**Toby Gifford and
Andrew R. Brown**

Queensland University of Technology
Victoria Park Road
Kelvin Grove, 4059
Australia
t.gifford@qut.edu.au
a.brown@qut.edu.au

The Ambidrum: Ambiguous Generative Rhythms

Abstract

This paper outlines a system for the real-time generation of rhythms that target a specified level of ambiguity with respect to a given metrical template. Ambiguity is measured via the statistical correlation of rhythmic variables to the specified metre. Applications are suggested in the area of machine improvisation as a means for the provision of rhythmic complementarity in a musical improvisation involving a human performer and a computational agent.

Introduction

This research forms a component of a larger research agenda into the construction of improvisational algorithms for performance collaborations between human musicians and computational agents. The broad aim is to construct a computational musical agent that displays rudimentary improvisational intelligence. In this paper we report on the development of the rhythmic component of this computational agent. In particular we describe a system, named the *Ambidrum*, which produces generative rhythms that posses a specifiable level of metrical coherence (or conversely, metrical ambiguity), given a pre-determined metre and tempo.

Machine improvisation has been an active area of research for many decades, and includes the work of Dannenberg (1989), Rowe (1993, 2001), Biles (2002), and others. We believe that the work presented here has the potential to be more broadly applicable across styles than previous work given that it is based more firmly in aural cognition and perception theories and relies less on stored databases or fixed musical structures than much of the earlier work.

Recent approaches to implementing agent-based improvisational intelligence are described in (Biles, 2002; Pachet, 2004; Raphael, 2003; Suzuki, 2002; Thom, 2003). In these approaches a statistical model is estimated by analysing a database of examples of a given musical style, and the estimated model is then used to generate novel musical material in real-time. The focus of these studies is primarily the production of melodic improvised lines given a chord progression, or of chordal accompaniment to a human produced melodic line. These systems model rhythmic and pitch elements jointly, and do not involve any musical 'knowledge' other than the database of examples used to train the systems.

This paper examines rhythmic improvisation independently of any pitch considerations. We outline a strategy for approaching machine improvisation by starting with the task of rhythmic complementarity. In particular we focus on the problem of maintaining an appropriate level of metrical ambiguity and show how this can be achieved with an algorithmic process based on statistical theories of expectation and coherence. Finally we discuss how these theories can be applied to real-time interaction with a human performer and discuss various potential mappings for interactions between human and machine in an improvisatory setting.

Rhythmic Coherence

An important facet of ensemble improvisation is a sense of complementarity: members of the ensemble strive to complement each other's musical output. The question arises as to how to assess the success of the resultant musical union. One criteria identified by numerous practitioners of improvisation (Borgo, 2004; Levine, 1995:191; Nunn, 1998:38) is striking a balance between novelty and coherence, as emphasised by Kivy.

> 'good' music ... must cut a path midway between the expected and the unexpected ... if a work's musical events are all completely unsurprising ... then the music will fulfil all of the listener's expectations, never be surprising – in a word, will be boring. On the other hand, if musical events are all surprising ... the musical work will be, in effect, unintelligible (2002:74).

The *Ambidrum* aims to address this aspect of complementarity by monitoring the level of rhythmic coherence of an improvisation in real-time, and generating complementary material that maintains an appropriate balance between the expected and the unexpected.

Expectation and Ambiguity

In designing the *Ambidrum* we have adopted the theory of musical expectations proposed by Leonard Meyer (1956) regarding expectations and affect. In this theory when listening to music the listener is constantly forming expectations of what is to come, and the fulfilment or frustration of these expectations stimulate an affective response in the listener. This theory is not without controversy (Jackendoff, 1992; Kivy, 2002) but is neverthe-

less widely regarded (Borgo, 2004; Dubnov *et al.*, 2006; Kivy, 2002; Pressing, 1998).

An important aspect of this theory is the role of ambiguity in musical affect.

> 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 (Meyer, 1956)

Taking Meyer's theories into account we have, as a first step, developed a measure of metric coherence that we hypothesise to have a direct relationship with ambiguity. The metric coherence tracks the degree to which rhythms imply a sense of particular metre.

