Computational Parsing of Melody (CPM): Interface Enhancing the Creative Process during the Production of Music

Andrew Blake and Cathy Grundy
University of Westminster
Cavendish School of Computer Science
115 New Cavendish Street
London W1W 6AL
blakea@wmin.ac.uk
grundyc1@wmin.ac.uk

Abstract *CPM is a practical and research driven project with an aim of designing and implementing a software system that can analyse the aesthetic content of a musical composition, in-particular its melody. The analysis is based on the way in which we cognitively process and perceive melody. It is expected the deliverables from CPM will have wide reaching implications within specific music contexts .For example, during the creation and production of musical compositions CPM is envisaged facilitating discourse between musicians, engineers and producers making value judgements about their music based on its aesthetics.*

Keywords Aesthetics, Melody, Cognition, Perception, Implication – Realisation

1. Introduction

This paper identifies two significant problems with discourse concerning the aesthetics of music between musicians, engineers and producers. It proposes the Sequencer, a software tool currently employed by professionals working within the music industry, as a potential solution to their problem. To support this view some simple but fundamental functions of the Sequencer are explained. Its shortcomings are also discussed and comparisons made between the Object-Orientated model used by software companies to develop Sequencer applications and Narmour's Implication-Realisation model [5], the theory underpinning CPM. One particular goal of CPM is to develop a computer interface that graphically represents aesthetical content of music from which value judgements can be made. The aim of this paper is to identify possible constraints for the design of such an interface and invite feedback regarding suitable interface design techniques.

1.1 Aesthetics

Simon Frith, in On Record (2000, p.180) [3], writes, "Academic researchers have failed to look in detail at how musical decisions are made, why sounds and images emerge the way they do". He goes on to explain the importance of musicians, engineers and producers adopting a musical discourse in order to make "value judgements" and "aesthetic decisions" about the music they create. He also adds, the audience or listeners' expectation of music is pivotal to their debate.

There are two fundamental problems with discourse between musicians, engineers and producers. Firstly, participants of the discussion often do not share a common perspective from which they can coherently discuss aesthetics of music. A musician is typically concerned with intonation between notes and the music they create while engineers and producers are often concerned with the clarity between timbre and the tones produced within the music. Secondly, it is evident the audience or listener are invariably absent from the discussion. Consequently, they have little opportunity to participate in debates that primarily occur to determine their expectations of music.

2. The Sequencer

There are a number of tools employed by professional musicians, engineers and producers to enhance the creative process and speed up production. These tools are often software applications that process audio data using a computer or specifically designed pieces of hardware with dedicated processors. Arguably one of the most widely used of these applications employed by professionals in the industry, and central to the recording, editing and production processes, is the Sequencer.

