, Kontakte 1960 includes a four track tape. The Philips pavilion for the 1958 Brussels world fair, designed by Iannis Xenakis and Le Corbusier included an 11 channel 425 loudspeaker sound system for which Edgar Vàrese composed Poème Electronique and Xenakis produced Concret PH. Stockhausen produced various pieces for the German pavilion at the Osaka EXPO 70. The pavilion was a geodesic dome with loudspeakers at nearly every vertex and integral sound control equipment. EXPO 70 also included the 800 speaker installation in the Japanese Steel pavilion. This system was used by Xenakis for his 12 channel composition Hibiki Hana Ma. The idea of introducing the spatial complexity of an orchestra into electroacoustic performance has been explored by the Groupe de Musique Expèrimentale de Bourges with their multi-loudspeaker Gmebaphone. A similar idea is seen in the Acousmonium of the Groupe de Recherches Musicales. Pierre Boulez utilised a multi-speaker system suspended in the auditorium for his Rèpons. In the States, John Chowning developed software to control spatial location and panning in a four speaker system.

Of course, I've left out lots of vital examples but over this period a range of practice evolved. At the same time in the more commercial field of music publishing, various methods were being employed to improve the fidelity of recordings. Engineers attempted to provide more and more accurate mapping of auditorium acoustics into home reproduction.

Composers such as Stockhausen, desperate to specify every aspect of the musical performance also did battle with the vagaries of auditorium acoustics. Repeatable presentation of spatial encodings continue to defy practitioners in all fields. Michael Gerzon's brilliant meta-theory of spatial encoding which supports the practice of ambisonic recording provides a perfect framework for spatial specification. However, it fails completely to overcome the effects of auditorium acoustics for replay and let's face it, for those who have done experiments with ambisonics, you can never have too many speakers or too little reflection in your performance space. These technologies, however, have become more subtle and refined and proven to very effective when employed carefully.

However, many composers and sound designers working with space at a musical level want more than subtlety. They want bold gestures, they need as much support for their fragile musical intentions as an electro-acoustic system can provide. Hence we find what appear to the acousticians as fairly crude multi-channel systems such as Dolby surround and those used in much of today's acousmatic and electro-acoustic performance.

Accurate spatial reproduction requires highly correlated, phase accurate, directivity controlled multi-channel speaker systems. Theatre sound designers of the past two decades have been working away at the margins, employing psycho-acoustic principals such as the Haas effect and utilising more and more sophisticated measurement systems, to trick audiences into believing the realism or transparency of their sound reinforcement systems.

Parallel research into psycho-acoustics has shown that our ability to understand and interpret musical information in polyphonic writing is aided by spatial de-correlation. Spatially separated steams can be processed in parallel retaining more of the discrete informational content of the individual streams than those collapsed into one spatially coherent stream.

Essentially, there are three approaches to spatial manipulation: Firstly, we can build a soundfield from a number of discrete, de-correlated, spatially separated streams or sources. This is what happens with an instrumental ensemble on stage. The acoustic interaction of the auditorium, however, produces diffusion, that is, a build up of discrete reflections which increases the correlation of the acoustic information coming from one location in space, and therefore fusing the instrument streams, blending timbres. The acoustics of an auditorium given by its overall volume and the absorptivity of its surfaces and their relative orientation thus producing differing reverberation times, ratios of direct to reflected sound and onset, density or diffusion of reflections, effects the correlation of the discrete streams and therefore the perceived spatialisation.

A second approach to spatialisation takes a single coherent stream and presents it from a number of spatially separated acoustic sources, here I'm talking specifically about loudspeakers. This might seem like a totally artificial type of spatialisation that is unlikely to occur in nature. But, I think in terms of our perception, it probably comes close to certain modes of listening. Our brains ability to perceive separate auditory streams is determined both by their number and spectro-morphology, and the type of attention that is paid to them. Later, I will introduce the concept of *modes of listening* or *listening relationships* which attempt to describe the broad overlapping categories of attention paid to musical material. In certain circumstances streams are fused in our perception, if these fused streams originate from separate acoustic sources we may tend to feel surrounded by a particular sound. Other instances of this type of perception are generated in highly diffuse sound fields and by the psycho-acoustic perception of low-frequency sounds.

The third type of spatialisation involves moving sources. The illusion of a moving source is usually achieve by proportional signal panning of a single stream across a number of spatially separated loudspeakers.

Each of these types of spatialisation is fairly distinct in terms of its coding or imaginative associations.

I have mentioned the term correlation which will be familiar to the mathematicians and engineers in the audience but probably needs a bit of conceptual unpacking. Broadly speaking, statistical correlation describes a continuous measure of matched groupings in a number of data sets. In signal processing in both discrete and continuous time systems correlation is a measure of the average coincidence of phase and amplitude values of two signals. It is the continuous average of the products of two or more signals. Correlation is used to describe a stereo signal to gauge its stereo spatial content. In this case, a correlation value of one indicates a mono signal. Minus one is 180 degrees out of phase, and a zero correlation indicates highly de-correlated material. A reasonably high positive values is usual in stereo music recordings.

Correlation is a useful concept to keep in mind when discussing music diffusion. Consider two examples: first a single audio stream is sent to eight loudspeakers surrounding the audience, the stream contains a solo violin. This is an artificial sound field, flat, strange, maybe dreamlike but highly spatially correlated. The second example uses eight separate streams, mono recordings of cicadas. Each stream is sent to one of eight loudspeakers. Here the soundfield is alive (in more ways than one), highly de-correlated, highly textured exciting and natural. Of course, there is a range of possibilities between these two extremes.

As I mentioned the diffusion of a sound field is directly related to the interaction of the electro-acoustic devices – the loudspeakers – and the architecture. This in turn influences the level of correlation that may be achieved in a given auditorium. Frequency dependant absorption and specular reflection of architectural materials inter-react with the frequency dependant directivity or beamwidth of loudspeakers. The unpredictability of these relationships has led the designers of diffusion systems to make a range of idiosyncratic choices in the selection and deployment of electroacoustic devices. A classic example being the Acousmonium of the GRM.

The choice of playback format and performance interface is further influenced by musicological factors, including distinctions between performance and composition, fixed/notated and improvised performance processes

The dichotomy between a compositional approach which begins from the intrinsic sonic properties of a sound object and develops through experimentation towards larger forms, and an opposite approach which begins from formal objectives and then moves to the sound material with which these forms might be realized, was identified by Pierre Schaeffer in his early researches into musique concrète. This dichotomy is characteristic of the divergent approaches to the composition and performance of electronic and electroacoustic music that have existed throughout its fifty year history. The culturally established norms of notated composition and interpretive performance are still influential in discussions of contemporary practice. Similarly, the distinction between composition and improvisation are still held in some quarters despite the blurring of these boundaries that has occurred through practice in the electronic music studio.

One interesting aspect of this is the traditional location of formal and constructional properties of a composition in the score. This represents the work of the *composer*. Conversely, in traditional performance practice the expressive nuance or gestural shaping of the sonic events that comprise the music are in the domain of the *performer*. Electronic and tape composition fuses these aspects of the music in the work of the composer in the studio. The composer is concerned both with the large scale formal aspects of the composition and the discrete gestural shaping of sonic events. The studio environment allows the composer to experiment, evaluate and modify each element of the musical material. With the appropriate direct manipulation control device the studio composer can *perform* elements of the composition and record the results.

This situation radically changes the traditional concept of the role of the composer in two related and fundamental ways. Firstly, the composer is working directly with the materials of his composition. The composer is manipulating the sonic material rather than merely encoding directions for its production in the score. Secondly, and as a result of the first point, the composer is *performing* the composition in the studio through a process of experimentation and evaluation. This results in a blurring of the distinctions between composition, performance and improvisation.

The unpredictability of spatial outcomes in different auditoria encourages the composer to adapt the spatialisation of a piece for each new performance. The lack of standardisation of replay systems also makes this a necessity. The ease of production and dissemination for two channel or stereo media dictates this choice as well. Behind all of this though, there is the need for reaffirmation of the composer as performer and the perceived need for theatricality in the concert space. The performed diffusion as has become common in the UK and Europe is the result of all these factors.