Metre and Rhythm

In the perception of rhythm, at least in post-Renaissance Western tonal music, there appears to be a strong tendency to interpret rhythmic material in terms of an underlying metre (Lerdahl & Jackendoff, 1983:17; Toiviainen & Snyder, 2000)

> Metrical structure provides listeners with a temporal framework upon which to build expectations for events (Large & Kolen, 1994)

Meyer (1960) defines metre as "the measurement of the number of pulses between more or less regularly recurring accents" whilst Yeston (1976) describes metre as "an outgrowth of the interaction of two levels – two differently-rated strata, the faster of which provides the elements and the slower of which groups them". Lerdahl and Jackendoff (1983) in their Generative Theory of Tonal Music (GTTM) propose a representation of metre which reflects a hierarchy of time-scales, each of which is an integral multiple of the fastest time scale (which defines the level of quantisation). In this representation, each quanta is assigned a strength according to the number of levels in the hierarchy that contain this quanta.

Notationally, the GTTM represents metre on a written score by drawing beneath each quanta a number of dots proportional to the strength of the quanta. We will utilise a similar representation, but with letters in place of dots, with alphabetic ordering of the letters indicating the relative strength of the quanta. For example, we will write the metre indicated by the 6/8 time signature as [*a c c b c c*] where each letter denotes a quanta and *a* indicates the strongest pulse, *c* the weakest.

Metre as we are describing it is not a directly observable property of a musical stream; it is a perceptual construct in the mind of the listener that is formed in response to the actual rhythmic patterns in the music.

Rhythm in this context refers to the manner in which accented beats are grouped with unaccented beats. Here by an accent we mean "a stimulus which is *marked for consciousness* in some way" (Cooper & Meyer, 1960), or what GTTM refers to as a phenomenal accent.

A number of authors (Arom, 1991; Cooper & Meyer, 1960; Epstein, 1995; Lerdahl & Jackendoff, 1983; Temperley, 2001) identify three primary means of phenomenal accent, which we will refer to as rhythmic markers:

(i)   Stress
(ii) Duration
(iii) Timbre

Where these markers are working in concert to imply a metre, we say that the music is metrically coherent, whilst if these markers are inducing contradictory senses of metre, we say the music is metrically ambiguous.

> Phenomenal accent functions as a perceptual input to metrical accent – that is, the moments of musical stress in the raw signal serve as "cues" from which the listener attempts to extrapolate a regular pattern of metrical accents. If there is little regularity to these cues, or if they conflict, the sense of metrical accent becomes attenuated or ambiguous. If on the other hand the cues are regular and mutually supporting, the sense of metrical accent becomes definite (Lerdahl & Jackendoff, 1983)

We utilise these markers as rhythmic attributes within the *Ambidrum* system, and correlate their values as a measure of metric coherence.

Coherence and Ambiguity

The *Ambidrum* ultimately uses any measurement of existing and proposed material in order to generate a new rhythmic pattern. As a step toward rhythmic complementarity the *Ambidrum* searches for a new rhythm that has a specified degree of coherence with, or similarity to, the currently specified metre. A metre is specified as a series of quanta strengths or emphases, as described in more detail later. The *Ambidrum* plays, as its next pattern, the rhythm that most closely matches a desired degree of coherence. We define coherence as a measure of the correlation between the strength of the rhythmic markers at each quanta. At one end of the scale a completely coherent pattern will match the underlying metre exactly, at the other end of the scale an incoherent pattern will have the inverse quanta values to those specified in the metre. As it turns out, rhythms at these two extremes of coherence provide a similar metrical stability and rhythms with a moderate degree of coherence are the least likely to imply a sense of metre. Therefore, we say that the rhythms with moderate coherence values are highly ambiguous with respect to metre and those with either a high or low coherence measure are less ambiguous. This relationship is shown in figure 1.
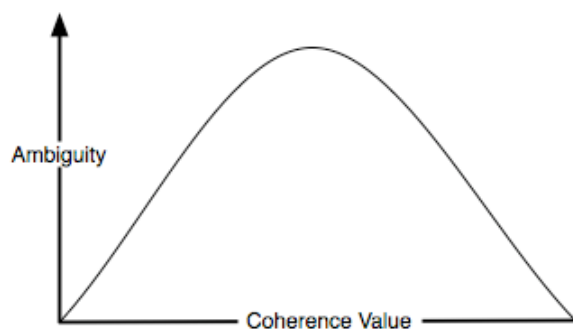
Figure 1. The relationship between metrical ambiguity and coherence

This relationship presents an interesting musicological or psychoacoustic relationship between statistical correlation and musical ambiguity and, again, reinforces the central insights of Meyer with regard to balance between forces at play in this computational generation of musical rhythms.

Another way of understanding the relationship between coherence and ambiguity in this context is to imagine that coherence and incoherence are magnetic forces attracting the rhythm into a pattern that moulds itself onto the specified metre. Ambiguity is introduced as these two forces pulling on the rhythm distort it. When the two forces are equally strong the ambiguity is highest because the rhythm bears least resemblance to the metre template. As the rhythm approaches one of the extremes of coherence it becomes less ambiguous by fitting closer to the metre or its inverse.

