Mixing Art with Science; Conceptual Framework for Applying Melodic Soundtracks to Visual Media

Andrew Blake & Catherine Grundy
University of Brighton, CMIS, Lewes Road, Brighton BN2 4GJ
a.l.blake@bton.ac.uk & c.grundy@bton.ac.uk

Keywords
Aesthetics, Melody, Grammar, Syntax, Implication, Realisation, Closure

Abstract
Computer science awards are evolving and with it an increasing need to acknowledge aesthetics along with the science. To do this emphasis for measuring student success also needs to evolve; focusing as much on the artefacts they create as the processes for their creation.

Introduction
In recent years applications for computing degrees have become unstable. An Institution enjoying consistent numbers of computing students' year on year is becoming an exception rather than the rule (UCAS, 2006). Consequently, for a computing department wanting to maintain or increase student numbers they continually evolve their diet of awards. This evolution typically involves broadening their definition of computing; describing a context in which computing is applied, for example BSc Computer Science (Games). If the right context is identified a wider range of applicants can be reached.

For the University of Brighton's computing Department, this evolution has broadened its students' expectations. Many of its undergraduate awards now offer modules providing access to industry standard / context dependent tools. These include 2D, 3D, video and sound. These tools have only been available for the last two years and already student focus is as much about engagement with the artefact as it is scientific rigor and process for its creation.

Melody
Computing students regularly submit coursework consisting of 3D animations or digital video. In an attempt to make the visual content more engaging a melodic soundtrack is added. The choice of soundtrack is often arbitrary with little attention paid to its melodic content. Able to articulate the aesthetics of the visual content they fail to do so with melody. This often results with a melody contrasting visual content.

To compliment visual content with melodic soundtracks many computing students need a language to articulate the aesthetics of melody. Computing requires the use of programming languages; syntax and grammars. It therefore seems reasonable to afford computing students, striving to employ melodic soundtracks successfully, a conceptual framework affording a language or grammar to facilitate this process.

Conceptual Framework
There is a great deal of research concerning sound from the field of computing. However, little attention is paid to melody. One area of research that does consider melody is 'Content Based Music Retrieval' (Yang, C. 2003). Its primary goal is to reliably copyright digital music for secure distribution over a network. A number of computing system have been developed facilitating this processes. One notable example is 'Query by Humming' (QBH). QBH considers melody. It allows

a person to retrieve music from a computer system by humming the melody of the music. The hummed melody is transcribed into pitches. The computer system stores digital music with an associated melodic transcription. The pattern of the hummed melodic transcription is compared with those stored in the system. The system attempts to find a match (Zhu, Y. 2003). This and similar research is ultimately concerned with matching patterns in melody but no information regarding the aesthetics of melody.

It seemed reasonable that music theorists would. Yet, few have addressed this issue underpinned with accessible or visible conceptual frameworks (Royal, M. 1995). Eugene Narmour, Professor at the University of Pennsylvania, addresses this problem. Author of "The Analysis and Cognition of Basic Melodic Structures: The Implication-Realization Model" (Narmour, E. 1999), describes a bottom up approach that parses melody from note to note based of the raw parameters of musical sound. With strong cognitive and perceptual leanings it provides a sophisticated symbology to analyse melody with the view to describe how it is perceived. To this end it is concerned with tracking a listener's changing expectancies over time, and the extent to which those expectancies are realised or denied. The theory measures listeners' ongoing levels of surprise as an aesthetic response to music as well as listeners' ongoing perceived structural closure (Royal, M. 1995).

Grammar
To apply the symbology requires a good understanding of music theory. Many computer science students do not poses this skill. Therefore, to facilitate its application, it has been abstracted into a grammar. Typically, the grammar consists of syntax, parameters and rules for application. To present a grammar reflecting the whole of the symbology is beyond the scope of this document. Therefore, a limited extract of the grammar is provided to facilitate the analysis that follows later. In order to help understand the grammar the premise of Narmour's theory is outlined first.