There are of course, other alternatives. Real-time digital signal processing is giving us the means to record and edit multi-channel diffusions with greater ease. Replay and mixing systems such as LCS, AudioBox and Protools allows for rapid experimentation and modification of existing diffusions. AudioBox control software allows for automatic translation of a diffusion from one physical speaker arrangement to another. This sort of processing brings the studio composition/performance practice of the composer into the concert hall. The diffusion and indeed the mix is no longer fixed on tape and the composers intentions may be more fully realised through rehearsal and refinement in the actual performance space. Of course these developments do not negate the validity of performed diffusion and the altered acoustics of a full auditorium will always encourage the corrective participation of virtuosic diffusion performance.

Analysis

The analysis of music diffusion or soundstage presentation in commercial music recordings has a long history and well established principals. The theoretical basis for analysis of diffusion in electro-acoustic and acousmatic music is also well established, however, its practice and the drawing together of the theoretical strands which must underlie its musicology are in their infancy. The topic has been developed by authors such as Denis Smalley and Trevor Wishart and Simon Emmerson but they have avoided some of the practical considerations outlined above. The field of the psychology of auditory scene analysis and the application of its insights to the development and presentation of scene based structures in musical composition is another area ripe for development. The principals of soundscape design and analysis should influence judgments on musical content and the intrinsic and extrinsic relations and structures inherent in its elements. The inter-relations between scene analysis and scene description in the soundscape should further be exposed when informed by an analysis of the spectro-morphological properties of the sonic materials. This approach should reveal aspects of the source and genesis of sound objects and more complex sound structures and their spatial inter-relations. Spectro-morphological analysis underpins the discussion of spatio-morphology (to borrow Smalley's jargon) and provides further theoretical basis for practical analysis. Finally, a poetics of audible space and an image analysis of spatial relations should

reveal a rich poetic source of resonance and identification. Image, like correlation, is a densely packed concept. It brings with it notions of spatial perception, literary analysis, and the shimmering inversions of subject and object which make the greatest examples of electronic music such rewarding listening.
Spatial motion as gesture is an obvious starting point. And an analysis of levels of gestural surrogacy, as Denis Smalley describes, provides a frame work. In Smalley's spectro-morphology low order gestural surrogates provide a direct audible link between the causal mechanism and its sounding outcome. As we move higher in the order of gestural surrogacy, possibly through the application of some signal processing treatment changing the amplitude envelope of a sound, we effectively disguise the causal source. So it is with spatial motion. The scale of the scene is determined by the possible spatial motion of the sources. The spectro-morphology must also relate to the spatial motion. A fly must be in motion otherwise it is silent.

In the wide-open sound world of electroacoustic music we need structured methodologies to facilitate analysis and I think a return to the philosophical origins of musique concrete may help to show the way. The phenomenology of Henri Bergson influenced writers such as Gaston Bachelard whose psychoanalytical approach to material imagination must surely have prompted Schaeffer in his typology of spectral forms. An attempt to measure the imaginative charge of spatial gesture in electroacoustic art would do well to start from Bachelard's "Poetics".

Following Schaeffer's analysis of *modes of listening*, Smalley approaches the analysis of extrinsic referents in terms of the listeners perception. Smalley identifies types of relationship that the listener may have with the musical material (Smalley 1996). These include the *indicative relationship* in which the sound carries information or a message for the listener. The listener may enter into this relationship actively or passively depending on whether the sound event is expected or not. The second relationship is the *reflexive relationship*. This is a fairly one-sided relationship in which the listener is mainly aware of the emotional or affective impact that the sound has on him. The final relationship described by Smalley is the *interactive relationship*. In this situation the listener is actively engaged in exploring the sound and its intrinsic properties and relationships to other sounds and events. According to Smalley most listening is characterised by the first two relationships. The interactive relationship is characteristic of the type of listening engaged in during a concert or other concentrated listening. During normal listening the relationship will shift between these three depending on the sound material and the conscious attention of the listener.

It is the scope for development of the subtle use of the indicative relationship that Smalley spends some time developing. He introduces nine classes within what he terms as the *indicative field* or *indicative networks*. These classifications include, in Smalley's terminology, *gesture, utterance, behaviour, energy, motion, object/substance, environment, vision, and space.*

The interplay of these listening relationships is what spatialisation does best. The lack of strong coding of spatial properties, should not and does not deter the composer or the critic from a thorough exploration of these principals.

Evaluation

The final area I would like to discuss suggests scope for the evaluation of diffusion as a musical property.

The analytical principals I have introduced should begin to help the critic to navigate in the unmapped territory of the wide-open-soundworld. Acousmatic music presents the paradox of a humanist tradition realised by material means. The materializing and abstracting force of technology must be harnessed and directed by the imaginative composer. The corruption of our listening culture can be exemplified by our acceptance of machine originated cyclic panning motions, made acceptable by our constant exposure to these patterns of motion in both commercial music and commercial machines.

Diffusion practice must start from the sonic material, from the perception of sonic phenomena. Just as the pattern forming processes of our minds revert to sounding bodies and gestural causes. In our imaginative life and in the world of dreams, spaces undergo endless transformations. These transformations parallel the poetic, mythic and psychic development represented by a rich and powerful acousmatic art.

Spatial expression that is too divorced from gestural sources will appear cold and sterile. These properties have often been exploited or exposed in the machine dominate culture of computer music and they have their place in the expression of these principals.

The successful critic will be aware of the physical limitations of performance and presentation systems at the level of the physical interface. The critic will also be concerned with the imaginative interface between the composers intentions and the listeners perceptions and subjective spatial images.

The need to develop and discuss the principles of diffusion, its realisation, analysis and evaluation have provided the motivation for this presentation and I thank you all for your attention.

References

Bachelard, G. 1994. *The Poetics of Space* Beacon Press, Boston USA.

Chion, M. 1994. *Audio-Vision - Sound on Screen* Columbia University Press, New York USA.

Dack, J. 1994. *Pierre Schaeffer and the Significance of Radiophonic Art* Contemporary Music Review 10:2 pp. 3-11 Harwood Academic Publishers, Amsterdam Netherlands.

Dahlhaus, C. 1982. *Esthetics of Music* Cambridge University Press, Cambridge U.K.

Emmerson, S. 1996. *Local/field: Towards A Typology Of Live Electronic Music* Journal of Electroacoustic Music Vol.9 January 1996. pp.10-12, Sonic Arts Network, London UK.

Griffiths, P 1994. *Modern Music: A Consice History* Thames and Hudson, London UK.

Griffiths, P 1995. *Modern Music and After* Oxford University Press, Oxford UK.

Griffiths, P. 1986. *The Thames and Hudson Encyclopedia of 20th-Century Music* Thames and Hudson, London.

Grout, D.J. & Palisca, C. V. 1988. *A History of Western Music* Fourth Edition J.M.Dent and Sons, London.

Maconie, R. 1976. *The works of Karlheinz Stockhausen* Oxford University Press, London UK.

Manning, P 1993. *Electronic and Computer Music* Oxford University Press, Oxford UK.

Minsky, M. 1981. *Music, Mind, and Meaning* in Roads, C. ed. 1989 *The Music Machine* MIT Press, Massachusetts USA

Smalley, D. 1986. *Spectromorphology and structuring processes* in Emmerson, S. ed. *The Language of Electroacoustic Music* pp.61-93 Macmillan Press, UK.

Smalley, D. 1994. *Defining timbre - refining timbre* Contemporary Music Review 10:2 pp. 35-48 Harwood Academic Publishers, Amsterdam Netherlands.

Smalley, D. 1996. *The Listening Imagination: Listening in the Electroacoustic Era* Contemporary Music Review 13:2 Harwood Academic Publishers, Amsterdam Netherlands

Smalley, D. 1997. *Spectromorphology: explaining sound-shapes* Organised Sound 2:2 Cambridge University Press, Cambridge UK.