Coherence Level

The *Ambidrum* is a real-time system that produces a rhythm one note at a time by analysing the coherence of its previous output and taking action to maintain the coherence of its output at a given target level. To this end it constructs a measure of rhythmic coherence, which we refer to as the coherence level.

The inputs to the *Ambidrum* process are a tempo, a metre, and a matrix of target coherence levels. The metre is defined as being a series of stress levels of quanta in a bar. For example 4/4 time could be represented by the series [*a c b c*] where each quanta is a quarter-note and *a* represents the strongest value and *c* the weakest value. At a higher quantisation level the same time signature could be represented with more quanta by [*a d d d c d d d b d d d c d d d*] providing sixteenth-note resolution. The *Ambidrum* takes the quantisation as being effectively determined by the quanta-length of the metre series relative to the time signature.

The *Ambidrum* considers three rhythmic variables: velocity, timbre and duration. When the process is running it generates MIDI messages which are sent to a drum machine. These variables are mapped to the velocity, pitch and duration parameters for a MIDI note-on/note-off pair. In the context of a drum machine the pitch parameter of the MIDI message is not directly related to frequency but rather to timbre, determining which drum sound is triggered.

At every quanta a value is set for these rhythmic variables. The variables take on discrete values selected from the range determined by the metre. So, for example, for the metre defined by [*a d d d c d d d b d d d c d d d*] the rhythmic variables may take the values *a*, *b*, *c* and *d*. Here *a* represents a strong rhythmic event, diminishing down to *d* which represents a weak rhythmic event.

For the variable of velocity, a strong value is mapped to a high velocity. For duration, a long duration is taken to be stronger than a short duration. For timbre (which really amounts to choice of drum) it is not always clear which timbres are stronger or weaker. In the case of a classic drum machine kit with kick-drum, snare-drum, high-hat and tom, probably the most obvious assignment would be

Timbre	Value
kick-drum	a
tom	b
snare-drum	c
high-hat	d

The generated rhythm is described by a series of values for each of the rhythmic variables. The *Ambidrum* selects values for these variables in an attempt to create a rhythm that is suitably coherent, as determined by the input target coherence matrix. Following the above discussion, the process considers the rhythmic ambiguities created by latent metrical dissonances induced by disparate metric suggestions of the different rhythmic variables.

A metrically unambiguous (e.g., completely coherent) rhythm would have all of the rhythmic variables matching the metre, as shown in figure 2.

```
metre     [a d d d c d d d b d d d c d d d]
velocity  [a d d d c d d d b d d d c d d d]
timbre    [a d d d c d d d b d d d c d d d]
duration  [a d d d c d d d b d d d c d d d]
```

Figure 2. A metrically unambiguous rhythm matrix.

However, let us consider a more ambiguous (less coherent) rhythm, shown in figure 3, where the rhythmic variables are not perfectly aligned to the metre, nor to each other.

```
metre     [a d d d c d d d b d d d c d d d]
velocity  [a c a c d b b d a c d d c d a d]
timbre    [b d d c b d d c b d a d c b d b]
duration  [c c d d c c c d a a b d c d d a]
```

Figure 3. A metrically ambiguous rhythm matrix.

The *Ambidrum* uses a measure of how closely aligned these sequences are to each other as a proxy for the coherence of the rhythm. The particular measure employed is a correlation statistic for each pair of these sequences. To calculate the correlation we assign each of the possible variable values a numeric value centred around zero. In the above example this would translate to mapping

a ⇒ 2
b ⇒ 1
c ⇒ -1
d ⇒ -2

Then considering each series of variable values as a vector we calculate the correlations via the formula

$$corr(x, y) = \frac{x^T y}{\sqrt{(x^T x)(y^T y)}}$$

for each pair of variables.

The correlation lies between 1 and -1. If the variables have identical values then their correlation will be equal to 1. When a pair of variables are inverse to each other, their correlation will be -1. When two variables are unrelated to each other (or orthogonal) their correlation will be zero.

The collection of pairwise correlations of the rhythmic variables to themselves and to the metre forms a correlation matrix. For example, the variable values in figure 3 yield the correlation matrix shown in figure 4.

