PLAYING WITH MUSIC

Investigating Emergent Music Composition for Games

Rodwin Malinga

Student number 1420512

SUPERVISED BY KIERAN REID AND TIMOTHY FLUSK
UNIVERSITY OF WITWATERSRAND
MA DIGITAL ARTS

WITS UNIVERSITY

Abstract

As video games have become more dynamic and complex, their individual components have become more emergent in nature. By increasing the potential for emergence in each component medium, the composite video game is likely to become more emergent, resulting in more meaningful play. Therefore, developing a responsive music system that can change based on the emotional state of the player is critical in building deeper video game experiences. The importance of music in video games cannot be overstated, as it plays a crucial role in creating more engaging and immersive experiences for players. This study aims to explore the development of a different type of emergent music for video games, that not only has the ability to invoke behavioural responses in players but can also complement and enhance the player's emotional journey within the game. The ultimate goal is to create a musical experience that elicits a more engaging and immersive experience for players, thereby contributing to the development of more affective and dynamic video games.

Keywords

Affect, Agency, Dynamism, Emergence, Musicality.

THIS DISSERTATION WAS WRITTIN IN PARTIAL COMPLETION OF A MASTER'S DEGREE

Contents

Intro

Over the past three decades, the video game industry has experienced a surge of growth and development, particularly with regards to the enhancement of visual fidelity, expanding interactivity, and creating more immersive gameplay (Gungormusler et al., 2015). Despite these advancements, video game music has remained mostly linear, often following the same scoring practices used in film music. This is concerning, given that video games require active participation from players, unlike films.

In light of this, it is essential to acknowledge the need for video game music to evolve and become more responsive to player actions. Just like the other components that make up a video game, music should be tailored to enhance the overall gaming experience by adapting to the actions of the player. By doing so, players can feel more engaged and connected to the game's storyline, characters, and world.

Although some technologies exist in contemporary video games, that allow for the creation of adaptive scores, they are more of the exception than the rule. These technologies, often referred to as dynamic music systems, typically rely on a vertical approach to composition that uses layering and branching segments (Larson, Hedges, & Mayer, 2010). This results in music that is variable using event triggers. However, these technologies are still based on pre-recorded audio loops, which are designed with a linear progression in mind, and lack the real-time flexibility afforded to the other elements of video games. Additionally, such approaches cannot respond to the emotional state of the player as their behaviours develop and emerge.

Therefore, this study seeks to investigate a different type of responsive music for video games. Music that not only has the ability to invoke behavioural responses in players, but can also change based on their emotional state. The goal is to create a musical experience that complements and enhances the player's emotional journey within the game, resulting in a more engaging and immersive experience. By developing music that responds in real-time to the player's emotions, I hope to induce a type of gaming experience that is more personal and compelling than ever before.

This paper will consist of four chapters, each exploring different aspects pertaining to the development of emergent music in video games. The first chapter will delve into the structures that govern behavioural change in players, providing an argument that video game scores should become emergent rather than linear or dynamic. By digging into the mechanisms of behavioural change, we should become more informed about how music can be designed to facilitate more engaging and immersive gameplay experiences.

The second chapter will focus on the relationships between players and sound in video games, providing a foundational understanding for how emergent music can be constructed and developed. By understanding how players interact with sound, we will be able to identify opportunities to create more responsive and personalised music systems that enhance the player's emotional journey within the game.

The third chapter will present a compendium of tools and strategies that can be used as a guide for the development of emergent music systems and technologies. We will explore various methods for creating music that responds in real-time to the player's emotional state and behaviour, providing developers with the tools they need to create more dynamic and engaging gaming experiences.

In the final chapter, we will detail the process of putting our toolkit to use by developing a video game where the core focus is the development of emergent music that elicits and responds to player affect. By applying the principles and strategies outlined in the previous chapters, we will demonstrate how

emergent music can be used to create more immersive and emotionally resonant gameplay experiences.

Lastly, I want to clarify that I am not formally trained in music production or composition. Thus, this paper will not be written from the perspective of a skilled professional. Instead, this investigation aims to serve as a tool for designers and composers to understand the relational needs of video games and music. By exploring the intersection of these two fields, I hope to provide insights and recommendations for how they can better aid one another in becoming more impactful and meaningful for players. Through this investigation, I aim to contribute to a broader understanding of the role that music can play in video game design. My intention is to inspire new and innovative approaches to game music composition, which can lead to more immersive and emotionally resonant gaming experiences. By bridging the gap between music and video game design, I believe that we can create games that are more engaging, more memorable, and more enjoyable for players.

Track 1: Structures of feeling

It is no secret that video games are an extremely moving and influential medium, that is immensely important for human beings. As systems that heighten emotional responses, video games effect player behaviour and generate meaning. Like a lot of media, they can function as a means of entertainment and relaxation. While this purpose may make them come across as trivial or unimportant, their presence is far more vital than one might assume. Video games exist as a space where people can behave and express themselves in interesting, and often experimental, ways. They employ emotion-producing techniques to support human behaviour (Sears & Jacko 2009), and ultimately act as bridges that connect players to experiences that they may not have otherwise had.

Video games then, are systems that produce affect in hopes of conveying meaning, encouraging behavioural responses in players. As such, they can be interpreted as structures of feeling, that give expression to emergent ways of being in the world (Anable, 2018: 7). This chapter seeks to explore the concept of video games as structures of feeling, and to examine the relationship between affect, emergence, and video games. Furthermore, this chapter will emphasise how music can contribute to conveying meaning, eliciting behavioural responses, and presenting carefully crafted experiences to players in video games.

1.1 Affect

Video game experiences are tremendously impactful. They are, after all, affective in nature. The intimate relationship between video games and affect means that video games themselves are “affective systems” (Anable, 2018). But what does it mean to be *affective*? Though affect has earned a reputation of being particularly vague, in this study, I use the word to refer to the feelings and forces that exist within, and circulate through bodies and objects. These forces inform our behaviours and induce our emotional states. The concept of Affect harkens back to that which is felt. Affect speaks to the relational forces that are attached and expressed through cultural objects and phenomena. It shapes the very being of subjects and objects as they come into contact with one another (Anable, 2018).

The design of video games is deeply rooted in affect. A fundamental goal of game design is to elicit specific emotional responses in players through the use of curated design practices. Video games are behaviourally driven and each game has its own unique aesthetic goals that it strives to achieve (Hunicke et al., 2004). The affective nature of video games is crucial to their design. Video games have been shown to be effective at reducing stress, uplifting moods, promoting social connections, and even encouraging cooperative play (Przybylski et al., 2010). Similarly, the affective nature of video games can

also lead to experiences such as sadness, heightened aggression, and competitive play (Anderson et al., 2010).

Nevertheless, the relationship between video games and affect remains a significant factor in the impactfulness of a video game. Video games have the ability to evoke desired aesthetics and behaviours in players, creating immersive and emotionally engaging experiences for players, that sets them apart from other forms of media. It remains apparent that the success of video games as a medium is in no small part due to the relationship between game design and affect.

By using techniques such as rubber-banding to manipulate speed values in Super Mario Kart (1992), ensuring that players stay close together, designers can increase emotional responses such as tension and exhilaration to create desirable (or undesirable) experiences for players. This is achieved by always presenting them with the feeling of a close race. Furthermore, such emotion amplifying methods are likely to result in meaning generation that can be specifically tailored for each video game. For example: it is plausible that the aforementioned rubber-banding technique will result in players feeling that they are evenly matched, regardless of their actual skill level, because races will never be won by a large enough margin to determine a definitive Super Mario Kart (1992) champion.

For video games then, being affective means that they are able to elicit specific emotional responses in players, shape their behaviours and tap into the forces and feelings they experience as they play. Over time, however, video games have become increasingly dynamic, responsive, and emergent. This means that rather than systems that merely produce affect they have evolved into systems that respond to and are even shaped by affect. When compared to film, a medium which invites its viewers to share in the joys and sorrows of the characters on screen (Sears & Jacko 2009), video games require far more participation from their audience. Players are required to engage with systems and be active contributors to the experience. A film without an audience can still be played out in its entirety without any input from an observer. Video games, on the other hand, require engagement from their players in order for any progression or meaning generation to occur. According to Sears and Jacko (2009), games can move beyond film to claim their true power by tapping into the player's emotions and providing a sense of reward for their actions. This highlights the interactive nature of games, where players can actively participate in the progression of a video game.

Unsurprisingly, in recent years, video games have evolved to become more interactive, engaging, and emergent. However, for video games to truly be affective they cannot merely induce affect but must in turn respond to it. In many contemporary games there is a clear focus on emotive storytelling, emergent narratives, reactive visuals, and player driven events; this is all in the name of creating games that are more immersive, engaging, and ultimately, more affective. Games are moving into the domain of dynamism and variation, where models of knowledge association and emotional effect are combined with multimodal, multiagent composition techniques (Hutchings & McCormack, 2020) to produce a play experience in which video games and their players have a symbiotic relationship of affect. In other words, video games are increasingly evolving into dynamic and responsive systems that engage with players in ways that are unique to each player; not only are games becoming more affective, but in turn they are reacting to the interactions and dynamics created through play (games are responding to affect).

As interactive systems, video games participate in a cyclical process (Crawford 2003 as cited in Collins et al., 2014), where agents act upon each other as mutual partners. Each agent acts both as a receiver and transmitter of affect. By playing a video game, players are engaged in a two-way relationship of reciprocal communication that allows for mutual adjustment and results in experiences that amplify the bespoke relationship between the game and a particular player. The reciprocal relationship

between video games and their players allows for a unique form of affective experience. Players are not passive recipients of feeling, rather, they are active participants in the creation of affect and meaning within the game world. As such, the affective quality of a video game is dependent, not only on the game's design, but also on the player's bespoke engagements with the game. This implies that, in video games, affect is not static or predetermined, rather it is constantly emerging alongside the player's behaviours and choices.

Ultimately, affect is a crucial component of video games that guides their design. Through their relationship with affect, video games are able to elicit specific emotional responses in players, shape their behaviours, and respond to the forces and feelings generated (and experienced) by them. By incorporating emotional rewards into gameplay, games have the potential to create a more engaging and memorable experience for players. The symbiotic relationship between video games and their players, enables a dynamic and constantly emerging form of affect that is carefully tailored and meaningful for players.

1.2 Emergence

The inclusion of variable features such as user generated content as seen in Little Big Planet (2008), or the branching dialogue options present in Detroit: Become Human (2018), have allowed games to include player-directed events (Hutchings & McCormack, 2020) and emergent narratives (Louchart et al., 2015). More than ever, player input is a requirement for the consumption of video games. Rather than passive observers, players are a crucial component of the play experience. Without player input, there would be no inciting occurrence to cause video game progression. Though a video game is a complete entity in and of itself, its value is far greater than the combination of its components. What may present itself as an amalgam of art and software, has the capacity to become an entire inhabitable world by the simple addition of player agency.

Video games are emergent systems after all. By emergence, I am referring to the occurrences in which a complete system exhibits properties that its component parts do not have. Emergence is a vital aspect of games that ties their systemic nature to their ability to elicit meaningful play (Salen & Zimmerman, 2003: 159). That is to say, the complex nature of video games is one of the things that makes them so unique, as there is a direct link between their rule-based nature and our ability to express ourselves and have experiences when we play.

