CUEING AND COMPOSING FOR LONG DISTANCE NETWORK MUSIC COLLABORATIONS

ALAIN B. RENAUD[1]

[1] *Experimental Media Research Group (EMERGE), Bournemouth University, Poole, United Kingdom*
arenaud@bournemouth.ac.uk

Long distance network music collaborations beyond the ensemble performance threshold (EPT) as exposed by Schuett in 2002 [14] where playability is affected beyond a roundtrip network delay of 50ms calls for the development of cueing mechanisms that are methodical and linked to musical parameters. The cueing strategies involved in such musical interactions will depend on the type of repertoire played and the network distance (ND) between the nodes involved in the performance. This paper proposes a semi-standardized cueing framework for real time collaborations over the network with latencies of more than 50ms. The paper also explores a compositional methodology for creating network centric performances, which couldn't occur outside of a networked situation.

INTRODUCTION

Performing in real time over a network presents exciting challenges and opportunities. The quality of networks in terms of bandwidth and reliability, at least on the current over provisioned research networks, is most of the time excellent. This means that nowadays the Network Music Performance (NMP) community can pretty much forget about solving connectivity issues that used to be prominent in setting up an NMP session and concentrate on how to better understand how to cue performers and the associated compositional processes.

The NMP community currently relies on two main approaches for playing over the network. The first relies on playing existing repertoire, which will not work for most repertoires beyond a 50ms roundtrip delay [14]. The second, for very long distance collaborations, which this paper is centred around, relies on the establishment and development of new repertoire, which considers the network as part of the compositional process. This paper covers the work undertaken with the development of network centric cues based on an initial classification proposed in 2010 [12] and how that classification has evolved. The paper also proposes a compositional methodology for long distance NMP collaborations.

1 BRIEF HISTORY

The history of NMP is relatively new. However, the quest to semi-standardize notation in distributed environments can be identified in systems such as Hajdu's Quintet.net [8], a modular software framework, which allows up to five players to compose and interact musically over a network. A system like Quintet.net allows the implementation of traditional and non-traditional notation to be distributed through a variety of nodes. Another notable initiative in this area is the Java Music Specification Language (JMSL) by Didkovsky [7]. JMSL has been thought from the ground up to be an API for the development of distributed networked compositions. Another initiative, started in 2006, is the Frequencyliator [11], an integrated visual environment for distributing abstract scores over a network of laptop musicians. The Frequencyliator was developed from the ground up to provide interactive music making over network distances (NDs) which exceed time lags of 50ms between interacting nodes.

The work undertaken with the Frequencyliator led to the initial development of more consistent and semi-standardized cueing mechanisms. The first public implementation of those cueing mechanisms happened initially with the Net Vs. Net collective [9], a network music ensemble started in 2007 by Juan-Pablo Cáceres and Alain Renaud. The collective performed intensively between Stanford University's CCRMA and Queen's Belfast SARC. The aim of the collective was and still is to demonstrate the potential of digital wide area networks for meaningful networked musical interactions. The first piece created and performed by Net Vs. Net in 2007, *"Divertimento Ritmico for Two Synthesizers, Two Locations, and One Acoustic Network of Four Channels"* [6] included synchronised visual cues and a dynamic acoustic spatialisation of audio based on the conditions of network delays. Based on the initial experience, it became quickly obvious that the development of a more robust set of strategies to make long-distance network music playing as interactive and innovative as possible was needed.

Since then, a basic set of cues has been developed along with an automated timeline, the Master Cue Generator (MCG) to manage the distribution of the cues in a linear way. Since the development of the MCG, several networked pieces have been composed and performed such as the distributed ensemble pieces *"Disparate Bodies 2.0"* in 2007 by Rebelo [5], "*Crossovers"* in 2009 by Renaud [3] and "*Renditions"* in 2010 by Renaud and McKinney [13]. *"Disparate Bodies 2.0"* was a three-site piece with a distributed ensemble (Belfast, Graz, Hamburg). The piece had a fixed set of cues, which were triggering a graphic score shared, by the three sites. *"Crossovers"* was described as a multichannel structured networked improvisation for two dislocated performers. A graphic score indicated the location of the performers in each space and allowed them to interact based on their physical location within the distributed virtual and physical spaces. The presence of the network and the two acoustically distinct spaces captured as ambisonic streams enhanced the sonic space in which the performers were interacting despite being more than 8000 kilometres apart (Belfast, Stanford). "*Renditions"* offered user generated visuals based on the interaction of three solo performers located in three different locations (Belfast, Graz, Hamburg). In addition, the cues were sonically represented across the three physical spaces using multi-channel spatialisation.

