Do Androids Dream of Electric Chimera?

Author

Gifford, Toby, Brown, Andrew R.

Published

2009

Conference Title

Improvise: The Australasian Computer Music Conference 2009

Gifford, T. and Brown, A. R. 2009. Do Androids Dream of Electric Chimera? Improvise: The Australasian Computer Music Conference, Brisbane, Australia: pp.56-63.

DO ANDROIDS DREAM OF MUSICAL CHIMERA?

Toby Gifford
Queensland University of Technology,
2 George Street, Brisbane, Australia

Andrew R. Brown
Queensland University of Technology,
2 George Street, Brisbane, Australia

ABSTRACT

This paper describes the use of the Chimera Architecture as the basis for a generative rhythmic improvisation system that is intended for use in ensemble contexts. This interactive software system learns in real time based on an audio input from live performers. The paper describes the components of the Chimera Architecture including a novel analysis engine that uses prediction to robustly assess the rhythmic salience of the input stream. Analytical results are stored in a hierarchical structure that includes multiple scenarios which allow abstracted and alternate interpretations of the current metrical context. The system draws upon this Chimera Architecture when generating a musical response. The generated rhythms are intended to have a particular ambiguity in relation to the music performance by other members of the ensemble. Ambiguity is controlled through alternate interpretations of the Chimera. We describe an implementation of the Chimera Architecture that focuses on rhythmic material, and present and discuss initial experimental results of the software system playing along with recordings of a live performance.

1. INTRODUCTION

"The Chimaera was a beast in Greek Mythology with the head of a lion, the body of a goat, and the tail of a serpent. We use the word Chimera metaphorically to refer to an image derived as a composition of other images. An example of an auditory Chimera would be a heard sentence that was created by the accidental composition of the voices of two persons who just happened to be speaking at the same time. Natural hearing tries to avoid chimeric percepts, but music often tries to create them. It may want the listener to accept the simultaneous roll of the drum, clash of the cymbal, and brief pulse of noise from the woodwinds as a single coherent event with its own striking emergent properties. The sound is chimeric in the sense that it does not belong to any single environmental object. To avoid Chimeras the auditory system utilizes the correlations that normally hold between acoustic components that derive from a single source and the independence that usually exists between the sensory effects of different sources. Frequently orchestration is called upon to oppose these tendencies and force the auditory system to create Chimeras" [1].

This work is a component of a broader research program into real-time improvisational computational agents, the ultimate goal being to create a 'robot' musician that can jam with human musicians playing live music on acoustic (or electronic) instruments. Conceptually this research program has been divided into a number of stages including signal processing to convert a raw audio stream into musical notes, analysis to convert the notes into some higher order musical representation, and generative processes for utilising the higher order representation to improvise appropriate musical output. Aspects of the signal processing have been reported elsewhere [2, 3], and a process for the generation of improvised rhythmic material, the Ambidrum, was outlined in [4]. We will similarly constrain our attention in this paper to unpitched rhythmic material for the purposes of clarity, but we expect that the Chimera Architecture can be generalised to include pitched material as well and will report on a more full implementation in future publications.

This paper is conceptually related to our earlier work on the Ambidrum system but extends it in a number of ways. The Ambidrum produced generative rhythms that contained a specifiable amount of 'metric ambiguity' given assumptions about the underlying metre. Although it was designed with an ensemble context in mind, at that stage of the research we had not implemented any machine listening algorithms and so the Ambidrum system did not 'listen'. Rather it introspectively used knowledge of its own output to ensure that in-and-of itself it produced appropriate rhythms.

The generative improvisational system described in this paper utilises a novel representational architecture dubbed the Chimera Architecture, which parses the musical surface into a collection of metric scenarios with associated confidences, and is able to 'listen' in an ensemble. The Ambidrum utilised a particular measure of metric ambiguity, which essentially measured how correlated the rhythmic variables (duration, dynamics, timbre) were with the actual metre, which was assumed to be known. The Chimera Architecture does not rely on the assumption of a known metre - rather it takes into account a distribution of metric scenarios that are contained within the Chimera. Like the Ambidrum, the generative system described here utilises metric ambiguity for musical effect.