> "When we open a video game, we are opening up a "form of relation" to the game's aesthetic and narrative properties, the computational operations of the software, the mechanical and material properties of the hardware on which we play the game, ideas of leisure and play, ideas of labour, our bodies, other players, and the whole host of fraught cultural meanings and implications that circulate around video games"(Anable, 2018: 9).

Put simply, video games are made up of many smaller components that merge into something that is greater than the sum of its parts. For instance, the rules of Gwent (2016) may consist of a number of simple instructions that dictate how each card in the game works and that state an overarching win condition. However, once the game is in motion new dynamics arise that alter our initial understanding of the game. These dynamics are not explained within the rules of the game, nor are they printed on the cards themselves. Rather, this dynamism[1] arises from the complexity of the game as a whole. New

[1] In this study, dynamism refers to the ability of an entity to respond to the affect applied to it by an external actor. An entity that can dynamically alter itself as a result of, or in response to, the behaviour (or state) of another, is an entity that exhibits a high level of dynamism.

meanings, implications and emotional responses arise as a result of the game being played, and both the game and its players are affected, in some way, by one another.

It makes sense then, that narratives, visuals, and interactions in video games have become increasingly dynamic and emergent, as emergence has a direct effect on a video game's ability to illicit and respond to affect. In order to synergise more effectively, the components that make up video games are becoming more and more emergent in and of themselves.

Games like Telling Lies (2019) and Her Story (2015) are key examples that illustrate the importance of an emergent structure for video games. In both games, the player is allowed to engage with the narrative in a non-linear fashion. These games allow players to interact with vignettes captured on video, that can be searched for and navigated through, to piece together the events of the narrative. Players can use the information gathered in one vignette to search for and view the next. While it is possible to engage with these vignettes linearly, players are encouraged to search for what is interesting to them. This allows players to engage with the games in a way that is unique to, and is governed by, them. Rather than mere observers, players are a crucial part of the meaning making process, as without their input the narrative would not progress. Furthermore, the player's choices directly affect the state of the game, and in turn the state of the game affects the player's choices. If a player is interested in an event that happened prior to the vignette they have just interacted with, the game allows them to search for said event; this results in a disruption of the chronological flow of content, however, it allows the player to continue pursuing that which is important to them and alters the state of play to support their desires.

Though both of the aforementioned games can be described as a search engine that is used to sift through collections of shortform videos, what they offer is far more than the combination of their components. While their structure is simple, through emergence they are able to present an infinite amount of narratives, emotions and meanings. Moreover, by allowing player agency to affect the state of the game, players are rewarded with more expressive and meaningful play. It is clear then, that by increasing a component's ability to dynamically respond to player affect, one directly influences the emergence of the video game as a whole. Furthermore, by increasing the game's emergence, one enhances its ability to affect its players which ultimately results in more meaningful play. This phenomenon can be observed throughout all of the media that make up the multimodal composite known as a video game, as the emergent nature of video games not only allows for the creation of unique play experiences, but also offers opportunities for players to engage with narratives and visuals in dynamic and meaningful ways.

Video game music, on the other hand, remains far less dynamic when compared to the other components of video games (Hutchings & McCormack, 2020: 270). If video games are emergent systems that promote dynamism through their complexity in hopes of producing more affective experiences, then surely as a component thereof, video game music should become less linear and more emergent to better synergise with its sister components (Visuals, Narratives, Coding Scripts, etc.). By doing so, video game music can aid in the overall enhancement of the play experience. Through the use of emergent music, developers and composers can induce more meaningful play, which will ultimately result in richer, more affective, experiences for players.

1.3 Emergent music

As video games have evolved, the many modes that exist within video games have also evolved. When the multi-modal composite called a video game becomes more emergent, it is as a result of its individual components becoming more emergent in nature. In modern video games, visuals and images seamlessly respond to the actions of players, user generated content allows games to surpass the

expectations of their creators, and narratives alter their course in response to decisions made, emotions felt, and actions taken by players. This phenomenon illustrates that there is a direct correlation between affect and emergence within video games and their components. The more a video game responds to player affect, the more it exhibits properties and behaviours that are unique to a specific player and their playthrough of the game.

Video game music, on the other hand, continues to remain linear in many cases. When compared to its sister media (Images, Videos, Narratives, Coding Scripts, etc.), little development has been made pertaining to the dynamism of the structural elements of music in video games. Few games make use of dynamic music to the same degree that they use dynamic visuals, narratives and so forth.

In the realm of triple-A game development there are inherent incentives to playing it as safely as possible when developing new titles (Collins et al., 2014). It makes sense then, that studios rarely make use of integrated composition techniques to create emergent music and rather opt for the more common practice of hiring composers to work remotely from images (Inglis 2012), or from a few lines of instruction written on a spreadsheet (Pham 2008 as cited in Collins et al., 2014: 9).

Music, however, is synesthetic (Collins, 2013) and should be designed in conjunction with the gameplay that it will support. When we hear music, we associate it to images and feelings that are inextricable from its sounds. This synesthetic nature is a result of our mental associations of sound and things that are affective. We experience sound in association with what is emotionally felt, hence, when a sound is heard, we might still "feel" the emotion linked to the aforementioned sound. Similarly, "We typically experience sound in association with image, and thus when the image is not apparent, we might still mentally "see" that image" (Collins, 2013: 26). In other words, the feelings that one has about the sounds they hear are directly related to the way they feel about the origin of the sound, the visuals and affective signifiers that they associate with said sound.

It seems then, that music is not so different to video games in its nature. Both are affective entities with a strong reliance on other media for meaning generation. In music we draw meaning by associating the sounds with images we envision and emotions that we feel, and in video games we draw meaning by analysing the agency afforded and communicated to us by each component of the video game (The visuals, narratives, sounds, inputs etc.). Furthermore, music is emotional and, as previously mentioned, affective. Music is played both passively and actively, that is to say, the act of playing music can be done in more than one way. One can play music with the intent of achieving a goal (as one would when attempting to complete a level in a challenging game), or one could play music to unwind or pass time (as one would when playing a mobile game during, or after, their commute home). Similarly, we play games in two ways; we play games like one might play a CD and we play games as one would play an instrument (Hocking, 2011). Interestingly, video games and music both act as motifs for describing one another. They are both inextricably linked by metaphor.

Surely then, of all the components that make up a video game, the music is the most suited to reflect the dynamic and emergent nature of video games. If components of video games such as their visuals and narratives can portray dynamism and responsiveness to player involvement, then why not music? Moreover, if increasing a component's ability to dynamically respond to player affect directly influences the emergence of a video game, then surely by increasing the overall dynamism of the video game's score one directly enhances our ability to express ourselves and have meaningful experiences when we play. Therefore, this research will serve as an investigation as to how emergent music can be crafted within video games, and why there are so few examples thereof.

1.4 Aim: Remastered

This study aims to investigate how to create emergent music in video games. Consequently, it aims to unearth why video game music has fallen behind in terms of dynamism when compared to other media that exist within video games. This is not to say that there has been no development regarding making video game music more responsive, interactive, adaptive, and dynamic. Rather, I believe that the developments made are not on par with those seen in the narrative, visual and systemic departments of video games.

Over the years, there have been significant advancements to video game music, and the ways in which video games incorporate their soundtracks to enhance affect and the overall play experience. Games like Final Fantasy XV (2016) make use of highly adaptive music systems (Iwamoto, 2017) to increase immersion in players. Such music systems use vertical layering and horizontal resequencing to cycle and transition between different configurations of songs (these are useful techniques for the creation of dynamic music). This allows for the presence of a notable amount of variation in the music that is heard by players. To some extent these music systems, such as the one present in Destiny 2 (2017), can be considered emergent. They do, after all, allow for the dynamic variation of in-game scores (Johnson, Lewin, & Sechrist, 2018).

This type of emergence, however, is not what this study hopes to investigate. While they do retain a high level of musical refinement, the aforementioned emergent music systems still - for the most part - make use of music that does not include players as part of the compositional process. Rather than assembling the music in real time, in response to player agency, these systems cycle and fade between music tracks that have already been recorded as finished pieces. In other words, these systems manipulate melodies that have compositional linearity, and seamlessly transition between them in order to promote the illusion of dynamism within the composition, without directly responding to player affect.

What if, however, there was a way to use the systemic nature of video games to allow for the creation of truly emergent music that responds to the dynamics of play, and enhances the emotional connection between games and their players. Music that is composed in real time, with the player as an active participant, that is complex enough to produce a number of diverse outcomes, while retaining the emotive quality that makes music so moving. This is the type of emergent music that this research is aiming to investigate, and ultimately create.

Unfortunately, this aim remains (to some degree) undefined. It is still unclear to what extent this will aim will be accomplished. As a result, the aim of this paper will remain flexible and subject to change as the study progresses. As more information is gathered, the aim of this paper will be remixed and remastered as necessary. This approach will allow this paper to adapt and adjust its methods and strategies as needed, in order to achieve the best possible outcome. As such this paper will become an embodiment of emergence, embracing uncertainty and exploring new possibilities in the pursuit of knowledge and understanding.

In order for this study to serve as a true investigation into the creation of emergent music in video games, it will be created in conjunction with a video game. This video game will act as a platform for showcasing and testing various methods of designing and implementing emergent music, and will aid in uncovering why video game music has fallen behind in terms of dynamism when compared to other media that exist within video games.

Track 2: Games, Players and Sound

It is commonplace for modern video games to include several unscripted moments that arise as a result of emergent design. As video game soundtracks are an inseparable facet of the play experience, it makes sense that developers and scholars alike have attempted to create emergent moments of gameplay that have unique musical identities to support them. Fortunately, many of these attempts have been documented and recorded. This means that there is an array of texts that can be used as a foundation, and framework, for investigating how to create emergent music in video games.

This Chapter will explore a handful of these texts and the concepts within them. While some of these texts and concepts do not directly relate to emergent music in particular, all of them are crucial for situating this study within the broader context of video game music. This chapter will unpack the concept of meaning generation through an exploration of the relationships between players, games, sound, and play. Furthermore, in order to understand the current landscape of video game music, this chapter will include a comparison of contemporary music systems and adaptive music systems to contextualise the existing forms of dynamic music in video games.

2.1 The relationship between players and games

It is important to understand the relationship between games and their players. Games are interactive and allow us to infer information about their player's intentions through the decisions they make when presented with choices. Through an investigation of interactive media, Sundar et al. (2015) provided evidence proposing that having the agency to change in interactive media has a significant psychological effect over one's sense of self, their perceived sense of control and the persuasiveness of messages presented to them. Being situated in an environment in which one has the ability to interact with objects and subjects freely, results in behavioural change (Hutchings & McCormack, 2020). Thus, it is clear that the player is more than a mere observer; they are a source of control, knowledge and affect.