Theory
Similar to the Gestalt principles of good continuation, similarity and proximity Narmour's theory proposes that any two successive pitches (one melodic interval) imply a third pitch (second interval). In its simplest form melodies can be divided into fundamental building blocks of three pitches. Pitches one and two of each block representing the antecedent interval of "implication" and pitches two and three representing the consequent interval of "realisation". Whether the antecedent interval is realised, partially realised or denied depends on the size and registral direction of the consequent interval. Interval size and registral direction is core to the theory but Narmour also observes that a listener's expectations are not purely determined by implications and realisations in the pitch domain (Royal, M. 1995).

Dynamic accent and note duration is also important in ascertaining a listener's expectations. Dynamic accent of more than three pitches infers grouping. Groups are either combinations or chains. These definitions are beyond the scope of this document however they both exhibit an important feature. The consequent interval of the first block of pitches forms the antecedent interval of the second block of pitches. Note duration is said to exhibit three fundamental relationships; additive, countercumulative and cumulative. Additive describes repetition of the same note duration partially realising the antecedent interval. Countercumulative describes a long note duration tending towards short note duration also partially realising the antecedent interval. Cumulative describes a short note duration tending towards long note duration and is said to suppress realisation of the antecedent interval. If the second note is greater than or equal to 1.5 times the duration of the first note it is considered cumulative. If a rest follows a consequent interval, what ever is implied by the antecedent interval is said to be suppressed. Consequently durational cumulation or the use of rests can cause individual pitches or pairs of pitches to be isolated and suppress realisation (Royal, M. 1995).

Narmour's theory pays little attention to musical style, culture or a listener's previous musical experiences (Royal, M. 1995). Obviously, these are very important in defining what music we listen to, the context in which it is heard and who we share the experience with. Further, profoundly influencing what we like and dislike (Frith, S. *et al*, 2000) However, it tells us little about how a listener might perceive music and melody therein. Namrour's approach is bottom-up parsing melodies from note to note employing the raw parameters of musical sound (Royal, M. 1995).

Syntax

Syntax	*Parameter*	*Rule*	*Response*
P	Process	Antecedent interval is the same size and registral direction as the consequent interval	Realisation
D	Duplication	Three repeated pitches	Realisation
R	Reversal	Large antecedent interval followed by small consequent interval with a change in registral direction	Realisation
IP	Intervallic Process	Antecedent interval followed by the same consequent interval with a change in registral direction	Partial Realisation
VP	Registral Process	Antecedent interval followed by a different consequent interval with the same registral direction	Partial Realisation
IR	Intervallic Reversal	Large Antecedent interval followed by a small consequent interval but no change in registral direction	Partial Realisation
VR	Registral Reversal	Antecedent interval smaller than the consequent interval with a change in registral direction	Partial Realisation
()	" " "	The inverse of the above denoted with brackets	Realisation Denied

Table 1

Table 1, above, identifies key syntax, their parameter, rules representing the foundation of the grammar and the listeners' response. The symbols are used to express a melody in terms of interval size and registral direction. Interval size is described as small or large. Registral direction is either up, down or lateral.

Realisation implies the listener's melodic expectations have been perceived and structural closure occurs. Partial realisation implies the listener's melodic expectations have been partly perceived and no structural closure occurs. Realisation denied implies the listener's melodic expectations have not been perceived and structural closure is denied. Analysis of melody using these symbols alone would not provide any information regarding implication and realisation outside the pitch domain.

Coursework

The following screen shots are from a student's final major project submitted in June 2006. The project's aim is to demonstrate how biped animation can be enhanced, using a physics engine, rendering believable animation. Kinematics is used to rig the biped determining how the biped will animate when forces are applied. This is used in conjunction with keyframe animation.

Figure 1

Figure 2

Figure 3

Figure 1, 2, and 3 illustrate the opening sequences leading up to the biped animation. The camera fades in, moves around to the front of the complex and tends towards the biped.