2. BACKGROUND

There are a wide range of roles for the computer to play in cybernetic musical partnerships (Brown 1999, 2000). As we consider our research into a system for collaborative improvisation it is useful to clarify the requirements for such a systems and how they differ from other computer music tasks, particularly in relation to the degree of autonomy required of the computer system. As a tool, computers are used for tasks such as recording, publishing and communicating. When using the computer as a tool the system is expected to follow directions unambiguously and show little initiative. As a technical assistant, for example in computer assisted compositional systems, the computer is often asked to complete tasks or to generate algorithmic material. Although CAC systems are generative, they less commonly achieve this based on analysis, with some notable exceptions (Cope 1996) and rarely operate in real time. Computers have played roles as real-time performers, from simple audio playback systems to score following accompanists. Another real-time application for computers in music is as instruments. This often involves the generation of synthesized sounds in response to gestural input, but can also involve symbolic manipulation and generative processes, for example during live coding performances. The computational processes here are usually highly directed, but non-linear or stochastic processes are frequently used to provide interest or surprise for the performer/user. Finally, computers can be used as improvisational collaborators where their role is to generate, at times based on ongoing evaluation, appropriate musical material as part of an ensemble performance. Early examples of these improvisational systems included those by Chadabe [5], Biles [6] and Rowe [7] with more recent examples utilising beat tracking [8] and adaptive feedback [9]. Our work reported here extends this latter tradition that combines machine listening and improvisation in collaborative performance situations.

3. ANALYSIS

In order for the improvisation system to make an informed contribution to the collaborative performance it first needs to form an impression about the musical context in which it is operating. To do this the Chimera Architecture performs a real-time analysis on the audio input it receives.

3.1. Audio Analysis

The Chimera Architecture aims to provide representational information to enable real-time percussive accompaniment in an ensemble context. To achieve this it utilises an analysis process to transform raw audio input into an abstract representation that it is used in generating appropriate musical accompaniment. In the current implementation the analysis process perceives only percussive onsets, but is designed to be able to extract percussive onsets from a complex audio stream, for example a full live band performing, or from the playback of an audio recording (so that it might be used by DJ's in performance).

A machine-improvisation system that can demonstrate some autonomy needs to represent its musical context in some way. Our approach assumes that musical understanding is codified as abstract structures that can be reorganised to form the basis for novel generative elaboration. It has been shown that these abstractions can be based on statistical abstraction (Huron, 2006) and successfully applied in improvisation systems such as as OMax, whose developers concur that "musical patterns are not stored in memory as literal chains, but rather as compressed models, that may, upon reactivation develop into similar but not identical sequences" [10:126].

In the current implementation of our analysis process we utilise a two level hierarchy of reduction to represent the musical context. The two levels are referred to as (i) salience and (ii) scenario. As the system listens only for percussive onsets, the representation of the musical context is limited to metric/rhythmic analyses (melodic and harmonic contexts are not considered). The first level of reduction is used to create a timeline of rhythmically salient moments, which involves the isolation of significant features in the signal and a measure of their relative significance. The second level, scenario, postulates possible metric abstractions, such as tempo and metre, from analysis of the salient features. We expect that similar and deeper hierarchical structures may be both possible and helpful as ways of elaborating this approach, especially given that hierarchical structures are commonplace in theories of musical analysis [11-13] and in contemporary theories of mind [14].

There are numerous benefits to be gained from the use of hierarchical abstractions for machine improvisation, including robustness and stability over time and situation, and reductions in data storage and processing requirements.

3.2. Salience

The first stage of the Chimera Architecture's perceptual hierarchy is the reduction of the raw audio signal into a timeline measuring the musical salience at each point in time. This is a low-level measure of salience - it is not intended to represent higher order musical features such as downbeats, phrase endings, and so on. Rather it aims to parse the audio stream into something similar to an event-based representation of just the percussive component of the audio signal. This is achieved using three novel percussive onset detection algorithms.

The three detection algorithms are, roughly speaking, looking for percussive onsets in the high, mid and low frequency bands; for example when listening to a drumkit the onset detection algorithms are tuned to discriminate between and kick-drum, a snare, and a hi-hat. The onset detection algorithms being used are not simply band-pass energy spike detectors. The following approaches are employed.

(i) The hi-hat detection algorithm is our SOD technique [3] for detecting 'noisy' percussive onsets.

(ii) The kick-drum detection algorithm utilises a novel time domain approach for detecting low-pitched percussive onsets.

(iii) The snare detection algorithm utilises a combination of the SOD technique and a traditional energy tracking technique.

All detection algorithms report on the amplitude of the onsets, which is interpreted as the salience at that point in time (the salience is zero at other times).

3.3. Scenario

The second tier of the Chimera Architecture's representation parses the salience into a metrical information in four steps.

(i) Firstly, it estimates the pulse.

(ii) Having estimated the pulse, it then estimates the number of beats in a bar.