Sequencers support the recording and processing of multiple tracks of audio data. This data typically represents part of, or an entire musical composition. Sequencers provide an interface to this composition through an Arrangement Window (see Fig. 1), an abstract graphical representations of all recorded data. The Arrangement Window typically represents two types of audio data in the form of MIDI and Wave files. Tracks numbered 1 to 8 (on the left hand side of illustration) represent Wave files while tracks 9 to 14 represent MIDI files. The arrangement of each track within the composition has two distinct linear properties namely a start point and end point. For example, track 1, illustrated below, starts on the 3rd beat of bar 2 and finishes on the 1st beat of Bar 6. Track 2 starts on the 3rd beat of bar 1 and finishes of the 1st beat of bar 6. Essentially, the start and end points of a track determine when the track is audible within the composition.
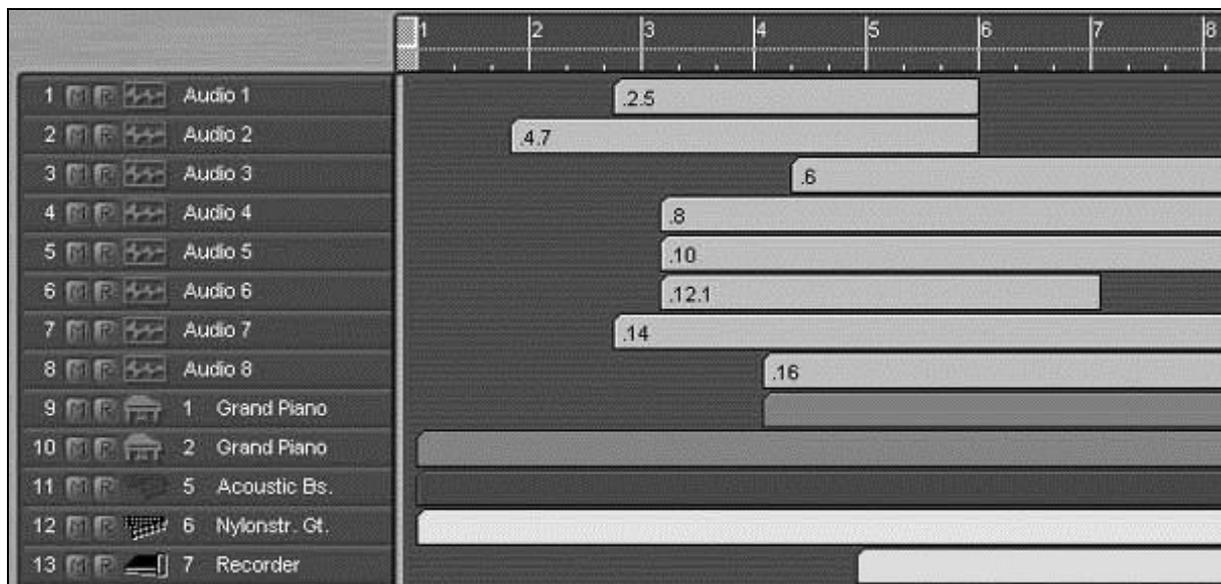

Figure 1. Arrangement Window

2.1 Editing Music

Sequencers afford great flexibility during the recording and production of music. Unlike analogue recording, digital formats allow a limitless degree of manipulation and modification to the recorded sound source. For example, if a vocal track has been digitally recorded and is perceived to be either out of tune, out of time or both, the sequencer can be easily employed to quickly rectify any of these perceived defects. An analogue recording would have to be re-recorded to rectify such defects. Further, if at some later date the digitally pre-recorded vocal, or parts of it, was needed for an entirely new composition a digital copy of the file can be quickly rendered and inserted into a new composition as a new audio track.