Truax, B. 1984. *Acoustic Communication* Ablex Publishing Corporation, Norwood New Jersey USA.

Wishart, T. 1986. *Sound symbols and landscapes* in Emmerson, S. ed. *The Language of Electroacoustic Music* pp. 41-60 Macmillan Press, UK.

Wishart, T. 1996. *On Sonic Art.* Harwood Academic Publishers, Amsterdam Netherlands.

Towards Melodic Extension Using Genetic Algorithms

Michael Towsey, Andrew Brown, Susan Wright, and Joachim Diederich
Queensland University of Technology
Victoria Park Rd.
Kelvin Grove, 4059
Brisbane, Australia.

Abstract

In this paper we will discuss the melodic features used as the basis for a Genetic Algorithm fitness function and mutation procedures. We will discuss the method of selecting the features and what they reveal about melodic writing and genetic musical processes. Genetic Algorithms are considered promising for music composition because they combine 'creativity' (ability to explore a large search space) with constraints (a fitness function 'weeds out' creative 'excess'). However, progress to date has been limited primarily because it is difficult to define fitness functions which capture the aesthetic qualities of the wide gamut of successful melodies. In this paper we report on research that addresses this problem for a modest compositional task, melodic extension. We will discuss how features were chosen, measured for significance, and might be combined to form the basis for a fitness function. We discuss how the features are implemented in knowledge-enhanced mutation and crossover operators.

Melodic extension

The use of heuristic principles for automated music composition is well established in computer music (Moorer 1972, Laske 1992) but the use of Genetic Algorithms in this task is less well established. They have been used previously in the harmonisation of melodies by McIntyre (1994), Horner and Ayers (1995), and Wiggens et al. (1997), an interactive system was developed by Biles (1994) for performance of jazz improvisations, and have been used for automated composition by Horner and Goldberg (1991) and Jacob (1995), whose systems attempted autonomous composition, rather than computer assisted composition.

This paper reports on analysis research that is part of an investigation into how secondary music students can be assisted in learning to compose melodies by working with a computer system that actively suggests extensions to, and can provide critiques of, a student's composed melody. The complete investigation will explore the suitability of Genetic Algorithms to the 'fuzzy' task of generating 'appropriate' melodic extensions. These melodic suggestions will provide 'timely' interactive support to the students as rated by the Genetic Algorithm fitness function. The first step in achieving an effective fitness function is the definition of the features of a 'good' melody. This paper outlines the process and results of that melodic feature analysis.

The recent work of Cope (1997) with his CUE software is the research most closely related to this project in that Cope's goal was similar—to provide generative extensions to partially composed material in common music notation. Our project differs from Cope's work in its use of Genetic Algorithms rather than Associative Transition Networks (ATN) for melodic generation. We feel that this approach will provide a system at least as customable as CUE with stable performance efficiency, rather than the problems of scalability inherent in increasing database size using ATNs.

Because the analysis study presented here is the first part of a project creating music which extends monophonic musical material, it is harmonically static and does not aim to accompany any predefined harmonisation, rather, the referent for melodic continuance is the temporal context where similarity and symmetry of features (continuance) is valued. The analysis project aimed to extrapolate features from existing musical material in an attempt to provide musically appropriate features which capture a balance of novelty and consistency.

The analysis is not intended to be an exhaustive study of melody writing but rather limited to the world of short melodies in the tradition of Western diatonic music. In keeping with this aim, some of the assumed restrictions in this study include that; the compositional activity is monophonic, the representational system is common practice notation, the style of music is assumed to be western, diatonic, rhythmically regular using only simple time signatures and no modulation of key, and the melodic suggestions need to be 'appropriate' not ideal. The features should reinforce the music theory 'rules' in the school curriculum, rather than being overly 'novel' or 'creative'.

The software environment for this project is the Java-based jMusic composition language which is an ongoing research project of the QUT music department. This language provides a music data structure and visualisation interface for notation and MIDI file reading and writing routines for accessing the sample melodies for analysis.

Possible melodic features

In order to have the analysis find features which will be reinforced in school music programs, a variety of educational publications on melody writing were reviewed when collecting the melodic features. From these sources the following list of features were defined as likely to indicate tendencies in melody. The melodic features were divided into several broad categories; counts (the number of discrete elements, such as pitches), proportion (the degree to which elements are represented), ranges (the diversity of elements), pitch movement (melodic stability and contour), rhythmic stability (degree of syncopation), and patterns (the amount of internal repetition or consistency).

The complete list of melodic features and their definitions are listed below. Each feature is represented as a ratio (x per y, or x over y) and the analysis output results in a real number between 0.0 and 1.0 for each feature of each melody.

Counts
======

(1) Note Density
The number of notes per quanta (Quantum is a unit representing the smallest allowed Rhythm Value—semiquaver beats). This feature indicates the degree of business and sparseness of a melody.

(2) Pitch Variety
The number of distinct pitches per note. This feature indicates the diversity of the pitch class set used in writing the melody.

(3) Rhythmic Variety
The distinct rhythm values per value. It measures the extent to which the 16 rhythmic values between semiquaver and semibreve are used. This feature gives some indication of the rhythmic coherence of the piece and the degree to which it is unsettled or varied.

(4) Climax Strength
How often the highest pitch is used. This value is divided by one such that reaching ht high point once score high and more often lessens its impact.

Proportion of total rhythmic time
=========================

(5) Rest Density

The number of rests per quanta (semiquaver beats). This feature indicates the degree of silence or sparseness of a melody. It is related to Note Density, (1) above.

(6) Tonal Deviation
The number of non-scale (ie., major or harmonic minor) pitches per quanta. This feature provides an indication of how strongly tonal the melody is.

(7) Key Centeredness
The number of primary note quanta per quanta. This feature checks the number of semiquaver beats allocated to the root and dominant notes. It provides an indication of how strongly the melody has a sense of key.

Ranges
======
(8) Pitch Range
The pitch span divided by the possible pitch span. This feature provides an indication of the extent of the melodic contour. For our purposes the maximum range was set to two octaves, which may seem narrow but greater values resulted in too narrow a standard deviation to be useful.

(9) Rhythmic Range
The rhythmic value range over the maximum possible rhythmic value range. This measures the extent to which the melody utilises a wide variance of rhythmic values. It is different from the rhythmic variety in that it takes account of the quantity of the different rhythm values, not just their occurrence. There were sixteen allowable rhythm values from semibreve to semiquaver.

Pitch Movement
============
(10) Repeated pitch
This feature counts the consecutive notes of same pitch. It measures the lack of movement in the melody. Note that for this feature rests are ignored ie., it considers all note to next note pairs.
(11) Repeated rhythmic value
This feature counts consecutive notes of same duration. It measures the rhythmic consistency in the melody. Note that for this feature rests are ignored ie., it considers all note to next note pairs.

(12) Melodic Direction Stability
The pitch steps in the same direction compared to the number of steps. This checks the degree of melodic contour variation. The feature measure pitch intervals, which means that three notes (two intervals) are required to measure a pattern. Rests are ignored in this feature.

(13) Overall Pitch Direction
The sum of rising intervals over the sum of all intervals. This feature checks the overall tendency of the melodic contour. A melody starting and finishing on the same note will score 0.5, higher indicates an overall rise, and lower scores indicate a descending contour. Rests are ignored in this feature.

(14) Movement by Step
The number of pitch movements by diatonic step over the total number of steps. This indicates the smoothness of the melody. A high score indicates a smooth melodic curve with few large leaps. A step interval is any diatonic interval and might be one or two semitones. Rests are ignored in this feature.

(15) Dissonant Intervals
The number of dissonant intervals over the total number of intervals. The dissonance rating of an interval for this feature can in be one of three classes:

0, 5, 7	0.0 (includes unison, 4th and 5th)
2, 3, 4, 8, 9, 10	0.5 (includes Maj 2nd, 3rds, 6ths, and min 7th)
1, 6, 11	1.0 (includes semitone, tritone, Maj 7th)

Rests are ignored in this feature.