(iii) Given the length of the bar, it then estimates which pulse is the downbeat; establishing the phase of metre.

(iv) Having estimated the downbeat, it accumulates the saliences of onsets that occur on each particular beat within the bar. The relative level of the salience at each bar position are interpreted as the metric weight of that beat position.

A given value for these variables (the pulse, number of beats in a bar, downbeat, and metric weights) is referred to as a *metric scenario*. A scenario represents the system's current understanding of the musical context; that is, the tempo, meter, and rhythmic density of the music it is listening to and performing with.

3.4. Chimera

Figure 1. A depiction of the Chimaera on an ancient Greek plate.

The analysis procedure from salience to scenario does not yield a single scenario, rather it yields a collection of plausible scenarios with associated confidences. This collection of scenarios we are calling a Chimera, referencing the beast of Greek mythology that Bregman [1] used as a metaphor for musical sounds that, whilst physically produced by of a number of disparate sources, are artistically combined to produce a new perceptual whole. In our case we are using the metaphor to refer to a combination of separate scenarios that may be combined to form a new hybrid scenario to artistic effect. This approach is inspired by psychological research that indicates that human musical perception maintains parallel, but not necessarily equal, interpretations and expectations during musical experiences.

During the first three steps of creating a Chimera the analysis yields a distribution of plausible values, and quantifies their plausibility as a confidence value between 0 and 1. The three steps are:

(i) From the salience is produced a collection of plausible pulses,

(ii) for each candidate pulse period, a collection of plausible bar lengths is produced,

(iii) for each candidate bar length, a collection of plausible downbeats is produced.

The confidences of the candidates at each stage are combined to calculate the confidence of each candidate scenario.

The Chimeric Architecture simultaneously tracks a collection of scenarios. To control the exponential explosion in the number of scenarios being tracked, we have limited the number of candidates at each stage of the analysis to two, so that at most there are 2*2*2 = 8 scenarios tracked at any one time. The architecture has 8 slots, that are filled with the most plausible scenarios. As time goes on, the confidence values of scenarios are updated. If a scenario becomes implausible, it is dropped from consideration, and if a new scenario is substantially more plausible than an existing scenario, the new scenario ousts the weakling. The Chimera's structure can be visualised as a tree:

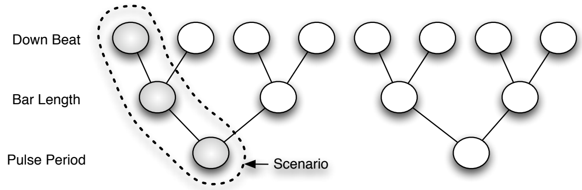

Figure 2. A Chimera data structure with one Scenario highlighted.

4. INTERPRETATION

The Chimera Architecture includes methods for reading, or interpreting, the Chimera data in a variety of ways when passing the scenario values to the generative processes that create improvisational content. The methods of interpreting data correspond to the degree of ambiguity in the generated content with respect to matching the performed input. This structure has been arrived at after careful consideration of theories relating to psychological studies of musical perception and expectation which we will now briefly outline.

4.1. Expectation

The importance of taking account of the dynamic nature of musical expectations when considering musical experience, either analysis of it (theory building) - or simulation of it (algorithmic composition/improvisation) has received attention in the music theory literature for some time [11, 15-17] but has only recently been translated into computational descriptions and rarely been the basis for algorithmic music systems. Meyer suggests that affect in music perception can be largely attributed to the formation and subsequent fulfilment or violation of expectations. His exposition is compelling but imprecise as to the exact nature of musical expectations and to the mechanisms of their formation.

A number of extensions to Meyer's theory have been proposed, which have in common the postulation of at least two separate types of expectations; structural expectations of the type considered by Meyer, and additionally dynamic expectations. Narmour's [16] theory of Implication and Realisation, an extension of Meyer's work, posits two cognitive modes; one of a schematic type, and one of a more innate expectancy type. Bharucha [17] also discriminates between schematic expectations (expectations derived from exposure to a musical culture) and veridical expectations (expectations formed on the basis of knowledge of a particular piece). Huron [18] has recently published an extensive and detailed model of musical expectations that builds further on this work. He argues that there are, in fact, a number of different types of expectations involved in music perception, and that indeed the interplay between these expectations is an important aspect of the affective power of the music. Huron extends Bharucha's categorisation of schematic and veridical expectations, and in particular makes the distinction between schematic and dynamic expectations. Dynamic expectations are constantly learned from the local context. Several authors have suggested that these dynamic expectations may be represented as statistical inferences formed from the immediate past [18-20]. Like Bharucha, Huron argues that the interplay of these expectancies is an integral part of the musical experience.