Figure 4

Figure 5

Figure 6

Figure 4, 5, and 6 illustrate the biped walking towards the camera, receives a massive force (gun shot) to the head, begin to fall and lose the gun.

Figure 7

Figure 8

Figure 9

Figure 7, 8, and 9 illustrate the biped crash against the wall, slide down the wall and bounce away. At this point in time the biped is falling with considerable force.

Figure 10

Figure 11

Figure 12

Figure 10, 11 and 12 illustrate force still pushing the biped forward, slump backwards, fall to the ground and come to rest. The force of the gun shot has finally dissipated the biped.

Figure 13

Figure 14

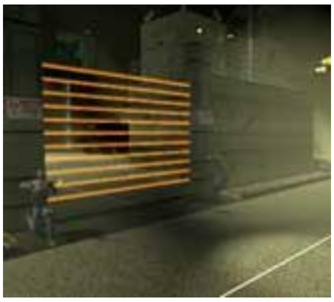

Figure 15

Figure 13, 14, and 15 illustrate two guards reacting to the gun shot. Figure 14 shows the left hand guard turn their head as the right hand guard receives a second gun shot and fall to the ground. The remaining guard and gun posts open fire in response.

Analysis
The animated sequence is accompanied by a sound track merging a simple orchestral score and sound effects. It is the score that is of interest. The first two bars of the score encompasses figure 1 to 4. It consists of 6 beats per bar changing from E to F on the first beat of each bar. Figure 5 to 15 encompasses bar three onwards. Each bar consists of 4 beats per bar changing from B to C on the first beat of each bar. The following analysis considers the first 6 bars of the score. Musical parameters include interval size, registral direction and dynamic accent. Symbols A consider interval size and direction only. Symbols B consider dynamic accent as well as interval size and direction.

```
Analysis:                VPVPIPIP

Symbols B:                     ______IR__________IP_____
                               |           |           |
Symbols A:               ______VP__________IP_____
                         |           |           |
Melody (register):       E(4)  F(4)  B(5)  C(5)  B(5)  C(5)

Bar:                     1     2     3     4     5     6
```

Evaluation
The melody is only partially realised by the listener and no structural closer. Both the sequence of notes and dynamic accent reinforce partial realisation (VP, IR and IP) moving between the pitch domain and dynamic accent. Together they reinforce a listener's expectation and therefore anticipation that something significant is to occur. The key interaction between the audio visual content is tension. The visual content is significantly dramatised when the biped is shot reinforcing visual closure, figure 5, while the melody is only partially realised with no melodic structural closure.

The analysis of melody can focus on any of its devices for example pitch domain, note duration or a combination. However, analytical focus is likely to be influenced by what the analyst deems as significant melodic devices. The above analysis ignored note duration as there was only one change in duration between the second and third bar. It is countercumulative (longer to shorter) reinforcing partial realisation. Note duration might be considered important with melodies utilising many notes within a bar for example an arpeggio.

Further Work
The above analysis is applied to a simple melody. Further examples need to be generated representing complex melodies. These should include realisation and denial of melody. In order for computing students to perform the analysis requires a methodology. This should include a definition and order of processes illustrating how to apply and evaluate the symbology.

References

Frith, S. *et al* "On Record", T J International, UK

Narmour, E. "The Analysis and Cognition of Basic Melodic Structures: The Implication-Realization Model", 1999, University of Chicago Press, USA

Royal, M.S. "Review of 'Analysis and Cognition of Basic melodic Structures and The Analysis and Cognition of Melodic Complexity by Eugene Narmour", 1995, University of Western Ontario, Canada

Yang, C. "Peer-to-Peer Architecture for Content-Based Music Retrieval on Acoustic Data", 2003, Stanford University, USA

Zhu, Y. *et al* "Query by Humming – In Action with its Technology Revealed", 2003, New York University, USA