Generally, video games are designed with this in mind; not only do they present players with the agency to make choices, but they respond and react to the decisions made by their players. By enforcing simple rules to produce explicit outcomes due to player action, video games create what Salen & Zimmerman refer to as "meaningful play" (2003). The notion of meaningful play emerges from the aforementioned relationship between the player and the game system. It is the process by which a player action occurs within the system and the game responds to said action (Salen & Zimmerman, 2003: 50). This implies that the meaning of an action in a game resides in the relationship between player action and system outcome. In their book *Rules of Play: Game Design Fundamentals* (2003) Salen & Zimmerman offer the following example to explore the relationship between games and their players:

> Think about an informal game of "Gross-Out" played during an elementary school recess. One by one, players tell a gross-out story, each tale more disgusting than the last. When a story is finished the group spontaneously and collectively responds, confirming or denying the player's position as master of the playground, until an even grosser story is told. If we look at Gross-Out from the perspective of meaningful play, we see that a player takes an action by telling a story. The meaning of the action, as a move in a game, is more than the narrative content of the story. It is also more than the theatrics used to tell the story. The outcome of the storytelling action depends on the other players and their own voting actions. Meaningful play emerges from the collective action of players telling and rating stories. The meaning of the story, in the sense of meaningful play, is not just that Hampton told a whopper about his big sister eating a live beetle it is that Hampton's story has beaten the others and he is now the undisputed Gross-Out king.

This example beautifully encapsulates the importance of player action while highlighting how crucial it is for games to respond to player agency. Salen & Zimmerman make an effort to indicate that

meaningful play emerges from the collective action of players telling and rating stories. The use of the word “emerges” is not merely aesthetic, rather, it is an intentional indication of the significance of emergence in games. By emphasising that the meaning of action in this game, is more than the narrative content of the story or the theatrics used to tell the story, they hint towards the understanding that meaning in games is ascertained through emergence.

Through this correlation between meaning and emergence, it becomes apparent that the nature of games as affective systems (Anable, 2018 as cited in Ruberg, 2020) relies not only on a game’s capacity to facilitate affect in players, but in its ability to be affected by said players. Hence, the relationship of affect between games and players is a symbiotic one. Therefore, in order for music in video games to become more meaningful and affective, it needs to respond to player action while remaining emotive enough to prompt player action.

2.2 Understanding sound

The concept of meaning generated through emergence is not exclusive to video games as a whole; it can be observed within the component media that make up video games as well. If the colour palette of a video game shifted hues as a result of a player’s input, a number of cognitive associations could be drawn connecting the input of the player to their understanding of the visuals presented on screen (Sundar et al., 2015). By increasing the potential for emergence in each component medium the composite video game is likely to become more emergent as a result, and ultimately elicit more meaningful play.

Bearing this in mind, it is reasonable to attempt to promote a deeper level of meaningful play through the use of emergent audio (or in the case of this research, emergent music). Sound is schizophonic (Schafer, 1969 as cited in Collins, 2013), meaning that it cannot be separated from its source in one’s mind. Sound is, therefore, multimodal as it has a strong semiotic relation to other modalities (Collins 2013). When we hear sound, our minds perceive it while imagining its source and the state thereof. Thus, sound exhibits similar qualities to video games. Depending on the audio heard and the perceived source thereof, listeners are engaged in a system of meaning generation which is ascertained through emergence. This relationship between games and sound makes music a suitable medium to reflect the dynamic and emergent nature of video games.

However, due to the schizophonic nature of sound, each individual sound holds a number of connotations which could infer a plethora of meanings depending on the listener and the context in which said listener hears the sound. In other words, when attempting to generate meaning through the use of sound, the context in which a sound is situated is equally as important as the sound itself. When attempting to create emergent music, it is important that the context in which said music is situated, does not permit interpretations of the music outside of those which were intended, while still allowing for the existence of dynamism and self-expression.

So, in order for emergent music systems to improve upon existing music systems in contemporary video games, they must not lose the ability to account for the schizophonic nature of sound. The symbiotic relationship of affect between games and players is similar to the relationship between emergent music and its listeners. When one engages with emergent music, they simultaneously take on the role of spectator as well as performer. It seems that for emergent music, it's not sufficient to just influence the listener; instead, the music should dynamically react to the listener's emotional state, almost as if the listener were the one composing it. Therefore, in order for music in video games to become more meaningful and emergent, it needs to respond to player action while remaining emotive enough to prompt player action.

2.3 Music in video games

Music is affective and emotionally driven. Mark Leman (2008: 139) proposed that sound involves all of our senses, it "moves the body, evokes emotional responses, and generates associations with spaces and textures." When we hear music we imagine spaces, see images, feel emotions, and draw connections between sounds to infer meaning. Because of this, it is important that video game music is scored in relation to its game worlds and the emotional intent thereof, as they are the key aspects of the compositional process of professional composers (Hutchings & McCormack, 2020: 271). As Howard Shore states: "I want to write and feel the drama. Music is essentially an emotional language, so you want to feel something from the relationships and build music based on those feelings" (Morton, 2010 as cited in Hutchings & McCormack, 2020).

Designing expertly crafted music to achieve specific emotional goals ensures that the schizophonic nature of sound is well accounted for, as composers are able to weave affect directly into their work to achieve the desired emotive response from their audience. Therefore, in the pursuit of emergent music (as a means to enhance the play experience), it is important that we do not dampen the experience by creating music that is inferior to the music that currently exists within video games. The goal of this study then, is to produce informed methods of composition that enable the creation of music that responds to the dynamics of play, while allowing expert composers to retain the quality present in contemporary music systems.

A) Contemporary music systems

In their paper *Adaptive Music Composition for Games* (2020) Hutchings and McCormack state that in most modern video games there is at least some demonstration of adaptive behaviour in music. As dynamic layering of instrument parts has become common in commercial games, tracks can be triggered and activated in conjunction with scripted events or modifiers to communicate gameplay state changes such as a loss or gain of player health.

In video games, sound can be assessed based on how well it supports play (Ekman, 2014); in other words, developers can implement sound according to the functional needs of the play experience. This explains why the current standard of assembling music for video games involves switching and fading between various layers of pre-recorded music tracks with predefined event triggers. In contemporary video game music systems, triggers such as entering a vehicle, encountering enemies, and a player's location in the game world, can add (alter, or remove) instrumental layers to the score. These layers affect the mood, energy or overall excitement of the game and help to contextualise certain state changes within the game.

There is a consensus in the video game music industry that higher levels of interactivity have a direct proportionality to modularity (Ashby 2008), so a focus on more modular forms of music are likely to allow for a quicker response to in game events while maintaining musical flow. Making use of scripted events that trigger under explicit conditions allows developers and musicians to score music for a number of anticipated scenarios with relative ease. In addition, this means that the music created will consistently support the gameplay experience, all while being of a reliable musical quality.

The problem with this, however, is that it introduces a high level of predictability and repetition, making it difficult to produce unique musical moments for a wide range of unscripted events (Hutchings & McCormack, 2020). Additionally, this method of scoring music – though emotionally driven – does not respond to the affective state of the player and, hence, does not support meaningful play (Salen & Zimmerman, 2003: 50). Rather, composers and developers make informed assumptions about the emotional state of the game and score accordingly. In contemporary music systems, if we consider the music and player as components of the system, we can see that the current practice for implementing

music in video games can be considered as simply reactive (Collins et al., 2014), as opposed to dynamic or emergent.

With this in mind, this study hopes to retain the consistency and reliability that contemporary music systems provide to developers, while adding a layer of complexity that allows for the scoring of unscripted/dynamic events.

B) Adaptive music systems

In their attempt to create emergent moments of gameplay that have unique musical identities to support them, Hutchings & McCormack (2020) developed an adaptive music system (AMS) based on cognitive models of emotion and knowledge organisation. In combination with a multiagent algorithmic music composition and arranging system, they were able to find suitable emotion models (Eladhari et al., 2006), and implemented them within their video game scoring system. They suggested that there is a crucial step missing in the production of affective music scores for video games. For them, this essential step is a communicative model for relating events, contents, and moods of game situations to the emotional language of music. As previously mentioned, the Schizophonic nature of sound means that it holds a number of connotations which could infer a plethora of meanings depending on the listener and the context in which said listener hears it. Composers need to be aware of the in-game context in which the music is going to be heard in order for them to effectively score moments of gameplay.

By developing a communicative model for relating game situations to the language of music they [Hutchings & McCormack] were able to create an AMS that supported the creation of unique musical moments that referenced past situations and established relationships without excessive repetition. The system used a spreading activation model (Collins & Loftus, 1975 as cited in Hutchings & McCormack, 2020) to assess the current state of gameplay and generated music (through the use of complex algorithms) that matched the affective state of the game.

This means that their system is in fact capable of creating truly emergent music that responds to the dynamics of play. Much like what this study aims to investigate, they were able to create music that is composed in real time and is complex enough to produce a number of diverse outcomes. However, though the AMS is capable of creating emergent music, in response to the probable[2] affective state of the player it falls short in its ability to appropriately affect the player. Unlike in contemporary music systems, there is a loss of consistency and reliability when it comes to the AMS's ability to produce music that is of the same calibre of that present in modern video games. "Listening to the music tracks, it becomes apparent that the overall music quality of the AMS is not that of a skilled composer" (Hutchings & McCormack, 2020 : 9). This summation offers some crucial insight into the potential reasons why video game music may have fallen behind in terms of dynamism when compared to other media that exist within video games, and serves as an excellent starting point for the unraveling thereof.

Reflecting upon the above, one can begin to understand some of the difficulties of attempting to add dynamism to music. Making video game scores more emergent is extremely challenging as there is a clear conflict between player agency and musical structure. While music in video games can heighten feelings of mastery and meaningful play by providing emotional rewards to player achievements, this type of music is typically more rigid in structure and coincides with the game's pace of action rather

[2] Though the AMS is emotionally driven, it can only respond to the probable emotional state of the game and player. The challenge of implementing an emotion model in a game to aid in music-scoring is that there is no universally accepted model of human emotion among cognitive science researchers and academics (Hutchings & McCormack, 2020).

than musically affective structures. “The confirmations or violations of musical expectancy that are effective in provoking emotional responses in the kind of tonal music typically utilised in games are difficult to achieve within the temporal uncertainty of a game, where the player has the autonomy to act at any time” (Collins et al., 2014 :1). Simply put, making music more responsive to player action makes it harder to create scores that affect players according to a specific emotional goal. More responsiveness to player agency introduces more variability resulting in more meaningful play, but less control over the generated meanings.

In order to promote a musical experience in which video games and their players have a symbiotic relationship of affect, The responsiveness of adaptive music systems, need to be combined with the emotive quality of contemporary music systems to produce a framework of musical composition that enhances meaningful play.

Track 3: Cultivating emergent audio

In chapters prior we have deduced that, of all the components that make up a video game, music is the most suited to reflect the dynamic and emergent nature of video games. Furthermore, by increasing the overall dynamism of a video game’s score, we directly enhance its ability to foster meaningful gameplay experiences. As interactive systems, video games participate in a cyclical process (Crawford 2003 as cited in Collins et al., 2014), where agents affect each other as mutual partners. By playing a video game, players are engaged in a two-way relationship of reciprocal communication between the game and themselves. By increasing a video game score’s ability to dynamically respond to player affect we directly influence the emergence of the video game as a whole.