The way in which an audio track can be edited is determined, largely, by the type of audio file it represents. For example, MIDI files are editable through either the Matrix or Score Edit Window (see Fig. 2 and 3), while wave files are editable through the Wave Edit Window (see Fig. 4). (All Figures are screen dumps form Emagic's Logic Audio Platinum. Fig. 2, 3 and 4 represent the first 8 bars of a piece of music entitled "Green Sleeves": specifically they represent the primary melodic device with a 3/4 time signature):
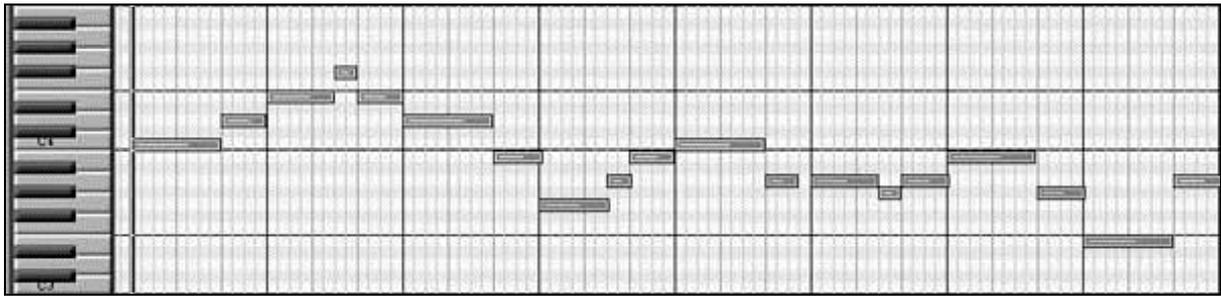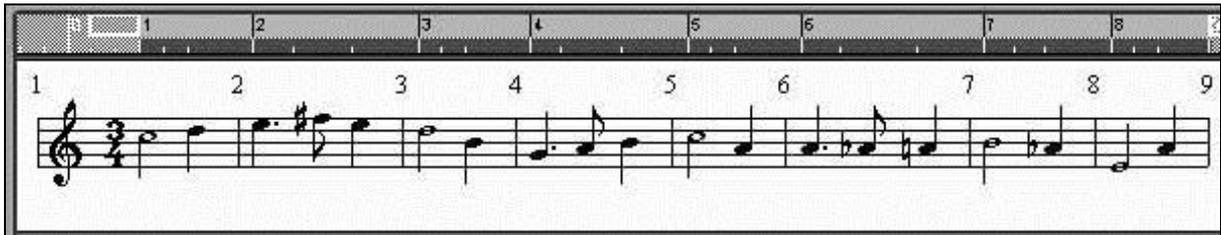
Figure 2. Matrix Edit Window
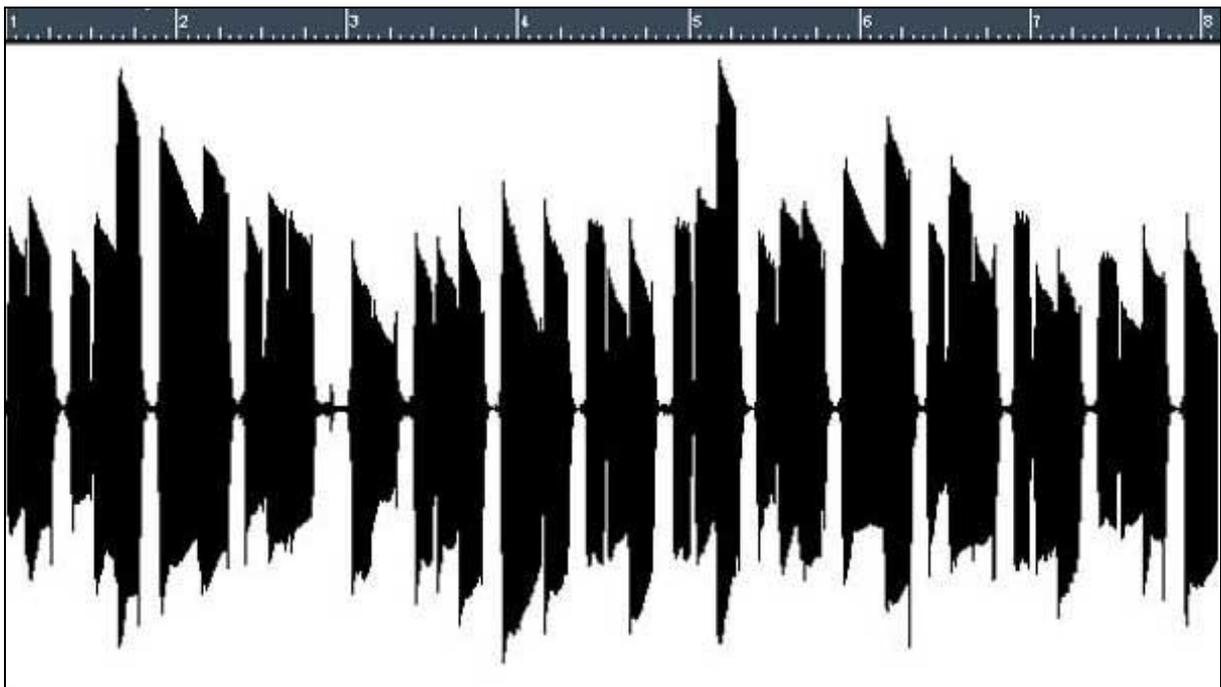

Figure 3. Score Edit Window

Figure 4. Wave Edit Window

2.2 Object-Orientated Model

Currently, Sequencers use a restricted model of music primarily focusing on instances of notes within a musical composition or metadata that describes the music as a whole. They typically generate a computerised visual representation of a musical composition in matrix form (see Fig. 2), notation form (see Fig. 3) or wave form (see Fig. 4). These abstractions are linear and punctuate notes, their duration and magnitude. The notes are defined by raw parameters of musical sound, for example pitch and velocity.