4.2. Metre as an Expectational Framework

Musical metre is frequently described as the pattern of strong and weak beats in a musical stream. From the point of view of music psychology, metre is understood as a perceptual construct, in contrast to rhythm, which is a phenomenal pattern of accents in the musical surface. Metre is inferred from the surface rhythms, and possesses a kind of perceptual inertia. In other words, once established in the mind, a metrical context tends to persist even when it conflicts with the rhythmic surface, until the conflicts become too great - "Once a clear metrical pattern has been established, the listener renounces it only in the face of strongly contradicting evidence" [11:17]. Jones [21] argues that metre should be construed as an expectational framework for predicting when salient musical events are expected to happen. This description of metre has been widely accepted within the music psychology community [18, 22, 23].

4.3. Ambiguity

Meyer [15] identifies ambiguity as a mechanism by which expectations may be exploited for artistic effect. In this context ambiguity is referring to musical surfaces that create several disparate expectations. The level of ambiguity in the music creates a cycle of tension and release, which forms an important part of the listening experience in Meyer's theory. An ambiguous situation creates tension; the resolution of which is part of the art of composition. "Ambiguity is important because it gives rise to particularly strong tensions and powerful expectations. For the human mind, ever searching for the certainty and control which comes with the ability to envisage and predict, avoids and abhors such doubtful and confused states and expects subsequent clarification" [15:27]. Temperley notes that ambiguity can arise as the result of multiple plausible analyses of the musical surface. "Some moments in music are clearly ambiguous, offering two or perhaps several analyses that all seem plausible and perceptually valid. These two aspects of music - diachronic processing and ambiguity - are essential to musical experience" [24:205].

4.4. Multiple Parallel Analyses

A number of researchers have posited systems of musical analysis that yield several plausible results as models of human musical cognition. Notably, Jackendoff [25:140] proposed the parallel multiple analysis model. This model, which was motivated by models of how humans parse speech, claims that at any one time a human listening to music will keep track of a number of plausible analyses in parallel. In a similar vein, Huron [18] describes the competing concurrent representation theory. Huron goes further to claim that, more than just a model of music cognition, "Competing concurrent representations may be the norm in mental functioning" [18:108].

4.5. Ambiguity in Multiple Parallel Representations

An analysis system that affords multiple interpretations provides a natural mechanism for the generation of ambiguity. In discussing their Generative Theory of Tonal Music (GTTM), Lerdahl & Jackendoff [11:42] observe that their "rules establish not inflexible decisions about structure, but relative preferences among a number of logically possible analyses", and that this gives rise to ambiguity. In saying this Lerdahl & Jackendoff are not explicitly referencing a cognitive model of multiple-parallel-analyses; the GTTM predates Jackendoff's construction of this model, and does not consider real-time cognition processes. Indeed it was considerations of the cognitive constraints involved in resolving the ambiguities of multiple interpretations that led Jackendoff to conclude that the mind must be processing multiple analyses in parallel [25].

Temperley has revisited the preference rule approach to musical analyses in a multiple parallel analyses model: "The preference rule approach [is] well suited to the description of ambiguity. Informally speaking, an ambiguous situation is one in which, on balance, the preference rules do not express a strong preference for one analysis over another ... At any moment, the system has a set of "best-so-far" analyses, the analysis with the higher score being the preferred one. In some cases, there may be a single analysis whose score is far above all others; in other cases, one or more analyses may be roughly equal in score. The latter situation represents synchronic ambiguity" [24:219]. In a similar spirit, Huron [18:109] argues that multiple-parallel-analyses (or competing concurrent representation, as he calls them) must all be generating expectations, and consequently must give rise to the kind of expectational ambiguity that was argued above to play a central role in producing musical affect. This is the model of ambiguity that we have adopted in the Chimera Architecture.

4.6. Interpreting the Chimera

The Chimera Architecture tracks multiple plausible scenarios. How can we use this information to create generative improvisations? The central hypothesis of this paper is that utilising all of the parallel scenarios for generative improvisation can be musically efficacious. Let us however consider other approaches. One obvious possibility is to select the most plausible scenario and generate material which is appropriate to this scenario. This approach is similar to that used by authors who have considered multiple parallel analyses models for musical analysis—the explicit assumption being that despite tracking multiple analyses, at any one time there is only one analysis which is perceived as being 'correct'. For example, Temperley (above) refers to the preferred analysis as the one with the highest score. Similarly the GTTM explicitly insists that only one analysis at a time can be 'heard', as they make clear; "Our hypothesis is that one hears a musical surface in terms of that analysis (or those analyses) that represent the highest degree of overall preference" [11].