(16) Leap Return Rule
The number of pitch interval direction reversals after large leaps over the total number of large leaps. A large leap is greater than or equal to eight semitones (minor 6th).

The returning interval must be less than the jump interval preceding it.

Syncopation
=========
(17) Syncopation
The beat quanta on which notes start compared with the total number of beats. This feature indicates the degree to which the beat pulse of the music is maintained. Rests are differentiated from notes in this feature.

Patterns
======
(18) Repeated Pitch Patterns
The number of pitch patterns which are repeated over the number of quanta. This feature compares intervallic relationships so that sequences count as repeats, but at present sequences are absolute rather than diatonic. Also, to cut computation, we only consider sequences of 3 and 4 intervals starting on the beat, but attempt to find a match at every quantum. Rests are taken into account such that they may be part of a pattern.

(19) Repeated Rhythm Patterns
The number of rhythm patterns with repetitions over the number of quanta. This feature measures the rhythmic coherence of the melody as developed through repetition. The rhythmic ratio is checked rather than absolute rhythms, therefore, a pattern crotchet, minim, crotchet will match quaver, crotchet, quaver. To cut computation, we only consider sequences of 3 and 4 intervals starting on the beat, but we attempt to find a match at every quantum so that offset or phased repetitions are not ignored. Rests are taken into account such that they may be part of a pattern.

A list of definitions for terms can be found at the end of the paper.

Analysis procedure

The analysis of features used MIDI files as a readily available source of melodic data. Code was written in the Java language, using the jMusic MIDI reading and writing libraries, to handle the analysis process. The interface for the analysis program is shown in figure 1. Melodies from the Western art music repertoire were sourced as MIDI files and were appropriately prepared for analysis by quantising rhythms, transposing into C major or A minor, and setting playback to 120 beats per minute. This normalised the data such that reasonable comparisons could be drawn between them.

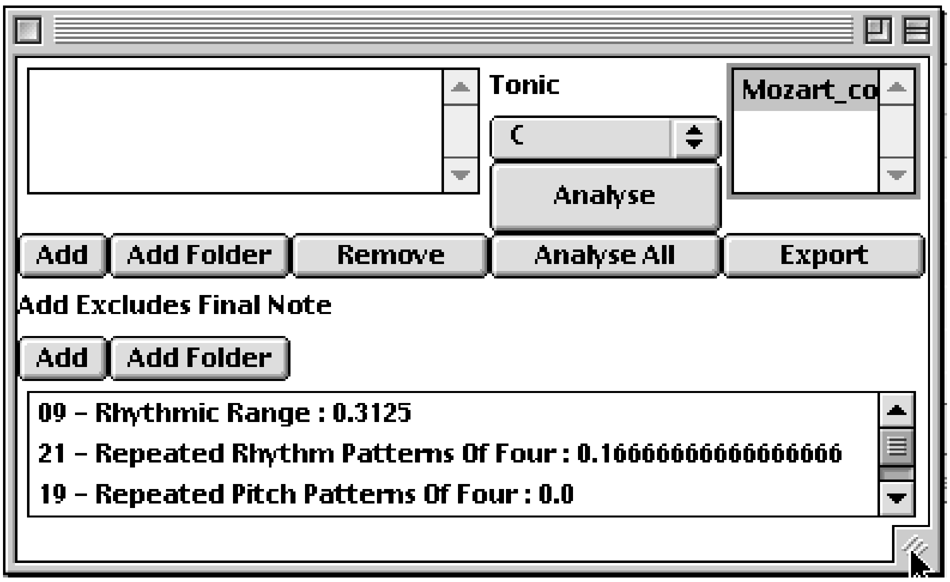


Fig. 1 – *Melodic Analysis Application Interface*

The repertoire list comprised melodies by the following composers;

- Bach
- Beethoven
- Du Fay
- Gesualdo
- Gibbons
- Montiverdi
- Mozart
- Palestrina
- Tchaikovsky
- Traditional Nursery Rhymes

Features were calculated on each melody and the average resalt and standard deviation of each feature was calculated for the whole group and various subgroups, as described below. In order to determine which features might be most important in identifying well formed melodies, T tests were run and significance assessed.

Results

We analysed a total of 31 melodies categorised according to style and mode. There were 17 melodies from classical compositions, 9 from early music compositions and 5 nursery rhymes. Eight of the melodies were in a minor key and the remainder were in major key. There were no significant differences in feature values between the major and minor categories. These categories were chosen because, to the researchers' ear, distinctive differences in melodies appear along these stylistic lines. In particular, there are differences from the harmonic and rhythmic simplicity of nursery rhymes and the modal 'flavour' of Early music. Another consideration was that the teaching of melody writing emphasises 'rules' derived from repertoire in these styles. There were some statistically significant differences between the three categories of melodic style.

Early music melodies had a significantly lower note density (Feature 1, t(df=24) = 2.89, significant @ 0.005) but a significantly higher rhythmic variety. This later characteristic was evidenced by a high value for rhythmic variety (Feature 3, t(df=12) = 2.12, significant @ 0.05), a high value for rhythmic range (Feature 9, t(df=12) = 1.93, significant @ 0.05) and the lowest rate of repeated note duration (Feature 11). Perhaps paradoxically, this greater rhythmic variety is due to more frequent use of long notes such as semibreves and breves in early music. By contrast, Classical melodies had the highest rate of repeated note durations.

There were no significant differences between the three categories of melody in their Pitch Variety (Feature 2) or in their Key Centredness (Feature 7). Tonic and dominant pitches (the pitches chosen as determining key centredness) constituted some 40% of quanta in all three categories.

Nursery rhymes used only scale notes in their melodies (Feature 6) and on average, repeated their climatic note twice (Feature 4). By comparison, the other two categories used non-scale notes for about 5% of quanta and the climatic note was repeated with less frequency. However, the rate of consecutive repeated notes (ie., having same pitch, Feature 10) was significantly lower in nursery rhymes than the other categories (t(df = 12) = 1.45, significance < 10%).

Pitch Range (Feature 8) was close to an octave for all three categories, slightly exceeding the octave for classical compositions and slightly under for nursery rhymes. Similarly, there were no significant differences between categories for stability of melodic direction (Feature 12), indicating that most melodies are generally symmetrical in their use of ascending and descending leaps. Although overall pitch direction (Feature 13) has an average value close to 0.5 for all categories, nevertheless, nursery rhymes have a significantly greater tendency to fall sightly than do Early Music melodies. (t(df=12) = 2.64, significance @5%).

With regard to Movement by Step (Feature 14), 93% of the intervals in nursery rhymes move by step which is significantly greater than the value of 75% for classical melodies and 78% for early music melodies. Nursery rhymes also have twice the rate of repeated 4 note pitch patterns (Feature 18) than the other two categories, (a significant difference) while classical melodies have the highest rate of repeated rhythmic patterns (Feature 19) although the differences for this feature are not significant due to high variability within melodic styles.

Towards a GA fitness function

The (normalised) quantitative melodic features we have described above were developed based on rules of best practice in melodic writing as taught to students of melody. While these features are 'obvious' in their application of established melody writing curricular, this analysis points toward the statistical integrity of the features and thus how reliable they will be in an computer-based judgement of melodic 'fitness'. Our use of the features is intended to modulate or constrain the application of a fitness function in a Genetic Algorithm. To be of practical use in this context, it is probably useful to have a variety of quantitative features that differ in their range of acceptable values, that is, which differ in their standard deviation. The lower the standard deviation of a feature the more severe the constraint it places on an acceptable melody. For example Feature 13, Overall pitch direction, has a comparatively low standard deviation and will place strong constraint on the selection and mutation of melodies. By contrast, Feature 19 has a large standard deviation, and may be useful only to remove extremely "dysfunctional" melodies from a population. The combination of two such features helps the Genetic Algorithm to gradually evolve a population of "fit" melodies.