	Metre	Velocity	Timbre	Duration
Metre	1	0.44	0.43	0.24
Velocity	0.44	1	-0.32	0
Timbre	0.43	-0.32	1	0.34
Duration	0.24	0	0.34	1

Figure 4. *A calculated correlation matrix.*

The *Ambidrum* considers its output each quanta based on a sliding window of its own historical output - generally a fixed number of bars. So, for example, using a metre of [*a c b c*] the process might find itself in the following situation depicted in figure 5.

```
metre       [a c b c] [ a …
velocity    [b c c b] [ ? …
timbre      [a c b b] [ ? …
duration    [c c b a] [ ? …
```

Figure 5. *Choosing the next note.*

The question marks signify that the *Ambidrum* must choose a value for each of these variables for the next quanta. The choice is made so as to have the resulting sequences as close as possible to the target coherence levels, determined by a target correlation matrix, which is an input to the generative process.

	Metre	Velocity	Timbre	Duration
Metre	1	1	1	1
Velocity	1	1	1	1
Timbre	1	1	1	1
Duration	1	1	1	1

Figure 6. *A coherent target correlation matrix*

For example using the target correlation matrix shown in figure 6 the *Ambidrum* would choose the velocity, timbre and duration [**v t d**] of the next note so as to make the series

```
metre       . [c b c a]
velocity    . [c c b v]
timbre      . [c b b t]
duration    . [c b a d]
```

have all pairwise correlations as close to 1 as possible. In this case the choice would be **v** = *a*, **t** = *a*, **d** = *a*.

The target correlation matrix in figure 6 is the completely coherent (totally unambiguous) target matrix. Any other choice of target matrix is possible and would result in different choices for the next note generated.

A useful metaphor for the coherence level is a VU metre, that constantly monitors the level of some property of an audio stream in real-time.

Figure 7. *A coherence level metre.*

The *Ambidrum* monitors the coherence of its generated rhythm and attempts to maintain it at a target level. This target level is externally controlled, and may be changed during the course of performance. In fact, the *Ambidrum* essentially monitors a coherence level bridge – comprising of a coherence level metre for each of the pairwise correlations of the rhythmic variables and the metre. The target correlations may be set independently, and comprise external control parameters that will affect the operation of the *Ambidrum* in real-time. The mute button on the picture in figure 7 alludes to the option of turning off tracking for any of the variable pairs.

Example Results

Rhythms generated by the system quickly locate a pattern that closely matches the target coherence values and then fall into a stable cycle which results in repeating that pattern indefinitely.

As an example we show the resulting patterns produced for a few target coherences using the metre [*a d d d c d d d b d d d c d d d*]. The completely coherent target matrix reproduces the metre exactly

Stable cycle for all target correlations = 1
```
metre       [a d d d c d d d b d d d c d d d]
velocity    [a d d d c d d d b d d d c d d d]
timbre      [a d d d c d d d b d d d c d d d]
duration    [a d d d c d d d b d d d c d d d]
```

However when we allow the rhythmic variables to be independent by setting the target correlations to zero we obtain a rhythm that is more ambiguous

Stable cycle for all target correlations = 0

```
metre       [a d d d c d d d b d d d c d d d]
velocity    [c a a a d b c d a b d d c d d d]
timbre      [d d d d a c a d c a d d a d d d]
duration    [c a a a d b c d a b d d c d d d]
```

Setting the target correlations to -1 results in

Stable cycle for all target correlations = -1

```
metre       [a d d d c d d d b d d d c d d d]
velocity    [d a a a c a a a d a a a c a a a]
timbre      [d d d d c d d d c d d d c d d d]
duration    [d a a a c a a a d a a a c a a a]
```

Target Automation

To create variation, and interest, in the generated rhythm pattern the target coherence values can be continuously adjusted. A simple way to do this is to modulate the target values by some simple function, for example a low frequency sine wave or selection of a random value.

Automating the target value by small degrees produces subtle and interesting variations that can sound almost evolutionary in nature, frequent large variations tend to produce unstable rhythmic behavior, while infrequent shifts from one value to another introduce sudden changes followed by periods of rhythmic stability.

The automation of the target cohesion value is an effective method for controlling the rate of change and the general interest of the generated rhythm patterns. However, the modulating functions usually become tiresome after some extended listening due to their lack of large-scale direction. It is more effective, and closer to the intention of this research, to have the rhythmic coherence controlled by a human performer.

Source Following

While it would be easy to have a performer directly control the coherence level via a dial or slider, we can utilise the existing coherence measuring techniques to follow a human rhythmic performance in real-time. This approach enables improvisation by the machine in direct response to the performance of the human, and is elegant in that the same rhythm coherence technique is used for both the performance tracking and the algorithmic generation.