2 CUES CLASSIFICATION

2.1 Initial classification

The initial cue classification as outlined in 2010 [12] included three main types:

- Temporal: A type of cue that is sent out as information from the server to the nodes that is related to timing. An example of such a cue would be a counter indicating the time left in a cue, or a warning signal.
- Behavioural: A type of cue that is sent with a certain scenario attached to it. This can, for example, include the triggering of a waveform, or the suggestion that a given node needs to play certain nodes only above the note C4.
- Notational: A type of cue able to display content that can be identified by the performers as being helpful in the good running of the performance. This can include the visualisation of waveforms from each site, the display of a cue number, a countdown or dynamic shapes that can be activated by various factors in the performance.

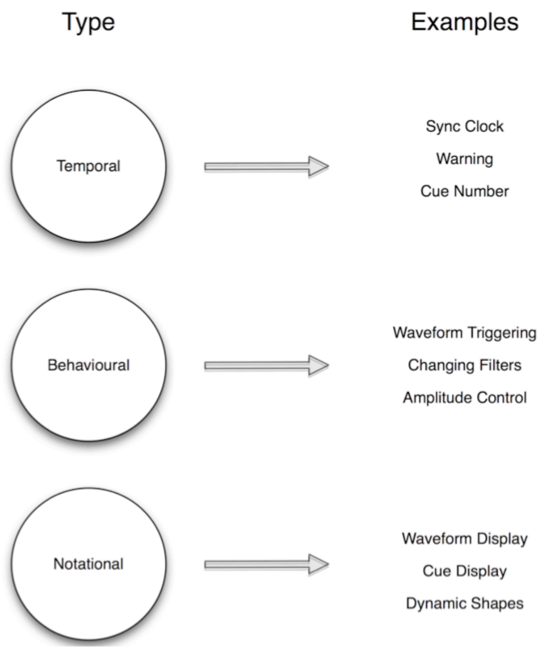

Figure 1: Initial cue classification

All the cues listed above can be active, where they actually act or change an element in the piece or passive, where they simply suggest an action to the performers.

2.2 Expanded classification

Based on the initial classification, which was established as a result of various performances, with some of them outlined earlier in the paper, an additional layer was added.

- Constant: Each cue, temporal, behavioural or notational can be delivered to the distant nodes as a constant stream of information. In this case, the latency between the main node and the distant nodes is ignored and the Master Cue Generator (MCG) sends information constantly to the distant nodes, regardless of their location and number.
- Punctual: Each cue, temporal, behavioural or notational is delivered to distant nodes through a mechanism of latency compensation. This means that the MCG will analyse the longest latency between the main node and the most distant nodes. The MCG will then compensate the latency to all the other distant nodes. For example, if the one-way latency from the main node to the most distant node is 150ms and the latency from the main node to a second closer distant nodes is 100ms, the MCG will automatically add 50ms to total 150ms. The result of this mechanism means that each distant node is getting information from the

main node at the same time, leading to an instant reaction from the distant performers.

- Countrapuntal: Each cue, temporal, behavioural or notational is delivered to distant nodes initially as a punctual cue. The latencies between the nodes can be tweaked as the cues go by to achieve countrapuntal effects, leading to various modes of interplay such as call and response or network drifts between the performers.

Each cue can be active or passive as outlined in the initial cue classification.

The new layer outlined above adds flexibility in the interplay between performers on the nodes as it allows a more dynamic attribution of cues and truly uses the inherent latency of the network as part of the musical process. The nature of each cue is determined as a flag at the beginning of each cue. Cue states can vary depending on the cue number issued. Being on the network offers quite a lot of flexibility; therefore there is also the option to send simultaneous cues of different nature to the distant nodes.