A similar hypothesis is widely held in a number of fields of psychology and neuroscience, under a variety of names. The Gestalt Psychologists refer to it as the Figure-Ground dichotomy [26], and in neuroscience it is called the Winner-Takes-All hypothesis [27:494]. The idea is that the mind can only be conscious of a single reality at any one time.

In music perception a number of authors have similar commented on the impossibility of consciously perceiving more than one musical analysis simultaneously: "It is true that we are conscious of only one analysis at a time, or that we can attend to only one analysis at a time. But this leaves open the possibility that other analyses are present unconsciously, inaccessible to attention" [11:141]. London, while discussing music psychology experiments on attending to metre by tapping along to a single stream of a polyrhythm notes that, "These studies ... indicate that while on any given presentation we tend to hear a passage under one and only one metric framework, it is possible to re-construe the same figure or passage under a different meter on another listening occasion. A polyrhythmic pattern may be heard "in three" or "in four," just as metrically malleable patterns may be set in different metric contexts. This should not surprise anyone familiar with the basic tenets of perception, as the need to maintain a single coherent ground seems to be universal ... Thus there is no such thing as polymetre" [23:50].

However, our research is at odds with this view, and we contrastingly suggest that all of the scenarios present in the Chimera may be used to artistic effect. Support for our multiple scenario approach can be found in Huron's [18:109] discussion of his theory of competing concurrent representations. He argues that each of the concurrent representations must all be creating expectations and so must all be contributing to the musical affect. Music psychology experiments (such as those above) that suggest that only one scenario can be consciously attended to at a time, may not be relevant to the acts of listening to, or improvising with, music since these activities do not necessarily involve conscious attention to musical representations.

More precisely, we propose that by interpreting the Chimera in different ways we may achieve different musical results, particularly in regard to the level of ambiguity present in the improvised generative material. For example, utilising a single scenario interpretation in which material is generated as appropriate to only the most plausible scenario should result in a decrease in the overall ambiguity of the ensemble's playing, since highlighting the most prominent scenario is likely to have the effect of further increasing its relative plausibility. Conversely, generating material that is equally appropriate to two scenarios regardless of their relative plausibility should increase the level of ambiguity present in the ensemble. We suggest three different strategies for interpreting the Chimera for the purposes of producing generative accompaniment, revolving around controlling the level of ambiguity present.

(i) Disambiguation - utilise only the most plausible of the scenarios in the Chimera.

(ii) Ambiguation - utilise all of the scenarios with equal weight, regardless of their plausibility.

(iii) Following - utilise the scenarios in the Chimera with weight according the their plausibility.

5. IMPLEMENTATION

We have implemented the Chimera Architecture in C++ on the Mac OS X platform as an Audio Unit. The Audio Unit receives audio input, and performs the analysis into the representation described. The generative improvisational processes (described below) were implemented in the Impromptu environment [28] which also acts as a host for our Audio Unit. Parameters defining the Chimera are stored as Audio Unit parameters, allowing the generative processes to query the Chimera at any time. The software system design is shown in figure 3.

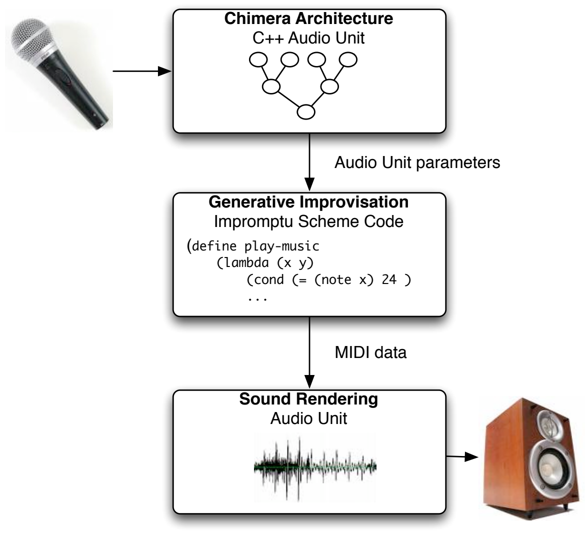

Figure 3. The schematic design of the system implementation.