For emergent music then, it is not enough to merely affect a listener, rather the music should in turn respond to the affective state of said listener, as if they were the composer. Therefore, in order for music in video games to become more meaningful and emergent, it needs to respond to player action while remaining emotive enough to prompt player action. That being said, we must not neglect music’s ability to affect players with clear intent. Designing expertly crafted scores, to achieve specific emotional goals, ensures that listeners interpret pieces of music as intended, since composers are able to weave affect directly into their work to achieve the desired emotive response from their audience. Therefore, in the pursuit of emergent music (as a means to enhance the play experience), it is important that we do not dampen the experience by creating music that is inferior to the music that currently exists within video games.

This chapter will explore the ways in which we can make use of the systemic nature of video games to create emergent music systems that respond to the dynamics of play, and enhance the emotional connection between games and their players. The goal is to produce music that is, to some extent, composed in real time, with the player as an active participant in the production process. The music created should be complex enough to deliver a number of diverse outcomes, while retaining the emotive quality that makes music so moving.

3.1 Compositional tools

In this study, emergence refers to a phenomenon in which a complete system exhibits properties that its basic parts do not have. Therefore, while attempting to create emergent music, it is important that we are aware of the tools available for use in our emergent music systems. The process of creating emergence is ultimately an act of fusion, as one can only control the components they use to build an emergent system, in hopes of achieving a desired outcome.

Bearing this in mind, the first step in making video game music more responsive to affect is the formation of a compendium of musical tools that promote dynamism, while allowing composers to retain the level of compositional control necessary for them to create powerfully emotional music, that enhances meaningful play while retaining that which makes music so uniquely affective. This subchapter will present a repertoire of tools (that are used in many contemporary and adaptive music systems) which will serve as components that, in conjunction, are likely to result in the creation of emergent music. To achieve our aim these components will be used in combination to create the score for a video game. Though the following tools may present themselves as simple, when combined they have the potential to produce scores that respond to and prompt affect in players while presenting unique musical experiences to individual players.

A) Horizontal resequencing

Horizontal resequencing is a compositional tool used to dynamically cue up, and transition between, different sections of music based on state changes (Sweet, 2014). Traditionally musical sections are played sequentially in an order defined by the composer. A composer might score each section of a song separately and position the sections in an order that is most in line with their desired outcome for the piece as a whole. When making use of horizontal resequencing, however, the musical order is dynamically altered to suit the requirements of a desired affective state. In the context of this study, horizontal resequencing can be viewed as the act of switching the order of musical sections to keep game scores synchronised with player affect. If the player encounters an enemy and the game switches from the exploration phase to the combat state, the music should transition to a section of the score that matches the aesthetics and intensity of the encounter. This allows for the existence of dynamism and variability while maintaining a cohesive gaming experience.

Unlike crossfading, where a piece of music immediately fades away while the next piece fades in, horizontal resequencing ensures that the musical transitions happen smoothly and remain intact. When making use of this tool, listeners will not be able to identify where the seams of the musical sections are positioned and will hear the combined segments as a unified entity, unaware that a transition has occurred.

A limitation of this tool, however, is that the score cannot respond to state changes immediately, unless they happen at the exact moment that a musical section reaches its end. Meaning it will likely take a number of seconds before a transition actually occurs. This is because horizontal reorchestration techniques, such as horizontal resequencing, alter musical material by switching between stems at predefined intervals or predetermined cue markers to ensure that transitions occur: on beat, at the end of a bar, at the start of a specific module, or at the end of a predefined musical phrase within the melodic structure of the score.

These transitions must transpire in this way, as violations of a listener's musical expectancy make it difficult to effectively provoke desired emotional responses in players. The type of scores typically associated with emotionally driven masterpieces, are difficult to achieve within the temporal uncertainty of a video game, where player affect can change the game state at any time (Collins et al., 2014). In other words, by making horizontal resequencing more responsive to player action, we make it harder to create scores that affect players according to a specific emotional goal. The more immediate the response to player affect, the more variability there is in the score. This results in more meaningful play, but less control over the generated outcomes (and the meanings thereof).

This, however, does not stop developers from cuing up sections of music before gameplay events have even occurred; meaning that, when creating an emergent music system, using horizontal resequencing, developers can pre-empt changes in the game state and trigger a reorchestration before an event has

even occurred. This will allow for musical shifts to happen in sync with game state changes, while ensuring that transitions happen smoothly and that the structural integrity of the score remains intact.

With all of this in mind, it makes sense that the most common method of horizontal resequencing involves slicing music into short segments, as the beginning of each segment serves as an entry point to other segments of the track (Evans, 2019). The smaller the pieces are, the more opportunities there are to transition from section to section and the easier it is for the music to respond to state changes quickly, without violating musical expectancy to a large degree.

In summation, horizontal resequencing is a powerful tool that can be used to create variety and interest within musical scores. Composers can dynamically adjust, reorchestrate, and create new versions (or mixes) of songs and melodies without sacrificing the cohesion of the piece. Though this tool does not allow for immediate feedback, it does permit composers to experiment with different combinations of musical elements and create emotively adaptive works that respond to player affect.

B) Vertical layering

The next tool, Vertical Layering (also known as vertical remixing or reorchestration), is a technique that makes use of different musical tracks that are layered above and below one another to alter a score's aesthetics, intensity, and overall musical feel (Sweet, 2014). In contrast to horizontal resequencing, tracks are not cued up to play in succession. Instead, each layer is played simultaneously in sync with the other layers. Vertical layering allows composers to control which layers of music are heard, and which are muted. In other words, vertical layering is a way of describing groups of synchronised music tracks that play back simultaneously, with one or more sounding while the rest are muted (Thomas, 2016).

Typically, the state of the layered tracks (and their intensity) is dependent on the preferences of the composer, however, these parameters can be linked to in-game events and state changes as well. For example, in a game where players are trying to prevent a bomb from exploding, the music could be altered based on how much time they have left before it detonates. By making use of vertical layering, developers could add, or remove, layers of percussion based on how many seconds remain on the clock. The less time the players have, the more percussive elements are added to the score, increasing the affective intensity of the scenario.

Many contemporary music systems make use of this tool to add dynamism to in-game scores. An example of this is demonstrated in Final Fantasy XV (2016). After making their way to the garage in Leide owned by Cid Sophiar, players are able to visit the Hammerhead Diner. When standing just outside of the diner, players can hear a soothing acoustic melody of instruments. Once the player enters the diner the instruments change and the feel of the ensemble transitions from smooth and tranquil to vibrant and upbeat. The acoustic guitars from the original mix are replaced by electric guitars, and a bass and drum kit are added to the mix. The musical structure of the tracks remains the same, however, the aesthetics and stylistic embellishments convey a different mood to the listener. By adding and subtracting layers from the mix, the music can seamlessly transition between different affective states without changing, sounding out of place, or violating the listener's musical expectancy. A limitation of vertical layering, however, is that the tempo and harmonic structure of the music must remain static and cannot be altered based on the game state (Evans, 2019).

Vertical Layering is useful as it allows for varied musical styles and seamless transitions. This tool allows for responsive alterations that provide immediate feedback to players, and it enables composers to create variations of scores that use different instruments and elements to produce a number of different emotive styles that can be transitioned between based on player affect. Provided that the

layers have similar temporal or harmonic frameworks, composers can dynamically alter, or transition between, mixes and tracks to induce and respond to player affect without sacrificing the cohesion of the piece.

C) Stingers

Stingers are short pieces of music or sound effects that are used in video games to punctuate events and transitions. These short bursts of music can be created to match specific in-game events and can be triggered to play from silence, or added to present musical content (Thomas, 2016). Unlike vertical layering stingers do not remain muted (and playing) while they are not in use. Instead, stingers are played at the exact moment that a particular event or state change occurs. They are crucial elements in the audio design of video games. Their use can greatly enhance the play experience of a video game, as they are effective in signalling gameplay changes to players in real-time. However, since stingers can be triggered at a moment's notice, developers and composers should remain aware that stingers can be triggered, and hence played, out of time. If not well accounted for, this is likely to create dissonance against the present harmony of the rest of the music (Evans, 2019).

Stingers serve as a form of auditory communication by letting the players know that certain events have been triggered in-game. The responsiveness of a stinger can induce a sense of immediacy and gratification in players, as it signifies that their actions have a tangible effect on the game world. As such, the use of stingers can help to keep players engaged and make them feel more immersed and emotionally invested in the game.

Stingers are excellent tools for elegantly conveying crucial information to players, without interrupting their gameplay experience. For example, a video game may be designed to play a stinger when players are low on health or when they have alerted a particularly challenging enemy. This can heighten the sense of danger and encourage the player to alter their behaviour without interrupting the visual experience of the game.

In addition to their ability to convey gameplay states, stingers can also be used to transition between different musical sections and tracks. For example, a composer may design a stinger that acts as an opening for a transition to occur. This helps to keep the player immersed in the game world even as they transition between different sections of the game and its accompanying score.

Overall, stingers are an important part of the audio design in video games. They help to punctuate events, communicate action, and smooth out transitions. By using stingers effectively, game designers and composers can greatly enhance the play experience and make the game more immersive, engaging and emergent.

3.2 Compositional strategies

In their paper *Designing a Game for Music: Integrated Design Approaches for Ludic Music and Interactivity* (2014: 2), Collins et al. argue that the evocation of "peak" emotions in response to music is directly linked to the confirmation, violation or creation of musical expectancy. For them, dynamic music provides a means to interactively shape the play experience by adding dynamism to scores, in hopes of further elevating the emotional responses of players. However, they stress that it is integral that dynamic music maintains a cohesive flow of musical expectancy while remaining interactive.

Though it is possible to trigger a number of emotions through genres that are not heavily expectation-based, many contemporary video games are expected to have an excellently-scored, emotionally driven, Hollywood-styled soundtrack (Jackson 2011). The type of music typically associated with well-crafted video games is distinctly expectation-based, and is unsuited for the temporal ambiguity of in-game state changes. This makes it difficult to achieve the level of dynamism needed to respond to the

affective state of players, without interfering with the evocation of peak emotions produced by well-scored music.

With this in mind, it is time to proceed to the second step of making video game music more responsive to affect. Now that we have a compendium of musical tools that promote dynamism within musical scores, we can formulate strategies to optimise the use of our compositional tools. Though each of the aforementioned tools have the capacity to increase interactivity and responsiveness, each of them is most affective when used in a way that maintains a high level of musical expectancy when added to in game scores. It is becoming increasingly clear that there is an inverse relationship between musicality and dynamism. The more musical systems adhere to musical expectancy, the longer the delay between game state changes and the music's response (Collins 2007).

In this subchapter, we will compile a number of strategies for maintaining musical expectations within our compositional tools, by restricting responses to state changes and score transitions to musically appropriate times. Through the use of curated musical transitions, and well-choreographed game state changes, it is possible to retain a high level of musical refinement while making use of our musical tools in ways that promote dynamism and enhance meaningful play by responding to player affect. The goal for this subchapter is to present ways of sustaining musical continuity while allowing for notable variations in instrumentation and texture; to provide players with scores that respond to, and are formed by, their play experience.

A) Adaptive quantisation

Quantisation is often viewed as the basis of rhythmic structure (Pichlmair & Kayali, 2007). It allows for the alignment of musical components to a predetermined grid. In digital audio production, quantisation helps to align the timing of musical events, such as notes played by a musician, to a grid of equally spaced musical beats or ticks. This process assists in the correction of unwanted timing deviations, inaccuracies, or inconsistencies in order to achieve a tighter, more consistent rhythm. While quantisation can limit a musician's expression, it greatly improves the music's adherence to structure and its listenability. The grid of musical beats can be adjusted to match different musical styles (such as swing or straight), and the level of quantisation can be altered to retain more, or less, of a human feel in the performance.