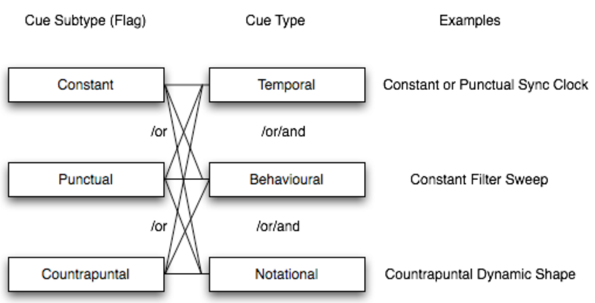

Figure 2: Cue subtypes linked to cue types

2.3 Naming convention

The MCG uses OpenSoundControl [16] formatted messages to broadcast cues over the network. The advantage of using OpenSoundControl is obvious as it is able to communicate between an ever-growing number of software applications, including those used as part of this practice based research such as Max/Msp [4], a graphical audio programming language, Processing [10], an open source graphical toolset for creating visualisations and Supercollider [15], an open source audio programming and synthesis language.

The semi standardised OSC syntax for a message broadcast by the MSC goes as follows:

cue_subtype/cue_type/cue_name_node/variable_type

Therefore, a cue representing a variable filter sweep which is imposed on a series of distant filters will be as follows, assuming that the MCG is communicating with three nodes with a range of values from 0 to 255 as integers.

Message 1:
constant/behavioural/filter_1/value_[0, 255]/[integer]
Message 2:
constant/behavioural/filter_2/value_[0, 255]/[integer]
Message 3:
constant/behavioural/filter_3/value_[0, 255]/[integer]

The approach above, even though quite straightforward, has allowed the establishment of a semi-standardised set of messages to create a structure for each of the networked pieces. As part of the practice a commonality of similar messages across the pieces has been achieved. Standardising the naming convention beyond the second layer of a message has proven difficult and is not necessarily desirable as each piece are by nature relatively different from each other.

3 COMPOSITIONAL METHODOLOGY

Deciding on a semi standardised structure for the development of network centric musical performances is not enough to establish a practice. Through the years of practice and performances over the network, a common composition methodology has been established. The process is mostly sequential and evolves as the design of the performance takes place. Four main stages for the establishment of the compositional methodology have been identified: joining up; requirements gathering; score writing; rehearsing. The joining up stage only occurs once whilst the other processes can be repeated indefinitely until the day of the performance.

3.1 Joining up

This stage is the most important stage as it happens only once and can generally not be modified because of the complex nature of performing over a network. It mostly involves deciding who will take part in the process. This is comparable to the process a traditional composer would go through when writing a score for an orchestra, such as the number and size of sections. The main questions that are being considered are mostly:

- How many site or nodes will take part and how far away are they from each other: this aspect is very important to consider as the latency values will grow in complexity between sites as the number of participants grows. It is also important to consider at this stage whether an audience will be present on all nodes or only a selection of nodes.
- Selection of sound sources: the composer needs to have a clear idea from the onset about

which sources will be used. Will they be acoustical, synthetic or both? Will some of the sound sources on distant nodes be autonomous (for example a saxophone player playing a graphic score), semi-autonomous (for example a sax player playing a graphic score which is also coupled with some sort of live electronics) or automated (for example a synthesizer or mechanical instrument that is triggered by a distant node or a centrally managed system such as the MCG).

- Bandwidth: this is mainly a technical issue. Within the current NMP community bandwidth is often dedicated but can range from a 10mb/s connection to a 1gb/s connection. Depending on the connection speed the overall requirements for composing and delivering a network centric piece will vary. Ideally each participant should have access to the same bandwidth.

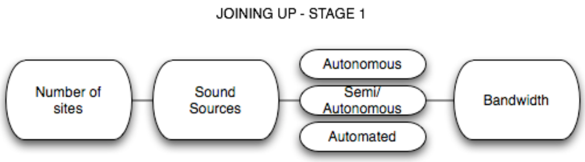

Figure 3: Joining up process

3.2 Requirements gathering

This process involves choosing the types of media involved in the performance.

- Audio: how many channels will be required? What sample rate and sample frame size will be used across the network nodes. It is an important step as the sample rate and sample frame size will have a direct impact on the overall latency values and are directly related to the condition of the network. The audio is generally transmitted over the JackTrip [1] application.
- Video: how many channels of video will be required and what quality needs to be achieved? There are no standard video solutions within the NMP community; however the solution chosen will have a direct correlation with the bandwidth available.
- Messages: messages that trigger various actions on distant nodes will generally be used. As stated earlier in the paper, OpenSoundControl messages are a de-facto standard.

3.3 Score writing

The activity of score writing, which aims to provide a structure to the network piece is directly related to the various cues outlined earlier. Common trends in the compositional process include:

- The number/type of cues present across the piece.
- The number of sections and how they are mapped to the different cues.