From these results, it appears that the following functions will be useful in constructing a Genetic Algorithm fitness function;
(2) Pitch Variety
(4) Climax Strength
(6) Tonal Deviation
(7) Key Centeredness
(8) Pitch Range
(12) Melodic Direction Stability

This next group of functions seems less significant but somewhat indicative, and might contribute to a fitness function with small weightings.
(10) Repeated pitch
(14) Movement by Step
(18) Repeated Pitch Patterns
(19) Repeated Rhythm Patterns

The possibilities of weighting and coordination of features into a fitness function are numerous and the topic of future extensions to this research project.

Summary

In this paper we have reported on the analysis of melodic features as a determinant for their application in a Genetic Algorithm fitness function designed to assess computer generated melodic extensions. We have specified a series of melodic features, based upon melody writing curriculum used in schools, and defined those features as normalised quantitative measures, as required for computer implementation. The process of analysing the features against a set of repertoire in western musical styles was described and the significant results reported. The features which appear useful for inclusion in a fitness function have been identified.

Further research is required to determine the effective weighting of the features and any significant co-occurrence of features which might be useful. A number of features in this process produced ambiguous results (including Features 1, 9, 7, and 13) and some changes to the analysis implementation for those features and the

testing a different melody repertoire may clarify the usefulness of those features.

It is clear that although the human application of melody writing ‘rules’ seems straight forward, the articulation of those rules for autonomous computer assessment, as required for a Genetic Algorithm fitness function, is less obvious. Through this research we have described those rules as features, and identified the feature descriptions which are statistically reliable against a set of melodies from respected composers. We look forward to continuing this research and reporting on how effective the features descriptions are in selecting respectable melodies of algorithmic than genetic origin.

Acknowledgments

Thanks to Adam Kirby for coding the analysis software in Java, using jMusic.

References

Biles, J. A. 1994. GenJam: A genetic algorithm for generating jazz solos. In ICMC Proceedings 1994. The International Computer Music Association.

Cope, D. 1997. “The Composers Underscoring Environment: CUE” in *Computer Music Journal*, Vol. 21, No. 3, Fall 1997, pp. 20-37

Horner, A. & Ayers, L. 1995. Harmonisation of musical progression with genetic algorithms. In ICMC Proceedings 1995. pp. 483-484. The International Computer Music Association.

JMusic. 2000. Sorensen, A and Brown, A. http://www.academy.qut.edu.au/music/newmedia/jmusic/

Laske, O. E. 1992. Artificial Intelligence and Music: A cornerstone of cognitive musicology. In, Balban, M., Ebcioglu, K. & Laske, O. Understanding Music with AI: Perspectives on music cognition. pp. 3-29. Menlo Park, CA: The AAAI Press.

Moorer, J. A. 1972. Music and Computer Composition. Reprinted in Schwanauer, S. & Levitt, D. 1993. Machine Models of Music. pp. 167-188. Cambridge, Massachusetts: The MIT Press.

McIntyre, R. A. 1995. Composition with genetic algorithms. Technical report, University of Michigan.

Tales, tellers, and taking part: The computer music/audience dialectic

Ian Whalley
Music Department
The University of Waikato,
Hamilton, New Zealand.
musik@waikato.ac.nz

Abstract

Different styles of Western music as narrative exist in an ongoing dialectic between an audience, composer, performance, the application of technology, and reception context. Through music as experience, identity is formed and maintained by audiences. Electronic popular music was/is largely successful because it combined points of familiarity with innovation. Acousmatic music removes many links to commonly understood musical gestures, language, instrumentation, and rhythmic cohesion familiar to audiences in the wider community. This discontinuity often results in confusion. If academic computer music composers are to create continuity with the musical experience that touches people in their daily lives, points of musical similarity with the languages of a wider audience and contexts of reception outside the academy need to be found.

Given that one of the main themes of the conference is the relationship between computer music and audience, this polemic allows the exploration of a topic being examined internationally: the dislocation of the computer music of the academy from a potentially wider audience (see Milicevic1988, Bridger 1993, Landy 1994, Tanzi 1999). It affords the opportunity to re-examine the 'problem' in light of identity theory and sonic/emotional response.

Of course, this is based unashamedly on the assumption that this dislocation is something to be addressed, and that communicating with wider audiences outside the academy is worthwhile pursuing.

The academic/wider audience division is more typical in New Zealand than Australia in that computer music as an exclusive style tends to be confined to the academy. The focus in the paper is then generally on the New Zealand scene, but includes reference to international stylistic movements as points of reference.

To illustrate the central thesis, academic computer music is narrowly defined here as the acousmatic idiom (Smalley 1997) based in the Anglo-French tradition, since it is one of the prevalent approaches to computer music in New Zealand Universities.

This does not mean it is exclusive in New Zealand Universities, and I acknowledge the diversity of approaches taken by my many colleagues. Nor does it mean that acousmatic music is entirely nonsensical to everybody outside the academy: it is simply a matter of degree. In practice, my argument should then be read as a continuum rather that a simple binary proposition.

Music and identity

The theoretical approach and general discussion here (Whalley 1998) combines aspects of recent writing on music as medium by Shepherd and Wicke (1997); notions of new ethnicity put forward by Hall (1991) and Back (1996); and the linking of musical aesthetics to identity in the recent work of Frith (1997).

Individual and group identity is 'formed at the unstable point where the unspeakable stories of subjectivity meet the narratives of history and culture (Hall (1996b: 115); and it is maintained and updated once established. Although identity can be fixed, which allows general discussion of it, it is a narrative used for experimentation or reaffirmation rather than an attribute of character.

Hall (1991: 44-45) and Back (1996: 8) suggests that for younger urban youth the older divisions of race, class, nation and gender are no longer homogeneous forms. People may mix aspects of these that once seemed in opposition to form new identities. In this sense, new ethnicities evolve at a local level resulting from the productive tension between global and local influences (Hall 1991: 55, Back 1996).

In discussing the r*eception* of music, Shepherd and Wicke (1997) point out that music is a *medium* which allows a dialectic between the sounds of music and people: it is a process not an object. The dialectic is both structured and structuring in interaction with the human body. It can restrict and facilitate simultaneously the range of meanings that can be constructed by the recipient.

As narrative the medium of music allows the creation and construction of experience. Frith (1996: 109) points out, 'music produces people since talking about identity is taking about experience'. Frith (1996: 111) further eloquently provides the link between music and identity when he writes:

> .. that social groups [do not] agree on values which are then expressed in their cultural activities (the assumption of the homology models) but that they only get to know themselves *as groups* (as a particular organisations of individual and social interest, of sameness and difference) *through* cultural activity, through aesthetic judgement. Making music isn't a way of representing ideas; it is a way of living them'

From a reception perspective, the use of different musical styles to construct or reaffirm identity can appear contradictory, since personal identity is often multifaceted. One may, for example, listen to Jazz for driving yet prefer Italian opera to prepare for skiing. This diversity is a part of global trends: musical taste is not necessarily a reliable indicator of status or wealth, and the consumption of cultural products has generally become more democratic.

Yet, the *production* of music in New Zealand has not gone through a similar diversification. For example, there are few Maori speed metal bands, white *kapa haka* groups, university music department staff punk bands, or RSA club symphony orchestras. Music production has perhaps become less experimental over the last fifteen years.

This may be due to number of factors influencing the distribution of professionally made music, many not typical of New Zealand. The use of the market mechanism has partly encouraged tightly focused niche marketing techniques. This is reinforced by the increasing commercialisation of leisure, and increasing competition for discretionary leisure dollars and time. Our ageing population whose taste becomes more conservative compounds this.

Apart from commercial distribution, a great deal of music *production* in New Zealand takes place outside commercial channels. Brass bands, folk clubs, pipe bands, garage bands Polynesian festivals, community orchestras, jazz festivals, and chamber music festivals are testament to this.

The question remains why and how various musical forms and associated compositional practices collectively reinforce difference and in professional and amateur activities.