These applications, for example Emagic's Logic Audio Platinum, have been developed using an Object-Orientated paradigm, and are primarily event driven. This developmental approach is suitable

for representing many instances of the same, or different, notes within an audio track, from which one note can inherit the properties of another e.g. timbre, velocity and gain. However, this developmental approach offers little information regarding the relativity of the order of notes and nothing pertaining to the relationship between one note and the next or groups of notes within a composition.

For a long time theorists investigating aesthetics, for example melody and timbre, have been trying to explain how, and in what way, these aesthetics influence a listener's appreciation of music. It is the relationship between notes, and groups of notes, that some music theorists propose as pivotal to a listener's cognition and perception of music and consider it fundamental to their appreciation. Fig. 1, 2, 3 and 4 provide powerful visual models regarding the arrangement and dynamics of instruments. Fig. 1 clearly and quickly affords visual information as to when particular instruments start and end within a composition. Fig. 2, 3 and 4 supplement this information with dynamic data for example, pitch, sustain and magnitude. However, such information reflects nothing pertaining to the aesthetical content of a composition, in particular melody and how a listener might perceive a melodic device. Consequently, there is no visual representation pertaining to what a listener's expectations might be on hearing a particular melodic device and whether, as musician, engineer or producer, the composition should or could be edited to realise or deny those expectations or only part of a listener's melodic expectation.

3. Implication – Realisation Model

Narmour, congruent with these notions, proposes the Implication – Realisation Model. The first volume describes the basic building blocks of the theory focusing on note-to-note relationships in melody cognition and analysis. The second Volume deals with more complex issues within larger groups of notes and non-contiguous tones. Together they boast a conceptual framework explaining the "'genetic code' of melody" (1990, p. 14) [5], to provide a nomenclature of archetypal melodic shapes.

The theory is primarily concerned with tracing the listeners' changing expectations over time, and the extent to which those expectancies are realised or denied. Consequently, based on bottom-up implications in the various parameters of sound, the theory measures the listeners' ongoing levels of surprise as an aesthetic response to music as well as a listeners' ongoing perceived structural closure. With strong perceptual and cognitive leanings the model affords a sophisticated symbology in order to analyse a musical composition with an aim to describe how listeners hear melody. Consequently, It parses melody from note to note on the basis of the raw parameters of musical sound. These parameters include interval size, direction, rhythmic duration, dynamic accent and so on. These parameters describe the relationship between notes and groups of notes.

Arguably a dated reference, Narmour's principles are wholly divorced from any influence of style, culture or musical ability / training on the part of the listener. It is anticipated top down implications, for example musical style and culture, will influence the listeners perception of music and will be integral to the successful outcome of the project. However, Narmour's hypotheses have recently undergone significant empirical evaluation among many notable music psychologists, validating many aspects of the model [7].

3.1 Basic Archetypes

Perhaps, the most far-reaching aspect of Narmour's theory is it offers a set of principles of melodic analysis and cognition that work on individual parameters of musical sound from the bottom up. Narmour's primary hypothesis is that any two successive pitches (1 melodic interval) imply a third pitch (second interval). Therefore, at the most basic level, melodies can be divided up into elemental building blocks of three pitches, with pitches one and two for the antecedent interval of "implication" and pitches two and three forming the consequent interval of "realisation". Whether the antecedent implicative interval is completely realised, partially realised or denied depends on the interval size and direction of the consequent interval. These two parameters, interval size and direction, lie at the heart of Narmour's theory.

Interval size is dichotomised into small and large intervals. Small intervals range from the unison to the perfect fourth and large intervals comprise the perfect fifth and wider. A tritone is considered

ambiguous and can be counted as either large or small dependent on its context. The relationship between successive intervals is characterised by calculating the difference between them and depending on their intervallic difference, antecedent and consequent intervals are considered "equal" (difference of a perfect unison between two intervals), "similar" (difference of less than a minor or major 3rd) or "different" (difference of more than a minor or major 3rd).