5.1. Generative Improvisational Processes

At this stage of the research we have implemented only a very simple generative process, designed to clearly demonstrate the musical results of the three interpretative strategies outlined above. For any given metric scenario, we want to create rhythmic material that is appropriate to, or stereotypical of, the scenario. To achieve this we simply play a percussive attack on each beat identified in the scenario, with an accent on the downbeat. The downbeat accent is created by utilising a different timbre (i.e., a different percussive instrument), whilst the other beats have the same timbre, but varying dynamics. The dynamics are chosen to correspond to the metric weights of the beats specified by the scenario.

The dynamics of the attacks on the beats are further modulated by the desired weighting of the scenario in the generated material. So, for example, if we are using the interpretative strategy for Ambiguation discussed above, the pulse streams corresponding to the different scenarios will be equally loud (on average), whereas if we are using the Following interpretative strategy then the pulse stream corresponding to the most plausible scenario will be loudest.

5.2. Experimental Results

We have provided examples of the system improvising rhythmic accompaniment to a short recorded loop of live drums to demonstrate its operation. The audio files discussed may be downloaded from http://explodingart.com/giffordbrown2009/. The original loop (MR10.aif) is a sample rock beat played on a standard drum kit. We perceived it as being in a 4/4 time signature at 110 bpm. The rhythm has regular hi-hats played on the quavers, with snare hits on the backbeat, and a triplet-feel kick-drum pattern. The kick-drum pattern contrasted with the snare and the hi-hat gives rise to metric ambiguity, as the triplet feel is metrically dissonant with the straight four/eight feel of the snare and hi-hat.

The results of the first stage of analysis, the parsing of the musical surface into salience streams for the kick-drum, snare and hi-hat, is pictured below. The spikes in salience correspond to detected onsets.

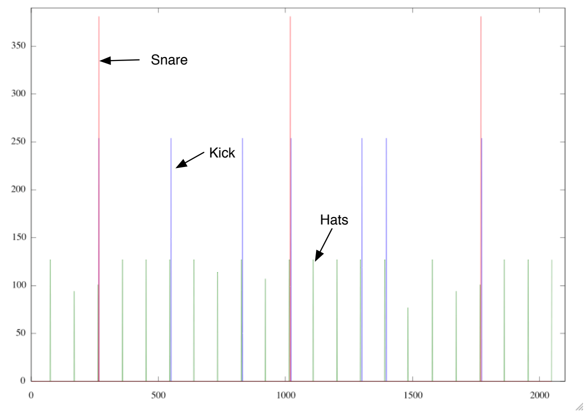

Figure 4. Salience streams for the original live music file MR10.aif

The second stage of analysis, the conversion of the salience streams into metric scenarios, yielded four plausible scenarios.

(a) The most plausible scenario has a pulse period of 0.27 seconds (~ 220 bpm), and 8 beats to the bar.

(b) The second most plausible has the same bar period as the first, but counts only 4 beats of 110 bpm.

(c) The third has 5 beats at 220 bpm

(d) The fourth has 3 beats at 220 bpm.

The Chimera (the collection of these four scenarios together with their relative plausibilities) changes through time, most notably change occurs in the confidence of scenarios with 5 and/or 3 beats to a bar. The top two scenarios (8 and 4 beats to the bar) are stable, whilst the others pop in and out of the Chimera as the system updates their relative plausibilities. An example print-out of the state of the Chimera data is given in Figure 5. In the print-out scenarios are labelled 'context dumps'. The letters S, B and H indicate the pulse period, bar period and number of beats in a bar respectively (periods are measured as a number of analysis windows, where an analysis window is 128 samples @ 44.1khz). Notice that the Chimera briefly entertains the notion of a bar of 10 before pruning it out when its plausibility becomes too low.

```
-------------------- Context Dump ----------------------
Context #0
S = 94 B = 282 H = 3 LDB = 6.79617 confidence = 0.136748
Context #1
S = 188 B = 752 H = 4 LDB = 2.25089 confidence = 0.224367
Context #2
S = 94 B = 751 H = 8 LDB = 2.62708 confidence = 0.357491
Context #3
S = 94 B = 939 H = 10 LDB = 2.50199 confidence = 0.1498
--------------------------------------------------------

*********** Pruning *********
S = 94 B = 939 H = 10 LDB = 2.5025 confidence = 0.0261448

-------------------- Context Dump ----------------------
Context #0
S = 94 B = 282 H = 3 LDB = 6.79757 confidence = 0.174113
Context #1
S = 188 B = 751 H = 4 LDB = 3.174 confidence = 0.353134
Context #2
S = 94 B = 751 H = 8 LDB = 2.62762 confidence = 0.553314
--------------------------------------------------------

-------------------- Context Dump ----------------------
Context #0
S = 94 B = 282 H = 3 LDB = 10.4628 confidence = 0.220562
Context #1
S = 188 B = 751 H = 4 LDB = 3.17417 confidence = 0.302642
Context #2
S = 94 B = 751 H = 8 LDB = 3.67601 confidence = 0.429445
--------------------------------------------------------
```

Figure 5. Data from three updates of the scenario values.