In game design, there is a concept dubbed the "threshold problem" by Clint Hocking (2012). The threshold problem refers to situations where discrete in-game state changes occur at arbitrary points during the play experience. These changes can result in frustrating experiences for players if they are not aware of them, or if they do not understand the rules that govern and/or trigger them. To mitigate these issues, Hocking suggests that game designers should strive to clearly communicate systemic boundaries to players or make those boundaries sticky (Hocking, 2012), so that the player can anticipate the change as well as where it will occur. While the threshold problem does not directly influence our ability to meaningfully score music for video games, it does highlight one of the key challenges of scoring music in real-time. To emergently score a video game, the music must reflect the current state of the game as well as the affective state of the player, both of which can change at any point regardless of whether it is a musically appropriate time for the score to transition or not. This makes it extremely difficult to adhere to musical expectancy as the state changes can occur outside of the temporal framework of a song.

To solve this problem, we can employ Hocking's suggestion of making state changes sticky, so that they magnetically snap to musically appropriate moments. By using the principles of quantisation to align the timing of in-game events and state changes (such as the introduction of a boss fight, or the death of the player character) to a musically influenced grid, developers and composers can ensure that state

changes only occur at musically appropriate moments. This allows for the emergent scoring of music that adheres to the player's musical expectancy. In other words, by adaptively quantising the temporal positions of in game events to musically appropriate moments, developers and composers can pre-empt affective state changes and ensure that they only occur at points that make sense within the rhythmic structure of the score.

If we design our games to perform state changes within variable ranges, rather than executing them according to strict parameter changes, we can enhance the evocation of peak emotional responses to in-game events by matching the affective aesthetics of the state change to the emotive properties of the score. For example, in an endless runner where players need to collect 50 orbs to unlock the next stage, the game could be designed to trigger a state change when the player has collected 50 glowing orbs. After triggering the event the player character could immediately begin to glow and a musical stinger could play, to signify that they have achieved their goal. Assuming there is no visual counter to display how many orbs they have collected, the glow around the player character as well as the musical stinger, would be the only indication to the player that they have collected enough orbs.

Now let's attempt to transition from the current section of our musical score, to a section that plays a triumphant brass melody when the player accomplishes their goal. While triggering the event at the exact moment the player collects their 50th orb is functional, it is not something that can be easily scored for without violating the player's musical expectancy (as the player could potentially collect their 50th orb at a moment that is musically inappropriate for a transition to occur).

To make scoring this hypothetical game easier, let's assume that the player character moves according to the rhythm of a 4/4 time signature[3]. If we continue to trigger the state change when the player has collected their 50th orb, it is still possible that the player might collect their final orb on the second, third or fourth beat of the bar (all of which would be an inappropriate time to trigger a horizontal or vertical transition).

Rather than finding a way to ensure that players can only collect their 50th orb at the beginning of a bar, we can make use of adaptive quantisation techniques to make the event trigger more malleable. By designing an adaptive quantisation system, instead of relying on rigid parameter governed state changes, we can produce emergent scores that recognise that a state change is about to occur, observe that the music is currently at the second beat of a bar and (assuming that it is aware that the best musical transition point would be on beat one), could delay the triggering of the state change until the player has collected an orb at a musically appropriate time. This would ensure that the ambient glow, musical stinger and triumphant brass melody cue at the exact same time.

Furthermore, the use of adaptive quantisation would allow for the score to even pre-empt the player picking up their last orb, meaning that, if the player picks up an orb at a musically appropriate time, the event could actually occur slightly earlier than the 50th orb (e.g., the 48th orb).

Though the method described above is not the only way adaptive quantisation can be used to align in game events to musically appropriate windows, it does raise the concern that Hocking's threshold problem may be further amplified, as the boundaries between state changes do become less rigid if they are coupled to variable parameters. However, there is a growing consensus which indicates that, as long as players can recognise that their actions have been acknowledged by the game, they are willing to accept a slight delay in feedback (Kastbauer 2011 as cited in Collins et al., 2014).

[3] The 4/4 notation indicates that each bar in the piece has four beats (numerator), and all of them are quarter notes (denominator) (Gungormusler et al., 2015).

In conclusion, by employing the principles of quantisation to align the timing of in-game events to a musically influenced grid, developers and composers can ensure that state changes only occur at musically appropriate moments, leading to the emergent scoring of music that adheres to the player's musical expectancy. The use of adaptive quantisation techniques can further enhance the evocation of peak emotional responses to in-game events by matching the affective aesthetics of the state change to the emotive properties of the score. This approach can even be used to allow the scores to pre-empt the player's affective state changes, ensuring that they only occur at points that make sense within the rhythmic structure of the score.

B) Sequenced play

Thus far, we have deduced that increased musicality (which refers to the adherence of video game systems to musical expectancy) can create a consistent and desirable musical experience for players. The elicitation of affect in response to in-game scores is directly linked to the confirmation or violation of a player's musical expectancy. Furthermore, an increase in dynamism can create a more interactive and engaging experience for players, as the game responds to their actions in real-time. However, there is an inverse relationship between musicality and dynamism, meaning that the more musical systems adhere to musical expectancy, the less responsive they are to the player's actions. This makes it extremely difficult to nurture a symbiotic relationship of affect between games and players using music. In order to evoke emotions in players, game systems need to be crafted with musicality in mind. However, in order for games to effectively respond to player affect, they need to react to player driven events and state changes in as little time as possible.

The challenge then, is to find a balance between musicality and dynamism, in order to create an immersive and emotionally resonant experience for players, that responds to their actions and behaviours. Fortunately, musicality and dynamism are not mutually exclusive. Though difficult, it is very possible to align state changes to musically appropriate moments, as demonstrated in the section above. This approach, however, is still quite limiting as adaptive quantisation techniques can only align state changes and musical transitions after an inciting event has occurred. This means that unless an event occurs on a musically appropriate transition point, or an event takes place before the player even performs the action to trigger said event, quantisation can only occur after player input has been registered. This ultimately results in a delay of feedback, which is not ideal for emergent music.

In the example of the hypothetical endless runner detailed in the section above, we assumed that the player character moved according to the rhythm of a 4/4 time signature. This is because 4/4 is the most common time signature in contemporary music (Gungormusler et al., 2015). Moreover, constraining the player character's movement to a format that identically matched the temporal structure of the accompanying music, made it far easier to respond to player affect in a timely manner while adhering to musical expectancy.

Through the use of the aforementioned design constraint, it is made clear that forming the mechanics of play in ways that mirror musical structure, allows for video game systems that can respond to player affect in a timely fashion, while still conforming to the temporal framework of a well-crafted score. In other words, if we design our systems of play to match the rules that govern our music, we can align player inputs to musically appropriate windows before they even occur. The player will never experience a delay in feedback due to quantisation, if they can only perform inputs at musically appropriate times.

Although sequencing the play experience to mirror the framework of a score does limit player expression, much like how quantisation can limit a musician's expression, it greatly improves a video game's adherence to musical structure and drastically increases said music's ability to respond to player affect. Furthermore, this type of play constraint is not an untested one. Games like Guitar Hero (2005),

Sing Star (2004) and Dance Dance Revolution (1998), all base their mechanics around the player's ability to play the game in a way that aligns with the formal structures of a song. Moreover, all three of these games offer limited amounts of expression to their players, yet players still feel engaged enough to play them repeatedly (and often for extended periods of time). Expressing yourself in Guitar Hero means playing flawlessly and showcasing a spectacular performance. This illustrates that though some video game systems may be concise and restrictive, there is often still enough agency within these systems to keep players engaged. Jenkins (2005) states that when children play games, they sway and move with the figures depicted on the screen; they bounce and remain totally engaged with the gameplay. Hence, it is fair to argue that such responses reflect the sense of control that players feel over what they experience in game (Jenkins, 2005). Furthermore, it showcases that video games can still present a strong evocation of affect, even when player input is constrained or restricted.

Making use of sequenced play to align player driven events to in-game scores, means that our compositional tools can be used to respond to player affect without unintentionally violating the player's musical expectancy. Take stingers for instance. Stingers are excellent tools for elegantly conveying crucial information to players, without interrupting their play experience, however, since stingers can be triggered at a moment's notice, they can be played out of time, which is likely to cause dissonance in contrast to the harmony of the rest of the score. By confining player input to sequenced intervals, however, we can ensure that stingers are always triggered within the temporal framework of the score, while retaining their ability to respond to player affect immediately.

What this implies is that, it is possible to reduce the amount of temporal ambiguity of in-game state changes, to increase the video game's adherence to musical structure. By constraining player input to methods that are in alignment with musical structure, it becomes more feasible to achieve a balance between musicality and dynamism.

In summation, emergently crafting an affective emotionally driven musical experience that responds to player affect, involves finding a balance between musicality and dynamism. While musicality can evoke player affect in response to in-game scores, and create an emotionally resonant experience for players, it can limit a video game's responsiveness to player actions. On the other hand, dynamism can increase player engagement and interactivity, but it reduces the overall adherence to musical structure. By aligning the mechanics of play with the frameworks that govern music, it is possible to ensure that state changes occur at musically appropriate moments and respond to player affect in a timely manner. Although sequenced play may limit player expression to some extent, it still offers a strong sense of control and engagement to players, while reducing the amount of temporal ambiguity in video game state changes. Overall, designing video game systems to mirror musical structure can lead to a more immersive and emotionally engaging experience for players, without enlarging the delay between player input and systemic feedback.

C) Dynamic layering

In the section above, we developed a means to achieve a balance between musicality and dynamism. Through making use of sequenced play, we were able to formulate a strategy for aligning player driven events to in-game scores. By restricting player input to windows that conform to the temporal requirements of a video game's score, we enable our compositional tools to respond to player affect in a timely manner (without unintentionally violating the player's musical expectancy). This greatly aids in creating an immersive and emotionally resonant experience for players, that responds to their affective state.

However, while sequenced play can provide a means of aligning player input with a game's musical score, it can also create a sense of rigidity that may detract from the immersive experience that we aim to provide. Additionally, it limits the types of gameplay that can be presented to players. While

restricting player input to designated windows can promote musicality, it can also limit the range of player expression and the types of gameplay that can be incorporated into a game. Many contemporary video game genres would struggle to pivot their mechanics to systems that only accepted player input at quantised intervals, as most video game mechanics cannot be altered to align with musical structure without extensive modifications or complete redesigns.

Restricting player input to designated windows, or confining state changes to quantised time intervals, promotes musicality but greatly limits the formats of player expression. This is particularly true for genres that rely heavily on player agency and the ability to make split-second decisions, such as racing games and first-person shooters. For these types of video games, the satisfaction produced by musical scores synchronising with game state changes, may not be enough to compensate for the frustrations that the restraint of player control is likely to induce (Collins et al., 2014).