Interval direction is divided into three categories: up, down and lateral (repetition of pitch). Up followed by up, down followed by down and lateral followed by lateral are considered continuations of registral direction otherwise they are considered changes in registral direction. Narmour proposes two basic three-pitch melodic archetypes called "Process" and "Duplication" based on similarity and proximity to the parameters of interval size and direction. If a listener hears a small interval, that interval implies a second similarly small interval continuing in the same registral direction. If this implication is realised by the third pitch, giving an up-up or down-down motion compromised of two small intervals then the melodic shape is a "Process". If the implication is a unison (lateral motion between two pitches) the listener will expect a second interval in the lateral direction, giving the shape of three repeated pitches. When this implication is fully realised the result is referred to as "Duplication".

In contrast when a large antecedent interval is heard a change in interval size and direction is expected. If the implications in both parameters of interval size and registral direction are realised, a third basic melodic archetype is generated, namely "Reversal".

"Process", "Duplication" and "Reversal" form the primal structures of Narmour's theory and are generated when melodic implications in both parameters of interval size and registral direction are realised. Narmour provides further detailed analysis on a group of three-pitch structures generated when an implication in only one or the other of the two parameters is realised. He goes to explain the notion of prospective and retrospective structures. This forms only a small part of his analysis on basic melodic structures and indicates little about his work on analysis and cognition of melodic complexity, all of which is beyond the scope of this document.

4. Conclusion

Given the basic primal structures of the Implication – Realisation Model it is anticipated aspects of CPM's proposed graphical interface should visually represent the following:

- Processes - small changes in interval within the same registral direction

- Duplications - repetition of the same interval

- Reversals - large change in interval followed by large change in interval in opposite registral direction

It is anticipated suitable interface design techniques will need to support two potential constraints. Firstly, graphical representation derived from the above primal structures. It is assumed these representations will be based on user's existing mental models of these concepts. A design technique should afford the ability to identify these mental models and their creation into graphical representations.

Secondly, the interface must indicate where and when within a composition a listener's realisation of melody occurs. For example, if a linear representation of these structures is deemed suitable, similar to conventional Sequencer interface design, the timeline must visually indicate at what point realisation happens. Further, extending this constraint coupled with concepts derived from Narmour's theory, realisation can also form implication at any point in a composition. This more complex concept implies the interface should afford modality. For example, mode A reflecting realisation and mode B reflecting implication. It is the experience of the author that modality, unless carefully designed into an interface, hinders performance. Put simply, user expectations in one mode are often very different from another. When modes are confused by the user, generally due to poor interface design, performance is greatly reduced. This is due to unanticipated user-feedback and the inability to

complete a task until the user's gulf of evaluation has been bridged i.e. the user realises the interface is in the incorrect mode. If modality is identified as an interface requirement any design technique employed must afford the ability to empathise with user expectations and clearly reflect, through the interface, the mode at any given point in time. Otherwise, the proposed interface must reflect realisation and implication simultaneously.

References

[1] Cavanagh, David. The Creation Records Story: My Magpie Eyes are Hungry for the Prize. London: Virgin Publishing, 2000.

[2] Crest, David. Music Theory: Intervals; 2001
http://library.thinkquest.org/15413/theory/intervals.htm [10/02/2003]

[3] Frith, Simon. On Record. London and New York: Routledge, last published 2000.

[4] Lucy, Charles E. H. Pitch, Pi, and Other Musical Paradoxes (A Practical Guide to Natural Microtonality). London: Music Teacher, 1988.

[5] Narmour, Eugene. Volume I: The Analysis and Cognition of Basic Melodic Structures: The Implication-Realisation Model. Chicago and London: University of Chicago, 1990.

[6] Narmour, Eugene. Volume II: The Analysis and Cognition of Melodic complexity: The Implication-Realisation Model. Chicago and London: University of Chicago, 1992.

[7] Roeder, John. Toward a Semiotic Evaluation of Music Analysis. Vancouver: A Publication for the Society of Music Theory, 1993.