Audio examples from the system generating material utilising the three interpretative strategies of disambiguation, ambiguation, and maintenance are given in MR10_click_disambiguate.aif, MR10_click_ambiguate.aif, and MR10_click_maintain.aif respectively. These examples have been transcribed as common practice notation below. From these examples, the alternative scenarios, especially relating to likely metres, is clear to see. The improvisation generated by this simple reflection of the analysis data indicates that the richness of material (multiple parts and poly metre) that arises from considering several scenarios (the Chimera) provides opportunities for much more interesting material from which generative processes can be derived. What is not so evident in the notated examples in figures 6-8 is that the relative dynamic levels are mapped to the plausibility of each pulse such that the less likely beats are downplayed in the audio mix providing an automatic depth and subtlety otherwise unavailable.

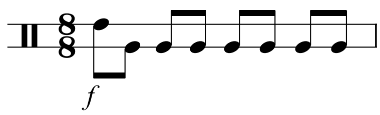

Figure 6. The generated 'disambiguated' music

Gifford, T. and Brown, A. R. 2009. Do Androids Dream of Electric Chimera? Improvise: The Australasian Computer Music Conference, Brisbane, Australia: pp.56-63.

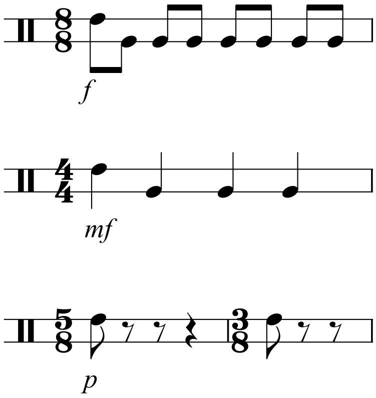

Figure 7. The generated 'following' music

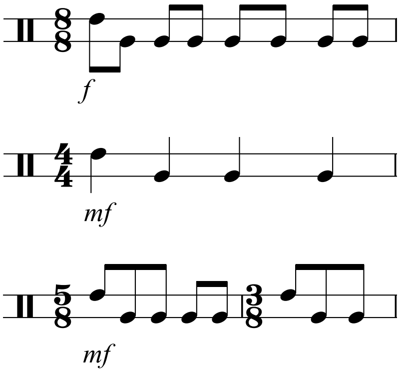

Figure 8. The generated 'ambiguated' music

6. CONCLUSION

In this paper we have reported on the Chimera Architecture and its use in generative machine improvisation that accompanies human performance. The implementation has focused on unpitched rhythmic music presented to the system as an audio signal. From this audio input the system performs a novel metric analysis that locates the pulse duration, tempo, bar length and down beat position, as well as metric weights at each beat of the bar. The system introduces an innovative approach in its maintenance of multiple analytical outcomes, described as scenarios that, with their associated plausibility weightings, we have called a Chimera and which is stored in a binary tree data structure. This approach provides flexibility for the system to interpret the analytical material in a variety of ways in order to inform its generative improvisational output. We have suggested that there is a connection between the interpretation of the Chimera and the level of rhythmic ambiguity in the generated material. When the most likely analytical scenario is used for generation, the music can reinforce the assumed metrical intent of the human musician; we labelled this a disambiguated response. When the scenarios in the Chimera are interpreted as a weighted average to inform the generative process, the musical result tends to maintain or reinforce the existing level of metrical ambiguity. When the scenarios in the Chimera are interpreted as being equally weighted the generated material tends to be more metrically ambiguous than the performed input.

Our in initial experimental results provide confidence in the ability for the Chimera Architecture to accurately track human performance via an audio stream, for the Chimera data representation to maintain several plausible interpretations of the metrical characteristics of the performance, and for different Chimera interpretations to enable a useful variety of appropriate musical responses.

In the future we plan to extend the implementation of the Chimera Architecture to incorporate pitched material that will inform the harmonic and melodic aspects of the systems generated improvisations. We also plan to test more throughly the correlation between human improvisational decisions and generative interpretations of the Chimera.

7. REFERENCES

[1] Bregman, A.S., *Auditory Scene Analysis: The perceptual organization of sound*. 1990, Cambridge, MA: The MIT Press.