There is no standard process for writing a network score apart from the codification of the different events as OpenSoundControl messages and how they are sequenced or choreographed across a piece.

3.4 Rehearsing

The process of rehearsing over the network is challenging mostly because such long distance collaborations take place across different time zones. Generally, rehearsing takes place a mutually agreed time and all the scores, software and code is sent in advance. The network is also tested in advance for open ports and bandwidth.

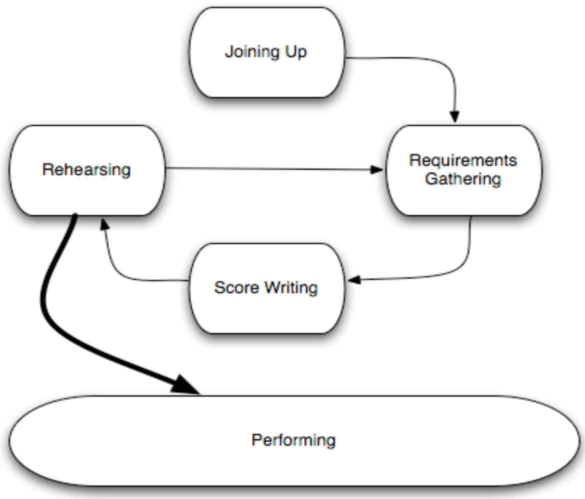

Figure 4: Composing for the network: a four stage process.

4 CONCLUSIONS AND FUTURE WORK

The aim of this paper was to offer a semi-standardised process for cueing performers over the network as well as proposing a methodology for composing for the network. Even though the approach can vary widely, the attempt to somehow codify cueing and offer a compositional methodology has so far proven useful for a variety of musical pieces that use the network as the core for interacting musically.

The classification of network cues is constantly evolving and is being codified as the practice continues. An important next step is to develop a more robust time-based OpenSoundControl sequencer as stand alone application for the MCG so that a wider audience can use it.

REFERENCES

[1] CACERES, J.-P. and CHAFE, C. “JackTrip/SoundWIRE Meets Server Farm.” Proceedings of 6th Sound and Music Computing Conference, pp. 95-98, Porto, Portugal, 2009.

[2] CCRMA Concert Series, 2008. URL http://ccrma.stanford.edu/concerts/ c_schedule.html.

[3] CROSSOVERS 2009 Available: http://www.youtube.com/watch?v=oVbwB41SOa4

[4] CYCLING ’74 2011 Available: http://www.cycling74.com

[5] DISPARATE BODIES 2.0 2007 Available: http://www.youtube.com/watch?v=k7qqCPh0HEw.

[6] DIVERTIMENTO RITMICO FOR TWO SYNTHESIZERS, TWO LOCATIONS, AND ONE ACOUSTIC NETWORK OF FOUR CHANNELS 2007 Available: http://vimeo.com/2497877.

[7] HAJDU, G. and DIDKOVSKY N. “On the Evolution of Music Notation in Network Music Environments”. Contemporary Music Review 28 395--407 (2009).

[8] HAJDU, G., Quintet.net. Available: http://www.quintet-net.org/ [22/08/2011, 2011].

[9] NET VS. NET, 2011 Available: URLhttps://ccrma.stanford.edu/~jcaceres/netvsnet/.

[10] REAS, C., FRY, B. and MAEDA, J., 2007. Processing: A Programming Handbook for Visual Designers and Artists. The MIT Press.

[11] REBELO, P., RENAUD, A. “The Frequencyliator – Distributing Structures for Networked Laptop Improvisation”. Proceedings of New Interfaces for Musical Expression (NIME), Paris, France, 2006.

[12] RENAUD, A., 2010. “Dynamic Cues for Network Music Interactions”. Proceedings of SMC 2010: 7th Sound Music and Computing Conference, 21-24 July 2010, Universitat Pompeu Fabra, Barcelona, Spain.

[13] RENDITIONS 2010 Available: http://www.qub.ac.uk/sonorities/old/2010/Sonorities2010Brochure.pdf.

[14] SCHUETT, N., 2002. The Effects of Latency on Ensemble Performance.

[15] SUPECOLLIDER 2010 Available: http://www.audiosynth.com/.

[16] WRIGHT, M. and FREED, A., “Open Sound Control: A New Protocol for Communicating with Sound Synthesizers.” Proceedings of 1997 International Computer Music Conference, Thessaloniki, 1997.