Fortunately, video games are made up of a plethora of systems that operate in union. While there are limitations to using sequenced play and adaptive quantisation on every one of a video game's mechanics, it is important to note that these techniques can still be effective when used sparingly. One potential strategy for incorporating musicality into video games without limiting player agency is to use these techniques selectively, focusing on specific moments or events in the game where musicality can enhance the play experience.

Nevertheless, the challenge remains: how can we score player driven events immediately without forfeiting player expression or adherence to musicality? For video game mechanics that require split-second decision-making, there must be a method of scoring gameplay without resorting to adaptive quantisation or sequenced play, which are not ideal as they may result in restrictive gameplay or delays in feedback.

To address this problem, we can refer back to our compositional tools. Vertical Layering is a technique that uses different musical tracks layered above and below one another to augment a score's aesthetics, intensity, and overall musical feel (Sweet, 2014). Unlike horizontal resequencing, vertical layering allows each layer to play simultaneously in sync with the other layers. This enables composers to control which layers of music are audible and which are muted. Moreover, as mentioned previously, the state of the layered tracks and their intensity can be linked to in-game events and state changes.

By employing vertical layering, developers can add or remove groups of instruments in real-time, enabling players to hear a single piece of music with a variety of different arrangements based on the game state. This adds dynamism to the score, evokes affective responses from players, and improves the overall play experience by presenting a meaningful response to game state changes. Additionally, vertical layering offers the advantage of crafting music without horizontal restrictions. The only limitation that vertical layering presents is that, any collection of layers must be able to play together without affecting the musicality of the score. Fortunately, this limitation is well accounted for, as the different instrumental parts of a musical piece are generally constructed to play at the same tempo, with complementary notes sounding at any given time (Plut & Pasquier, 2020).

However, scoring player-driven events immediately without forfeiting player expression or adherence to musicality remains a challenge, merely turning layers on or off when state changes occur is not enough to address the issue. In order for vertical layering to solve this problem, musical layers would need to be dynamically muted and unmuted based on player input, rather than state changes, allowing players to retain their freedom of expression while preserving the musicality of the score. For example, in a racing game, rather than unmuting a percussive layer as the player crosses the finish line, developers can turn the percussive layer on when the player holds down the accelerate button. This

approach of dynamically muting and unmuting musical layers based on player input can be more deeply integrated into the video game's systems. This could aid in creating a more personalised experience for players, where the game's score adapts to their playstyle, and the sounding layers alter based on their affective state.

Overall, the use of dynamic layering to respond to player input offers an effective solution to the challenge of scoring player-driven events. By using vertical layering selectively and dynamically muting and unmuting layers based on player input, developers and composers can create responsive scores that react to the actions of the player without sacrificing musicality or freedom of expression. Through the incorporation of these techniques, developers and composers can create a more immersive and emotionally resonant experience for players, that responds to their affective state and promotes player agency.

D) Dynamic Mixing

In the three sections above, we developed tools to aid in the construction of music that responds to player action. The first tool explored is Horizontal Resequencing, which involves dynamically cueing and transitioning between different sections of music based on state changes. This helps maintain cohesion in the gaming experience while accommodating shifts in player affect. The second tool, Vertical Layering, showcased how different musical tracks layered simultaneously can alter a score's aesthetics and intensity. This allows for immediate feedback to players and seamless transitions between different affective states. Lastly, we discussed Stingers, short bursts of music or sound triggered during specific in-game events. Stingers serve as auditory cues, enhancing player engagement and signalling important changes.

While it is not as intricate as the three above, there is one more tool that can be used in tandem with all three of the aforementioned. Dynamic Mixing, our final compositional tool, is one that focuses on the real-time manipulation of audio levels (volume and loudness), to create an evolving sonic landscape within our scores. As opposed to leaving the audio levels of tracks static, Dynamic Mixing allows composers to shape the auditory experience based on player actions and in-game events.

Employing Dynamic Mixing allows composers and developers to modulate the volume and loudness of individual elements within the music in response to the changing dynamics of the game. For instance, during intense gameplay moments, one could dynamically amplify percussion and instrumental layers to heighten the sense of urgency. Conversely, during quieter exploration phases, the mix could delicately balance softer elements, fostering a more contemplative atmosphere.

Dynamic Mixing is extremely useful because of its adaptability. Composers can define specific triggers or conditions within the game that prompt changes in the mix. This may include shifts in player emotion, alterations in the environment, or significant narrative events. By doing so, the music becomes not just an accompaniment but an active participant in the player's emotional journey.

One challenge, however, is finding the right balance between responsiveness and cohesiveness. An abrupt change in volume may disrupt the overall listening experience for the player, interrupting their immersion. Therefore, careful calibration is necessary. This tool requires keen experimentation, encouraging composers to craft a mix that not only complements the game's narrative and emotional journey but also responds dynamically to the player's actions.

In summary, Dynamic Mixing empowers composers to dynamically shape the auditory landscape, using changes in volume and loudness to synchronize the music with the evolving emotional and gameplay states, contributing to a more engaging and responsive video game music experience.

Interlude: The story so far…

So far in our investigation, we have deduced that video games and music are both affective and emotionally driven. Video games can elicit specific emotional responses from players, while music involves our senses, moves the body, and generates associations with spaces, images, and textures. Affect is rhythmic (Anable, 2018) and requires both motion and stillness, much like music, which informs our behavioural responses and enables us to convey and interpret meanings through what we feel. Unlike video games, however, music is synesthetic, and we experience sound in association with what we emotionally feel.

Through this investigation, we have inferred that sound is schizophonic and has a strong semiotic relation to other modalities. Each individual sound holds a number of connotations that could infer a plethora of meanings depending on the listener and the context. Therefore, when attempting to create emergent music, it is important to consider the context in which the music will be heard. We must ensure that the score only permits interpretations of the music that were intended, while still allowing for responses to dynamism and player expression.

To create a cohesive and immersive gaming experience, music must be designed in conjunction with the gameplay it will support. This requires paying close attention to the rhythm, melody, instrumentation, emotional content, and associations of the music. Additionally, video game music must be scored in relation to its game worlds and the emotional intent thereof, as they are key aspects of the compositional process.

However, we have found that adding dynamism to video game music is challenging, as player agency often conflicts with musical structure. Making music more responsive to player action makes it harder to create scores that affect players according to a specific emotional goal. This is because music follows its own structural formulae which are unsuited for the arbitrary rate that in-game state changes occur in. Although it is possible to create scores that align perfectly with the temporal flow of game events, many contemporary video games are expected to have emotionally driven soundtracks that align with the player's understanding of conventional musical tendencies. This makes it difficult to achieve the level of dynamism necessary to respond to the player's affective state, without interfering with the transmission of affect produced by well-scored music. To help video game scores become more emergent and responsive to affect, we forged musical tools that promote dynamism in scores. Although each tool has the capacity to increase interactivity and responsiveness, each is most effective when used in ways that keep the player's musical expectancy intact. This helps to maintain a balance between predictability and variability. Emergent music should respond to player action while retaining its ability to prompt player action.

At the same time, designing skilfully crafted music to achieve specific emotional goals is the only way to ensure that the schizophonic nature of sound is well accounted for. When aware of the context in which their music will be heard, composers are able to weave affect directly into their work to achieve a desired emotive response from their audience. In the video game music industry, there is a widely accepted belief that higher levels of interactivity correspond to greater requirements for modularity in musical structure. As such, a focus on segmented forms of music can facilitate more rapid responses to in-game events while maintaining a cohesive musical experience. This approach ensures that the music consistently supports gameplay and remains of a reliable quality. A cohesive musical experience, however, does not equate to an emotionally driven soundtrack. This method of composition introduces

predictability and repetition, making it difficult to create affective musical moments for unscripted gameplay events.

In short, creating video game music that is responsive to player action while achieving specific emotional goals requires a delicate balance between predictability and variability. Our musical tools can help promote dynamism in scores, but it is important to maintain a cohesive musical experience and ensure that the schizophonic nature of sound is well accounted for. Designing segmented forms of music can facilitate more rapid responses to in-game events, but can also introduce predictability and repetition, making it difficult to create affective musical moments. In the final chapter of this paper, we will apply our findings by creating a video game with emergent music and analysing the merits and shortcomings thereof.

Dear Reader,

I would like to thank you for making it this far into my paper. If you'd like to experience the game itself. You can do so by simply following one of the links below:

https://rodwinmalinga.itch.io/allegory

https://drive.google.com/file/d/1mAoGOhnn1-JYQdepmX7FG65DPqMNL2Zs/view?usp=sharing

Track 4: Creating emergent music

Up until this point, the goal of this paper has been to explore the creation of emergent music in video games. Moreover, it has focused on developing systemic and compositional methods for in-game scores to respond to the dynamics of play, while maintaining the quality present in contemporary video game music. This, however, is only one aspect of my investigation. This study was conducted in conjunction with the development of a video game called Allegory (2022), which serves as a platform for testing various methods of forming and implementing emergent music.

Allegory is a puzzle platformer that places a deep emphasis on music. The game employs a fusion of 2D hand-drawn animation and 3D environments to create a unique visual style that matches the emotional aesthetics of the score. The game was developed, in collaboration with Daniel Hartwell and Erik Prinz, as part of our Master's research.

When creating Allegory, my primary goal was to produce music that is, to some extent, composed in real time, with the player as an active participant in the production process. The music created needed to be complex enough to deliver a number of diverse outcomes, while retaining the emotive quality of statically scored music. To achieve this, our team used the Unity game engine, and built a custom audio engine to serve as an emergent music system. All of the music and sound effects heard in Allegory were created using GarageBand on iOS.

By reflecting on the development of Allegory, I aim to identify where emergent music struggles as an alternative to contemporary music systems in video games, as well as unravel why video game music has fallen behind in terms of dynamism when compared to other media that exist within video games. By using the game as a testing ground for different methods of constructing and composing emergent music, I hope to contribute to a broader understanding of the role that music can play in video game design, and to help pave the way for new and innovative approaches to game music composition.

One of the main advantages of using emergent music in Allegory is the dynamic and interactive nature of the music, which can create a deeper sense of player agency and engagement. The music in Allegory is responsive to player action, which helps to create a unique experience for each individual player. Additionally, the use of emergent music allows for a more seamless integration of music and gameplay, as the music is dynamically orchestrated and can adapt to changes in the game state. This chapter will serve as a discussion on the process of creating emergent music for Allegory, and will analyse the strengths and weaknesses of this approach.

4.1 Emergent music system

In the previous chapter, we formulated compositional tools and strategies for the creation of emergent music. When developing Allegory, we employed these tools and techniques to develop Allegory's Emergent Music System (EMS). They were modified, enhanced, and combined into systems that we utilised alongside our video game mechanics. Both the music and game mechanics were designed from scratch to function in tandem, for a unified and seamless play experience.

The compositional tools we used, were applied in combination to create the score for the video game. Our curated musical transitions, choreographed game state changes, and alignment of game mechanics and musical structure contributed to a video game experience where the music is an integral element of play, rather than an afterthought. Though some of the techniques appear simplistic, their integration served to produce a score that elicits affective responses from players while simultaneously providing them with unique musical experiences.

This subchapter will provide a detailed account of the methods and techniques used in creating Allegory's EMS, which allowed for a dynamic and interactive musical experience that responds to the actions and choices of the player. Allegory's EMS highlights the intersection of musicality and dynamism, showcasing the potential of our compositional tools to create engaging and immersive gameplay experiences.