This is where Frith's (1996:111) notion above on music as process and activity also provides significant insights: musical style is part of group identity *maintenance*.

Stark examples are found professionally in the divisions of race and class, where these divisions are reinforced and reinvented. The general concentration of many university music departments mainly on 'art music', or the recent growth of Maori performing arts festivals illustrate this.

Beyond this, amateur music groups create and recreate musical styles and practices as individuals to express commonality as a group, and demarcate the group's difference from others. On a more intimate level, this allows individuals to demarcate their own subgroup within a larger identifiable style, which may seem obscure than to all but initiates of a style.

Production allows the mixture of global and national styles as an integral part of the dynamic process of Hall's (1996b) notion of identity *formation*, particular in popular music where aspects of international styles are mixed with local influences in new bands. This is evident in many Maori popular music groups being heavily influenced by black American styles.

In these terms, a good deal of what makes music 'work' as practice, is its usefulness to particular audiences made possible because its meaning is mediated through extra-musical associations, and other group activities one *participates* in.

To Electronic Popular Music

All this is made possible because of particular characteristics that Western popular and 'classical' music has as a *sonic* medium.

A factor that has made tonal music stable is its shared system of meaning that allows for many styles to be recreated and understood despite limited aural production quality and sometimes performance ability as a transmission medium.

Further, many tonal styles are based on instrumental gestures and patterns of form that are collectively shared and understood. A parallel will quickly illustrate these points. If the person on the other end of the telephone line understands English but it is a very bad line, you may catch the drift of the conversation but not understand the subtlety of it all. If you cannot understand Japanese grammar or vocabulary, the best sound in the world will not help.

One of the ways that western popular and 'classical' music developed as narrative in the last century, was through its basis in acoustic performance as a referent in the ongoing dialectic between composer, performance, and the application of new technology.

A factor that made popular 'electronic' music seem new but was relatively quick to grasp, was that it did not abandon simultaneously all aspects of acoustic music making as points of sonic reference.

The introduction of electrically recorded and generated sound resulted in the dislocation of composer/performer, real-time performance and acoustic sound, live venues and reception contexts. Despite the increasingly vague connection between the music as an expression of the body through performance value represented in acoustic instrumental and vocal gestures, electronic pop remained within the familiar gambit of instrumental *performance.* The work of Carlos with his focus on new sounds to provide interesting orchestrations of 'classical' works provides a fitting example.

The introduction of electronic instruments such as the synthesiser saw the weakening of the bonding between known clusters of instrumental technology and recognised sound, yet kept many links to tonal music traditions. They may have presented new sounds, but were largely treated as an old instrument for the most part (Whalley 2000a). Synthesiser based rock bands such as Tangerine Dream, King Crimson and Kraftwerk portray this. This may also explain the popularity of new age and world music styles (Deep Forest, Enigma) amongst younger audiences.

Perhaps the most significant new style in popular electronic music in the last decade has been the widespread adoption in youth culture of various forms of *techno* that superficially seem dislocated from the older acoustic performance tradition. This is through largely removing traditional melodic performance gestures and to some extent the traditional performer from the music, even in many live events. At first, it seems odd that such a mechanical art form has become such a signifier of bodily expression through dance.

Yet a continuity with the older Western acoustic tradition remains by focusing on elements of music that make the body integral to music as a group experience: symmetrical rhythm (often layered by cross-rhythms) and low sonic registers within implicit tonal structures. Through a concentration on these elements, the use of dance music as a *collective expression* of *individual experience* returns this music to a new tribal activity: its ubiquity as a form of youth identity is undeniable.

Popular electronic music's original communicative success was then partly due to it being adapted into known sonic frameworks and structures and replicating recognisable gestures, forming a bridge between the familiar and the unfamiliar. It was/is also useful in diverse social settings, and accessible under playback situations of widely varying quality.

Music, expectation and the academy

On this basis, unabsorbed new music styles based on *entirely* new languages and idioms present any general western audience with something of a double dilemma: a lack of extra musical (social) reference as a way of mediating musical meaning as identity through participation, however surrogate the intention of the composer; and a lack of sonic frames of reference as a foundation to find a way into them. This new experience can be emotional confrontational at a fundamental level even to the musically educated and literate.

Extending the theoretical argument helps to further explain this.

Drawing parallels between music and emotion as absolutes is fraught with danger, and the sustainability of some of the more popular arguments are often found wanting when applied to non-western or experimental music styles. Meyer 's (1956) idea that music creates and fulfils expectation seems to hold some weight when applied to traditional Western music, but lacks generality because of its tonal bias. Jourdain's (1988) selective approach to evidence seems carefully crafted to fit a predetermined view of what music should be.

However, when applied to western tonal and rhythmically symmetrical music, the general musical language of our wider community, these perspectives afford valuable insights. Jourdain (1998:302) talks of internal modelling in the construction of musical experience:

> Everything we do, including grasping a moment of music, commences with a kind of fleeting hypothesis that is confirmed or disconfirmed; every subtle mismatch is countered by adjustments to the next anticipation. We perceive music only as well as we can predict what's coming, for to predict is to model the deep relations that hold music together.

This suggests that traditional westerns music's role in the structuring of experience for many then lie in the appeal of sound as a vehicle to structuring response within shared meaning systems that are combinations of surprise and confirmation. The composer in musical work creates it with reference to common experience. This allows a continuity and exploration of identity.

In these terms, new idioms that are a subtle balance of the familiar and unfamiliar, that manipulate the creation and fulfilment of expectations, stand more chance of being absorbed by segments of the wider community, even if this takes a number of hearings.

Of course, this does not mean that the basic frame of reference of common experience does not change. Experimental popular electronic music of the 1970s now seems comparatively conservative. As stated, music is a narrative, not a given.

Given the current basis of musical language in the wider community, academic computer music in New Zealand will run into immediate problems in both musical and extra-musical terms.

One of the primary notions of acousmatic music in its purest form (Smalley 1997) is to not only to remove instrumental and performance gesture as a primary mode of communication, but in its purest form, directly familiar timbral source bonding.

In conjunction with the removal of traditional music forms based on tonality and/or static or symmetrical rhythms, any general audience is going to be left with few points of familiarity as identity, or as a basis for emotional modelling.

The sonic quality of many acousmatic pieces is highly dependent on quality sound diffusion systems and the context of listening is often critical: no place to look for a musical language that is robust independent of replication quality.

Finally, being primarily a listening rather than active participatory art such as dance music, this idiom lacks *social* generality as a robust cultural *meme* (Dawkins 1976) that will be able too freely replicate and mutate because of its wide social and individual usefulness.

My experience of introducing acousmatic music that I am fond of to those who seem liberal in their tastes, is often a reaction of bemusement.

Similarly, in introducing students to this idiom for the first time, the theory of why it is significant is the best introduction to the aural experience. Sometimes the theory creates interest than the sonic experience.

If the primary province of most music is seduction and manipulation in the first instance, introducing music theoretically can seems somewhat inverted. As an implicit art, music needs to grab people on a more fundamental level, to make some connections with common experience, if it is to find a wider use outside the academy.

The problem confronting contemporary academic computer music if it is to escape accusations discontinuity with the many musical styles that touch people in their daily lives, is to find points of similarity without descending to the banality that popular music is often accused of, and degenerating in endless pastiche of 'classical' or know innovative patterns (Whalley 1999). This is perhaps the difference between creativity and constant innovation (Milicevic 1998).

As reactionary as this approach may seem, allowing the evolution of message through extending traditional sonic references and contexts rather than adding complexity to negate it may prove to be the far more radical step. Without this, computer music in the academy runs the risk of remaining an art form that reflects the experience/identity of all but a few academy bound technicians, artists and theorists, and that needs the academy to survive.

Conclusion

What is not claimed here is that popularity equals artistic credibility, or its opposite. Such arguments usually hide more than they reveal. It is also accepted that the argument here could apply to some other experimental styles, but it is a matter of degree if some musical links to traditional acoustic music practice are kept.