[2] Gifford, T. and A.R. Brown. *Polyphonic Listening: Real-time accompaniment of polyphonic audio*. in *Trans: Boundaries / Permeability / Reification. Australasian Computer Music Conference '07*. 2007. Canberra, Australia: ACMA.

[3] Gifford, T. and A.R. Brown. *Stochastic Onset Detection: an approach to detecting percussive attacks in complex audio*. in *Sound : Space - The Australasian Computer Music Conference*. 2008. Sydney: ACMA.

[4] Gifford, T. and A.R. Brown. *The Ambidrum: Automated Rhythmic Improvisation*. in *Medi(t)ations: Australasian Computer Music Conference*. 2006. Adelaide: ACMA.

[5] Zicarelli, D., *M and Jam Factory.* Computer Music Journal, 1987. **11**(4): p. 13-29.

[6] Biles, J.A. *GenJam: A genetic algorithm for generating jazz solos*. in *International Computer Music Conference*. 1994. San Francisco: ICMA.

[7] Rowe, R., *Interactive Music Systems: Machine listening and composing*. 1993, Cambridge, MA: MIT Press.

[8] Collins, N., *Algorithmic Composition Methods for Breakbeat Science*, in *Music Without Walls*. 2001, De Montfort University: Leicester.

[9] Cont, A., S. Dubnov, and G. Assayag, *Anticipatory Model of Musical Style Imitation Using Collaborative and Competitive Reinforcement Learning*, in *Anticipatory Behavior in Adaptive Learn-*

ing Systems. 2007, Springer: Heidelberg. p. 285-306.

[10] Assayag, G., et al., *Omax Brothers: A Dynamic Topology of Agents for Improvization Learning*, in *ACM Workshop on Audio and Music Computing for Multimedia, International Multimedia Conference*. 2006, ACM: Sanata Barbara. p. 125-132.

[11] Lerdahl, F. and R. Jackendoff, *A Generative Theory of Tonal Music*. 1983, Cambridge, MA: MIT Press.

[12] Schenker, H., *Harmony*. 1980, Chicago: University Of Chicago Press.

[13] Cope, D., *Virtual Music: Computer synthesis of musical style*. 2001, Cambridge, MA: MIT Press.

[14] Hawkins, J. and S. Blakeslee, *On Intelligence*. 2004, New York: Times Books.

[15] Meyer, L.B., *Emotion and meaning in music*. 1956, Chicago: The University of Chicago Press.

[16] Narmour, E., *The Analysis and Cognition of Melodic Complexity: The Implication-Realization Model*. 1992, Chicago: University of Chicago Press.

[17] Bharucha, J.J., *MUSCAT: A connectionist model of musical harmony*, in *Machine Models of Music*, S. Schwanauer and D. Levitt, Editors. 1993, MIT Press: Cambridge, MA. p. 497-509.

[18] Huron, D., *Sweet Anticipation: Music and the psychology of expectation*. 2006, Cambridge, MA: The MIT Press.

[19] Temperley, D., *Music and Probability*. 2007, Cambridge, MA: The MIT Press.

[20] Pearse, M. and G. Wiggins, *Expectation in Melody: The Influence of Context and Learning.* Music Perception, 2006. **23**(5): p. 377-405.

[21] Jones, M.R., *Dynamic patterns structure in music: Recent theory and research.* Perception & Psychophysics, 1987. **41**: p. 621-634.

[22] Large, E.W., *Dynamic Representation Of Musical Structure*. 1994, Ohio State University: Ohio.

[23] London, J., *Hearing in Time: Psychological Aspects of Musical Meter*. 2004, Oxford: Oxford University Press.

[24] Temperley, D., *The Cognition of Basic Musical Structures*. 2001, Cambridge, MA: The MIT Press.

[25] Jackendoff, R., *Languages of the mind: Essays on mental representation*. 1992, Cambridge, Mass. ; London: MIT Press. ix, 200p : ill : music ; 24cm.

[26] Boring, E., *Sensation and Perception in the History of Experimental Psychology*. 1942, New York: Appelton-Century.

[27] Kandel, E.R., J.H. Schwartz, and T.M. Jessell, *Principles of neural science*. 2000, New York: McGraw-Hill Professional.

[28] Sorensen, A., *Impromptu: An interactive programming environment for composition and performance*, in *Australasian Computer Music Conference 2005*, A.R. Brown and T. Opie, Editors. 2005, ACMA: Brisbane. p. 149-153.