A) Sequenced stingers

Systemically stingers act as an excellent tool for auditory communication in games. They effectively inform players when specific events have been triggered or activated. Since they are short, stingers can be played immediately in response to player action, providing an instant sense of feedback. In Allegory, we utilised stingers to signify when a significant event had occurred, such as when a key[4] was activated or when all keys in the level were activated to unlock the next stage door. These stingers afforded players with a sense of immediacy and gratification, indicating that their actions have a tangible effect on the game world.

The use of stingers not only kept players informed about events but also helped present key information without unnecessary visual clutter on the screen. By playing these short, distinct sounds, we avoided the need to rely on visual cues alone, ensuring players could remain fully immersed in the game. Overall, the implementation of stingers in Allegory greatly enhanced the player experience by providing timely feedback, increasing engagement, and supporting a more streamlined gaming experience.

However, we also recognised that incorporating musicality into the game could enhance the player experience. Research has shown that in-game scores can elicit emotional responses in players by

[4] By key I am specifically referring to the jewel shaped objects found in Allegory, that need to be activated in order to unlock doors. These should not be confused with musical keys, as I am referring to the game's mechanical elements in this instance.

confirming or violating their musical expectations. To achieve this, we introduced Sequenced Stingers, which combine stingers with sequenced play.

Sequenced play allowed us to ensure that stingers were always triggered within the temporal framework of the score, while still allowing for immediate response to player actions. However, we did not want to limit player expression, as this could greatly impact gameplay in a puzzle platformer like Allegory. Instead of using sequenced play to align player actions with the in-game score, we developed a terminal mechanism that aligned a player-driven event system to the score. This allowed players to retain control over their actions while still experiencing the benefits of a consistent musical experience.

In Allegory, terminals are interactive level elements that allow players to interact with nodes that appear when the terminal is approached. The terminal cycles between nodes at a fixed BPM in accordance with the tempo of the song. When the terminal cycles to a node, it becomes active (and enlarges) so that players can toggle it on or off. Each node that is toggled on plays a stinger when the terminal cycles to it, which aligns with the temporal structure of the current song and sounds harmonically with the rest of the score. To ensure that player actions are registered, their inputs are made sticky (Hocking, 2012), meaning that their input toggles the active node regardless of whether their input was on beat or not. This addresses the threshold problem and ensures that the player's intentions are always acted upon.

After leaving the terminal, the nodes remain as they were, and play melodies that were constructed by the player. This approach allows sequenced stingers to respond to player affect without unintentionally violating the player's musical expectancy or interrupting the player's expression. In fact, it increases player expression as players can customise a piece of the score to match their affective state without causing any melodic dissonance.

In Allegory, the terminal serves a crucial role in promoting gameplay. Many of the level elements in the game have variable states. Platforms for instance, can be solid or intangible, and can change positions based on their state. These level elements can be linked to a terminal, with their variable states tied to the nodes in the terminal. When a node in the terminal is turned on and becomes active, it can trigger a state change in any of the level elements connected to that terminal. For example, a platform can switch from being solid to intangible and vice versa when a sequenced stinger plays. This allows players to stand on the platform whenever the music causes it to toggle to the tangible state. The exciting part is that all of these events occur on beat and are fully customisable by the player themselves (who can choose which nodes to turn on or off).

It is important to note that turning off a node does not trigger any effects. It has been intentionally designed this way, as affect and music are rhythmic (Anable, 2018) and require both motion and stillness to invoke behavioural change. This deliberate design ensures that the gameplay experience is both dynamic and emotionally resonant for players.

Crafting an emotionally driven musical experience that responds to player affect requires striking a balance between musicality and dynamism. Through the use of sequenced stingers, we were able to achieve this balance. Though a portion of the gameplay experience may be restrictive while the player is interacting with the terminal, once they leave, the game becomes just as expressive and responsive as a conventional puzzle platformer. The added benefit of our approach is that it features a musical system that responds to the player's actions and affect, making the experience even more emotionally resonant. Overall, the integration of sequenced stingers into the game design allowed us to create a unique and engaging gameplay experience that was both musical and dynamic.

B) Zoning: Adaptive quantisation reorchestrated

By utilising adaptive quantisation, game developers can synchronise in-game events with musically appropriate moments to create a more cohesive and emotionally impactful gaming experience. This involves aligning the affective state in the score with the player's in-game actions to ensure that musical transitions occur only at points that make sense within the rhythmic structure of the score and preserve the player's musical expectancy.

Furthermore, by designing games to perform state changes within variable ranges, rather than executing them according to strict parameter changes, developers can enhance the induction of deep emotional responses to in-game events by matching the emotive aesthetics of a state change to the affective properties of the score. By employing the principles of quantisation to align the timing of in-game events to a musically influenced grid, developers and composers can ensure that state changes only occur at musically appropriate moments, leading to the emergent scoring of music that adheres to the player's musical expectancy. This approach creates a more immersive and emotionally engaging gaming experience for players, as the music and game mechanics work together to evoke a range of emotions throughout gameplay.

With this in mind, our team utilised adaptive quantisation in our EMS to enhance the player's emotional engagement with the game. By predicting the player's current affective state and preparing the music accordingly, we were able to create a more immersive and emotionally impactful gaming experience. To achieve this, we developed a technique called Zoning, which involved highlighting specific areas within each level that would trigger specific events as the player moved through them.

Each zone was carefully selected and analysed to ensure that the EMS had enough time to queue a musical transition, as the game state changed, thus providing a seamless and musically appropriate window for the transition to occur in. This approach not only allowed us to respond to player affect linked to mechanical state changes, but also to player affect that was not linked to such changes, thereby increasing the emotional depth of the game.

In order to illustrate how our zoning technique worked, let's take a closer look at a specific example. We positioned certain zones in areas that the player could only reach if they failed to make a crucial jump. In such a situation, the player would have no choice but to backtrack and retry the jump. When this occurred, the EMS was prompted to queue a transition to a more sombre section of the score as the player landed, since they would likely be experiencing a great sense of disappointment upon missing the jump.

Conversely, the zones were also designed to prompt the EMS to transition to brighter and more uplifting sections of the music as the player approached the jump on subsequent tries. As the player began to feel a sense of determination upon approaching the designated jumping area, the music would reflect this emotion and create a more engaging and emotionally impactful experience.

By using the zoning technique and adaptive quantisation, our team was able to create a more deliberate and emotionally impactful play experience. This approach allowed us to respond to the player's emotional state in real-time and enhance their affective engagement with the game. These techniques can be applied to other games, affording developers the potential to create more immersive and emotionally impactful experiences for players.

C) Dynamic layers

Video games are complex creations that rely on numerous components working in unison. Although there are limitations to utilising sequenced play and adaptive quantisation, these methods can still

prove effective if used with clear intent, as demonstrated in the sections above. However, they sometimes lack the dynamism needed to respond to player actions immediately.

One successful approach we took to incorporate more dynamism into Allegory's score involved using dynamic layering. Though far simpler to implement than the aforementioned tools, adaptive quantisation enabled us to confidently score undefined moments of gameplay without compromising player agency. By adding and removing musical layers in real-time, players were able to experience different arrangements of the score based on their actions. These dynamic layers not only added interactivity to the score, but also helped evoke emotional responses from players, ultimately improving their emotive experience.

When developing the EMS, we composed a vertical layer that could be muted and unmuted dynamically based on player input, rather than state changes. This allowed players to retain their freedom of expression while preserving the musicality of the score. The dynamic layer featured a female voice that sang in harmony with the other tracks in the game music. Regardless of when it was faded in or out, the dynamic layer remained melodic and in harmony with the rest of the music, giving players a cohesive and enjoyable audio experience.

Mechanically the dynamic layer was linked to player input to maximise the effect of dynamism on the score. Specifically, we linked the dynamic layer to the player's interact ability. By holding down the interact button, players could unmute the dynamic layer and have Aria, the player character, sing along with the music. When the player released the button (or Aria ran out of breath), the dynamic layer would be muted once again.

This approach added a sense of personalisation to the game's score, as it allowed players to control the layer's presence based on their continued input. Players could activate and release the button at any time, giving them a sense of agency over the music. Additionally, the singing would change in response to any changes in the score caused by adaptive quantisation. This was made possible by composing each section of the soundtrack with an accompanying dynamic layer that could be toggled on or off at any time. Despite the dynamic nature of the layer, it would remain melodically harmonic with the rest of the score, ensuring a cohesive and enjoyable listening experience for players. Overall, the use of dynamic layers in Allegory greatly enhanced the affective experience for players.

4.2 Reflecting on emergent audio

The subchapter above highlights evidence, albeit anecdotal, that the creation of emergent scores is achievable. These scores can respond to player agency and manipulate affect, resulting in deep emotional responses from players. While this approach to video game audio design is possible and realisable, it is important to acknowledge its difficulties, shortcomings, drawbacks, and limitations.

With that said, this final subchapter aims to reflect on the creation of Allegory, and its role in our investigation, providing a window into the current state of adaptive audio. It also aims to shed light on the distance that still needs to be covered before video games can make use of emergent audio to the fullest degree.

A) Challenges and limitations

The creation of emergent scores in video game audio has evolved significantly over the years. Games such as Tetris Effect (2018), Rez Infinite (2016), No Straight Roads (2020), and Hi-Fi Rush (2023) showcase a growing interest in developing video game scores with emergent elements. However, it is crucial to acknowledge that there are still several challenges and limitations that need to be overcome before truly emergent audio can be achieved.

The first major challenge is the complexity of designing EMSs that can respond to the wide range of player actions and choices that occur in a game. It is an immense task to create a system that can adapt to the vast number of possible in-game scenarios while still producing coherent and emotionally resonant music. This requires a deep understanding of both game design and music theory, and the systems need to be developed alongside game mechanics from the beginning of the design process. EMSs cannot be added to video games at the end of the development process, as is typically the case for contemporary video game audio design.

Additionally, there is the challenge of integrating emergent audio into the broader game design. Video games are multifaceted systems with numerous interlocking parts, and creating an EMS that fits seamlessly into this ecosystem can be a daunting task. In some cases, emergent audio can even interfere with gameplay if not well accounted for. Therefore, it is essential for us to carefully consider the integration of emergent audio within the broader game design to achieve the desired result.

Secondly, there are a limited number of resources available to developers. Emergent audio requires a great deal of investment to implement. Creating a robust EMS is likely to be time-consuming and expensive, especially for smaller studios with limited budgets. This implies that, in order for truly emergent audio to arise in a game, it needs to be well invested into from the game's inception. As a result, emergent audio is likely to be reserved for larger, more ambitious projects, or smaller games with a focus on this technology, and little else.

Finally, there are limitations to the types of emotions that can be evoked, and gameplay experiences that can be had, through emergent audio. While it is possible to create music that responds to player actions and elicits emotional responses, there are limits to the complexity of these responses. It is also difficult to create music that can respond to multiple emotions at once, which can limit the range of emotional experiences that can be offered to players. Additionally, some video game genres incorporate gameplay that is unsuitable for emergent audio. For example, games that rely heavily on scripted events or narrative-driven gameplay may be better off without the use of emergent audio. In such games, the music is usually carefully composed to match the mood and tone of the story, and too much of an adherence to dynamism may result in a rupture of player immersion.