Given that the metre and tempo are specified in advance, sections of the human performance can be captured and their coherence value calculated. These values can be used to adjust the machine's coherence value and thus the generated rhythms. The mapping between human and machine coherence values is a matter of choice depending upon the desired musical outcome. Two obvious mappings include a) that the machine use the same coherence values as the performer which results in the reinforcement of the coherence or ambiguity dictated by the performer, or b) that the machine use an inverse coherence mapping such that as the performer played less metrically obvious rhythms the machine would tighten-up and play quite 'straight' or conversely as the human

played regular metrical patterns the computer would provide greater rhythmical interest and freedom. This latter scenario demonstrates a type of rhythmic complementarity in the *Ambidrum's* behaviour, in the sense that it is acting so as to attain a musical goal for the ensemble as a whole. Further scaling and offsetting of the coherence mappings could increase the range of interactions and the adjustment of the mappings over time would provide even greater interest and variety.

Conclusion

We have outlined a method for the real-time creation of rhythmic material with a specified level of metrical ambiguity, and suggested applications to rhythmic improvisation between a human performer and a computational agent. This method has been implemented as the *Ambidrum* system in the Impromptu environment (Sorensen, 2005). At the current stage of this research a number of assumptions need to be maintained about the improvisation, in particular the metre and tempo are assumed to be given, but within these constraints the *Ambidrum* is a robust interactive rhythmic improvising system. In future research on the *Ambidrum* system we plan to utilise beat induction techniques to remove the need for the tempo and metre assumptions and will also examine control structures for larger scale organisation of musical structure so that the evolution of the improvisation is not solely controlled by the human performer.

References

Arom, S. 1991. *African polyphony and polyrhythm*. Cambridge: Cambridge University Press.

Biles, J. 2002. "Genjam: Evolutionary computation gets a gig" Paper presented at the *3rd Conference on Information Technology Curriculum*. Rochester.

Borgo, D. 2004, April 2004. "Sync or swarm: Group dynamics in musical free improvisation" Paper presented at the *Conference of Interdisciplinary Musicology*. Graz, Austria.

Cooper, G. and Meyer, L. 1960. *The rhythmic structure of music*. Chicago: University of Chicago Press.

Dannenberg, R. 1989. "Real-time scheduling and computer accompaniment" *Current directions in computer music research*. ed, M. Matthews. Cambridge, MA: MIT Press.

Dubnov, S. McAdams, S. and Reynolds, R. 2006. "Structural and affective aspects of music from statistical audio signal analysis". *Journal of the American Society for Information Science and Technology*.

Epstein, D. 1995. *Shaping time: Music, the brain, and performance*. New York: Schirmer Books.

Jackendoff, R. 1992. *Languages of the mind*. Cambridge, MASS: MIT Press.

Kivy, P. 2002. *Introduction to a philosophy of music*. Oxford: Oxford University Press.

Large, E. and Kolen, J. 1994. Resonance and the perception of musical meter. *Connection Science. 6,* 1.

Lerdahl, F. and Jackendoff, R. 1983. *A generative theory of tonal music*. Cambridge, Massachusetts: MIT Press.

Levine, M. 1995. *The jazz theory book.*

Meyer, L. 1956. *Emotion and meaning in music.* Chicago: University of Chicago Press.

Nunn, T. 1998. *Wisdom of the impulse: On the nature of musical free improvisation.* San Fransisco: Published by Author.

Pachet, F. ed. 2004. *On the design of a musical flow machine*: IOS.

Pressing, J. 1998. "Psychological constraints on improvisational expertise and communication" In *the course of performance.* ed, B. Nettl. Chicago: University of Chicago Press.

Raphael, C. 2003. *Orchestra in a box: A system for real-time musical accompaniment.* IJCAI.

Rowe, R. 1993. *Interactive music systems*: MIT Press.

Rowe, R. 2001. *Machine musicianship*: MIT Press.

Sorensen, A. 2005. "Impromptu: An interactive programming environment for composition and performance" Paper presented at the *Australasian Computer Music Conference.* QUT, Brisbane.

Suzuki, K. 2002. "*Machine listening for autonomous musical performance systems*" Paper presented at the *International Computer Music Conference.* Gothenburg.

Temperley, D. 2001. *The cognition of basic musical structures.* Cambridge, Massachusetts: MIT Press.

Thom, B. 2003. "Interactive improvisational music companionship: A user-modelling approach" *The User Modelling and User-Adapted Interaction Journal.*

Toiviainen, P. and Snyder, J. 2000. "The time-course of pulse sensation: Dynamics of beat induction" Paper presented at the *"Sixth International Conference on Music Perception and Cognition".* Keele University.

Yeston, M. 1976. *The stratification of musical rhythm.* New Haven: Yale University Press.