The classic counter argument, that Harbison (1980) aptly deconstructs, asserts that although experimental art works might be beyond the understanding of most of the current audience, one day they will be understood by a general audience. This view is demonstrably historically as not always being so: the serial works of Arnold Schoenberg from the 1930s, for instance, have no bigger an audience today than they did half a century ago.

As previously stated, all this relies on the assumption that it is worthwhile for academic composers to want to communicate to diverse audiences, and that they may have a fair degree of control over the way any audience may respond. Many popular music practitioners have attempted this for years to no avail, and even professional marketers seem to have doubtful control over the process. Besides, many pieces of music have found their usefulness in social settings in ways that they were not originally intended for.

Neither reason is sufficient in itself to stop many people attempting this, even if the success rate is low.

Most musical composition is borne out of the dialectic between private expression and group identity: between the personal and the public. These two are not necessarily contradictory: there may be an aesthetic synergy between them.

The desire to communicate to wider audiences in the process of composition through crossing stylistic boundaries to reach audiences one may not be sympathetic with is then perhaps the key to beginning to 'jump the fence'. It is a process already happening in the many hybrid styles of popular music where local and global influences meet.

If there is to be a bridge to a larger audience, the main theoretical thrust that supports current academic computer music composition seems focused in the wrong area. Tanzi (1999:106) notes:

> By stressing the rational components of compositional practices, research applied to the treatment of musical information has stimulated experimentation in works basis on new algorithms and compositional models……both the removal of dialogue with the audience and the difficulties in finding a balance between emotional aspects of music have become critical factors that together have substantially weakened the relationship between the acts of production and listening….in order for innovations to be validated, the basic grammar constituted by the phenomenological rules of musical communication needs to be respected.

Rather than greater theoretical input going into forms of production, this suggests the need for a greater emphasis on modelling points of commonality with general audiences (Whalley 2000b). Research also needs to be coupled with empirical study of reception (Bridger 1993) as the basis of more widely accessible and unique computer music styles, and the exploration of different contexts for reception.

We may then arrive at unique hybrids that typify the fluid, debated, and at times hotly contested aspects of identities current in

New Zealand, the way that the blues and Jazz defined aspects of America in the last century.

References

Back, L. 1996. *New Ethnicities and Urban Culture.* London: UCL Press.

Bridger, M. 1993. 'Narriativation in Electroacoustic Music and Computer Music - Reflections on Empirical Research into Listeners Response'. *ICMC Proceedings*, pp. 296-99.

Dawkins, R. 1976. *The Selfish Gene.* London: Oxford University Press.

Frith, S. 1996. 'Music and Identity' in, Hall, Stuart & Du Gal, Paul (eds.) *Questions of Cultural Identity.* London: Sage.

Hall, S. 1991. 'Old and New Identities' in *Culture, Gobalization and the World System.* King, Anthony (ed) London: McMillian.

Hall, S. 1996a. 'Introduction: Who Needs Identity' in Hall, S. & Du Gal, P. (eds.) *Questions of Cultural Identity.* London: Sage.

Hall, S. 1996b. 'Minimal Selves, New Ethnicities' in *Black British Cultural Studies: A Reader.* Baker , A. et al (eds). Chicago: University Press.

Harbison, R. 1980. *Deliberate Regression: The Disastrous History of Romantic Individualism in Thought and Art, form Rousseau to Twentieth-Century Fascism*. New York: Alfred A. Knoff Inc.

Jourdain, R. 1998. *Music, the Brain & Ecstasy : How Music Captures Our Imagination.* Avon Books.

Landy, L. 1994. 'The Something to Hold onto Factor in Timbral Composition'. *Contemporary Music Revue,* 10 (2): 49-60.

Meyer, L. 1956. *Emotion and Meaning in Music.* London: University of Chicago Press.

Milicevic, M. 1998. 'Deconstructing Musical Structure'. *Organised Sound,* 3(1) 27-34.

Shepherd, J. & Wicke, P. (1997) *Music and Cultural Theory.* Cambridge: Polity Press.

Smalley, D. 1997. 'Spectromorphology: Explaining Sound-shapes'. *Organised Sound* 2(2):107-26.

Stokes, M. (ed.).1994. *Ethnicity, Identity and Music: The Musical Construction of Space.* Oxford: Berg.

Tanzi, D. 1999. 'The Cultural Roles and Communicative Properties of Scientifically Derived Compositional Theories' *Leonardo Music Journal*, 9:103-106.

Whalley, I. 1998. 'Music For A New World: Towards A Greater Synthesis In New Zealand Composition' Unpublished paper presented at: *Pacific Spaces/Global Marketplaces: Cultural Studies In Pacific Contexts* 13-14 July, Victoria University Of Wellington.

Whalley, I. 1999. 'Towards a New Aesthetic in Electroacoustic Music Practice in New Zealand Universities'. *Enculturation* 2(2):1-15.

Whalley, I. 2000a. (Forthcoming) 'Unarted Territory: Digital Music Techniques and the Development of Neo-World Music Styles' *KlangArt 99*. Osnabrueck: Univerisity of Osnabrueck Press.

Whalley, I. 2000b. Forthcoming. ' Emotion, Theme and Structure: Enhancing Computer Music Through System Dynamics Modelling'. *ICMC Proceedings*, Berlin.

Abstracts

Brisbane Powerhouse Opening

Andy Arthurs
Queensland University of Technology
Brisbane, QLD, Australia

Brisbane has been fortunate in witnessing the opening of an innovative space for "Live Arts". It is backed by Brisbane City Council and is therefore relatively well funded compared to most spaces of this type in Australia.

Its mission is to develop and present bold new work from the Brisbane Community as well as to "import" significant cutting edge work from the rest of Australia (and sometimes the world)

In early May QUT Music staff & students were involved with the opening ritual, which was a multimedia event (additionally featuring Rock n Roll Circus and Arterial Contemporary Public Arts). It was a confluence of youth, community and new work. The process and the work itself will be presented in this talk.

Environmental Infrasound Recordings

Densil Cabrera
Department of Architectural & Design Science
University of Sydney NSW

The earth, and indeed the universe, vibrate across a very large frequency range, only a small portion of which is recordable using conventional audio equipment. Nevertheless, very low frequency sound and vibration is routinely recorded for purposes other than our listening pleasure, using other transducers and data formats.

In this talk I discuss three sources of very low frequency sound recordings, what was involved in converting them to audible sound, and the results.

One month of seismic data was generously supplied by Arie Verveer, who monitors seismic activity at Kalamunda in Western Australia. As the data format was 16-bit signed integer, only a little manipulation was required to convert this to a soundfile format. The original sample rate of 40 Hz was converted to 44100 kHz. Seismometers have a dipole sensitivity pattern: the North-South data and the East-West data were used for the two audio channels. Within the period of the recording, several large events can be heard, along with a nuclear test, and varying local micro-seismic noise.

Fifty years of barometric data were obtained from the Australian Bureau of Meteorology, measured at Laverton and Williamtown Air Force Bases. The sample period was usually 6 hours. It is doubtful that any antialiasing filtering was applied. The data were supplied in text format, so a simple computer program was written to convert it into a sound file. The recording was sped up so that one year's record corresponded to one second of playback (a factor of approximately 32 million). The atmospheric “tide”, the yearly cycle, and sudden pressure changes can all be heard in this short recording.

Tidal spectra, published in the Australian National Tide Tables, were resynthesised (but sped up 32 million times). Of course, spectra are a much cleaner (and less interesting) source of data than waveforms and interest is created in this recording by simulating a circumnavigation of the Australian mainland in a virtual space. Tidal spectra involve more than twenty significant frequency components, the amplitudes of which vary from port to port. Most tidal components have a period of approximately 12 hours (semi-diurnal) or 24 hours (diurnal). However these components are all “out-of-tune” with each other, creating prominent beats. The south-west of Australia has very weak tides. The north-west has a very strong semi-diurnal tide, with the strongest tide at Derby.