In conclusion, while emergent audio has progressed significantly in recent years, there are still immense challenges and limitations that need to be overcome before truly emergent audio can be achieved. That being said, with continued research and development, it is possible that emergent audio could become an even more powerful tool for game designers, enabling them to create deeply immersive and emotionally resonant experiences for players, by tapping into the affective nature of music.

B) Assessing the sound

The music in Allegory was designed to respond to player actions while maintaining its musicality. Our use of emergent audio in the game was exciting, as it had the potential to enhance the player's experience in innovative ways. In many ways this goal was achieved. However, despite my best efforts, there are times when the music falls short and negatively affects the player's experience. This is partly due to the rudimentary nature of our EMS, which has limitations in terms of its complexity and adaptability.

Moreover, my lack of formal training as a musician or composer, and limited understanding of musical structure, has also contributed to the failings of Allegory's EMS. Without the expertise to create intricate musical compositions, I had to rely on simpler musical elements and practices to create emergent audio in the game.

As a result, while the music in Allegory showcases the potential of emergent audio, it also highlights the struggles of developing such a system without the necessary skills and resources. Nonetheless, the use of emergent audio in Allegory is a significant step forward in the development of interactive audio experiences in video games. As developers continue to refine and improve upon emergent audio technologies, and as more composers and musicians with formal training and expertise become involved in the process, it is likely that we will see even more impressive uses of emergent audio in future games.

In future, as the use of emergent audio in video games becomes more widespread, it is important for developers to consider the limitations and challenges associated with the technology. Investing in the development of robust EMSs from the beginning of the game design process, and integrating the emergent audio seamlessly into the broader game design, can help ensure that the audio enhances rather than detracts from the player experience.

Furthermore, in order to fully realise the potential of emergent audio, it may be necessary for developers to collaborate with experienced composers and musicians who possess a deep understanding of musical theory and structure. Such collaborations could result in more nuanced and sophisticated emergent audio experiences, capable of responding to a wider range of player actions and evoking more complex emotional responses.

Outro

Throughout our investigation, we have gained valuable insights into the process of creating emergent music for video games. We have also uncovered the reasons why video game music has lagged behind other forms of video game media in terms of dynamism. It turns out that creating emergent music is a challenging and resource-intensive task, which explains why it has not yet been widely adopted.

Despite all of that, this investigation has been a fruitful journey that has shed light on the intricacies of producing an integrated approach to emergent music composition. More importantly, this paper has made it abundantly clear that there is merit in pursuing and developing emergent music for video games.

In my opinion, this paper has been successful in contributing to a broader understanding of the role that music can play in video game design. Though it may be boastful to say, I believe that the findings and recommendations in this paper can inspire new and innovative approaches to game music composition, which may lead to more immersive and emotionally resonant gaming experiences for players. In short, I am optimistic that this investigation can help push the field of video game music towards greater dynamism and interactivity in the future.

References

Anable, A. (2018). *Playing with feelings: Video games and affect*. Minneapolis, MN: University of Minnesota Press.

Anderson, C. A., Shibuya, A., Ihori, N., Swing, E. L., Bushman, B. J., Sakamoto, A., ... & Saleem, M. (2010). Violent video game effects on aggression, empathy, and prosocial behavior in Eastern and Western countries: A meta-analytic review. Psychological Bulletin, 136(2), 151-173.

Ashby, S. (2008). *Interactive Audio for Video Games*. Paper presented at Concordia Electroacoustic Studies Student Association, March 20, 2008, Concordia University, Montreal, Canada.

Collins A. M., Loftus, E. F. (1975). *A spreading-activation theory of semantic processing*, Psychol. Rev., vol. 82, no. 6, pp. 407–428.

Collins, K. (2007). *An Introduction to the Participatory and Non-Linear Aspects of Video Games Audio.* In Essays on Sound and Vision, ed. Stan Hawkins and John Richardson, 263–298. Helsinki: Helsinki University Press.

Collins, K. (2013). *Playing with sound: A theory of interacting with sound and music in video games*. The MIT Press.

Collins, K., Kapralos, B., Tessler, H., Stevens, R., & Raybould, D. (2014). Designing a Game for Music. In K. Collins, B. Kapralos, & H. Tessler (Eds.), *The Oxford Handbook of Interactive Audio*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199797226.013.009

Crawford, C. (2003). Chris Crawford on Game Design. Berkeley, CA: New Riders.

Eladhari, M., Nieuwdorp, R., & Fridenfalk, M. (2006). *The soundtrack of your mind: Mind music-adaptive audio for game characters, in Proc*. ACM SIGCHI Int. Conf. Adv. Comput. Entertainment Technol., p. 54, doi: 10.1145/1178823.1178887.

Evans, R. (2019). *Analyzing and Designing Dynamic Music Systems for Games.* PhD thesis, University of York.

Gungormusler, A., Paterson, N., & Haahr, M. (2015). *barelyMusician: An Adaptive Music Engine*

Hocking, C. (2012). In the Click of It: Living on the Edge. Edge Magazine 241: 152.

Hunicke, R., LeBlanc, M., & Zubek, R. (2004). MDA: A Formal Approach to Game Design and Game Research. In Proceedings of the AAAI Workshop on Challenges in Game AI (Vol. 4, No. 1).

Hutchings, P. E., & McCormack, J. (2020). *Adaptive Music Composition for Games*. IEEE Transactions on Games, 12(3), 270–280. https://doi.org/10.1109/TG.2019.2921979

Inglis, S. 2012. Music and Sound Effects for Videogame Mass Effect 3: Interview, Rob Blake (Bioware). Sound on Sound, June.

Jackson, L. (2011). *Nobuo Uematsu: Interview with a Legendary Video Game Composer.* G4tv, September 9, 2011.

Jenkins, H. (2005). *Games, the New Lively Art,* in Jeffrey Goldstein (ed.) Handbook for Video Game Studies. Cambridge: MIT Press, 2005.

Kastbauer, Damian. (2011). Audio Implementation Greats #10: Made for the Metronome. Designing Sound, January 3, 2011.

Larson, K., Hedges, J., & Mayer, C. (2010). An adaptive, generative music system for games. In Game Developers Conference.

Leman, M. (2008). *Embodied Music Cognition and Mediation Technology*. Cambridge: MIT Press.

Louchart, S., Truesdale, J., Suttie, N., & Aylett, R. (2015). *Emergent Narrative, Past, Present and Future of an Interactive Storytelling Approach*. Interactive Digital Narrative: History, Theory and Practice, 185–199.

Morton, M. (2010). *Howard shore discusses the passion-play within the twilight saga: Eclipse score*. [Online]. Available: https://bit.ly/2SpYWst

Pham, A. 2008. Their Scores Can Be Huge. Los Angeles Times, December 8.

Pichlmair, M., & Kayali, F. (2007). *Levels of Sound: On the Principles of Interactivity in Music Video Games*.

Plut, C., & Pasquier, P. (2020). Generative music in video games: State of the art, challenges, and prospects. Entertainment Computing, 33, 100337. https://doi.org/10.1016/j.entcom.2019.100337

Przybylski, A. K., Weinstein, N., Ryan, R. M., & Rigby, C. S. (2010). Having to versus wanting to play: Background and consequences of harmonious versus obsessive engagement in video games. Cyberpsychology, Behavior, and Social Networking, 13(6), 581-584.

Ruberg, B. (2020). *Empathy and Its Alternatives: Deconstructing the Rhetoric of "Empathy" in Video Games*, Communication, Culture and Critique, 13(1), pp. 54–71. doi: 10.1093/ccc/tcz044.

Salen, K., & Zimmerman, E. (2003). *Rules of Play: Game Design Fundamentals*. MIT Press.

Schafer, R. M. (1969). *The New Soundscape*. Toronto: Berandol.

Sears, A., & Jacko, J. A. (Eds.). (2009). *Human-computer interaction: Designing for diverse users and domains*. CRC Press.

Sundar, S. S., Jia, H, Waddell, T. F., & Huang, Y. (2015) *Toward a theory of interactive media effects* (time), The Handbook of the Psychology of Communication Technology. Hoboken, NJ, USA: Wiley, pp.4

Sweet, M., 2014. *Writing interactive music for video games*. Upper Saddle River: Pearson Education, Inc.

Thomas, C., 2016. *Composing music for games: the art, technology and business of video game scoring*. Boca Raton: CRC Press.

Audiovisual References

AI and Games. (2021, May 30). *How Forza's Drivatar Actually Works* | AI and Games [Video]. YouTube. https://www.youtube.com/watch?v=JeYP9eyIl4E

Allegory (PC version) [Video game]. (2022). Hartwell, D., Malinga, R., Prinz, E. University of Witwatersrand https://rodwinmalinga.itch.io/allegory

Dance Dance Revolution [Video game]. (1998). Japan: Konami. Konami Digital Entertainment.

Destiny 2 (PC version) [Video game]. (2017). North America: Activision.

Detroit: Become Human (PlayStation 4 version) [Video game]. (2018). Quantic Dream, Sony Interactive Entertainment.

Final Fantasy XV (PlayStation 4 version) [Video game]. (2016). Japan: Square Enix.

Guitar Hero (2005). [Video game]. RedOctane. Activision.

Gwent (PC version) [Video game]. (2016). Warsaw, Poland: CD Projekt RED.

Her Story (PC version) [Video game]. (2015). Sam Barlow.

Hi-Fi Rush (PC version) [Video game]. (2023). Tango Gameworks. Bethesda software.

Hocking, C. (2011). *Dynamics: The State of the Art*. [video] Available at: https://www.gdcvault.com/play/1014597/Dynamics-The-State-of-the [Accessed 5 June 2021].

Hudson, K. (2011) *Player-Driven Stories: How Do We Get There?* [Video]. GDC. https://www.youtube.com/watch?v=qie4My7zOgI

Iwamoto, S. (2017, June). *Epic and Interactive Music in Final Fantasy XV* [Video]. GDC. https://www.youtube.com/watch?v=jtoyunrO62I&list=PLGpkw3fwGgX6UvGihXZ_dZ3w_pexF7_Z0&index=18&t=1891s

Johnson, P., Lewin, S., & Sechrist, M. (2018) *Destiny 2: Linear Music in an Open World* [Video]. GDC. https://www.youtube.com/watch?v=kGlhtQzi-HA&t=748s

Little Big Planet (PlayStation 3 version) [Video game]. (2008). North America: Media Molecule, Sony Computer Entertainment.

No Straight Roads [Video game]. (2020). Hazmer, D., Daim, M., & Metronomik. Sold Out Ltd.

Rez Infinite (PC version) [Video game]. (2016). Enhance, Resonair, Monstars,

SingStar (2004). [Video game]. London, UK: Sony Computer Entertainment Europe.

Super Mario Kart (Super Nintendo Entertainment System version) [Video game]. (1992). Japan: Shigeru Miyamoto, Nintendo.

Telling Lies (PC version) [Video game]. (2019). Sam Barlow, Half Mermaid, Furious Bee Limited, Annapurna Interactive.

Tetris Effect (PC version) [Video game]. (2018). Enhance, Monstars, Resonair, Sony Interactive Entertainment, Stage Games.