

Jazz and the art of anticipation

Dr Kenneth B. McAlpine

This is the authors' final version of this article. The final publication is available at Springer via <http://dx.doi.org/10.1007/s40869-015-0008-6>

The Computer Games Journal
Editorial

Jazz and the art of anticipation
Dr Kenneth B. McAlpine

Music and the moving image have cohabited quite amicably almost since film was first committed to celluloid – the first recorded instance reportedly occurred on 28th December 1895, when a pianist provided accompaniment to a series of Lumière Brothers' shorts in the Grand Café, Paris¹. Ever since, music has been a core element of the cinematic experience: sometimes driving the narrative; sometimes giving the audience privileged insight into characters' motivations and behaviours, and often being the vehicle through which much of the film's emotional impact is delivered. Done well, the soundtrack has the power to emote without the viewer consciously registering that it's even had an effect.

One reason that this marriage has been quite so successful is because both film and music are temporal media with fixed, fairly linear structures: once a film sequence has been cut, it provides a narrative framework around which can be built the musical underscore. It provides definite hitpoints to which the composer can write, and is suggestive of tempo and intensity, creative building blocks that translate well into the musical domain, and which can be built upon in the soundtrack.

Although there are problems in simply translating directly notions from film theory and semiotics without adaptation, there are sufficiently many similarities in function and affective purpose between gaming and film to draw comparisons.

Two significant areas of departure, at least from the perspective of the soundtrack, are scale and interactivity. The enormity of the combinatorial possibilities offered by many games makes it impossible for the composer to anticipate and score for every eventuality, and interactivity, coupled with multi-threaded or multi-linear game narratives, means that predicting the game-state at any point in time is very difficult, particularly if, to misquote Niels Bohr, you try to do it in advance. As a result, traditional approaches to scoring for film are largely redundant for in-game music sequences, and a range of adaptive music techniques have grown up to replace them.

The simplest form of adaptive music is the event-driven cue. Here, the game engine monitors the current game state for a specific event. With each of these is associated a specific musical cue, which is triggered by the game event. That could be as simple as the musical sting that is triggered when Mario collects a gold coin, or a more subtle weaving together of transitional game states and music as is seen in Taito Corporation's *Space Invaders Part II*. Released in 1980,

¹ Fielding, R (1979). A technological history of motion pictures and television: an anthology from the pages of the Journal of the Society of Motion Picture and Television Engineers. Oakland: University of California Press. pp. 49–51.

the game uses a descending dorian scale pattern as the main in-game music track which is triggered as the level begins. The sequence is repeated and gradually speeds up as the alien invaders are picked off one-by-one.

Horizontal resequencing, whose name derives from visualising the soundtrack extending along a horizontal timeline, is essentially a refinement of the event-driven cue, and uses similar gameplay markers to rearrange pre-recorded sequences of music in real time. Turning again to Mario – since, of course, Mario provides a metaphor for all of life – the technique is most easily seen when he collects the super star. At that point, the ubiquitous Mario theme is swapped for the faster-paced invincibility theme for the duration of the power-up. The Mario intro is then used to segue back into the main Mario theme.

Vertical reorchestration similarly monitors the game state, but effectively remixes musical material in real-time, either dropping out elements of the soundtrack or building in intensity depending upon the current gameplay state. The ‘vertical’ monicker here relates to the discrete musical or sound layers that are introduced or removed from the music track as time progresses.

Finally, procedural or algorithmic generation composes ‘on-the-fly’ in response to current or anticipated game states. A good analogy here can be found in live jazz ensembles. A typical ensemble may consist of 5 musicians, each with slightly different roles to play within the group, and, although as a collective, the band will have a rough idea of how each piece they play is going to sound, each of the musicians effectively composes on-the-fly, creating the musical detail as the piece develops. The key to successful jazz musicianship is for each musician to listen and adapt to what each of the other musicians are doing, varying in intensity and complexity to ensure that the performance is balanced and musical. Although there is no predetermined ‘script’ for the music, save, perhaps, for the loose framework provided by the chord chart and a sketch of the melodic ‘head’, good jazz musicians will pick up on cues from their fellow players, giving the music direction and resulting in a coherent output that exhibits the classical musical elements of tension and release.

Typically, games will use some combination of the above methods, sometimes to great effect: LucasArts’ iMuse system, for example, works beautifully in *Monkey Island 2*, creating a rich and evocative evolving soundtrack which – crucially – is subtle enough to be genuinely affective without drawing attention to itself. In effect, the music engine is performing a task similar to that of a good club DJ, who will create a complete set by triggering breaks, layering and crossfading, raising and dropping the intensity and tempo of the music to both match and guide the atmosphere in the club.

There is, however, a crucial difference. The techniques outlined above are largely reactive. Music changes in response to game events rather than in anticipation of them. The importance of this should not be underestimated. In the same way that the club DJ will set the mood and tone of the dance floor – not just react to it – so too a movie soundtrack will swell and build to anticipate the emotional and climactic peaks of the film. Think of the moment in the *Godfather* when Michael Corleone shoots Sollozzo and McCluskey in Louis’ Restaurant. In the moments

before the shooting, there is a close-up of Corleone's eyes accompanied by the exaggerated rumble of a subway train. Walter Murch, the sound designer for the scene is quoted as saying that the sound is "metaphorical, in that we've never established the train tracks and the sound of the train is played so abnormally loud that it doesn't match what we're looking at, objectively. For a sound that loud, the camera should be lying on train tracks."² That combination, of the visual and the soundtrack conspires to create one of the great moments of cinematic tension, so that the emotional state of the audience, when the shooting occurs, is already beyond heightened. The soundtrack leads, rather than follows, the action.

In games, the key to this is perhaps not to be found in the mechanics of adaptive music. Rather, it is to be found in the finessing of their application, using a sense of artificial musical intelligence to handle musical transitions sensitively and coherently, and, perhaps more challengingly, systems of analysing gameplay and extrapolating from current game states to likely future game states so that the music might lead the gameplay.

These problems are surely not intractable, but if ideas such as these are to gain traction and currency, they must prove that they bring real added value to a game. There is a computational cost associated with the implementation of a sophisticated music engine as well as a production cost. The real risk, though, is that if the soundtrack does its job well, nobody will notice.

² Ondaatje, Michael, and Walter Murch. *The Conversations: Walter Murch and the Art of Editing Film*. Toronto: Vintage Canada, 2002.