These recordings are not just data transformed into sound. They are reproductions of real acoustic events, occurring on a very large scale in our acoustic environment. In March these recordings were used in a dance performance by Tess de Quincey, entitled “SkyHammer”, at The Performance Space in Sydney.

LaTrobe CD Archive

David Hirst

A selection of recordings form La Trobe Music's collection of tapes have been transferred to CD over the last year. We have tried to put together a mixture of materials on each CD consisting of works from 1976 to 1999, electroacoustic & instrumental/vocal. Most of the works are original pieces by staff and students, and the electronic works constitute about 50% of the works. There are now 10 CDs of material. Our object is to preserve some of the sounds from La Trobe for future listening & research and deposit 4 complete sets in 4 archives/libraries in Sydney, Canberra, & Melbourne.

An Approach to Spatiotemporial Music

Peter McIlwain
School of Music - Conservatorium
Monash University

The artists talk will involve a discussion of the SNet software program which generates music for multispeaker sound systems. The software generates spatiotemporal events (notes) for multispeaker arrays from 8 to 32 speakers. These events are generated by a real time neural network consisting of 32 nodes which develop and evolve spatiotemporal patterns. Notes are generated using a flocking algorithm which determine MIDI values for; pitch, velocity and note duration.

The talk will centre on the following:

- compositional problems in relation spatiotemporal music and the SNet program as one approach to solving these problems.
- a technical overview of the SNet program and its implementation
- an analysis of the behaviour of the neural network
- an analysis of the behaviour of the flocking algorithm
- a demonstration of the program with reference to the "Bird Speaker Pieces" which is submitted for the concert program in the conference.

Programmable Composition

Alistair Riddell

Composition using computers has inherently promoted the idea of a private composition discourse. Since the composer works directly with the computer to produce sound such an idea is often simply taken for granted. It is only when the composer is confronted by a new and explicit language for composition that there arises a concern about whether the language will facilitate the kinds of compositional statements that the composer hopes to make. The primary concern is the language of music rather than the language of musical production.

The use of general computer languages has always been a starting point in any creative endeavour based on the computer but traditionally these languages have been superseded or augmented by more musically orientated versions which in themselves impose a specific compositional paradigm.

Also most electronic composers confront systems which hide or disguise the detail of the musical semantic structure by providing cognitively easier modes of access (graphics or an instrumental interface). Additionally, examples facilitate use of the system by making the logic of the process much more evident. It would be exceedingly frustrating to work out a compositional procedure from the basis of the syntax alone.

So what have the general computer languages in their immediate form got offer the composer? Is there a trend to using these languages in conjunction with special musical functions that is going to result in different musical paradigms?

Distributed Audio Sequencer

Kenny Sabir

The aim of the DASE project was to create a Distributed Audio Sequencer (DASE) that would allow electronic musicians to play together over the Internet. DASE is audio based and is broken into 2 programs. The Castle Application is for playing the audio, while the peasant application is for plotting the beats. These 2 programs can be combined and networked in any configuration. An example of an application of the DASE is for a musician in Sydney to create and edit a drum pattern over the Internet with a musician in London who is creating a bass riff at the same time. The architecture of DASE assumes that the music to be created will be mainly loop based, that is, the song is made up of fixed duration patterns that are often repeated. Therefore, DASE will suit the creation of modern electronic music.

Interactive interface design: Tools and processes to enable the development of alternate input devices

Gavin Sade

This paper will discuss tools and techniques used by practitioners working with interactive media, with a view to developing an understanding of how these tools and techniques can be used to enable the creative process.

The material will be discussed in relation to the field of human computer interface design and focus on the creation of input devices to support the exploration of alternate interaction techniques.

A Survey to the Theory and Practice of Just Intonation

Greg Schiemer
Sydney Conservatorium of Music

This presentation is an introduction to the theory and practice of just intonation showing how it provides a unified theoretical basis for music found not just in the western tradition but in many non-western cultures past and present. It summarises the contribution of just intonation advocates, explains theoretical concepts such as harmonic identity, tuning limits, Partch's tonality diamond, Wilson's combination product sets and Chalmer's tri- and pentatriadic scales, and examines the role of computer technology in the on-going development of just intonation. Musical examples will be limited to the time available.

An Examination of Selected Algorithmic Compositional Processes

Andrew Troedson

This paper will look at several algorithms which have historically been used in the algorithmic composition of music, and will discuss the usefulness of each of these in the creation of music to accompany computer entertainment software. The algorithms will be classified according to their functional basis and it will be shown that each group has particular aesthetic characteristics and therefore are different in their compositional tendencies and potential application.

Musical examples of each of these algorithmic methods will be given to highlight the aesthetic successes and failures of each when assigned to particular musical parameters such as pitch, rhythm, and dynamics.

These observations will then be summarised to suggest a possible direction for the creation of an adaptive algorithmic composition system.

Musical Environments

Richard Vella

Referring to models of perception, *Musical Environments* is a discussion of musicianship which avoids the reliance on traditional music notation or performance skills and discusses music and sound on equal terms. Computer, instrumental, pop, classical music, etc, are examined in terms of sounds in time and space. Part One discusses sounds in space whereas Part Two, using texture as its main thrust, focuses on sounds in time.

Musical Environments is an approach to develop a systematic understanding of musical thinking at the secondary and tertiary level at a time when the Western European art tradition exists on an equal footing with pop, jazz and world music traditions. Referring to a wide range of musical examples including Jimi Hendrix, Beethoven, Australian composers, jazz, pop, computer music, etc, *Musical Environments* offers a generic approach to the discussion of musical structures.

Composition as revelation

David Worrall

"When I look at a starry sky, I love it in a certain way because I know it in a certain way; ... Consequently, I can handle the concepts of things themselves without being in direct possession of them, under the condition that I may conceive of them and feel them from within in some way.... [E]ven if I am incapable of dominating a certain phenomenon, I am capable of obtaining a truth which is inherent to the conceived or observed phenomenon, thanks to a kind of immediate revelation. Henceforth, I can accept and use this, in and as itself." [Iannis Xenakis]

The use of algorithms as formal compositional devices can provide mechanisms for revealing an inner nature: of sound, of material, of process, of our perceptions, our systems of symbols and metaphors. By placing an emphasis on active listening rather than trying to convey/receive messages, they can free an artist from the bounds of mere self-expression; in this sense they are revelatory.

Workshop

AudioBox & ABControl: A platform for diffusion design and realisation

Ian Stevenson
Audile[TM] Design
2/46 Pitt Street
Ringwood Victoria 3134
Audile@bigpond.com

Ringwood, Victoria, Australia.
The Richmond Sound Design AudioBox and Thirdmonk ABControl software provide an ideal platform for the design and presentation of multi-channel, multi-speaker electro-acoustic music diffusion.

The AudioBox is a hardware system which provides 8 channels of asynchronous hard disc audio playback and a 16 X 16 real-time DSP based matrix mixer. The matrix provides level control on all inputs and outputs and individual level control of each of the 256 crosspoints. In addition, delay and equalisation is available for all inputs and outputs. All level control parameters are dynamically adjustable in real-time providing the facility for complex spatial gestures and dynamic moves.

The ABControl software provides an intuitive user interface for the Audiobox. This software has been designed with electro-acoustic music diffusion in mind and has been employed on a number of electro-acoustic concert programs in Canada and elsewhere. ABControl provides facilities for diffusion cuelists, real-time diffusion control with movement capture, algorithmic diffusion generators, and vector based physical-speaker-independent speaker maps. Speaker maps allow for some degree of translation of a diffusion from one speaker set-up to another.

This workshop presentation will introduce the AudioBox and ABControl as a one-box solution for concert diffusion. The presentation will also consider the application of this system to soundscape installations, and sound effects and music playback for theatre. The workshop will address issues of performance, improvisation and compositional process. The related topics of human computer interaction, acoustics, psycho-acoustics and spatialisation will also be considered. Samples of various electro-acoustic compositions realised with the AudioBox will be presented during the workshop.